

# Advances in Science, Technology & Engineering Systems Journal



VOLUME 5-ISSUE 4|JUL-AUG 2020

## EDITORIAL BOARD

### Editor-in-Chief

**Prof. Passerini Kazmerski**  
University of Chicago, USA

### Editorial Board Members

**Prof. Rehan Ullah Khan**  
Qassim University, Saudi Arabia

**Prof. María Jesús Espinosa**  
Universidad Tecnológica  
Metropolitana, Mexico

**Dr. Hongbo Du**  
Prairie View A&M University, USA

**Dr. Nguyen Tung Linh**  
Electric Power University,  
Vietnam

**Tariq Kamal**  
University of Nottingham, UK  
  
Sakarya University, Turkey

**Dr. Mohmaed Abdel Fattah  
Ashabrawy**  
Prince Sattam bin Abdulaziz  
University, Saudi Arabia

**Mohamed Mohamed Abdel-  
Daim**  
Suez Canal University, Egypt

**Dr. Omeje Maxwell**  
Covenant University, Nigeria

**Prof. Majida Ali Abed Meshari**  
Tikrit University Campus, Iraq

**Dr. Heba Afify**  
MTI university, Cairo, Egypt

### Regional Editors

**Dr. Hung-Wei Wu**  
Kun Shan University, Taiwan

**Dr. Maryam Asghari**  
Shahid Ashrafi Esfahani, Iran

**Dr. Shakir Ali**  
Aligarh Muslim University, India

**Dr. Ahmet Kayabasi**  
Karamanoglu Mehmetbey  
University, Turkey

**Dr. Ebubekir Altuntas**  
Gaziosmanpasa University,  
Turkey

**Dr. Sabry Ali Abdallah El-Naggar**  
Tanta University, Egypt

**Aamir Nawaz**  
Gomal University, Pakistan

**Dr. Gomathi Periasamy**  
Mekelle University, Ethiopia

**Dr. Walid Wafik Mohamed Badawy**  
National Organization for Drug Control  
and Research, Egypt

**Dr. Abhishek Shukla**  
R.D. Engineering College,  
India

**Abdullah El-Bayoumi**  
Cairo University, Egypt

**Ayham Hassan Abazid**  
Jordan university of science and  
technology, Jordan



## Editorial

**A**dvances in Science, Technology and Engineering Systems Journal (ASTESJ) is an online-only journal dedicated to publishing significant advances covering all aspects of technology relevant to the physical science and engineering communities. The journal regularly publishes articles covering specific topics of interest.

Current Issue features key papers related to multidisciplinary domains involving complex system stemming from numerous disciplines; this is exactly how this journal differs from other interdisciplinary and multidisciplinary engineering journals. This issue contains 98 accepted papers in Network and Information system domain.

**Editor-in-chief**

*Prof. Passerini Kazmersk*

# ADVANCES IN SCIENCE, TECHNOLOGY AND ENGINEERING SYSTEMS JOURNAL

Volume 5 Issue 4

July-August 2020

## CONTENTS

<i>Offline Signature Recognition and Verification Using ORBKey Point Matching Techniques</i>	01
Aravinda Chickmaglore Venkataramu, Atsumi Masahiko, Akshaya, Amar Prabhu Gurupura, Udaya Kumar Reddy Kyasambally Rajashekar	
<i>A Novel Representative k-NN Sampling-based Clustering Approach for an Effective Dimensionality Reduction-based Visualization of Dynamic Data</i>	08
Dharamsotu Bheekya, Kanakapodi Swarupa Rani, Salman Abdul Moiz, Chillarige Raghavendra Rao	
<i>Citizen Behavior: The Evaluation of Complaint Application that Connected to Smart City</i>	24
Evaristus Didik Madyatmadja, Hendro Nindito, Debri Pristinella	
<i>Ontologic Design of Software Engineering Knowledge Area Knowledge Components</i>	30
Bulat Kubekov, Anar Utegenova, Leonid Bobrov, Vitaliy Naumenko, Aibek Ibraimkulov	
<i>Adapting to Individual Differences (ATID) For Inductive Thinking and Learning Purpose</i>	35
Nia Rohayati, Deni Darmawan	
<i>A Mobile Application Design to Prevent Criminal Acts in Lima, Peru</i>	40
Alexi Delgado, Enrique Lee Huamaní, Elizabeth Jenny Cortez-De La	
<i>Selection of Top-Notch Global Circulation Models that Properly Simulate the Climate Change of Egypt Based Upon the Spatial Assessment Metrics</i>	Withdrawn
Khaled Kheireldin, Mahmoud Roushdi, Mostafa Aboelkhear	
<i>Energy Recovery Equipment and Control Strategies in Various Climate Regions</i>	47
Rand Talib, Alexander Rodrigues, Nabil Nassif	
<i>Fraud Detection Call Detail Record Using Machine Learning in Telecommunications Company</i>	63
Ma'shum Abdul Jabbar, Suharjito	
<i>Robust Static Output-Feedback Fault Tolerant Control for a Class of T-S Fuzzy Systems using Adaptive Sliding Mode Observer Approach</i>	70
Slim Dhahri, Essia Ben Alaia	

<i>Maximum Power-Point Tracking and Stall Control with Eddy Current Brake System on Small-Scaled Wind Turbines and its Application on Agricultural Harvesting</i>	81
Anupa Koswatta, Faramarz Alsharif, Yasushi Shiroma, Shiro Tamaki, Junji Tamura	
<i>Multi-Agent Data Recognition System Based on Received Signal in Antenna on Board Telecom System</i>	94
Chafaa Hamrouni	
<i>Smart Grid Users Benefits Based on DSM Algorithm Mathematical Optimization Problems Studied</i>	99
Chafaa Hamrouni	
<i>Contribution in Private Cloud Computing Development based on Study and KPI Analysis</i>	105
Chafaa Hamrouni, Slim Chaoui	
<i>Optimization of Dual Motion Mechanism with Double Grooved Cams for High-voltage Gas Circuit Breaker</i>	109
Masanao Terada, Yuki Nakai, Hiroaki Hashimoto, Daisuke Ebisawa, Hajime Urai, Yasunobu Yokomizu	
<i>Design of an EEG Acquisition System for Embedded Edge Computing</i>	119
Kanishk Rai, Keshav Kumar Thakur, Preethi K Mane, Narayan Panigragi	
<i>Composition of Methods to Ensure Iris Liveness and Authenticity</i>	130
Ali Al-Rashid	
<i>A Survey on 3D Hand Skeleton and Pose Estimation by Convolutional Neural Network</i>	144
Van-Hung Le, Hung-Cuong Nguyen	
<i>Nearest Neighbour Search in k-dSLst Tree</i>	160
Meenakshi Hooda, Sumeet Gill	
<i>Modelling and Simulation of Aerodynamic Parameters of an Airship</i>	167
Anoop Sasidharan Pillai, Venkata Ramana Murthy Oruganti	
<i>Industry 4.0 Operators: Core Knowledge and Skills</i>	177
Olayan Alharbi	
<i>Contextualization of the Augmented Reality Quality Model through Social Media Analytics</i>	184
Jim Scheibmeir, Yashwant Malaiya	
<i>Cause of Time and Cost Overruns in the Construction Project in Nepal</i>	192
Bishnu Prasad Khanal, Sateesh Kumar Ojha	



<i>High-Order Thinking Skills: The Educational Treasure Hunt Game</i> Yogi Udjaja, Sasmoko, Jurike V. Moniaga, Kevin Zulfian Bay	196
<i>Entertainment Technology: Dynamic Game Production</i> Yogi Udjaja	203
<i>Malware Classification Based on System Call Sequences Using Deep Learning</i> Rizki Jaka Maulana, Gede Putra Kusuma	207
<i>The Importance of Sustainability Audit Report in Go Public Companies Sector, in Indonesia</i> Bambang Leo Handoko, Ang Swat Lin Lindawati	217
<i>Dynamic Decision-Making Process in the Opportunistic Spectrum Access</i> Mahmoud Almasri, Ali Mansour, Christophe Moy, Ammar Assoum, Denis Lejeune, Christophe Osswald	223
<i>Nonlinear <math>1_{2,p}</math>-norm based PCA for Anomaly Network Detection</i> Amal Hadri, Khalid Chougali, Raja Touahni	234
<i>Buffering Supercapacitor Mechanism based on Bidirectional DC/DC Converter for Mini All-Terrain Vehicle Application</i> Syifaul Fuada, Braham Lawas Lawu, Bommegowda Kabbala Basavarajappa	244
<i>Utilization of Data Mining to Predict Non-Performing Loan</i> Yosaphat Catur Widiyono, Sani Muhamad Isa	252
<i>Diffusion of Technology for Language Challenges in the South African Healthcare Environment</i> Phathutshedzo Makovhololo, Tiko Iyamu	257
<i>Exploring the Performance Characteristics of the Naïve Bayes Classifier in the Sentiment Analysis of an Airline's Social Media Data</i> Mba Obasi Odim, Adewale Opeoluwa Ogunde, Bosede Oyenike Oguntunde, Samuel Ayodele Phillips	266
<i>Coastal Risk Modelling for Oil Spill in The Mediterranean Sea</i> Abdellatif Soussi, Chiara Bersani, Roberto Sacile, Dounia Bouchta, Ahmed El Amarti, Hamid Seghioeur, Driss Nachite, Jaouad Al Miys	273
<i>Optimization of the Procedures for Checking the Functionality of the Greek Railways: Data Mining and Machine Learning Approach to Predict Passenger Train Immobilization</i> Ilias Kalathas, Michail Papoutsidakis, Chistos Drosos	287

<i>Cluster Centroid-Based Energy Efficient Routing Protocol for WSN-Assisted IoT</i>	296
Nalluri Prohess Raj Kumar, Josemin Bala Gnanadhas	
<i>Deep Learning Model for A Driver Assistance System to Increase Visibility on A Foggy Road</i>	314
Samir Allach, Mohamed Ben Ahmed, Anouar Abdelhakim Boudhir	
<i>Logistics Solutions in the Preparation Phase for the Appearance of Disasters</i>	323
Erika Barojas-Payán, Diana Sánchez-Partida, Miguel-Josué Heredia-Roldan, Victorino Juárez-Rivera, Jesús Medina-Cervantes	
<i>Distributed Microphone Arrays, Emerging Speech and Audio Signal Processing Platforms: A Review</i>	331
Shahab Pasha, Jan Lundgren, Christian Ritz, Yuexian Zou	
<i>Computational Intelligence and Statistical Learning Performances on Predicting Dengue Incidence using Remote Sensing Data</i>	344
Nittaya Kerdprasop, Kittisak Kerdprasop, Paradee Chuaybamroong	
<i>Design and Optimization of a Three Stage Electromechanical Power Unit using Numerical Methods</i>	351
Yashwant Kolluru, Rolando Doelling, Lars Hedrich	
<i>Bilateral Communication Device for Deaf-Mute and Normal People</i>	363
Raven Carlos Tabiongan	
<i>Survey Analysis: Enhancing the Security of Vectorization by Using word2vec and CryptDB</i>	374
Hana Yousuf, Said Salloum	
<i>Development of Soil Moisture Monitoring by using IoT and UAV-SC for Smart Farming Application</i>	381
Sarun Duangsuwan, Chakree Teekapakvisit, Myo Myint Maw	
<i>The Role of KM in Enhancing AI Algorithms and Systems</i>	388
Hani AlGhanem, Mohammad Shanaa, Said Salloum, Khaled Shaalan	
<i>Potential of Solar Energy in Residential Rooftop Surface Area in Semarang City, Indonesia</i>	397
Djoko Adi Widodo, Purwanto Purwanto, Hermawan Hermawan	
<i>Keyword Driven Image Description Generation System</i>	405
Sreela Sreekumaran Pillai Remadevi Amma, Sumam Mary Idicula	

<i>Tidal Propagation Based On Co-Phase Chart and Co-Range Chart in Sunda Strait, Indonesia</i>	412
Denny Nugroho Sugianto, Harjo Susmoro, Khoirol Imam Fatoni, Virginia Stephanie Claudia, Haris Djoko Nugroho	
<i>Fuzzy-logical Control Models of Nonlinear Dynamic Objects</i>	419
Siddikov Isamidin Xakimovich, Umurzakova Dilnoza Maxamadjonovna	
<i>A Framework of E-Procurement Technology for Sustainable Procurement in ISO 14001 Certified Firms in Malaysia</i>	424
Pratik Kumar Singh, Fadillah Binti Ismail, Chan Shiau Wei, Muhammad Imran, Syed Ashfaq Ahmed	
<i>Effects of Oversampling SMOTE in the Classification of Hypertensive Dataset</i>	432
Nurhafifah Matondang, Nico Surantha	
<i>A CNN-based Differential Image Processing Approach for Rainfall Classification</i>	438
Roberta Avanzato, Francesco Beritelli	
<i>Artificial Intelligence Approach for Target Classification: A State of The Art</i>	445
Maroua Abdellaoui, Dounia Daghouj, Mohammed Fattah, Younes Balboul, Said Mazer, Moulhime El Bekkali	
<i>Assessing the Operator's Readiness to Perform Tasks of Controlling by the Unmanned Aerial Platforms</i>	457
Dmytro Kucherov, Olha Sushchenko, Andrii Kozub, Volodymyr Nakonechnyi	
<i>Short-Term Dynamic Exchange Rate Model: IFEER Concept Development</i>	463
Anton Kuzmin	
<i>The Sound of Trust: Towards Modelling Computational Trust using Voice-only Cues at Zero-Acquaintance</i>	469
Deborah Ooi Yee Hui, Syaheerah Lebai Lutfi, Syibrah Naim, Zahid Akhtar, Ahmad Sufri Azlan Mohamed, Kamran Siddique	
<i>Study Of The Effect Of Abnormalities In The External Ear Inducing Hearing Problems</i>	477
Jihane Melloui, Jamila Bakkoury, Omar Bouattane	
<i>Neural Network-Based Fault Diagnosis of Joints in High Voltage Electrical Lines</i>	488
Marco Bindi, Igor Aizenberg, Riccardo Belardi, Francesco Grasso, Antonio Luchetta, Stefano Manetti, Maria Cristina Piccirilli	



<i>Remote Control of Garden Plantation Water Pumps using Arduino and GSM Mobile</i>	499
Beza Negash Getu, Mohamed Abdulkadir, Michael Tous	
<i>Decision Making System for Improving Firewall Rule Anomaly Based on Evidence and Behavior</i>	505
Suchart Khummanee, Phatthanaphong Chomphuwiset, Potchara Pruksasri	
<i>IoT Based Human Activity Recognition System Using Smart Sensors</i>	516
Deepti Sehrawat, Nasib Singh Gill	
<i>Damage Accumulation Model for Cracked Pipes Subjected to Water Hammer</i>	523
Zakaria Mighouar, Laidi Zahiri, Hamza Khatib, Khalifa Mansouri	
<i>Development of an Adaptive HVAC Fuzzy Logic Controller for Commercial Facilities: A Case Study</i>	531
Hamidi Meryem, Bouattane Omar, Raihani Abdelhadi, Khalili Tajeddine	
<i>Human-Robot Multilingual Verbal Communication – The Ontological knowledge and Learning-based Models</i>	540
Mohammed Qbadou, Intissar Salhi, Hanaâ El fazazi, Khalifa Mansouri, Michail Manios, Vassilis Kaburlasos	
<i>Development Trends of Smart Cities in the Future – Potential Security Risks and Responsive Solutions</i>	548
Vu Nguyen Hoa Hong, Luong Tuan Anh	
<i>Application of EARLYBREAK for Line Segment Hausdorff Distance for Face Recognition</i>	557
Chau Dang-Nguyen, Tuan Do-Hong	
<i>Organizational Agility Assessment of a Moroccan Healthcare Organization in Times of COVID-19</i>	567
Fadoua Tamtam, Amina Tourabi	
<i>A Support Vector Machine Based Technique for Fault Detection in A Power Distribution Integrated System with Renewable Energy Distributed Generation</i>	577
Katlho Moloi, Yskandar Hamam, Jacobus Andries Jordaan	
<i>Overview of Solar Radiation Estimation Techniques with Development of Solar Radiation Model Using Artificial Neural Network</i>	589
Amar Choudhary, Deependra Pandey, Saurabh Bhardwaj	

<i>Customer Satisfaction Recognition Based on Facial Expression and Machine Learning Techniques</i>	594
Moulay Smail Bouzakraoui, Abdelalim Sadiq, Abdessamad Youssfi Alaoui	
<i>Real-Time Traffic Sign Detection and Recognition System for Assistive Driving</i>	600
Adonis Santos, Patricia Angela Abu, Carlos Oppus, Rosula Reyes	
<i>Performance Analysis of Go-Back-N ARQ Protocol Used In Data Transmission Over Noisy Channels</i>	612
Fayza Ahmed Nada	
<i>Modeling and Transformation from Temporal Object Relational Database into MongoDB: Rules</i>	618
Soumiya Ain El Hayat, Mohamed Bahaj	
<i>Modelling and Simulation of 3-DOF Lower Limb Rehabilitation Robot using Force Feed Forward Control</i>	626
Pham Van Bach Ngoc, Le Thi Hong Gam, Dam Hai Quan, Bui Trung Thanh, Nguyen Luong Thien	
<i>Earthquake Response of Multi-Storey Infilled Reinforced Concrete Buildings</i>	633
Miloud Mouzzoun, Abdelkader Cherrabi	
<i>Machine Learning for Network Intrusion Detection Based on SVM Binary Classification Model</i>	638
Anouar Bachar, Nouredine El Makhfi, Omar EL Bannay	
<i>COVID-19 Crises: Global Economic Shocks vs Pakistan Economic Shocks</i>	645
Naveeda K. Katper, Muhammad Nawaz Tunio, Niaz Hussain, Attaullah Junejo, Faheem Gul Gilal	
<i>Social-Interactive Learning Concept Used for Electronic Educational Resource "Post-Graduate Foreign Language" and the Obtained Learning Curve</i>	655
Natalya Chernova, Victor Chernov, Margarita Emelianova, Raisa Akhunzianova, Danil Sukhopluev	
<i>Correlation-Based Incremental Learning Network for Gas Sensors Drift Compensation Classification</i>	660
Panida Lorwongtrakool, Phayung Meesad	
<i>Design of a Flapping Wings Butterfly Robot based on Aerodynamics Force</i>	667
Kanjanapan Sukvichai, Kan Yajai	

<i>Clustering of Mindset towards Self-Regulated Learning of Undergraduate Students at the University of Phayao</i> Pratya Nuankaew	676
<i>Fuzzy Recognition by Logic-Predicate Network</i> Tatiana Kosovskaya	686
<i>Deep Learning Approach for Automatic Topic Classification in an Online Submission System</i> Tran Thanh Dien, Nguyen Thanh-Hai, Nguyen Thai-Nghe	700
<i>Bathtub-Shape Failure Rate Life Time Model in a New Version with the use of Bayesian Prediction Bounds for the Presence of Outliers</i> Ayed Rheal A. Alanzi	710
<i>Interpolatory Projection Techniques for H2 Optimal Structure-Preserving Model Order Reduction of Second-Order Systems</i> Md. Motlubar Rahman, Mohammad Monir Uddin, Laek Sazzad Andallah, Mahtab Uddin	715
<i>Towards Classification of Shrimp Diseases Using Transferred Convolutional Neural Networks</i> Nghia Duong-Trung, Luyl-Da Quach, Chi-Ngon Nguyen	724
<i>Determination of ERP Readiness Assessment Using Agile Parameters: A Case Study</i> Santo Fernandi Wijaya, Harjanto Prabowo, Ford Lumban Gaol, Meyliana	733
<i>A Review on Cross-Layer Design Approach in WSN by Different Techniques</i> Sultana Parween, Syed Zeeshan Hussain	741
<i>Design and Implementation of a PLC Trainer Workstation</i> Matthew Oluwole Arowolo, Adefemi Adeyemi Adekunle, Martins Oluwaseun Opeyemi	755
<i>An Explanatory Review on Cybersecurity Capability Maturity Models</i> Adamu Abdullahi Garba, Maheyzah Muhamad Siraj, Siti Hajar Othman	762
<i>A Method for Detecting Human Presence and Movement Using Impulse Radar</i> Young-Jin Park, Hui-Sup Cho	770
<i>Ethics as a Motivation Indicator in Second Language Vocational Digital Teaching</i> Natalya Viktorovna Matveeva, Ludmila Vladimirovna Makar	776



<i>The Effects of Transmission Power and Modulation Schemes on the Performance of WBANs in on-Body Medical Applications</i> Marwa Boumaiz, Mohammed El Ghazi, Mohammed Fattah, Anas Bouayad, Moulhime El Bekkali	783
<i>Non-Ridership Presumption Toward New Bus Rapid Transit Purwokerto-Purbalingga Executio</i> Fauzan Romadlon	795
<i>Mentoring Model in an Active Learning Culture for Undergraduate Projects</i> Wongpanya Nuankaew, Kanakarn Phanniphong, Sittichai Bussaman, Direk Teeraputon, Praty Nuankaew	805
<i>Performance Analysis of Grid-Connected PV Rooftop, at Sakon Nakhon Province, Thailand</i> Supalak Sathiracheewin, Patamaporn Sripadungtham, Settakorn Kamuang	816
<i>Fine Tuning the Performance of Parallel Codes</i> Sanaz Gheibi, Tania Banerjee, Sanjay Ranka, Sartaj Sahni	824
<i>Generic Decision Support Model for Determining the Best Marketer</i> Ditdit Nugeraha Utama, Sherly Oktafiani	841
<i>Design of Aerial Panoramic Photography: Contrast between Industrialized and Natural Zones</i> Wei Zhong Feng, Yu-Che Huang, Fang-Lin Chao	849

# Offline Signature Recognition and Verification Using ORB Key Point Matching Techniques

Aravinda Chickmaglore Venkataramu<sup>\*1</sup>, Atsumi Masahiko<sup>2</sup>, Akshaya<sup>1</sup>, Amar Prabhu Gurupura<sup>1</sup>, Udaya Kumar Reddy Kyasambally Rajashekar<sup>1</sup>

<sup>1</sup>Department of Computer Science and Engineering, N.M.A.M Institute of Technology, Nitte 574110, India

<sup>2</sup>Department of Electronic and Computer Engineering, Ritsumeikan University, Kusatsu, 5258577, Japan

---

## ARTICLE INFO

Article history:

Received: 08 May, 2020

Accepted: 17 June, 2020

Online: 06 July, 2020

---

Keywords:

Image matching

Robust independent elementary features (BRIEF)

Oriented FAST, Rotated BRIEF (ORB).

---

## ABSTRACT

An extensive work has been carried out in the field of human transcribe-verification and transcribe-recognition by extensive scholars across the globe from past decades. In order to demeanour immense experiments for considering the performance of the newly intended models and to substantiate the efficacy of the proposed model which is moderately required. This paper monologue the problem of signature-verification and recognition using diverse ORB key points and Convolution Neural Network. This method reckons on descriptors for detection and matching. The systematic approach is tested more on an few real and few fake signatures. Many features and combination of features were proposed for signature substantiation and acknowledgement. Numerous experiments are conducted to determine the capability of the proposed models in selective genuine and forgery signature. In this context, a large signature corpus comprising of 29950 offline signatures from 605 persons is created during the course of the research work. Finally the achievement was achieved about 90% of correct accuracy of the given original signature.

## 1 Introduction

In modern society there is a greater demand for reliable methods of personal authentication because authentication is a crucial aspect of our daily life activities such as document authentication, having access to computing facilities or information, seeking entry to secured zone, etc. These activities require fast and effective automatic personal authentication systems. Existent classes in handwritten signatures: 1) Online/dynamic signatures 2) Offline/static signatures. A specially designed electronic device like digitizing tablet is used to record online-signature along with features like position, pressure, elevation etc., whereas, an offline signature is a digitized signature image taken using electronic devices like scanner, camera etc.[1]-[3] During the present day situation the digital images play a very basic role in many applications and have become a important source of information. As we all knew that feature detection and feature matching will plays the vital role in image processing applications, like image retrieval, object detection and many more. The offline signature and verification are related to probing the components of the signature into integral elements for extracting features by matching techniques is as shown in the Figure 1. The SURF technique

is a feature finder which estimate of SIFT and performs faster than SIFT without reducing the quality of identified points[4]-[5]. These two techniques are thus based on descriptor and a detector. One more well-known technique BRIEF is an alternative for SIFT which requires less complexity with similar matching performance. Rublee et al. proposed Oriented FAST and Rotated BRIEF (ORB) as another efficient alternative for SIFT and SURF. In this paper the main concentration was on the ORB feature rather than SIFT and SURF. The experiment was carried out for the robustness of the signature key point techniques against rotation, scaling, and deformity due to horizontal or vertical shears.

## 2 Related work carried out so far

As we all knew that signature verification and recognition are considered to be most complex and challenging problem in pattern recognition area by virtue of large intraclass divergence between individual and less interclass deviations between signatures of individuals in toady's' global field.

---

\*Aravinda C.V, N.M.A.M Institute of Technology Nitte INDIA , +91-9742260265 & aravinda.cv@nitte.edu.in,0077ev@ed.ritsumei.ac.jp

## 2.1 Template Matching

This is one of the oldest analogous artistry used for pattern-matching. This technique is used for a signature class denoted by a template and is being compared with the original to establish its authentication. As referred in the "Dynamic time warping", (DTW) is the materialistic technique used to establish a match between a test signature and a template[6].

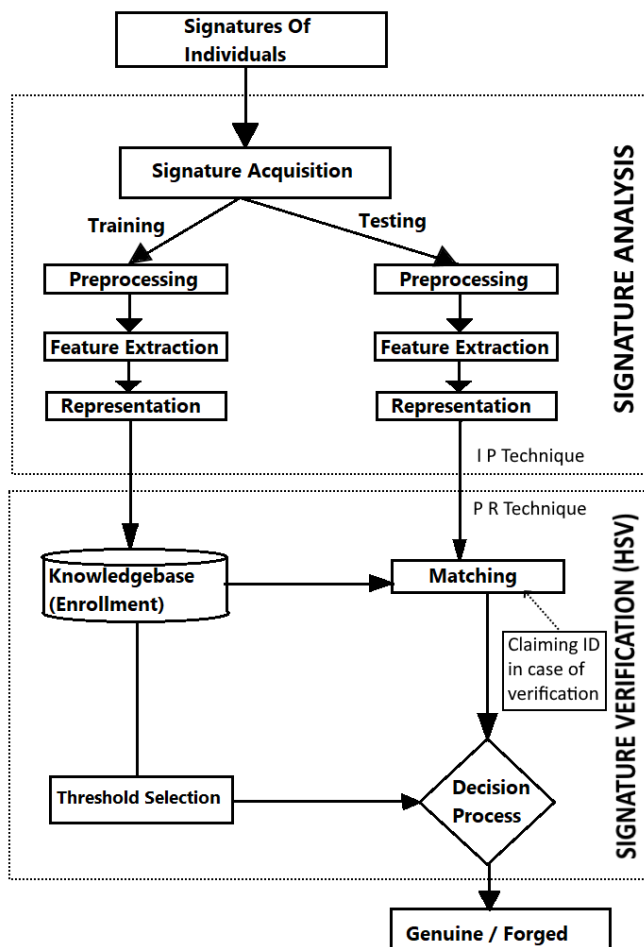


Figure 1: Depiction of HSVR system

## 2.2 Distance Classifiers

By using this technique such as Euclidean distance experiments were carried out for tracing the solutions and considered this technique by taking comprehensive features such as "heights", "width", "slant angles", "maximum horizontal projection shift as measure of similarity". Fang et al. used Mahalanobis distance for computing similarity among the local and global feature vectors. This work used a data set which was consisting of 800 signatures from 20 individuals. In this method the nearest neighbor and a threshold classifier are used[7].

## 2.3 Bayesian Classifiers

As per the author considered the pertinence of the refitted version of Bayesian network for offline. This kind of network is different

from the traditional Bayesian. This network node are splits into two classes namely common and alternative hypothesis, to build the network in tree structure. This captures the conditional probability of every node and topological relations between the components with the nodes of the network. this method uncertainties in structure description between various components which are encoded. The FRR of 20% and the FAR of 14% were reported on a small database of 160 signatures from eight individuals [8].

## 2.4 Support Vector Machines

As per the author measured a Human Signature Verification and Recognition system as 2 class pattern-recognition problem. By this method "DTW technique", "SVM classifier", "PCA" and "Bayes, concepts are used. First the test signature is combined with referred signature using DTW to find authenticity of test signatures. This feature is then classified as authentic or forgery. For this case linear classifier in combination with PCA, Bayes, SVM classifier are used for classification purposes. The FRR of 1.64% and the FAR of 3.85% are obtained on a dataset of 306 signatures.

## 3 Overview of Feature matching techniques

### 3.1 SIFT

The SIFT mainly projected for solving the image rotation, transformations and opinion point in matching features, as per the author Lowe. The main key functionality is as follows (a) Evaluating the scale space using Gaussian techniques (b) point localization (c) orientation (d) generator.

### 3.2 SURF

The SURF is nearly alike to DoG features. The replacement of Gaussian average, the squares are used for approximation. It uses Binary Large Object detector based on Hessian matrix to find the point of interest and wavelet response for feature description as mentioned [9]-[10].

### 3.3 ORB

The ORB is an amalgamation of FAST key point and BRIEF descriptor with minor changes. The main technique of FAST is to identify the key points along with Harris corner to find top N points. This will not figure out the orientation and rotation variant but computes the weighted centroid of the located point. The resultant gives the corner point to centroid and this leads to orientation. The ORB rotation matrix is calculated using the orientation of patch and BRIEF descriptors are controlled to orientation[11].

### 3.4 Feature matching and representation

Even though there has been widely published papers and solved problem based on this offline signature, still there is lot of research problems which is still being carried out for analysis purpose. The

vital issue of this analysis is to resolve the most pertinent features. Special features like characterization of signatures based on feature and formulate the corresponding between the insinuation and test scribbles are the significant issue in authentication[12]-[13]. This study mainly relies on the effectual representation of signatures and learning similar features. These features are derived from geometrical properties for offline signatures as shown in the Figure 2.

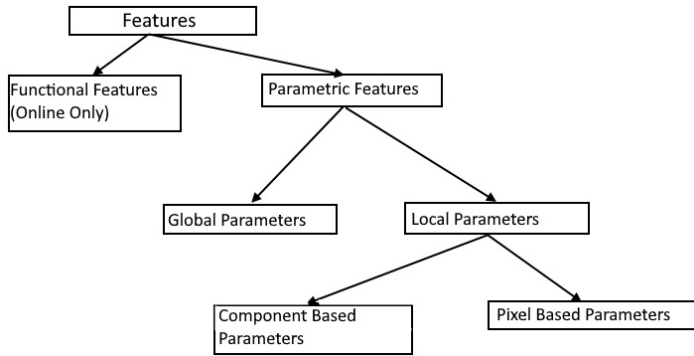


Figure 2: Organization of Offline and Online signature

## 4 Proposed Technique

In the proposed method the process of verification and recognition is carried on a huge database which consists of both genuine and forgery signatures. These signature samples features are extracted from a few samples of original specimen and rest original samples and forgery samples are used for verification process. The flow of the process is as shown in the Figure 3 The steps carried is as follows

1. CNN classifier is used to recognize which person the signature belongs to
2. If the recognition threshold is  $>80\%$  feature matching technique is used, else it will be rejected.
3. Extracting the signature part from signature database and resize the same
4. Calculating the keypoints using ORB technique.
5. Comparing the key points with original sample within database.
6. If the number of matches between the given sample and the original image is  $>$  than the average number of matches between the original signatures it will be accepted else rejected.
7. Carried out the experiment for matching and classifying the genuine and forgery.

## 5 Feature Detection And Matching Using ORB

The most relevant for problem solving the task relating to application as mentioned in the paper [14]. The offline signature samples

specific structures like point, edges were considered. The resultant neighbourhood operation is applied on the sample. The ORB featured was measured for the experimental purpose, whereas the SIFT performed better detection on the same problem. This builds well-known "Features from accelerated segment test", "Features from accelerated segment test", key-point detector and "Binary Robust Independent Elementary Features descriptor".

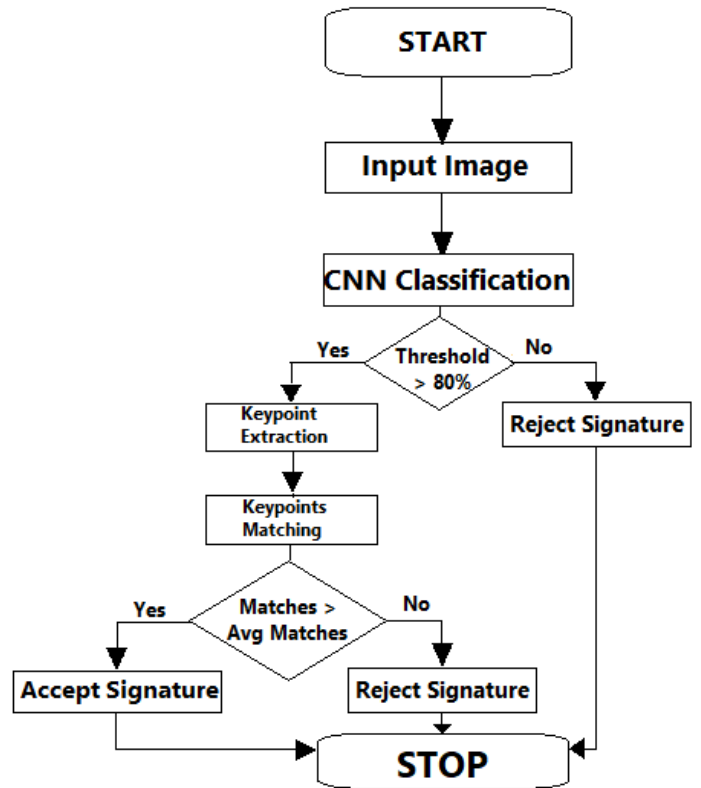


Figure 3: Architecture flow of Human Signature Verification Recognition

"Features from accelerated segment test", is a corner detection method, which could be used to extract feature points is as shown in Figure 5. The "Features from accelerated segment test", Algorithm applied on the signature sample to find the key points is as follows

1. Chosen a pixel  $p$  on the signature sample which is to identified as key point or not a key point. This intensity be  $I_p$ .
2. Applied the appropriate threshold value  $t$ .
3. After this the pixel  $p$  was considered as a corner set of  $n$  contiguous pixels in the circle and then all the brighter pixels than  $I_p + t$  was considered.
4. To improve the accuracy of the algorithm, first the intensity of the pixel was compared 1,5,9,13 of the circle with  $I_p$ . Finally all least three pixels out of these four pixels were considered because of satisfying the threshold criterion.
5. The least three of the four pixel values  $I_1, I_5, I_9, I_{13}$  were not found above nor below the  $I_p + t$ , then  $p$  is not an interest point was discarded.

6. These steps were repeated to all the pixels in the signatures. The BRIEF performance is used to find the key points which was recognized by the prior FAST algorithm where this method was considered for processing and converting this into binary feature vector to epitomize the image of the given sample [15]-[17]. This consist of 0 and 1 were every key-point is designated by a feature vector which is 128-512-bit string and the selected random pair of pixels by default was identified as neighbourhood key point. The discern neighbourhood around pixel is known as a patch, which is a square of some pixel width and height. The first pixel in the random pair was drawn from Gaussian and centred around the key point with standard deviation. This creates a vector for each key-point in an signature sample is as shown in the Figure 7.

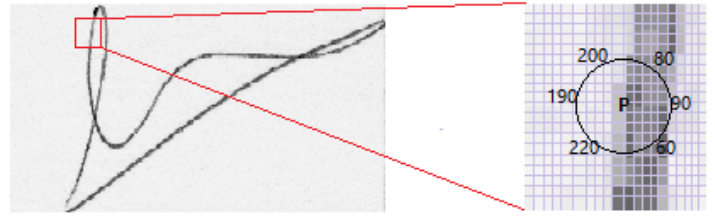


Figure 5: FAST Algorithm work flow

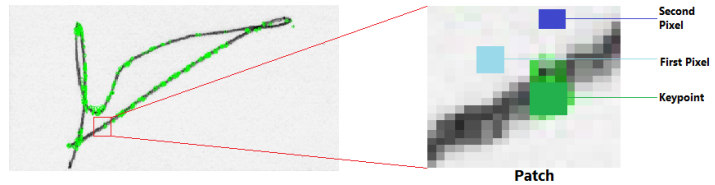


Figure 6: BRIEF Algorithm work flow

**Algorithm 1:** The rBRIEF() algorithm for ORB is as follows

```

Result: Greedy Search
Run each test against all training patches obtained from the
sample;
Order the tests by their distance from a mean of 0.5, forming
the vector T;
Greedy search;
while  $R > threshold$  do
    Put the first test into the result vector R and remove it
    from T;
    Take next test from T and compare with R;
    if discard it then
        If its absolute correlation of R  $\geq$  threshold;
        else;
    else
        add it to R;
    end
    if  $length(R) < 256$  then
        Repeat until length of R = 256 ;
    else
        raise threshold and try again;
    end
end
    
```

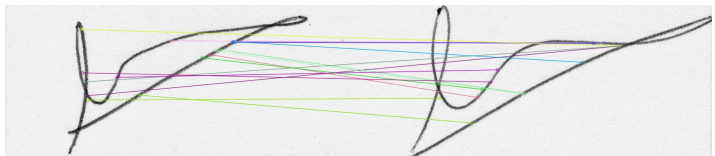


Figure 7: Good matches of keypoints from 2 signature samples

## 6 Experimental Results Conducted

**F.A.R** Choosing the counterfeit signature sample and discerning it is a genuine signature. It is as given by the equation 1

$$\frac{n(n+1)}{(n-1)(n-2)(n-3)} \left( \frac{s^4}{V(x)^2} \right) - 3 \frac{(n-1)^2}{(n-2)(n-3)} \quad (1)$$

**False Acceptance Rate**=Number of Forgery Signature Accepted/Number of Forgery Tested

**False Rejection Rate (F.R.R)** Snubbing the signature though it is frank sampled signature discerning it is a forgery sample. This is given by the equation 2

$$Skewness = \frac{n\sqrt{n-1}}{n-2} \frac{\sum_{i=1}^n (X_i - X_{avg})^3}{n-2 \left( \sum_{i=1}^n left(X_i - X_{avg})^3 \right)^{3/2}} \quad (2)$$

Skewness can also be defined as ratio of the 3rd moment and 2nd moment raised to the power of 3/2 ratio of the 3rd moment and the standard deviation as mentioned in the equation.

3 and 4

$$Skewness = \frac{m_3}{m_2^{3/2}} \quad (3)$$

$$Skewness = \frac{m_3}{\sigma^3} \quad (4)$$

**F.R.R**=Number of Genuine Signature Rejected/Number of Genuine Tested

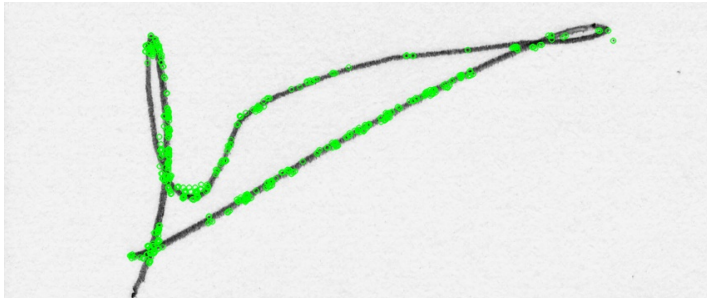


Figure 4: Detected keypoints



Table 1: Online signature dataset for research purpose

Datasets	No. of Signers	No. of Genuine	No. of Forgeries	total no. of signatures
MCYT-75[151]	75	75*15	75*15	2250
GDPS-signature	160	160*24	160*30	8640

Table 2: Summary of DB-605 dataset

Datasets	No. of Individuals	No. of Genuine	No. of Forgeries	total no. of signatures
DB-200	200	200*28	200*28	11200
MCYT-300	330	330*25	330*25	16500
DB-605	605	14975	14975	29950

Table 3: Representation of C.N.N Architecture

Layer(type)	Output-Shape	Parameter
conv1d.1 (Conv1D)	(None, 173, 173, 32)	320
conv1d.2 (Conv1D)	(None, 169, 169, 64)	51264
batch_normalization.1	(Batch (None, 169, 169, 64)	256
activation.1 (Activation)	(None, 169, 169, 64)	0
max_pooling2d.1	(MaxPooling2 (None, 84, 84, 64)	0
global_average_pooling2d.1	( (None, 64)	0
dense.1 (Dense)	(None, 70)	4550

Total-parameters: 56,390  
 Trainable-parameters: 56,262  
 Non-trainable-parameters: 128

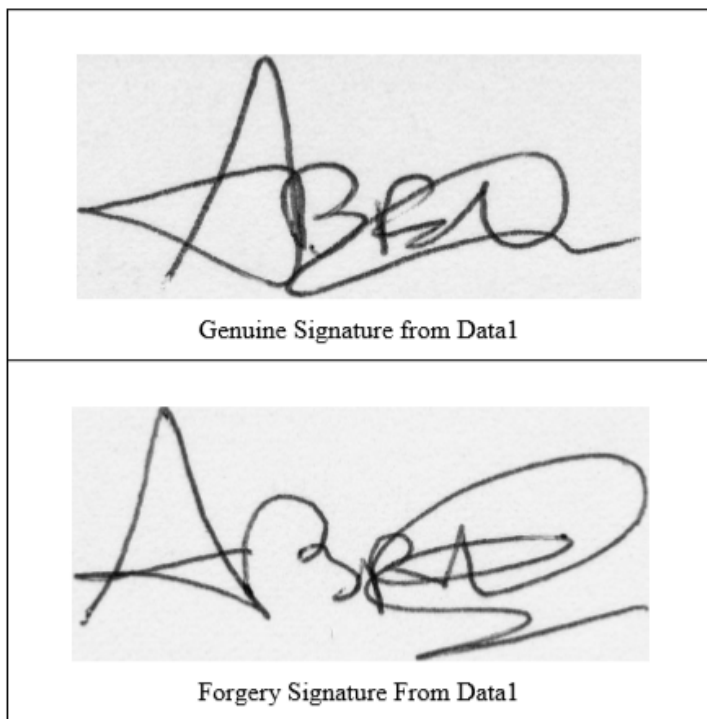


Figure 9: Sample of Feature extraction

### 6.1 Neural Network setup for experiment

**Initialize position:** This is the first stage, where NN is put up by describing the Network architecture (i/p and o/ps) . Further, the additive the number of patterns which will be used to train the network has to be considered along with the number of units in output layers are decided **Generate Training Set:**The quality of cognitive process is deliberated as important to the neural network; this gives an ability to recognize any changes in the signature images of the same person. Many individual learning algorithms have been developed for the verification/recognition purposes.

**Create Neural-Network Set:**In the antecedent period, the signature samples of every persons are divided into training and testing set. The training set are fed into a multilayer network. Multilayer NN is nothing but network with multiple hidden layers. In this work, data flow in the forward direction that is from input layer to output layer.



Figure 8: Images of MCYT-75: (a) Accurate signatures (b) Corresponding fraudulence signatures



Table 4: Summary of datasets used

Sample	FAR	FER	Accuracy
	0	0	100
	0	20	90
	0	20	90
	0	40	80
	0	0	100
	0	20	90

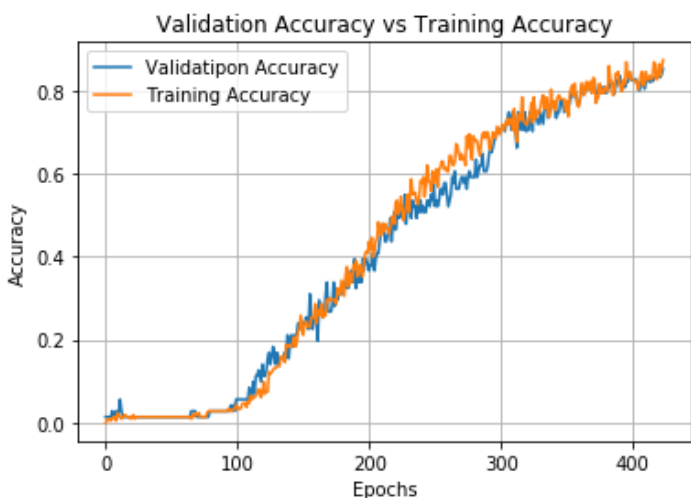


Figure 10: Training Accuracy and Validation Accuracy

In this phase, weights are stated for every construction using the simple weighting equation which is preferred to be forwarded that is, data flows from the input layer to the output layer. This stage specifies the weights which are available on every connection using the simple weighting equation.  $Weight(t + 1) =$

$weight(t) + \eta(\omega_{weight}(t))$  the weight  $t+1$  is equal to the weight at time  $t$  plus the learning rate times the calculated weight change.

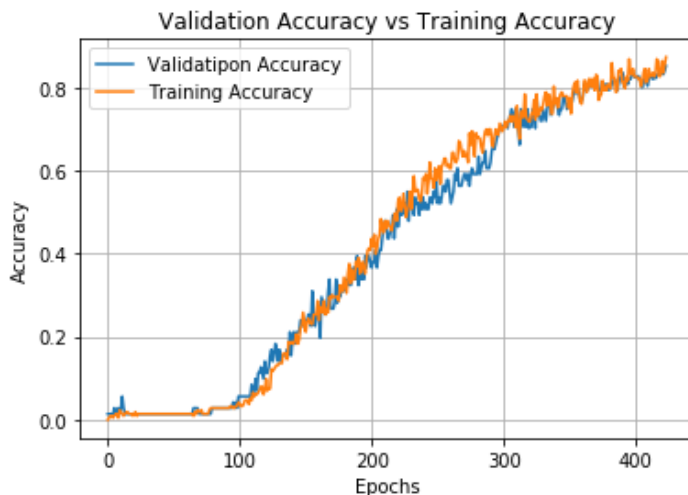


Figure 11: Training Accuracy and Validation Accuracy



Figure 12: Training Loss and Validation Loss

## 7 Conclusion

In the paper Contour based technique was used to classify signature into 2 classes (genuine and fake) the accuracy achieved was around 81% though this method used was fast and robust, the main disadvantage posed by the method was it required 2 classes of each signature and about 15 samples from each class. By applying the Multiple classifiers technique using Hu and Hog features, it was able to achieve an accuracy of around 90%. In this experiment, we set out to find various methods which proved effective for signature classification. Though the method was fast, it posed the same disadvantage of having to use 2 classes per person (genuine and fake) and consider about 15 samples each. This was far from a practical approach as collecting 15 samples for each person was not possible. In this paper, we sought out to find a practical approach

to classify signatures into real and fake considering practicality of the problem. Hence as a practical approach, we collected 3 genuine signatures from each person and used a combination of deep learning and machine learning to classify signatures. The accuracy in this case was found to be around 92% but this method gets rid of the disadvantages faced in the previous 2 cases and makes the classification task suit the real world problem.

**Conflict of Interest** The authors declare no conflict of interest.

**Acknowledgment** We would like to thank to N.M.A.M Institute of Technology, Nitte, INDIA and Ritsumeikhan University Japan, for their support towards this paper.

## References

- [1] T.S. Wilkinson, T.S. Pender, J.W. Goodman, "Use of synthetic discriminant functions for hand-written signature verification". *Appl. Opt.* **30** (23), 33453353, 1991.
- [2] J.E. Ward, R.E. Wendell, "A new norm for measuring distance which yields linear location problems". *Operation Research*, (28) 836844, 1980.
- [3] Xiao, "Signature verification using a modified Bayesian network", *Article in Pattern Recognition* **35**(5) 983-995, May 2002. 10.1016/S0031-32030100088-7
- [4] M.B. Yilmaz, B. Yanikoglu, C. Tirkaz, "Offline signature verification using classifier combination of HOG and LBP features- 2011 international joint 2011. DOI: 10.1109/IJCB.2011.6117473
- [5] David G. Lowe, "Distinctive Image Features from Scale-Invariant Key-Points", *International Journal of Computer Vision*, **60**(2), pp. 91-110, 2009.
- [6] M. Ramakrishna, S.S. Shylaja, "Is ORB efficient Over Surf for Object Recognition?", *International Journal of Advanced Research in Science Engineering and Technology*, 2014, **3**(8), pp. 2783-2788, ISSN: 2278-1323
- [7] J. Li, X. Li, N. Yang, Xingming, "Segmentation Based Image Copy-Move Forgery Detection Scheme", *IEEE Transactions on Information Forensics and Security*, 2015, **10**(1), pp.507-518, 2015. DOI: 10.1109/TIFS.2014.2381872
- [8] D.G. Lowe, "Distinctive Image Features from Scale-Invariant Key-Points", *International Journal of Computer Vision*, **60**(2), pp.91-110, 2009. DOI <https://doi.org/10.1023/B:VISI.0000029664.99615.94>.
- [9] B. Moghaddam, C. Nastar and A. Pentland, "A Bayesian similarity measure for deformable image matching", *Image and Vision Computing*, **19**(5), pp. 235-244, 2001. [https://doi.org/10.1016/S0262-8856\(00\)00059-7](https://doi.org/10.1016/S0262-8856(00)00059-7)
- [10] M. Gzel, "A Hybrid Feature Extractor using Fast Hessian Detector and SIFT", *Technologies*, **3**(2), pp. 103-110, 2015. <https://doi.org/10.3390/technologies3020103>
- [11] X. Bo, W. Junwen, L. Guangie, D. Yuewei, "Image Copy-Move forgery Detection Based on SURF", *International Conference on Multi-media Information Networking and Security*, 2010, pp. 889-89. DOI: 10.1109/MINES.2010.189
- [12] E.R.W. Garage, M. Park, "ORB: An efficient alternative to SIFT or SURF", *IEEE International Conference on Computer Vision*, 2011, pp. 2564-2571. DOI: 10.1109/MINES.2010.189.
- [13] T. Mohmood, "A Survey on Block Based Copy-Move Image Forgery Detection Techniques", *International Conference*, pp.1-6, 2015. DOI: 10.1109/ICET.2015.7389169
- [14] P. Sykora, P. Kamencay and R. Hudec, Comparison of SIFT and SURF Methods for Use on Hand Gesture Recognition based on Depth Map, *AASRI Procedia*, vol. 9, pp. 19-24, 2014. <https://doi.org/10.1016/j.aasri.2014.09.005>.
- [15] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, "ORB: an efficient alternative to SIFT or SURF", *IEEE International Conference on Computer Vision*, 2011. DOI: 10.1109/ICCV.2011.6126544.
- [16] C.V. Aravinda, L. Meng, Akshaya, U. K. Reddy, "Signature Recognition and Verification Using Multiple Classifiers Combination of Hu's and HOG Features", 2019 International Conference on Advanced Mechatronic Systems (ICAMEchS) 2019. DOI: 10.1109/icamechs.2019.8861686 ISBN: 9781728134802
- [17] C.V. Aravinda, L. Meng, Akshaya, U.K. Reddy, "An approach for signature recognition using contours based technique", 2019 International Conference on Advanced Mechatronic Systems (ICAMEchS), 2019. DOI: 10.1109/icamechs.2019.8861686 ISBN: 9781728134802

# A Novel Representative $k$ -NN Sampling-based Clustering Approach for an Effective Dimensionality Reduction-based Visualization of Dynamic Data

Dharamsotu Bheekya<sup>\*</sup>, Kanakapodi Swarupa Rani, Salman Abdul Moiz, Chillarige Raghavendra Rao

School of Computer and Information Sciences, University of Hyderabad, Hyderabad, Telangana, 500046, India.

## ARTICLE INFO

Article history:

Received: 13 May, 2020

Accepted: 19 June, 2020

Online: 06 July, 2020

Keywords:

Dimensionality Reduction

Exploratory Analysis

Visualization

Clustering

Sampling

$t$ -Distributed Stochastic Neighbor Embedding

Cluster Validation Index

Interpolation

## ABSTRACT

Visualization plays a crucial role in the exploratory analysis of Big Data. The direct visualization of Big Data is a challenging task and difficult to analyze. Dimensionality Reduction techniques extract the features in the context of visualization. Due to the unsupervised and non-parametric nature, most of the dimensionality reduction techniques are not evaluated quantitatively and not allowed to extend for dynamic data. The proposed representative  $k$ -NN sampling-based clustering, determines the underlying structure of the data by using well-known clustering techniques. The external cluster validation index determines the order sequence of clustering techniques from which the appropriate cluster techniques are recommended for the given datasets. From the recommended set, the samples of the best clustering technique are considered as representative samples which can be used for generating the visual representation. The  $t$ -Distributed Stochastic Neighbor Embedding ( $t$ -SNE) algorithm is applied to generate a low-dimensional embedding model of representative samples, which is more suitable for visualization. The new data samples are added to the generated model by using the interpolation technique. The low-dimensional embedding results are quantitatively evaluated by  $k$ -NN accuracy and trustworthiness. The performance analysis of representative  $k$ -NN sampling-based clustering results and embedding results accomplished by seven differently characterized datasets.

## 1 Introduction

Exploratory analysis of Big Data is ubiquitous in an increasing number of fields and vital to their progress. Visualization plays a paramount role in an exploratory study. Data visualization is applicable for the limited number of dimensions, which depends on the perceptual capability of the analyst. For exploratory analysis, traditional visualization techniques may not provide useful visual insights of high-dimensional data and they are restricted for a limited number of dimensions. The conventional feature selection methods may not provide helpful visual insights for exploratory analysis, which happens due to the inappropriate feature selection. There is a requirement of the feature extraction technique, which shows the correlation between the original features of the data. The Dimensionality Reduction (DR) technique transforms the data and extract new features, which makes data analysis tractable. From the last few decades, researchers have proposed various linear as well as non-linear DR techniques in the context of visualization. The linear DR techniques like Principal Component Analysis (PCA) [1], Multidimensional Scaling (MDS) [2] and Factor Analysis (FA) [3]

deals with the simple structured data. It is challenging to extract valuable information from complex structured data using linear DR. In contradiction to linear DR, the non-linear DR techniques like Isomap [4], Local Linear Embedding (LLE) [5], Laplacian Eigenmap (LE) [6] and Stochastic Neighbor Embedding (SNE) [7] deals with non-linear data. The paper [8] provides a complete comparative study of DR techniques. The methodology of the DR technique depends on its feature extraction criteria. The feature extraction depends on the characteristics of interest in the data such as inter-point distances, reconstruction weights, variation, linear subspace, geodesic distances, linear tangent space, neighborhood graph and conditional probability distribution.

Among all the DR techniques,  $t$ -Distributed Stochastic Neighbor Embedding ( $t$ -SNE) [9], is an improved version of SNE, introduced by Laurens Van Der Maaten and Geoffrey Hinton in 2008. The  $t$ -SNE has gained impressive attention and enormous popularity in several fields [10–12]. The empirical study states that the low-dimensional visual representation of  $t$ -SNE is more robust than any other DR technique. The  $t$ -SNE algorithm most commonly used to preserve the original structure of the high-dimensional data in

<sup>\*</sup>Corresponding Author: Dharamsotu Bheekya, +919573466965, & 15mcp21@uohyd.ac.in

very low-dimensional embedding (i.e., either 2D or 3D). From past several years the t-SNE has been explored in various aspects such as optimization [13, 14], scalability [14–17], dealing with non-numeric data [18], outliers separation [17] and many more.

The t-SNE is a non-parametric technique that provides flexibility in learning and reduces computational complexity. The non-parametric nature limits t-SNE applicability to the out-of-sample extension, which means the addition of new data samples into the existing t-SNE environment is not possible. If we want to add a new sample, then we should re-run the entire t-SNE model by including a new sample. When the addition of new data points increases, the computational cost of t-SNE also increases monotonically. Therefore, it does not apply to time-series and streaming data. The LION-tSNE [17] of Boytsov et al. addressed the problem of adding new data into the existing t-SNE environment using Local Inverse Distance Weighting Interpolation (LIDWI) without re-running. The outlier's handling is also addressed by LION-tSNE using outlier placement heuristic, which assumes that some percentile of outliers present in the designed t-SNE and determine the outliers from the newly added data points. In LION-tSNE, the sample t-SNE model is designed based on the random sample selection, which may cause the non-representativeness of the data. The representative samples are selected by our earlier approach called  $k$ -NN sampling [19] and the results are statistically significant which is measured by statistical method pairwise t-test [20].

This paper is an extension of our earlier paper presented in High Performance Computing, Data and Analytics (HiPC) [19], which deals with the preservation of the underlying cluster structure of high-dimensional data in low-dimensional t-SNE embedding with a representative sample. The underlying cluster structure preservation is measured in terms of a quantitative metric. In the existing methods, the low-dimensional embedding of t-SNE describes the quality of the structure preservation. Still, it is an open problem for giving the quantitative proof for the number of clusters that exist in the original data. The proposed novel representative  $k$ -NN sampling-based clustering approach for effective dimensionality reduction-based visualization finds the solution. In the first step, the proposed approach determines several distinct data samples using our earlier proposed  $k$ -NN sampling algorithm. The number of samples depends on the range of  $k$  (i.e.,  $1 \leq k \leq m$ ), which gives the neighborhood representation.

In the second step, the effective number of clusters existing in the original dataset is determined by the sampled data using clustering techniques such as  $k$ -Means [21], Agglomerative Hierarchical Clustering (AHC) [22], Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH) [23], Fuzzy  $c$ -means (FCM) [24], Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [25], Ordering Points To Identify the Clustering Structure (OPTICS) [26], Mean Shift (MS) [27], Spectral Clustering (SC) [28], Expectation-Maximization Gaussian Mixture Model (EMGMM) [29], Affinity Propagation (AP) [30] and Mini Batch  $k$ -Means (MBKM) [31]. Each clustering technique generates cluster labels for sampled data. The cluster labels of remaining data (i.e., data points other than selected samples) are labeled by the  $k$ -Nearest Neighbor ( $k$ -NN) algorithm. The  $k$ -NN of each remaining data sample is subject to the selected samples. After assigning the labels to the remaining data samples, the representative sample of each clus-

tering technique is determined by their external cluster validation index called Fowlkes-Mallows Index (FMI) [32]. In our contribution, we are also recommending the order sequence of clustering techniques among the selected techniques for a given dataset. We are also presenting the Compactness (CP) [33], Calinski-Harabaz Index (CHI) [33] and Contingency Matrix (CM) [34] of clustering techniques for gaining a more detailed analysis about clustering. The first technique in the order sequence denotes the best clustering technique and the representative samples of it considered as the samples for t-SNE model design for a given dataset. The cluster validation index provides the comparison between representative  $k$ -NN sampling-based clustering and the aggregate clustering (i.e., clustering on the whole dataset). The proposed representative  $k$ -NN sampling-based clustering is scalable to all clustering techniques which are suitable for numerical datasets. Also, we can apply any clustering technique to the large scale dynamic data with representative  $k$ -NN sampling-based clustering. Due to the paper limitation, we have chosen the most popular clustering techniques from different groups.

In the third step, the sample t-SNE model is designed on a representative sample of best clustering techniques, which transforms the high-dimensional data into low-dimensional embedding. The remaining data samples are added to the sample t-SNE environment with the help of LIDWI. The outliers from the remaining samples are identified and controlled by proposed heuristic and the identified outliers are placed into the t-SNE environment using outlier placement heuristic of Boytsov et al. [17]. In the fourth step, the t-SNE embedding of input data is quantitatively evaluated by  $k$ -NN accuracy in the context of clustering and trustworthiness. The quantitative evaluation answers the question, how much structure of high-dimensional data is preserved by the low-dimensional t-SNE embedding. The  $k$ -NN accuracy of t-SNE embedding is measured in two ways: baseline accuracy and sampling accuracy. For quantitative performance evaluation, the  $k$ -NN accuracy of t-SNE embedding of representative  $k$ -NN sampling-based clustering and aggregate clustering results are compared with the ground truth class labels, which is measured in our earlier paper. In our earlier approach, we used ground-truth class labels for obtaining the  $k$ -NN accuracy of the t-SNE embedding but here we are using for checking the derived cluster purity of representative  $k$ -NN sampling-based clustering. The  $k$ -NN sampling-based clustering results and t-SNE embedding of it are analyzed by seven differently characterized toy and real-world datasets. In summary, our contribution consists of

- The order sequence of applicable clustering methods with representative  $k$ -NN sampling-based clustering using cluster validation index. The set of recommended techniques for the given dataset using a threshold. The comparison between representative  $k$ -NN sampling-based clustering and aggregate clustering results.
- The outliers from the remaining samples are identified and controlled by the proposed heuristic before adding into the t-SNE embedding. The embedding structure quantitatively evaluated by  $k$ -NN accuracy in the context of clustering and trustworthiness.

The organization of the paper is as follows. In section 2, we present the background knowledge that gives the basic mathemati-

cal intuition required for understanding the clustering techniques, t-SNE and other supplements. Section 3 presents the related work. In section 4, we present a detailed description of the proposed  $k$ -NN sampling-based clustering for the visualization of dynamic data. In section 5, we are providing the details of datasets, experimental setup and analysis of results. Finally, section 6 gives directions for future work and conclusions.

## 2 Background Knowledge

The following section describes the few techniques for understanding the background formulation of our work. Section 2.1 gives a detailed description of the selected clustering techniques for representative  $k$ -NN sampling-based clustering. Section 2.2 describes the mathematical intuition of the t-SNE algorithm. The intuition behind the addition of the inlier data point into an existing t-SNE environment is explained in section 2.3. In section 2.4, we are giving a detailed description of metrics that are used for the quantitative evaluation of embedding.

### 2.1 Clustering and their performance measures

Section 2.1.1 describes the selected clustering techniques which are used for representative  $k$ -NN sampling-based clustering. 2.1.2 provides detailed mathematical intuition of cluster validation indexes such as Fowlkes-Mallows Index (FMI), Compactness (CP), Calinski-Harabaz Index (CHI) and Contingency Matrix (CM).

#### 2.1.1 Clustering Techniques

Clustering is unsupervised learning, where the similarly characterized data objects are grouped. The clusters of data objects can be represented as a set  $C$  of subsets  $C_1, C_2, \dots, C_k$  such that  $\cap_{i=1}^k C_i = \phi$ . The different clustering algorithms are proposed based on their measures of similarity: partitional, hierarchical, fuzzy theory-based, distribution-based, density-based, graph partition-based, grid-based, model-based and many more. The user decides the number of clusters present in the dataset by using a heuristic, trail-and-error and evolutionary approaches such as density and probability density. From the groups mentioned above, we have chosen the most frequently and popularly used algorithms for experimental evaluation. But proposed representative  $k$ -NN sampling-based clustering approach is scalable to all the clustering techniques which are suitable for the numerical datasets. We explored traditional clustering algorithm [35–37] for our work.

In partitional clustering, data points assigned to any one of the  $k$ -clusters using distance similarity measures such as Euclidean distance.  $k$ -means [21] clustering is one of the simplest, best-known and benchmarked partition-based clustering. The  $k$ -means clustering classifies the given data points through a user-defined number of clusters. The main goal of the  $k$ -means clustering is the initialization of an appropriate  $k$ -centroids, one for each cluster. The objective of  $k$ -means is the minimization of the sum of square distance which can be defined as follows

$$SSD = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (1)$$

where  $\|x_i^{(j)} - c_j\|^2$  is a L2 norm between a data point  $x_i^{(j)}$  and the cluster center  $c_j$ . In AP [30], at the initial stage, all data points are considered as centroids and nodes of the network. The clusters and their centroids are measured by transmitting the similarity message recursively. The Mini Batch  $k$ -Means (MBKM) [31] is an improved version of  $k$ -means, which performs clustering on batches instead of considering each point. Therefore, MBKM requires less computational time and is also applicable to large datasets.

In hierarchical clustering, groups are formed by iteratively dividing the data objects either in the bottom-up or top-down approach. The bottom-up approach is also known as Agglomerative Hierarchical Clustering (AHC) [22], in which initially each data object is considered as a separate group, then merging these small groups into larger and larger groups until a single group formation or certain threshold. The top-down approach is also known as Divisive Hierarchical Clustering (DHC), and it works in reverse order of the bottom-up approach. Both hierarchical approaches depend on the linkage criteria such as single, complete and average linkage. The linkage criteria determines the metric for merging two similar characterized small clusters. For example, single linkage criteria determines the minimum distance pair from the neighbor clusters  $\min\{d(x, y) : x \in C_i, y \in C_j\}$ . The BIRCH [23] method has proposed to deal with large datasets, outliers in robust and also to reduce the computational complexity. The BIRCH method works on the idea of Cluster Features (CF) which is a height-balanced tree.

The basic idea of fuzzy theory-based clustering is that the discrete labeling is converted to continuous intervals, to describe the belonging relationship among objects more reasonably. The Fuzzy  $c$ -means (FCM) [24] is an extension of  $k$ -means where each data point can be a member of multiple clusters with membership value. The main advantage of FCM is that the formed groups are more realistic.

The density-based clustering finds the clusters based on the density of data points in a region. The principal objective of density-based clustering is that for each instance of a group, there should be at least a minimum number of neighbor instances within the given radius. The DBSCAN [25] is the most well known density-based clustering. In DBSCAN clustering, the data objects fall into three groups: core-object, border-object and noise-object. The data points of the core-object group have enough number of neighbors in the given radius, and these data points are from the higher density region. The data points of the border-object group have fewer neighbors than the required number of neighbors in the given radius and they are in the neighborhood of the core object group. The data points of the noise-object group are not in either core or border-object group. The advantage of DBSCAN clustering is that the generated clusters are in arbitrary shape based on the given parameters such as radius and number of minimum instances. The improved version of DBSCAN is OPTICS [26], which overcomes the limitations of DBSCAN. Mean-shift [27] algorithm determines the mean of offset of the current data point, the next point is identified based on the current point and the offset. The process will continue until some criteria is satisfied.

The spectral clustering [28] is a graph partition-based technique in which clustering is obtained by similarity graph partition. The paper [28] describes the different ways of constructing the similarity graph. While building the similarity graph, the data objects are con-

sidered as vertices and the similarity (i.e., the similarity distance) between the data objects is a weighted edge. The critical idea of similarity measure is to find a method of graph partition by making edges between different groups have low edge weight and the edges of the same group have more edge weights. In similarity graph partitioning, the cluster analysis is carried out by feature space obtained by the eigenvectors corresponding to k largest eigenvalues of the laplacian matrix [38] of similarity graph. The spectral clustering is suitable for the datasets with arbitrary shape, high-dimensions and outliers.

The EMGMM [29] is a model as well as a distribution-based clustering technique in which cluster data obey the same independent Gaussian distribution which exists in the original data distribution. The generated clusters are more realistic because a probability distribution obtains them.

The performance and applicability of clustering depends on the used dataset characteristics. Therefore, the researchers empirically state that none of the clustering technique is generalized for any datasets.

### 2.1.2 Cluster validation Measures

In our experimentation, we have used two internal and one external cluster validation measure to evaluate the performance of clustering techniques. The internal measure computes the performance of clustering without using the ground-truth class labels. The compactness (CP) [33] is an internal cluster validation measure which can be computed as follows

$$CP = \frac{1}{n} \sum_{l=1}^k n_l \left( \frac{\sum_{x_i, x_j \in C_l} d(x_i, x_j)}{n_l(n_l - 1)/2} \right) \quad (2)$$

where  $d(x_i, x_j)$  is Euclidean distance between two objects in cluster  $C_l$  and  $n_l$  is the number of objects in  $C_l$ . The smaller CP value of clustering is more compact and gives better clustering results. The Calinski-Harabaz Index (CHI) [33] or Variance Ratio Criterion is defined as the ratio of between-clusters dispersion mean and the within-cluster dispersion mean. The CHI is an internal cluster validation measure which is given by

$$CHI = \frac{\sum_{l=1}^k n_l d^2(C_l, C) / (NC - 1)}{\sum_{l=1}^k \sum_{x \in C_l} d^2(x, C_l) / (n - NC)} \quad (3)$$

Where  $n$  is the dataset size,  $C$  is the dataset center,  $n_l$  is the size of  $l^{th}$  cluster,  $C_l$  is the  $l^{th}$  cluster centroid and  $NC$  is the number of clusters.

The external validation measure knows the ground truth class labels. The primary purpose of the external validation index is to choose an optimal clustering algorithm for a given dataset. The external validation measure also checks the cluster purity. In our experiment, we used Fowlkes-Mallows Index (FMI) [32] is the geometric mean between precision and recall. The FMI defined as follows

$$FMI = \frac{TP}{\sqrt{(TP + FP)(TP + FN)}} \quad (4)$$

The FMI value is bounded between 0 and 1, 1 denotes that the obtained clusters and the given ground truth classes are the same. The large FMI indicates that the obtained clustering is purer.

We also measured the contingency matrix that reports the intersection cardinality for every true/predicted cluster pair. The contingency matrix gives sufficient statistics for all clustering metrics. But it is hard to interpret the contingency matrix of extensive clustered data.

### 2.2 Student t-Distributed Stochastic Neighbor Embedding (t-SNE)

The t-SNE [9] algorithm introduced by Laurens Van Der Maaten and Geoffrey Hinton in 2008, based on the SNE algorithm. The principal objective of SNE is to preserve the underlying structure of high-dimensional data in low-dimensional embedding space. Lets assume the given input data  $X = \{x_1, x_2, \dots, x_N\}$  where each  $x_i \in \mathbb{R}^D$  is a D-dimensional vector. The t-SNE computes the embedding  $\mathcal{Y} = \{y_1, y_2, \dots, y_N\}$  of  $X$  where each  $y_i \in \mathbb{R}^d$  is a d-dimensional vector, where  $d \ll D$  and most commonly  $d=2$  or 3. The similarity between  $x_i$  and  $x_j$  of  $X$  is calculated by conditional probability  $p_{ij}$  is given by

$$p_{j|i} = \frac{\exp\left(\frac{-d(x_i, x_j)^2}{2\sigma_i^2}\right)}{\sum_{k \neq i}^N \exp\left(\frac{-d(x_i, x_k)^2}{2\sigma_i^2}\right)}, \quad p_{i|i} = 0, \quad p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N} \quad (5)$$

where the bandwidth  $\sigma_i$  of Gaussian kernel, is obtained by binary search by matching the perplexity of  $P_i$  and the user-defined perplexity ( $\mu$ ) which is given as a parameter. The perplexity is a smooth measure of an adequate number of neighbors for each data point. The equality and perplexity of  $P_i$  is as follows

$$\mu = 2^{H(P_i)} \quad \text{where} \quad H(P_i) = - \sum_j^N p_{j|i} \log_2 p_{j|i} \quad (6)$$

where  $H(P_i)$  denotes the entropy and  $P_i$  is the conditional probability distribution across all data points for the given  $x_i$ . The  $y_i$  and  $y_j$  are the corresponding low-dimensional values of  $x_i$  and  $x_j$  (i.e., the values of  $\mathcal{Y}$  are initialized by Gaussian or uniform distribution). The similarity between  $y_i$  and  $y_j$  is defined as

$$q_{ij} = \frac{(1 + d(y_i, y_j)^2)^{-1}}{\sum_{l=1}^N \sum_{k \neq l}^N (1 + d(y_l, y_k)^2)^{-1}} \quad (7)$$

where  $d(y_i, y_j)$  is defined as distance similarity measure such as Euclidean distance. In low-dimensional embedding, the similarity  $q_{ij}$  is obtained by student t-distribution with one degree of freedom. But, the similarity  $p_{ij}$  of high-dimensional data uses the Gaussian distribution. The cumulative function curvature of Gaussian distribution is flatter than the cumulative function curvature of student t-distribution with one degree of freedom. The principal idea of using student t-distribution in low-dimensional embedding is to overcome the crowding problem [7].

If  $p_{ij} \sim q_{ij}, \forall i \neq j \in N$ , then the given data is perfectly embedded into the low-dimensional space. Otherwise, compute the KL-divergence (i.e., error) between  $p_{ij}$  and  $q_{ij}$  that is equal to the cross-entropy in Information Retrieval System (IRS). The cost function (C) or objective of t-SNE is defined as follows

$$C = KL(P \parallel Q) = \sum_j^N p_{ij} \log \frac{p_{ij}}{q_{ij}} \quad (8)$$



The simple gradient descent applied on cost function  $\mathbf{C}$  for obtaining the optimization or minimization of it. The simple gradient descent of cost function  $\mathbf{C}$  given by

$$\frac{\partial \mathbf{C}}{\partial y_i} = 4 \sum_{j \neq i} (p_{ij} - q_{ij}) q_{ij} Z(y_i - y_j) \quad (9)$$

where  $Z = (1 + d(y_i, y_j)^2)^{-1}$  is a normalization term of student t-distribution. The Equation 9 produce either positive or negative values depending on the  $p_{ij}$  value. If an Equation 9 gives positive value then there is an attractive force among the  $y_i$  and  $y_j$  of embedding space. Otherwise, there is a repulsive force among the  $y_i$  and  $y_j$  of embedding space. The degree of repulsion depends solely on the closeness of points in the embedding space. In the optimization process, the early exaggeration coefficient  $\alpha > 1$  plays a paramount role in forming groups of similar objects of high-dimensional data in low-dimensional embedding space. In the early exaggeration process, the elements of similarity matrix (P) (i.e.,  $p_{ij}$ 's) multiplied by the early exaggeration coefficient, which is measured by the intuition given by George C. Linderman and Stefan Steinerberger [39]. Therefore, similar data objects bring near to each other in low-dimensional embedding space. This process can happen at the early stage of optimization. The gradient descent of  $\mathbf{C}$  after early exaggeration is

$$\frac{\partial \mathbf{C}}{\partial y_i} = 4 \sum_{j \neq i} \alpha p_{ij} q_{ij} Z(y_i - y_j) - 4 \sum_{j \neq i} q_{ij}^2 Z(y_i - y_j) \quad (10)$$

The momentum and learning rate parameters improve the optimization process of the cost function. The momentum parameter reduces the number of iterations of the cost function optimization. At the initial stage of iteration, the momentum value is small until the map points have become moderately well organized. The optimization is improved by input approximation and tree-based algorithms [13], which reduce the memory and computational complexity. The updated values of  $\mathcal{Y}$  at iteration  $t$  is obtained by

$$\mathcal{Y}^{(t+1)} = \mathcal{Y}^{(t)} + \eta \frac{\partial \mathbf{C}}{\partial y_i} + \alpha(t)(\mathcal{Y}^{(t-1)} - \mathcal{Y}^{(t-2)}) \quad (11)$$

where  $\alpha(t)$  is momentum at  $t^{\text{th}}$  iteration,  $\eta$  is learning rate. The work-flow of t-SNE is shown in the Figure 1.

The well-separated clusters of input data are well preserved in the low-dimensional embedding by setting the early exaggeration coefficient  $\alpha$  and learning rate  $\eta$ . The intuition for setting the above parameter is derived by George C. Linderman and Stefan Steinerberger in [39]. According to the George C. Linderman and Stefan Steinerberger observations the early exaggeration coefficient  $\alpha$ , learning rate  $\eta$ , and minimum probability  $p_{ij}$ 's of same cluster objects (i.e.,  $x_i$  and  $x_j$  belong into the cluster  $C_l$  for all  $i \neq j$ ) is defined as follows.

$$\alpha \sim \frac{n}{10}, \quad \eta \sim 1 \quad \text{and} \quad p_{ij} = \frac{1}{10n |\pi^{-1}(\pi(i))|} \quad (12)$$

where  $\pi : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, k\}$  assigns each data point to one of the  $k$  clusters.

## 2.3 Local Inverse Distance Weighting Interpolation (LIDWI)

The LIDWI [40] maps new data point  $x \in \mathcal{R}^D$  into the existing embedding  $y_i \in \mathcal{R}^d$ , where  $\{i = 1, 2, \dots, m\}$ . The LIDWI determines the value of  $x$  as a weighted sum of values  $y_i$ , where weight is proportional to inverse distances. The LIDWI of  $x$  is

$$LIDWI(x) = \sum_{\|x-x_i\| \leq r_x} w_i(x) \cdot y_i, \quad w_i(x) = \frac{\|x-x_i\|^{-p}}{\sum_{\|x-x_i\| \leq r_x} \|x-x_i\|^{-p}} \quad (13)$$

for instance, when the data point  $x \rightarrow x_i$ , the inverse distance  $\|x-x_i\|^{-1} \rightarrow \infty$ , the corresponding weight  $w_i(x) \rightarrow 1$  (i.e.,  $\forall_{j \neq i} w_j(x) \rightarrow 0$  due to the normalization) and  $LIDWI(x) \rightarrow y_i$ . The neighbor points selection is obtained by a radius  $r_x$  parameter. The parameter  $r_x$  value is calculated by the heuristic proposed by Andrey Boytsov et.al. [17]. In LIDWI, the power parameter  $p$  plays an important role. For instance, very small value of  $p$  predicts the value of  $x$  around the center:  $y \approx \text{mean}(y_i)$  (unless  $x = x_i$ ) even the distance  $\|x-x_i\|$  is low because the weight distribution is close to uniform. When the power parameter is high and the distance  $\|x-x_i\|$  is low, the weight  $w_i(x)$  of very first nearest neighbor is dominating all other neighbors, therefore  $y \approx y_i$  where  $i = \text{argmin} \|x-x_j\|$ . The overfitting suffers from either too small or too large values of power parameter  $p$ . In LOIN-tSNE, the authors proposed a generalization for obtaining power parameter by using leave-one-out cross-validation of the training sample. The computation of the generalized power parameter is obtained by applying the LIDWI for each training sample that produces the estimation of each  $y_i$ . Then the mean square distance between the estimated  $y_i$ 's and real  $y_i$ 's is computed. The optimal power parameter is obtained by optimizing the mean square error (i.e., the mean square distance is minimum). The obtained power parameter is considered as a metric. However, this metric is heuristic, not an exact criterion.

## 2.4 Performance evaluation metrics

### 2.4.1 k-NN accuracy

The existence of the cluster structure of high-dimensional data in low-dimensional embedding is quantitatively measured by  $k$ -NN accuracy of the t-SNE embedding in the context of clustering. The  $k$ -NN accuracy is defined as the percentage of the neighbors having the cluster label equivalent to the observational point cluster label.

### 2.4.2 Trustworthiness

Trustworthiness [41] is one of the measure to evaluate the quality of the t-SNE embedding. Trustworthiness is defined as any unexpected nearest neighbors in the output space are penalized in proportion to their rank in the input space.

$$T(k) = 1 - \frac{2}{Nk(2N-3k-1)} \sum_{i=1}^n \sum_{x_j \in U_k(x_i)} (r(x_i, x_j) - k) \quad (14)$$

where  $U_k(x_i)$  is the  $k$ -NN of  $x_i$  in embedding space,  $r(x_i, x_j)$ ,  $i \neq j$  the rank of  $x_j$  when the data vectors are ordered based on their

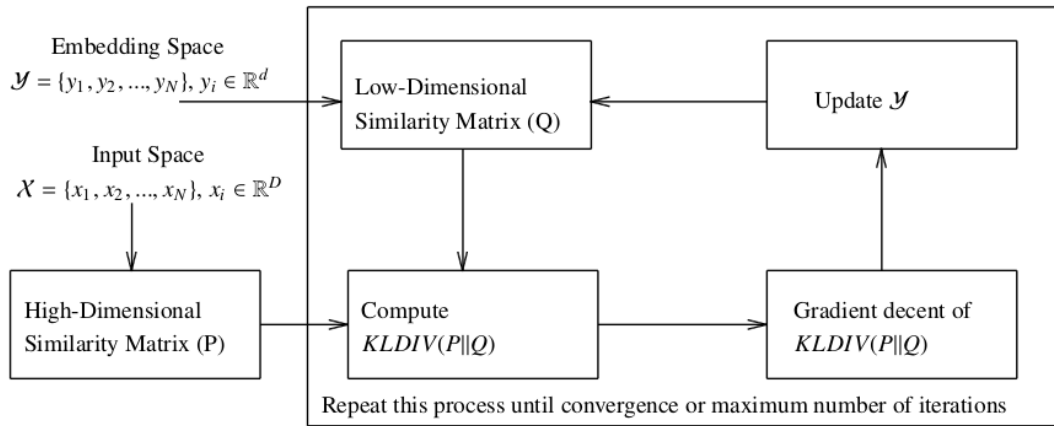


Figure 1: Work-flow of t-SNE algorithm [19]

Euclidean distance from the data vector  $x_i$  in the original data space. It is bounded between 0 and 1, where 1 represents the complete structure of the data preserved in the embedding space.

### 3 Related Work

Our earlier proposed  $k$ -NN sampling based tSNE mainly concentrates on the preservation of the underlying structure of high-dimensional data in low-dimensional embedding space with representative samples. The obtained low-dimensional embedding structure describes the quality of the structure using ground-truth class labels. But the low-dimensional embedding does not give the quantitative proof for the number of clusters that exist in the original data. In [39] authors gave the theoretical observations for well-separated clusters of high-dimensional data in low dimensional embedding space.

The researchers have proposed various methods to incorporate new data samples into the existing t-SNE environment. Most of the existing techniques designed a mapping function  $f: \mathcal{X} \rightarrow \mathcal{Y}$ , which accepts multi-dimensional data and returns its low-dimensional embedding. The designed mapping functions are used for incorporating the new data samples into the existing t-SNE environment. In [14–16] authors have proposed different approaches for adding a new data sample or scaling up the t-SNE algorithm.

Andrey Boytsov et.al. [17] proposed the LION-tSNE algorithm based on local IDWI for adding new data sample into an existing t-SNE environment. It also addresses the outlier handling approaches. Our earlier work extended the idea of the LION-tSNE algorithm by proposing a  $k$ -NN sampling method for designing a representative sampling based t-SNE model. It allows the selection of the sample concerning their  $k$ -nearest neighbors instead of random sampling. In this paper, we are proposing the novel representative  $k$ -NN sampling-based cluster approach for effective dimensionality reduction-based visualization of dynamic data, which determines the underlying cluster by using the most popular clustering techniques. The obtained cluster structure is quantitatively evaluated by  $k$ -NN accuracy in the context of clustering and trustworthiness.

## 4 Proposed representative $k$ -NN sampling-based clustering for effective visualization Framework

The proposed representative  $k$ -NN sampling-based clustering for effective dimensionality reduction-based visualization of dynamic data framework is shown in the *Figure 2*. It has four phases: at phase 1, data preprocessing is done for removing the redundant data points and filling the empty variables with appropriate values. At phase 2, the proposed  $k$ -NN sampling-based clustering is applied to determine: (i) The generation of  $k$  distinct samples using mutual  $k$ -NN sampling with static graph updation algorithm. (ii) The list of clustering techniques applicable for the given dataset. (iii) The optimal sample size produces the best grouping for each concerned clustering technique. (iv) The ordered sequence of recommended clustering techniques for the given dataset from which the best clustering technique is selected by using FMI scores. The detailed representative  $k$ -NN sampling-based clustering is presented in section 4.1. At phase 3, the dimensionality reduction-based visualization is obtained by t-SNE algorithm, which derives the low-dimensional embedding of data and the LIDWI algorithm is used to interpolate the new data points into the learned t-SNE model. The outliers from the remaining data points are identified and controlled by the proposed heuristic. The dimensionality reduction-based visualization is described in section 4.2. In the final phase, the quantitative measure of the t-SNE embedding is computed by  $k$ -NN accuracy and trustworthiness. The process of quantitative metric derivation is described in section 4.3.

### 4.1 Representative $k$ -NN sampling-based clustering

Initially, the  $k$  distinct samples are generated by the modified algorithm of our earlier proposed method. The sample generation depends on the parameter  $k$  of  $k$ -NN sampling. Each  $k$  generates one distinct sample; the boundary of  $k$  is denoted as  $1 \leq k \leq m$ , where  $m$  is the maximum number of neighbors required for preserving the behaviour of any data point of the given dataset. Our earlier approach has given two different  $k$ -NN sampling strategies



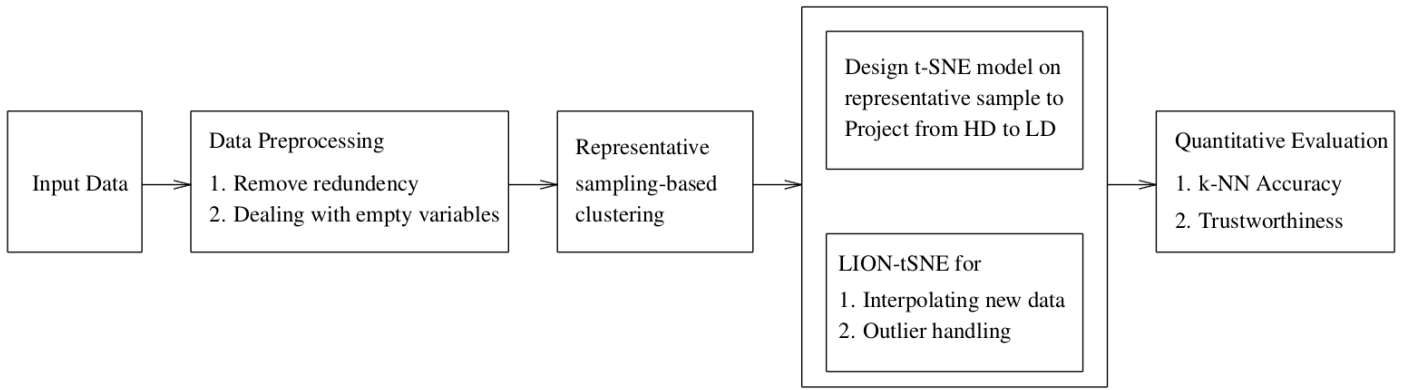


Figure 2: Proposed framework of representative  $k$ -NN sampling-based clustering for an effective dimensionality reduction-based visualization of dynamic data using LION-tSNE with quantitative measure of cluster accuracy

such as static and dynamic  $k$ -NN sampling based on  $k$ -NN graph modification. In this paper, we are considering only static  $k$ -NN sampling in which the  $k$ -NN graph is modified statically that means no new edge is added among any two vertices of the  $k$ -NN graph. The main reason for selecting static  $k$ -NN sampling is that the t-SNE embedding results of it are more statistically significant than the dynamic one. The Nearest Neighbor score (NN\_score) and Mutual Nearest Neighbor score (MNN\_score) of each node of the  $k$ -NN graph are used for selecting the samples. Lets assume  $k$ -NN graph is a directed graph  $G = (V, E)$ , the edge  $E(v_1, v_2)$  gives  $v_2$  as a neighbor of  $v_1$  and neighborhood of  $v_1$  is denoted by  $N_{v_1}$ . The *out-degree* of each vertex is equal to  $k$ , and the *in-degree* of a vertex depends on the neighborhood property of other vertices (i.e., the data point  $x_i$  belong into the neighborhood of other datapoints). In our method, each data point is a vertex of the  $k$ -NN graph and  $k$  is a parameter for deriving the edges between neighboring vertices. The NN\_score of data point  $x_i$  is equal to the *in-degree* of  $x_i$  which is defined as follows

$$NN\_score(x_i) = |\{x_j \mid x_i \in N_{x_j}\}|, \forall_{j \neq i} x_j \in \mathcal{X} \quad (15)$$

where  $\mathcal{X}$  denotes whole data set,  $N_{x_j}$  denotes the neighborhood of  $x_j$ . The MNN\_score of data point  $x_i$  is at most  $k$  which is given by

$$MNN\_score(x_i) = |\{x_j \mid x_i \in N_{x_j}\}| \text{ where } x_j \in N_{x_i}, \forall_{j \neq i} x_j \in \mathcal{X} \quad (16)$$

From the NN\_score and MNN\_score matrix the Representative Sample (RS)  $x_i$  selection and the sample  $x_j$  Represented by Representative Sample (RRS)  $x_i$  is given by

$$RS(x_i) = first\_index\{argmax_{x_i \in \mathcal{X}}\{NN\_score(x_i)\} \cap argmax_{x_i \in \mathcal{X}}\{MNN\_score(x_i)\}\} \quad (17)$$

$$RRS(x_j) = \{x_j \mid x_j \in N_{x_i}\} \text{ where } x_i \in N_{x_j}, \forall_{j \neq i} x_j, x_i \in \mathcal{X} \quad (18)$$

The updated mutual  $k$ -NN sampling with a static graph updation algorithm is shown in *Algorithm 1*. Initially, the Train\_samples (i.e., representative samples set) and Rep\_samples (i.e., the sample represented by the selected training sample that is the first

nearest mutual neighbor of train sample) are null sets. For each iteration, the data point  $x$  which has maximum NN and maximum MNN score appended to the Train\_samples set. The data point  $y$  is appended to the Rep\_samples if  $y \in N_x$  and  $x \in N_y$ , that denotes the point  $x$  and  $y$  are more similar (i.e., the distance  $d(x, y) < d(x, z)$  where  $\forall z \in N_x$ ) to each other. The data point  $x$  and  $y$  are deleted from  $\mathcal{X}$  and their corresponding vertices are deleted from the  $k$ -NN graph. After deleting the vertices of  $x$  and  $y$ , the corresponding in-edges and out-edges are removed from the  $k$ -NN graph and the graph is updated accordingly. The elements of Train\_samples and Rep\_samples are obtained iteratively, the iteration repeats until the NN\_score of the remaining  $\mathcal{X}$  is equal to zero.

**Algorithm 1:** Mutual  $k$ -NN Sampling with a static graph updation

```

Data: data set  $\mathcal{X} = \{x_1, x_2, \dots, x_N\}$ , parameter  $k$  for minimal training sample selection
Result: Return Train_samples, Rep_samples
Train_samples =  $\emptyset$ 
Rep_samples =  $\emptyset$ 
begin
  Compute  $k$ -NN graph of  $\mathcal{X}$ 
  repeat
    Compute  $NN\_Score(\mathcal{X})$ 
    Compute  $MNN\_Score(\mathcal{X})$ 
    index =  $[NN\_Score(\mathcal{X}) == argmax\{NN\_Score(\mathcal{X})\}]$  /* gives all index which are having same NN.Score */
    if len(index) > 1 then
      | train_index =  $argmax\{MNN\_Score(x_i)\}$  where  $i \in index$ 
    end
    else
      | train_index = index
    end
     $x = \mathcal{X}[train\_index]$ 
    Train_samples = Train_samples  $\cup$   $x$ 
    Determine the mutual neighbors of train_index
     $y = Most\_NMN(x)$  /* return first mutual nearest neighbor of  $x$  */
    if  $y \in N_x \& x \in N_y$  then
      | Rep_samples = Rep_samples  $\cup$   $y$ 
    end
    Delete  $x$  and  $y$  from  $\mathcal{X}$  and update the  $k$ -NN graph
  until ( $NN\_Score(\mathcal{X}) = 0$ );
end

```

In our earlier approach of static graph updation, we have considered the whole mutual neighborhood set of  $x$  as the samples represented by  $x$ . The whole neighborhood set selection causes

loss of information due to that reason the earlier approach select Train\_samples set at an early stage of  $k$  value. The data points of the remaining  $X$  (i.e., the data point do not belong into either Train\_samples set or Rep\_samples set) is handled in two different ways. Handling of the remaining samples is described in subsection 4.4. The mutual  $k$ -NN sampling with a static graph updation algorithm generates a various set of Train\_samples and Rep\_samples with different  $k$ -values.

The clustering techniques which are described in subsection 2.1.1 are applied on each Train\_samples set of mutual  $k$ -NN sampling with a static graph updation algorithm that is called a sampling-based clustering. The cluster labels for the data points other than the Train\_samples are assigned by using  $k$ -NN algorithm. For example, the dataset size is  $N$ , the Train\_samples size is  $n$  and then the unmarked sample size is  $N - n$  (i.e., the samples are not assigned with any labels). The  $k$ -NN of each unmarked sample is derived from the associated Train\_samples set. The label assignment of each unmarked sample depends on the labels of its  $k$ -NN, a label with maximum occurrence in  $k$ -NN that is assigned to a sample. From each clustering technique, the optimal Train\_samples set is obtained by the FMI score. The FMI is an external validation index that uses the ground truth class labels. Therefore, the number of clusters is defined as a constant that is equal to the number of ground truth classes. The FMI score of optimal Train\_samples of clustering techniques generates an order sequence of the clustering techniques. From the order sequence, we can recommend the most desirable techniques for a given dataset from the selected set of techniques. From this recommendation, the best suitable technique is selected and its optimal Train\_samples set is considered as the best representative sample. The threshold parameter is used to derive the recommended techniques, which is defined as the FMI score difference between two adjacent techniques of order sequence. The three cluster validation index such as FMI score, CHI score and CP of representative  $k$ -NN sampling-based clustering are compared with their aggregate clustering (i.e., clustering on whole dataset) validation index of chosen clustering techniques. The result comparison is discussed in subsection 5.3. The algorithm of representative  $k$ -NN sampling-based clustering is shown in *Algorithm 2*. The embedding of selected optimal Train\_samples and addition of other data samples into an existing t-SNE environment is described in the following section.

## 4.2 Dimensionality reduction-based visualization

The subsection 4.2.1 describes the low-dimensional embedding of a representative sample with t-SNE algorithm. Subsection 4.2.2 describes the addition of new data samples into an existing t-SNE environment that is called out-off-sample extension.

### 4.2.1 Low-dimensional embedding of a representative sample with t-SNE algorithm

Barnes-Hut t-SNE (BH-tSNE) [13] algorithm is an optimized version of the t-SNE algorithm. BH-tSNE optimizes the t-SNE objective function by input similarity approximation and gradient descent approximation. Therefore, it generates low-dimensional embedding of data with minimum computational and memory complexity than

original t-SNE. In our approach, the BH-tSNE algorithm is used in two different ways for calculating the low-dimensional embedding space: 1. Baseline t-SNE embedding, 2. Sampled t-SNE embedding. The baseline t-SNE embedding is obtained by applying BH-tSNE on the whole dataset. In contrast, the sampled t-SNE embedding is obtained by applying BH-tSNE on the best representative sample, which is selected from the representative  $k$ -NN sampling-based clustering. The Baseline t-SNE embedding results analyze the overall structure of the data in low-dimensional embedding. The sampled t-SNE embedding results analyze the data structure with sampled data and allows the addition of new data samples into an existing t-SNE environment, which solves the scalability issue of the t-SNE. For obtaining a well-separated cluster in low-dimensional t-SNE embedding, the value of early exaggeration coefficient  $\alpha$ , learning rate  $\eta$  and input similarity probability  $p_{ij}$ 's are adjusted according to the George C.linderman and Stefan Steinerberger intuition. In our experimentation, the initial solutions of t-SNE is assigned in three different ways, such as random, PCA based and MDS based initial solutions. The random initial solution takes many iterations for convergence. The PCA and MDS based initial solutions overcome the problem of random initialization and they produce better accuracy results, but their initial solution is cost-effective. Adding new data points into a designed t-SNE model is discussed in the below section.

### 4.2.2 Out-off-sample extension: Interpolation and Outlier handling

The addition of new data point to t-SNE embedding depends on the parameter  $r_x$ ,  $r_y$  and  $r_{close}$  values. The value of parameter  $r_x$ ,  $r_y$  and  $r_{close}$  is obtained from the best representative sample and the t-SNE embedding of it. The parameter  $r_x$  is defined as the percentile of the 1-NN distance of a representative sample that decides whether the given new data point is either inlier or outlier. In our proposal, we came to know that the objects of the representative sample set are representing at least one sample of the dataset. Therefore, the representative sample set does not contain any outlier object and we have considered the parameter  $r_x$  as the maximum 1NN distance of it. If the new data point  $x$  has at least one data point within the  $r_x$  from the representative sample set, then  $x$  is an inlier, otherwise outlier. The dilation factor (df) is used to derive a heuristic  $r_x = (1 + df) * r_x$  which controls the consideration of outliers percentage. The LIDWI interpolation technique is used for adding an inlier data point to the t-SNE embedding of the representative sample set. The outlier placement depends on the parameter  $r_y$  and  $r_{close}$ . The outliers placed into the t-SNE embedding of the representative sample set according to the heuristic of Boytsov et.al. [17]. The  $k$ -NN accuracy and trustworthiness of t-SNE embedding quantitatively evaluate the existence of a high-dimensional cluster in the low-dimensional embedding. The quantitative evaluation described in the following section.

### 4.3 Quantitative metric derivation: $k$ -NN accuracy

In our experimentation, the  $k$ -NN accuracy is calculated in two different ways: baseline accuracy and sampling accuracy. The  $k$ -NN accuracy of t-SNE embedding of aggregate data is known as

**Algorithm 2:** Representative  $k$ -NN Sampling-based clustering

---

**Data:** data set  $X = \{x_1, x_2, \dots, x_N\}$ , parameter  $k$  for minimal training sample selection

**Result:** Order sequence of recommended techniques, Representative samples, samples represented by representative samples and best cluster technique

*Recommended\_tech* =  $\emptyset$

**begin**

**for**  $k \leftarrow 1$  **to** *Max\_kval* **do**

    train\_sample, rep\_sample =  $k$ -NN Sample( $k$ )

**for**  $i \leftarrow 0$  **to** *NumberOfClusterTechniques* **do**

      train\_labels = Clustering(train\_sample,  $i$ ) /\* apply  $i^{\text{th}}$  clustering technique on train\_sample set \*/

      remaining\_samples =  $X - \text{train\_sample}$

      Compute labels of remaining\_samples using  $k$ -NN algorithm (i.e., remain\_labels)

      cluster\_labels = train\_labels  $\cup$  remain\_labels

      Compute cluster validation indexes such as FMI, CHI and CP using cluster labels, ground truth class labels

**if**  $FMI > \text{optimal\_FMI}[i]$  **then**

        /\*  $\text{optimal\_FMI}[i] = 0$ , where  $i = 0, 1, \dots, \text{NumberOfClusterTechniques}$  \*/

$\text{optimal\_FMI}[i] = FMI$

$RS[i] = \text{train\_sample}$

$RRS[i] = \text{rep\_sample}$

**end**

**end**

**end**

  Sort optimal\_FMI list:  $\text{optimal\_FMI}[1] > \text{optimal\_FMI}[2] > \dots > \text{optimal\_FMI}[I]$

  Swap the RS and RRS lists values and maintain the tech\_index list of clustering techniques according to the sorted optimal\_FMI

**for**  $i \leftarrow 1$  **to** *size(tech\_index)* **do**

**if**  $\text{optimal\_FMI}[i] - \text{optimal\_FMI}[i + 1] \leq \text{threshold}$  **then**

$\text{Recommended\_tech} = \text{Recommended\_tech} \cup \text{tech\_index}[i]$

**end**

**end**

  The First element of Recommended\_tech represents the best clustering technique among the selected clustering techniques and corresponding RS is chosen as sample for t-SNE embedding.

**end**

---

baseline accuracy. The sampling accuracy is computed in three different forms such as training accuracy, test accuracy and overall accuracy. The  $k$ -NN accuracy of t-SNE embedding of the representative sample is known as training accuracy. The  $k$ -NN accuracy of interpolated samples (i.e., the samples other than the representative sample) is known as test accuracy. The  $k$ -NN accuracy of combined low-dimensional space (i.e., the integration of t-SNE embedding of sampled data and interpolated data) is known as overall accuracy. The  $k$ -NN accuracy typically depends on the parameter  $k$ , which is considered as fixed in our experimentation which is discussed in subsection 5.2. For instance, the smaller  $k$  will give good performance accuracy and while increasing the  $k$ , performance accuracy will decrease. The selection of parameter  $k$  also plays a paramount role in  $k$ -NN accuracy measure. The relationship between the parameter  $k$  and accuracy is shown in the Figure 3. For quantitative performance evaluation, the  $k$ -NN accuracy of representative  $k$ -NN sampling-based clustering and aggregate clustering are compared with the  $k$ -NN accuracy of ground truth class labels, which is discussed in subsection 5.3.

#### 4.4 t-SNE sample selection criteria

In our experimentation, we have considered two different sample selection criteria for designing the t-SNE model. In the first criteria, the samples are mostly representative of at least one or more other samples. The samples are obtained by a mutual  $k$ -NN sampling with a static graph updation algorithm and the best representative sample is selected by the representative  $k$ -NN sampling-based clustering. We observe that most of these samples were chosen from the dense region because they produce good NN\_score and MNN\_score.

Therefore, these samples do not contain any outliers. In this criterion, the outliers from the remaining sample addition are not sufficiently identified by the adequate 1-NN distance. The outliers consideration is controlled by the proposed heuristic, which is discussed in subsection 4.2.2. In the second criteria, the data points added to the representative sample from the remaining data points using  $(t, m, s)$ -Nets sampling [42]. For the addition of  $(t, m, s)$ -Nets samples, we used the same procedure of our earlier work. The  $(t, m, s)$ -Nets select samples randomly in a uniform distribution. These samples may not represent any other samples and the samples may change among the executions due to the randomness. Therefore, the samples of this criteria may contain outliers. The outliers of this criteria are handled in similar ways as they are handled in our earlier approach.

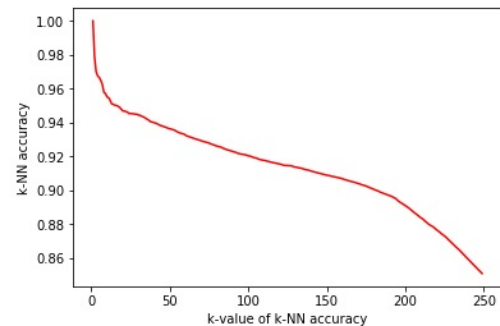


Figure 3: Relationship between  $k$ -value of  $k$ -NN accuracy and  $k$ -NN accuracy which is obtained from the baseline accuracy of  $k$ -NN sampling-based clustering.

## 5 Experimental result analysis

Here we are briefing the datasets, experimental setup and the result analysis. Section 5.1 describes the numerical datasets of different characterizations. The details about the experimental setup is given in section 5.2. The results analysis of each dataset is described in section 5.3.

### 5.1 Datasets

In our experimental evaluation, we have considered differently characterized datasets such as IRIS, Breast-Cancer, Leukemia, Wine, MNIST, Olivetti-Faces, and COIL-20. The *Table 1* provides a detailed description of all the datasets which are downloaded from the UCI Machine Learning repository [43]. For our experimentation, the features of IRIS, Breast-Cancer and Wine datasets are normalized between zero and one, which improves the computational complexity. Initially, the KernelPCA [44] is applied to reduce the dimensionality of the high-dimensional datasets such as Leukemia, MNIST, Olivetti-Faces and COIL-20. The Leukemia dataset is a micro-array of gene expression. The MNIST, Olivetti-Faces and COIL-20 are image datasets that are represented in pixel orientations. The MNIST is a handwritten digits dataset. The Olivetti-faces dataset consists ten face images of 40 individuals with small variation in viewpoint, the addition of glass and large variation in expression. The COIL-20 is an image of 20 group objects such as animals, furniture and etc.

Table 1: Overview of datasets along with their size, dimensions, and classes

Dataset Name	Size	Dimensions	# Classes
IRIS	150	4	3
Breast Cancer	569	30	2
Leukemia-ALL-AML	72	7129	2
Wine	178	13	3
MNSIT	70K	784	10
Olivetti faces	400	10304	40
COIL-20	1440	1024	20

### 5.2 Experimental configuration

In our experimentation, The parameter  $k$  of *Algorithm 1* is initially considered as  $1 \leq k \leq 50$ . The upperbound of  $k$  is equal to largest perplexity value from the literature study [9]. The perplexity is set between 5 and 50 for a fairly good visual representation of any real-world data. In our proposal, the samples of any clustering technique depends on the parameter  $k$ . Also, we observed that when there is an increment in parameter  $k$  then there is an increment or no change in FMI score of clustering until certain  $k$  value. The FMI score becomes stable afterwards. In our experimentation at most of the cases, the selected clustering technique generates representative sample with  $k$  value less than or equal to 20. In other cases, the cluster technique generates representative sample with  $k$  value greater than 20. But there is a small increment in FMI score comparatively FMI score of clustering with  $k$  value less than or equal to 20. Therefore, we have generalized the  $k$ -value as less than or equal to 20. The sensitivity of parameter  $k$  needs to be investigated more in future. The number of clusters is defined as a

constant that is equal to the number of ground truth classes. The original class labels are not used anywhere in the experimental evaluation. The ground truth class labels are used only for measuring the FMI score that determines the cluster purity. The threshold parameter is set to 0.05 that provides the intuition for selecting the recommended clustering techniques from the order sequence. The threshold parameter derived from the statistical method where the maximum allowable difference between two consecutive values of either increasing or decreasing order sequence is 5%. The recommended set size and threshold parameters are inversely proposional to each other. The sensitivity of threshold parameter needs to be investigated further. From the recommended set, the best technique is chosen and the optimal sample of it being considered as the best representative sample for designing the t-SNE model. The representative sample is embedded in a 2D space using the BH-tSNE algorithm. The parameters of BH-tSNE are set up according to the paper [13] experimental setup. In addition to that, we initialized the embedding space  $\mathcal{Y}$  by sampling the point  $y_i$  from a uniform distribution with  $[-0.02, 0.02]^2$  for obtaining well-separated clusters in embedding space. The datasets with more than 50 dimensions, their dimensionality is reduced to 50 by kernel PCA. The dimensionality reduction speeds up the computation of the probability distribution of the input similarity and suppresses some noise. The results of the BH-tSNE algorithm are shown in 2D scatter-plot representation. The minimum value of  $p_{ij}$  of clustered data, the early exaggeration factor  $\alpha$  and the learning rate  $\eta$  values are assigned similar to the George C.linderman and Stefan Steinerberger paper.

The data points other than the representative samples are interpolated into the BH-tSNE of a representative sample using LIDWI. The parameter  $r_x$  is obtained by either proposed heuristic or intuition of LION-tSNE. In the proposed heuristic dilation factor is bounded between 0 and 1. The parameter  $r_y$ ,  $r_{yclose}$  and power are measured similar to LION-tSNE algorithm.

Table 2: Parameters setting for the experimental setup

Parameter	Value
k of $k$ -NN sampling	$1 \leq k \leq 20$
Threshold	0.05
Perplexity	5 - 50
Early exaggeration coefficient	$\sim \frac{N}{10}$
Adaptive learning rate	1 - 200
Dilation Factor	$0 \leq df \leq 1$
$r_x$ at <i>dist.perc</i>	95 - 100
$r_y$ at <i>dist.perc</i>	100
$r_{yclose}$ at <i>dist.perc</i>	10
Fixed $k$ -value for $k$ -NN accuracy	3 - 10

*Table 2* provides the parameter settings of our experimental evaluation. The parameter *perplexity* represents the effective number of neighbors for each data point. For instance, the small value of *perplexity* creates subgroups within the same cluster of t-SNE results. In contrast to small, the large value of it does not maintain the clear separation between two clusters of t-SNE results. Both cases suffer from either under-fitting or over-fitting problem that causes a lack of visual clarity. The empirical studies state that the *perplexity* value between 5 to 70 gives a good visual representation of t-SNE results. The parameter *dist.perc* represents the overall percentile of the representative sample that needs to be considered as inliers. For

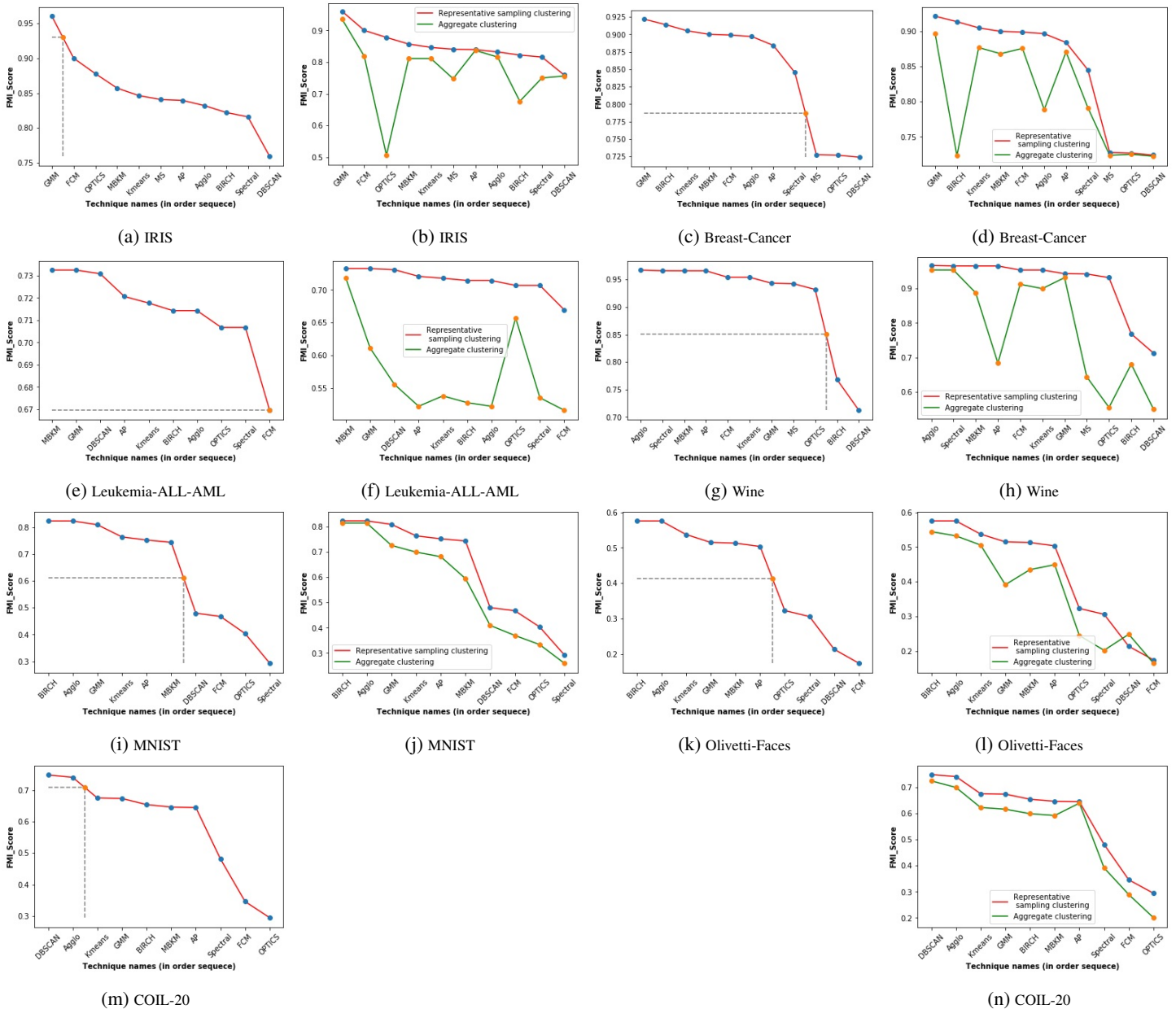


Figure 4: The column one and column three (i.e., sub-figure ((a), (c), (e), (g), (i), (k) and (m)) shows the threshold derivation (i.e., the integration of dotted horizontal and vertical lines) of all seven datasets from the order sequence curve. The column two and column four (i.e., sub-figure (b), (d), (f), (h), (j), (l) and (n)) shows the relationship curve of all datasets from the order sequence, which provides the relationship between the FMI scores of proposed representative  $k$ -NN sampling-based clustering and aggregate clustering.

example, if we take  $dist\_per$  as 95th percentile, that means out of 100 points, 95 points are considered as inliers and the remaining 5 points are outliers. The parameter  $k$  plays an important role in the computation of  $k$ -NN accuracy of the data. The effect of parameter  $k$  is shown in the Figure 3. It is clear that when there is an increment in  $k$  value, accuracy decreases monotonically. The result evaluation of representative  $k$ -NN sampling-based clustering is discussed in the next section.

5.3 Result analysis

In our experimentation, we have considered most frequently and popularly used clustering techniques. The Figure 4 shows the

curves of all seven datasets of subsection 5.1. The Figure 4 gives the threshold derivation for the recommended order sequence and relationship between the FMI scores of proposed representative  $k$ -NN sampling-based clustering and aggregate clustering. In the Figure 4, the x-axis represent the clustering technique names in the order sequence of FMI\_scores (i.e.,  $FMI(t_1) > FMI(t_2) > \dots > FMI(t_i)$  where  $t_i$  represents  $i^{th}$  best clustering technique) and y-axis represents the corresponding FMI\_Scores. Column one and column three of Figure 4 (i.e., sub-figure (a), (c), (e), (g), (i), (k) and (m)) shows the threshold derivation for the recommendation of clustering techniques for all the given seven datasets. The column two and column four of Figure 4 (i.e., sub-figure (b), (d), (f), (h), (j), (l) and (n))

Table 3: The applicable clustering and their recommended order sequence for the given seven datasets.

Dataset	Applicable cluster methods	Recommended order sequence of cluster methods
IRIS	k-Means, Agglomerative, BIRCH, Spectral, EMGMM, FCM, DBSCAN, AP, MBKM, OPTICS, MS	EMGMM
Breast-cancer	k-Means, Agglomerative, BIRCH, Spectral, EMGMM, FCM, DBSCAN, AP, MBKM, OPTICS, MS	EMGMM, BIRCH, k-Means, MBKM, FCM, Agglomerative, AP, Spectral
Leukemia	k-Means, Agglomerative, BIRCH, Spectral, EMGMM, FCM, DBSCAN, AP, MBKM, OPTICS	MBKM, EMGMM, DBSCAN, AP, k-Means, BIRCH, Agglomerative, OPTICS, Spectral, FCM
Wine	k-Means, Agglomerative, BIRCH, Spectral, EMGMM, FCM, DBSCAN, AP, MBKM, OPTICS, MS	Agglomerative, Spectral, MBKM, AP, FCM, k-Means, EMGMM, MS, OPTICS
MNIST	k-Means, Agglomerative, BIRCH, Spectral, EMGMM, FCM, DBSCAN, AP, MBKM, OPTICS	BIRCH, Agglomerative, EMGMM, k-Means, AP, MBKM
Olivetti-faces	k-Means, Agglomerative, BIRCH, Spectral, EMGMM, FCM, DBSCAN, AP, MBKM, OPTICS	BIRCH, Agglomerative, k-Means, EMGMM, MBKM, AP
COIL-20	k-Means, Agglomerative, BIRCH, Spectral, EMGMM, FCM, DBSCAN, AP, MBKM, OPTICS	DBSCAN, Agglomerative

shows the relationship curve that provides the relationship between the FMI\_Score of representative *k*-NN sampling-based clustering and aggregate clustering for all the given seven datasets. The Figure 4 clearly shows that the proposed representative *k*-NN sampling-based clustering results are superior to the aggregate clustering results.

The sets of applicable and recommended order sequence of clustering techniques for the given seven datasets is listed in the Table 3. The CP and HCI of representative *k*-NN sampling-based clustering and aggregate clustering of all clustering techniques are listed in the Table 4 and 5. The Table 4 and 5 clearly states that the CP of the proposed representative *k*-NN sampling-based clustering is smaller than the aggregate clustering for some techniques. In contrary techniques, the compactness is in reverse order, but the compactness difference is very minute in both situations. The CHI is also deriving the relationship between both clustering as same as CP where the CP is small, and then there is a large CHI score.

The contingency matrices of representative *k*-NN sampling-based clustering and aggregate clustering of IRIS, Wine, Breast-cancer and Leukemia datasets are shown in the Table 6, 7, 8 and 9. If the number of classes of the given dataset is larger, then it is difficult to analyze the data with the contingency matrix. Therefore, the contingency matrix of MNIST, Olivetti-Faces and COIL-20 datasets are not addressed. The Table 6, 7, 8 and 9 clearly states that the proposed representative *k*-NN sampling-based clustering classifies the labels much similar to the ground truth class labels compared to aggregate clustering. The Tabel 9 represents the contingency matrix of the Leukemia dataset, which indicates that none of the selected technique gives the best clustering results. Still, representative *k*-NN sampling-based clustering produces better results than the aggregate clustering.

For Quantitative evaluation, the baseline *k*-NN accuracy of proposed representative *k*-NN sampling-based clustering is compared with baseline *k*-NN accuracy of aggregate clustering and baseline *k*-NN accuracy of ground truth class labels. The baseline *k*-NN accuracy and trustworthiness of representative *k*-NN sampling-based clustering, aggregate clustering and ground truth class labels of seven datasets are listed in Table 10. The Table 10 clearly indicates that the proposed method produces more robust results than others.

Table 4: Compactness and CHI score of representative *k*-NN sampling-based clustering, overall sampling clustering (i.e., assigning the cluster labels based on the KNN of remaining samples in the context of representative samples) and the aggregate clustering of IRIS, Breast-Cancer, Leukemia and Wine datasets for the selected clustering techniques.

Method	Samp. size	Compactness			CHI Score		
		Samp. ctr	Overall ctr	Aggr. ctr	Samp. ctr	Overall ctr	Aggr. ctr
<b>IRIS</b>							
<i>k</i> -Means	73	0.4945	0.5601	0.5557	221.54	347.56	358.56
Agglomerative	63	0.5143	0.5579	0.5606	178.78	353.21	348.03
Birch	73	0.5291	0.5646	0.8507	183.85	335.22	193.97
Spectral	60	0.5234	0.5606	0.5709	167.457	348.03	322.48
EMGMM	63	0.5316	0.5754	0.5789	161.44	317.59	307.76
FCM	72	0.5107	0.5633	0.5576	197.71	340.79	355.71
DBSCAN	73	0.6379	0.7047	0.6723	104.84	175.96	124.78
AP	73	0.52	0.557	0.5588	196.99	354.54	353.22
MBKM	74	0.4834	0.5613	0.5562	234.32	343.95	356.28
OPTICS	73	0.5894	0.6071	1.0919	120.53	259.45	15.99
MS	70	0.527	0.557	0.6002	188.2	354.45	289.52
<b>Breast-Cancer</b>							
<i>k</i> -Means	291	1.5608	1.6064	1.6013	203.91	355.42	364.09
Agglomerative	290	1.5735	1.6073	1.6322	199.73	356.80	319.01
Birch	291	1.5698	1.6083	1.9509	193.53	352.67	53.96
Spectral	289	1.5694	1.6177	1.6449	197.34	351.45	328.74
EMGMM	276	1.5744	1.6096	1.6168	178.38	350.08	337.00
FCM	276	1.5645	1.6048	1.6001	186.30	358.22	363.04
DBSCAN	291	1.9493	2.0004	1.9735	21.27	19.66	38.11
AP	213	1.5146	1.6088	1.6088	156.95	356.82	358.62
MBKM	268	0.8824	1.6047	1.6007	180.883	359.63	363.83
OPTICS	268	1.9635	2.0179	2.0069	21.72	24.65	38.41
MS	213	1.9138	2.0054	1.9412	19.79	17.65	68.97
<b>Leukemia-ALL-AML</b>							
<i>k</i> -Means	22	3.4413	3.9367	3.9279	1.567	1.173	1.685
Agglomerative	19	3.3494	3.9315	3.9222	1.874	1.269	1.756
Birch	19	3.3494	3.9315	3.919	1.874	1.269	1.921
Spectral	19	3.3076	3.915	3.92	1.707	1.279	1.845
EMGMM	20	3.4012	3.9415	3.9237	1.337	0.953	1.591
FCM	20	3.3743	3.939	3.9271	1.622	1.117	1.619
DBSCAN	17	3.2654	3.9436	3.9225	1.351	0.959	1.326
AP	14	2.8993	3.895	3.9355	1.548	0.882	1.292
MBKM	22	3.3088	3.8957	3.9366	1.102	0.859	1.122
OPTICS	19	3.315	3.922	3.9071	0.859	0.717	1.317
MS	...	....	....	....	....	....	....
<b>Wine</b>							
<i>k</i> -Means	82	1.3275	1.4298	1.4289	47.255	82.828	83.373
Agglomerative	87	1.3207	1.4356	1.4328	45.964	80.465	81.327
Birch	82	1.555	1.6567	1.629	19.859	34.395	42.564
Spectral	87	1.3532	1.4348	1.4298	42.372	81.014	82.828
EMGMM	71	1.3481	1.4331	1.4348	39.846	81.796	81.698
FCM	85	1.3473	1.4317	1.4305	45.029	82.346	83.135
DBSCAN	64	1.6181	1.6656	1.9606	15.719	37.502	3.102
AP	82	1.5534	1.3368	1.4458	45.362	80.714	80.828
MBKM	87	1.3562	1.4848	1.4278	42.472	81.784	81.523
OPTICS	85	1.4484	1.4531	1.4248	38.866	82.796	81.698
MS	71	1.3489	1.4371	1.4748	49.846	81.797	80.598

Table 5: Compactness and CHI score of representative  $k$ -NN sampling-based clustering, overall sampling clustering (i.e., assigning the cluster labels based on the KNN of remaining samples in the context of representative samples) and the aggregate clustering of MNIST Handwritten digits, Olivetti-Faces and COIL-20 datasets for the selected clustering techniques.

MNIST		Compactness			CHI Score		
Method	Samp. size	Samp. cptr	Overall cptr	Aggr. cptr	Samp. cptr	Overall cptr	Aggr. cptr
$k$ -Means	4481	3816.4	3897.8	3877.8	233.97	477.45	492.48
Agglomerative	4900	3936.2	3974.2	3974.2	206.16	409.22	395.43
Birch	4900	3936.2	3974.2	3974.2	206.16	409.22	395.43
Spectral	4735	4662.2	4694.5	4695.0	1.048	1.0308	1.0459
EMGMM	4206	3862.2	3930.6	3977	200.8	443.59	400.32
FCM	4835	3962.2	4374.5	4555.	1.248	1.308	1.59
DBSCAN	4735	4662.2	4694.5	4695.	1.048	1.0308	1.0459
AP	4496	3832.5	3978.6	3897.3	213.97	457.78	472.34
MBKM	4783	4216.4	4597.1	4577.3	224.17	467.65	452.18
OPTICS	4625	3962.5	4094.3	4195.7	1.258	1.13	1.045
MS	...	....	....	....	....	....	....
<b>Olivetti-Faces</b>							
$k$ -Means	200	1568	1784.5	1565.8	10.384	14.009	21.210
Agglomerative	178	1495.9	1798.8	1546.8	10.718	13.634	22.008
Birch	178	1495.9	1798.8	1546.8	10.718	13.634	22.008
Spectral	188	2235	2545.9	2608.4	0.6858	0.6839	0.9924
EMGMM	165	1528.9	1812.8	1591.5	9.211	13.168	20.272
FCM	178	2479.3	2659.4	2732.6	4.321	4.874	4.178
DBSCAN	194	2679.4	2750.5	2745.1	4.021	3.804	5.782
AP	194	1628.9	1852.8	1691.5	10.321	14.168	20.872
MBKM	178	1668.6	1852.8	1671.5	9.711	12.168	21.275
OPTICS	165	1598.4	1932.8	1891.5	8.217	13.468	22.728
MS	...	....	....	....	....	....	....
<b>COIL-20</b>							
$k$ -Means	679	10.333	10.362	10.180	87.911	183.571	188.454
Agglomerative	701	10.239	10.237	10.294	90.220	184.703	181.461
Birch	701	10.577	10.562	11.1165	87.377	180.283	162.538
Spectral	676	14.852	14.944	17.447	28.585	57.964	27.4288
EMGMM	679	10.273	10.243	10.356	87.832	186.73	187.686
FCM	689	12.872	13.645	15.745	34.784	56.768	34.58
DBSCAN	679	18.843	18.962	19.769	19.532	38.902	38.548
AP	701	10.573	10.253	10.856	83.83	188.63	186.656
MBKM	679	10.253	10.143	10.366	80.832	188.73	187.656
OPTICS	679	10.243	10.233	10.326	87.432	185.738	184.286
MS	...	...	...	...	...	...	...

Table 6: Contingency matrix of  $k$ -NN sampling-based and original clustering on IRIS dataset with best  $k$ -NN sampling-based clustering technique (i.e., EMGMM).

		Class labels generated by					
		$k$ -NN sampling-based clustering			Aggregate clustering		
		C1	C2	C3	C1	C2	C3
Original	C1	50	0	0	50	0	0
	C2	0	49	1	0	45	5
	C3	0	2	48	0	0	50

Table 7: Contingency matrix of  $k$ -NN sampling-based and original clustering on Wine dataset with best  $k$ -NN sampling-based clustering technique (i.e., Agglomerative).

		Class labels generated by					
		$k$ -NN sampling-based clustering			Aggregate clustering		
		C1	C2	C3	C1	C2	C3
Original	C1	0	0	59	2	0	57
	C2	69	2	0	69	2	0
	C3	1	47	0	0	48	0

The sampling  $k$ -NN accuracies such as train, test and overall of t-SNE embedding of a representative sample and interpolation of

other samples are listed in Table 11 and the overall trustworthiness is also covered. The best representative sample of representative  $k$ -NN sampling-based clustering and random sampling of IRIS dataset is shown in Figure 5. The 2D scatter-plots shown in Figure 5 are the first two coordinates of IRIS dataset. The Figure 5(a) shows the best representative sample of  $k$ -NN sampling and the Figure 5(c) shows the representative sample of  $k$ -NN sampling along with samples of (t,m,s)-Nets of remaining data points. The Figure 5(b) and 5(d) shows the random sampling of size equal to the sample size of Figure 5(a) and 5(c). The samples of Figure 5(a) are constant and consistent comparatively other sampling.

Table 8: Contingency matrix of  $k$ -NN sampling-based and original clustering on Breast-Cancer dataset with best  $k$ -NN sampling-based clustering technique (i.e., EMGMM).

		Class labels generated by			
		$k$ -NN sampling-based clustering		Aggregate clustering	
		C1	C2	C1	C2
Original	C1	21	191	16	196
	C2	352	5	340	17

Table 9: Contingency matrix of  $k$ -NN sampling-based and original clustering on Leukemia dataset with best  $k$ -NN sampling-based clustering technique (i.e., MBKM).

		Class labels generated by			
		$k$ -NN sampling-based clustering		Aggregate clustering	
		C1	C2	C1	C2
Original	C1	0	47	9	38
	C2	2	23	6	19

Table 10: Baseline  $k$ -NN accuracy of  $k$ -NN sampling based clustering, original clustering and ground truth class labels

Dataset Name	Trust	Baseline $k$ -NN Accuracy		
		Sampling Clustering	Aggregate Clustering	Ground-truth class labels
IRIS	0.9861	0.9735	0.9666	0.96
Breast-Cancer	0.958	0.9876	0.9862	0.9577
Leukemia	0.6799	0.9875	0.6643	0.7485
Wine	0.9552	0.9774	0.9828	0.9717
MNIST Digits	0.9895	0.9259	0.9483	0.9512
Olivetti face	0.9494	0.6775	0.8287	0.88
COIL-20	0.9972	0.9554	0.9709	0.9743

The Figure 6 shows the 2D scatter-plot visualization of IRIS data with outliers projection. In Figure 6, we are also showing the outliers (i.e. denoted by large grey color circle) of original data after finding the outliers from the addition of remaining data points with the radius  $r_x$  which is obtained by the proposed heuristic  $r_x = (1 + df) * r_x$  with  $df = 0.2$ . It clearly states that sampled data obtain the outliers of original data.

The 2D scatter-plot visualization of embedding space of representative  $k$ -NN sampling-based clustering of the MNIST dataset is shown in the Figure 7. The Figure 7 shows the baseline 2D scatter-plot of t-SNE embedding of whole MNIST data of size 10K



Table 11: Quantitative measure using  $k$ -NN accuracy of  $k$ -NN sampling based clustering, original clustering and ground truth class labels

Optimal Sample			Sampling $k$ -NN Accuracy								
Dataset Name	Sample Size	Trust	Sampling-based clustering			aggregate clustering			Ground truth class labels		
			Train	Test	Overall	Train	Test	Overall	Train	Test	Overall
IRIS	73	0.9594	0.9726	0.9594	0.9661	0.9753	0.9405	0.9652	0.9589	0.9324	0.9585
Breast Cancer	291	0.9239	0.9896	0.9856	0.9820	0.9690	0.9856	0.9876	0.9452	0.9676	0.9546
Leukemia	33	0.5665	0.9416	1.0	0.9875	0.6714	0.4102	0.6266	0.6363	0.5897	0.6247
Wine	87	0.9299	0.9377	0.956	0.9379	0.966	0.967	0.9604	0.9511	0.945	0.9548
MNIST	4900	0.9635	0.9346	0.8956	0.9120	0.9201	0.91	0.9206	0.9335	0.8972	0.9170
Olivetti-Faces	200	0.968	0.6076	0.5128	0.6045	0.6356	0.7948	0.8270	0.6334	0.7794	0.8578
COIL-20	701	0.9941	0.9841	0.9417	0.9509	0.9529	0.9634	0.9749	0.9426	0.9688	0.9727
<b>Random Sample of size equal to optimal sample</b>											
IRIS	73	0.9654	0.9726	0.9859	0.98	0.9808	0.9802	0.977	0.9616	0.9661	0.9728
Breast Cancer	291	0.9337	0.9759	0.9712	0.9753	0.9828	0.9892	0.9929	0.9553	0.9496	0.9472
Leukemia	33	0.5495	1.0	0.9722	0.9866	5654	0.6111	0.6133	0.5771	0.4444	0.6087
Wine	87	0.9316	0.9639	0.9662	0.9517	0.9431	0.9662	0.949	0.9671	0.9775	0.9547
MNIST	4900	0.9599	0.9122	0.8918	0.9001	0.9157	0.8977	0.9081	0.9249	0.8956	0.9048
Olivetti-Faces	200	0.9578	0.5080	0.5050	0.5860	0.7235	0.7121	0.7869	0.6967	0.6767	0.7842
COIL-20	701	0.9935	0.9543	0.9362	0.9449	0.9699	0.9606	0.9652	0.9273	0.9375	0.9515
<b>Optimal Sample plus Lhs Sample</b>											
IRIS	79	0.965	0.9746	1.0	0.9797	0.9733	0.9852	0.9862	0.9594	0.9852	0.9632
Breast Cancer	305	0.9348	0.9736	0.9772	0.9806	0.9934	0.9848	0.9912	0.9608	0.9810	0.9701
Leukemia	41	0.5979	0.9777	1.0	0.9723	0.5361	0.6551	0.5994	0.5964	0.6551	0.6285
Wine	95	0.9356	0.956	0.9512	0.9488	0.9461	0.9634	0.9717	0.9671	0.9512	0.9545
MNIST	5327	0.9684	0.9289	0.9092	0.9144	0.9195	0.9158	0.9210	0.9276	0.9073	0.9180
Olivetti-Faces	211	0.9641	0.6591	0.5271	0.6175	0.6418	0.7826	0.8492	0.6888	0.7989	0.886
COIL-20	729	0.9942	0.9705	0.9464	0.9536	0.9425	0.9690	0.9743	0.9263	0.9760	0.9785
<b>Random Sample of size equal to Optimal Sample plus (t,m,s)-Nets Sample</b>											
IRIS	79	0.9626	1.0	0.9705	0.9661	0.9772	0.9558	0.9657	0.9866	0.9705	0.9728
Breast Cancer	305	0.9339	0.9802	0.9734	0.9806	0.9902	0.9924	0.9894	0.9542	0.9772	0.9683
Leukemia	41	0.5557	0.9492	1.0	0.9732	0.6333	0.6451	0.6365	0.7289	0.4516	0.679
Wine	95	0.9434	0.9444	0.9518	0.949	0.9456	0.9638	0.9493	0.9666	0.9518	0.9487
MNIST	5327	0.9616	0.9326	0.8977	0.9138	0.9316	0.8977	0.9172	0.9433	0.8959	0.9177
Olivetti-Faces	211	0.9708	0.6642	0.4867	0.6011	0.6938	0.7671	0.8108	0.6673	0.7724	0.85
COIL-20	729	0.9949	0.9838	0.9381	0.9558	0.9381	0.9620	0.9721	0.9330	9648	0.9666

with representative  $k$ -NN sampling-based clustering labels as the colors of scatter point groups.

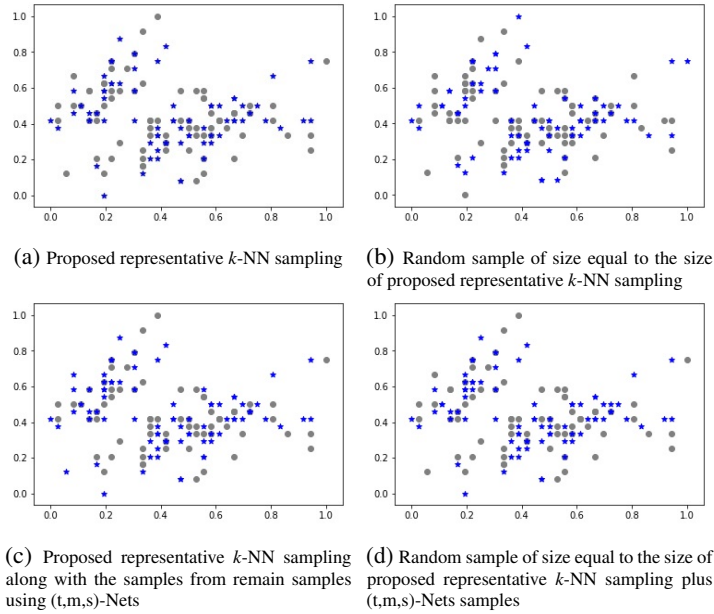


Figure 5: Training sample selection (i.e., represented by blue star scatter-plot point) from IRIS dataset with four different strategies.

The Figure 8 shows the four different 2D scatter-plot representation of MNIST dataset. In Figure 8(a), we are representing the

2D scatter-plot of sampled t-SNE along with the interpolation of inliers and placement of outliers. The sampled t-SNE is designed based on the best representative sample of the representative  $k$ -NN sampling-based clustering concerning the best clustering technique. The inliers of remaining samples (i.e., other than representative samples) are interpolated to t-SNE with the parameter  $r_x$  that is obtained from the proposed heuristic. The outlier of remaining samples are placed into an existing t-SNE environment based on the Boystov et.al heuristic.

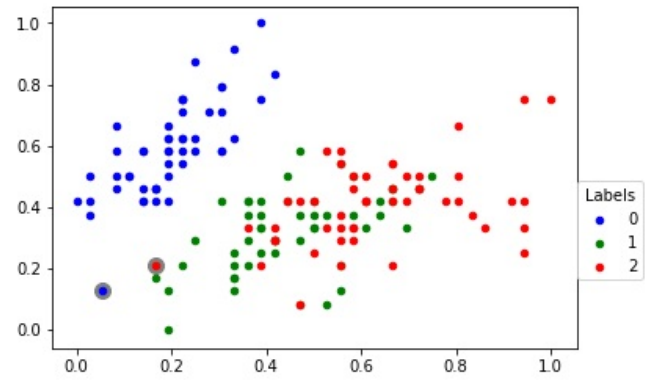


Figure 6: The 2D scatter-plot of IRIS data which denotes the outliers from the newly added sample with proposed heuristic.

In Figure 8(a), the outliers are placed separately from other embedding points. The Figure 8(b) shows the 2D scatter-plot of



(t,m,s)-Nets based t-SNE along with interpolation of other samples. The (t,m,s)-Nets based t-SNE is designed with a combination of representative  $k$ -NN sampling-based clustering and (t,m,s)-Nets of remaining samples. The (t,m,s)-Nets samples are selected from the data samples, which are having NN-score as zero after selecting the samples using a mutual  $k$ -NN sampling with a static graph updation algorithm. The inliers of remaining samples (i.e, other than representative samples plus (t,m,s)-Nets samples) are interpolated to t-SNE with the parameter  $r_x$  that is obtained from the Boystov et.al heuristic. The Figure 8(c) and 8(d) shows the 2D scatter-plots of a random sampled t-SNE along with the addition of new data samples. The random sample size is equal to the sample size of Figure 8(a) and 8(b). The inliers and outliers of remaining samples are placed according to the intuitions of Figure 8(a) and 8(b). The representative  $k$ -NN sampling-based results are statistically significant than the random sampling based results which is derived in our earlier paper [19]. The following section describe the conclusion and future work.

## 6 Conclusion and Future Work

In this paper, we have proposed a representative  $k$ -NN sampling-based clustering approach, which generates cluster results on a sampling basis. The most frequently used clustering techniques are applied to obtain the sampling-based cluster results. Initially, we determine the applicable set of techniques for the given dataset. From the applicable set, sampling-based clustering results of each technique are evaluated by an external cluster validation index called FMI-score. The applicable techniques are arranged in an orderly sequence of their FMI scores. Some threshold parameter derives the recommendation of clustering techniques for the given dataset. From the recommended set, the first technique is selected as the most desirable clustering for the given dataset.

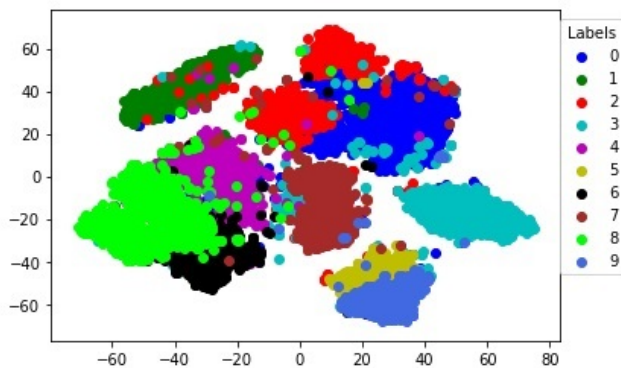


Figure 7: The baseline 2D visual representation of Representative  $k$ -NN sampling-based clustering t-SNE embedding where t-SNE applied on whole data

The samples of this technique are used for generating the low-dimensional embedding of input data. The embedding results are visualized and quantitatively evaluated in the context of sampling-based cluster results. The proposed approach is expanded to identify all the applicable set of clustering techniques for the given dataset, which can be done as future work. There is scope for implementing

the proposed algorithm in a distributed environment that can be applied to Big Data.

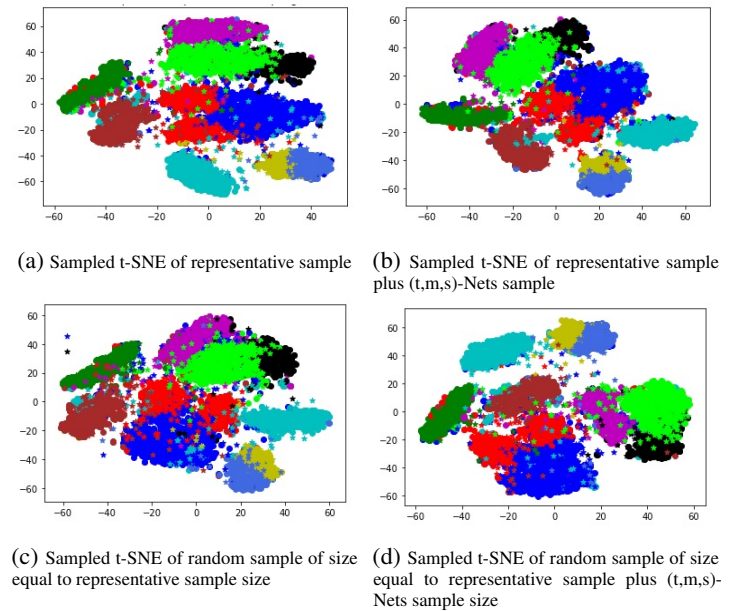


Figure 8: 2D representation of four different t-SNE embedding of MNIST dataset, (a) t-SNE embedding of representative train sample and the interpolation of inliers as well as placement of outliers with proposed heuristic, (b) t-SNE embedding of representative train sample which includes the samples from (t,m,s)-Nets and the interpolation of inliers as well as placement of outliers with A Boystov et.al heuristic, (c) t-SNE embedding of random sample of size equal to the size of sub-figure (a) and the interpolation of inliers as well as placement of outliers with proposed heuristic, (d) t-SNE embedding of random sample of size equal to the size of sub-figure (b) and the interpolation of inliers as well as placement of outliers with A Boystov et.al heuristic.

**Acknowledgment** I would like to express my sincere gratitude to Visveswaraya PhD Scheme for providing financial support to pursue my PHD degree.

## References

- [1] M. Partridge, R.A. Calvo, "Fast dimensionality reduction and simple PCA" Intelligent data analysis, **2**(3), 203-214, 1998.
- [2] T.F. Cox, M.A. Cox, Multidimensional scaling, Chapman and hall/CRC, 2000.
- [3] J. Shi, J. Malik, "Normalized cuts and image segmentation" IEEE Transactions on Pattern Analysis and Machine Intelligence, **22**(8), 888-905, 2000, doi: 10.1109/34.868688.
- [4] J.B. Tenenbaum, V.D. Silva, J.C. Langford, "A global geometric framework for nonlinear dimensionality reduction" Science, **290**(5500), 2319-2323, 2000. DOI: 10.1126/science.290.5500.2319
- [5] S.T. Roweis, L.K. Saul, "Nonlinear dimensionality reduction by locally linear embedding" Science, **290**(5500), 2323-2326, 2000. DOI: 10.1126/science.290.5500.2323
- [6] M. Belkin, P. Niyogi, "Laplacian Eigenmaps for dimensionality reduction and data representation" Neural computation, **15**(6), 1373-1396, 2003. <https://doi.org/10.1162/089976603321780317>
- [7] G.E. Hinton, S.T. Roweis, "Stochastic neighbor embedding" in Advances in neural information processing systems, pp. 857-864, 2003.
- [8] L.V.D. Maaten, E. Postma, J.v.d. Herik, "Dimensionality Reduction: A Comparative" J Mach Learn Res, **10**(13), 66-71, 2009.
- [9] L.V.D. Maaten, G. Hinton, "Visualizing data using t-sne" J Mach Learn Res, **9** 2579-2605, 2008.

- [10] W. Li, J.E. Cerise, Y. Yang, H. Han, "Application of t-sne to human genetic data" *Journal of Bioinformatics and Computational Biology*, **15**(4) 1750017, 2017. <https://doi.org/10.1142/S0219720017500172>
- [11] M. Nguyen, S. Purushotham, H. To, C. Shahabi, "m-tsne: A framework for visualizing high-dimensional multivariate time series" *arXiv preprint arXiv:1708.07942*, 2017. [arXiv:1708.07942](https://arxiv.org/abs/1708.07942)
- [12] W.M. Abdelmoula, B. Balluff, S. Englert, J. Dijkstra, M.J.T. Reinders, A. Walch, L.A. McDonnell, B.P.F. Lelieveldt, "Data-driven identification of prognostic tumor subpopulations using spatially mapped t-sne of mass spectrometry imaging data" in *Proceedings of the National Academy of Sciences*, **113**(43) 12244-12249, 2016. <https://doi.org/10.1073/pnas.1510227113>
- [13] L.V.D. Maaten, "Accelerating t-sne using tree-based algorithms" *J Mach Learn Res*, **15**, 3221-3245, 2014.
- [14] N. Pezzotti, B.P.F. Lelieveldt, L.V.D. Maaten, T. Höllt, E. Eisemann, A. Vilanova, "Approximated and user steerable tsne for progressive visual analytics" *IEEE transactions on visualization and computer graphics*, **23**(7) 1739-1752, 2016. DOI: 10.1109/TVCG.2016.2570755
- [15] L.V.D. Maaten, "Learning a parametric embedding by preserving local structure" in *Artificial Intelligence and Statistics*, pp 384-391, 2009.
- [16] A. Gisbrecht, A. Schulz, B. Hammer, "Parametric nonlinear dimensionality reduction using kernel t-sne" *Neurocomputing*, **147**, 71-82, 2015. <https://doi.org/10.1016/j.neucom.2013.11.045>
- [17] A. Boytsov, F. Fouquet, T. Hartmann, Y. LeTraon, "Visualizing and exploring dynamic high-dimensional datasets with lion-tsne" *arXiv preprint arXiv:1708.04983*, 2017. [arXiv:1708.04983](https://arxiv.org/abs/1708.04983)
- [18] S. Ingram, T. Munzner, "Dimensionality reduction for documents with nearest neighbor queries" *Neurocomputing*, **150** 557-569, 2015. <https://doi.org/10.1016/j.neucom.2014.07.073>
- [19] B. Dharamsotu, K.S. Rani, S.A. Moiz, C.R. Rao, "k-NN sampling for visualization of dynamic data using LOIN-tSNE" in *2019 IEEE 26th International Conference on High Performance Computing, Data, and Analytics (HiPC)*, pages 63-72, 2019. DOI: 10.1109/HiPC.2019.00019
- [20] L.J. Williams, H. Abdi, "Fishers least significant difference (lsd) test," *Encyclopedia of research design*, **218**, 840-853, 2010.
- [21] T. Kanungo, D.M. Mount, N.S. Netanyahu, C.D. Piatko, R. Silverman, A.Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**(7) 881-892, 2002. DOI: 10.1109/TPAMI.2002.1017616
- [22] W. Zhang, D. Zhao, X. Wang, "Agglomerative clustering via maximum incremental path integral" *Pattern Recognition*, **46**(11) 3056-3065, 2013. <https://doi.org/10.1016/j.patcog.2013.04.013>
- [23] T. Zhang, R. Ramakrishnan, M. Livny, "BIRCH: an efficient data clustering method for very large databases" *ACM Sigmod Record*, **25**(2) 103-114, 1996. <https://doi.org/10.1145/235968.233324>
- [24] J. C. Bezdek, R. Ehrlich, W. Full, "FCM: The fuzzy c-means clustering algorithm" *Computers and Geosciences*, **10**(2) 191-203, 1984.
- [25] E. Schubert, J. Sander, M. Ester, H.P. Kriegel, X. Xu, "DBSCAN revisited, revisited: why and how you should (still) use DBSCAN" *ACM Transactions on Database Systems (TODS)*, **42**(3) 1-21, 2017. <https://doi.org/10.1145/3068335>
- [26] J.C. Dassun, A. Reyes, H. Yokoyama, P.B.E. Jireh, M. Dolendo, "Ordering points to identify the clustering structure algorithm in fingerprint-based age classification" *Virtutis Incunabula*, **2**(1) 17-27, 2015.
- [27] D. Comaniciu, P. Meer, "Mean shift: A robust approach toward feature space analysis" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**(5) 603-619, 2002. DOI: 10.1109/34.1000236
- [28] U.V. Luxburg, "A tutorial on spectral clustering" *Statistics and Computing*, **17**(4) 395-416, 2007.
- [29] D.A. Reynolds, "Gaussian mixture models" *Encyclopedia of biometrics*, **741**, 2009.
- [30] B.J. Frey, D. Dueck, "Clustering by passing messages between data points" *Science*, **315**(5814) 972-976, 2007. DOI: 10.1126/science.1136800
- [31] David Sculley, "Web-scale k-means clustering" in *Proceedings of the 19th international conference on World Wide Web*, pages 1177-1178, 2010. <https://doi.org/10.1145/1772690.1772862>
- [32] E.B. Fowlkes, C.L. Mallows, "A method for comparing two hierarchical clusterings" *Journal of the American Statistical Association*, **78**(383) 553-569, 1983.
- [33] Y. Liu, Z. Li, H. Xiong, X. Gao, J. Wu, S. Wu, "Understanding and enhancement of internal clustering validation measures" *IEEE Transactions on Cybernetics*, **43**(3) 982-994, 2013. DOI: 10.1109/TSMCB.2012.2220543
- [34] S. Tsumoto, "Contingency matrix theory: Statistical dependence in a contingency table" *Information Sciences*, **179**(11) 1615-1627, 2009. <https://doi.org/10.1016/j.ins.2008.11.023>
- [35] A. Saxena, M. Prasad, A. Gupta, N. Bharill, O.P. Patel, A. Tiwari, M.J. Er, W. Ding, C. Lin, "A review of clustering techniques and developments" *Neurocomputing*, **267** 664-681, 2017. <https://doi.org/10.1016/j.neucom.2017.06.053>
- [36] A.K. Jain, M.N. Murty, P.J. Flynn, "Data clustering: a review" *ACM computing surveys (CSUR)*, **31**(3) 264-323, 1999. <https://doi.org/10.1145/331499.331504>
- [37] P. Berkhin, "A survey of clustering data mining techniques" in *Grouping multidimensional data*, pp 25-71. Springer, 2006.
- [38] A.Y. Ng, M.I. Jordan, Y. Weiss, "On spectral clustering: Analysis and an algorithm" in *Advances in neural information processing systems*, pp 849-856, 2002.
- [39] G.C. Linderman, S. Steinerberger, "Clustering with t-sne, provably" *SIAM Journal on Mathematics of Data Science*, **1**(2) 313-332, 2019. <https://doi.org/10.1137/18M1216134>
- [40] D. Shepard, "A two-dimensional interpolation function for irregularly-spaced data" in *Proceedings of the 1968 23rd ACM national conference*, pp 517-524, 1968. <https://doi.org/10.1145/800186.810616>
- [41] J. Venna, S. Kaski, "Neighborhood preservation in nonlinear projection methods: An experimental study" in *International Conference on Artificial Neural Networks*, pp 485-491, Springer, 2001.
- [42] T. Kollig, A. Keller, "Efficient multidimensional sampling" *Computer Graphics Forum*, **21**(3), 2002.
- [43] D. Dua, C. Graff, *UCI machine learning repository*, 2019.
- [44] B. Schölkopf, A. Smola, K. Müller, "Kernel principal component analysis" in *International conference on artificial neural networks*, pages 583-588, Springer, 1997.

## Citizen Behavior: The Evaluation of Complaint Application that Connected to Smart City

Evaristus Didik Madyatmadja<sup>1</sup>, Hendro Nindito<sup>1</sup>, Debri Pristinella<sup>2,\*</sup>

<sup>1</sup>Information Systems Department, School of Information Systems, Bina Nusantara University, 11480, Indonesia

<sup>2</sup>Faculty of Psychology, Atma Jaya Catholic University of Indonesia, 12390, Indonesia

### ARTICLE INFO

Article history:

Received: 21 October, 2019

Accepted: 22 February, 2020

Online: 06 July, 2020

Keywords:

e-government

complaint

Citizen

Smart City

### ABSTRACT

SIARAN is application which is created by Government of South Tangerang City, Indonesia. The application is e-government services social media based. It is designed for citizens to be able to report problems that occur around South Tangerang City. The research was conducted to find out the factors that affect the intention of SIARAN's users. The research model in this research was Theory of Planned Behavior which was combined by several variables of several acceptance models. The methods which were used to collect the data were literature review and questionnaire. The findings of the research were going to be developed and proceeded by using SEM-PLS (Structural Equation Model - Partial Least Square). The results of this research indicated that Information Design, Architecture, Security, Esthetical Design, give significant positive effects towards attitude. Whereas, Perceived Usefulness and Perceived Ease of Use did not give significant effects towards Attitude. Facilitating Condition and Participation Efficacy gave significant positive effect towards perceived behavior control. Perceived Behavioral Control and Attitude gave significant positive effect towards intention.

## 1. Introduction

As implementation steps of the smart city in South Tangerang City, the government of the city together with South Tangerang City Communication and Information Office (DISKOMINFO) developed South Tangerang Live application. The application is able to have features that which support an integrated smart city concept. One of the features is the Complaints Service (SIARAN) which is the object of this research. SIARAN is one of the parts of smart city, that is a smart government in which the citizens are able to participate to give opinions in term of public services. Due to it has social media based, SIARAN becomes e-government service which can be used on android platform. The features on SIARAN provide opportunities for citizens to make complaints related to problems which occur around South Tangerang City. The complaints will be known immediately by South Tangerang official, such that they can be followed up immediately. In the process of developing and implementing in Indonesia, it has been carried out in several cities. The cities which have applied smart city are DKI Jakarta, Bandung, Balikpapan, Makassar, and Surabaya. Smart city has 6 (six) dimensions, which are Smart

Economy (innovation and competition), Smart Mobility (transportation and infrastructure), Smart People (creativity and social capital), Smart Environment (sustainability and resource), Smart living (quality of life and culture), Smart Government (empowerment and participation) [1]. There are various reasons which cause an organization invests on information system. There are pressures for cutting costs, increasing productivity without increasing costs, and increasing product or service quality [2], [3] various organizations are beginning to see the power of social media technology in creating new information technology capabilities that will help them to develop new strategies for services as well as for products [4].

Internet and social media are technology which is predicted can change the communication process between the government and the citizen. An innovative mode of communication and open government can encourage public participation in government decision making, the statement is also supported by several studies [5]-[9]. Online Media is believed can facilitate citizen to access on government information. It will encourage to share the information and public participation on the process of decision making which is government's responsibility [7], [9]. This research is related to online communication media which takes a

\*Debri Pristinella, Jl Jend, Sudirman No. 51, Setiabudi, Jakarta Selatan, 12930 Indonesia, Email: [debri.pristinella@atmajaya.ac.id](mailto:debri.pristinella@atmajaya.ac.id)

role as public space, such that it has an impact on political communication between government and citizen.

The research on acceptance of information technology has become the focus of development model and testing in estimating intentions [2], [10], [11]. The research study evaluation of social media-based e-government service is conducted using several merging models. Based on the result of the analysis of the merging model, it will be finding out the factors which affect the level of interest in using social media-based e-government application. Previous research has focused on e-Government in general, while the novelty of this research lies in the intention to use citizens to support government-based social media.

## **2. Smart City Definition**

Smart City is able to create the components and urban infrastructure services smarter, interconnected, and efficient through the use of intelligent computer [12]. In understanding according to [13], Smart City is city that guarantees the freedom of speech, access to information and public services which are strengthened by information and communication technology. Smart City has a concept that is through smart city order, thus it facilitate the citizen in terms of getting information quickly and precisely.

IBM defines “Smart City” that using information and communication technology to understand, analyze, and integrate information from prime system in running a city [14]. Smart City is well-defined geographies, where technology such as ICT, logistic, energy production, and others works together to create benefits for citizen in term of inclusive welfare and participation, environment quality, smart development; it is managed by a well-defined set of subject, able to state the rules and policies for development and city government [15].

## **3. Social Media Based Government**

In the use of social media web and tools by citizens, business and public organization, in term of the government needs to look for the ways to take benefits from the social media’s users, in case of participation and level of involvement [16]. As a public e-government site, this is certainly important in its efforts so that citizens always visit the site. Meanwhile in a sense, social network is a site where people interact and enter the site in order that the government and the citizen are able to have closer relationship [17]. The increase of the performance and welfare of government employees or civil servants who have worked serving the citizens is one of the objectives of e-government applications [18]. The performance of public service is getting more professional and able to satisfy the customers, is citizens’ expectation of as the customers of public services [19].

The government changes the way of communication to the citizens by using communication technology, such as internet and social media. It is because, the citizens have decreased the level of trust in the government. The increase of government transparency makes citizens more motivated to participate in government’s decision-making process, therefore, several studies suggest the innovative communication mode [9]. It is expected that the efforts will strengthen public trust and democratic values, effective and operational efficiency. Transparency and government response, as well as citizens’ participation based on

innovative communication technology will contribute for the main foundation of democracy [9]. In terms of the main foundations of democracy, government transparency and government-citizen interaction are important in increasing citizens’ trust in government [6], [8]. In this case, social media is considered as the right solution. By direct interaction, the citizens will trust in government increasingly. Therefore, the citizens have a better chance to interact directly through social media channels which are managed by the government [9].

The foundation of democracy is citizens’ participation. In the process of decision making, the citizens are able to give suggestions and comments to the government by following government’s social media account. The citizens’ perception towards transparency, efficiency, and corruption is able to be influenced by social media and cellular technology. Hence, the technology becomes important [20]. The result of this research indicates that the citizens’ perception towards transparency, efficiency, and corruptions have increased due to interactions between the government and the citizens which is supported by information technology through social media, websites, and other technologies.

## **4. Research Methodology**

The main model that was used in this research was TPB (Theory of Planned Behavior), that was added with several variables. Several theories and models which support, were TAM, TRI, UTAUT, TPB, and Delone. The questions on questionnaire distributed contained certain values based on variables and indicators in the TPB (Theory of Planned Behavior). TPB is a theory in social psychology that predicts and refutes behavior that can be done by someone [21]. This theory states that one's intention to conduct behavior (behavior intention) can be a strong predictor of actual behavior (actual behavior).

The TPB theory suggests that a person's behavior intention can be a powerful predictor of the actual behavior. BI can be understood as the degree that a person is willing to try or perform certain behaviors and is determined by three conceptually independent determinants: Attitude, Subjective Norm, and Perceived Behavioral Control.

TRI measures individual readiness to use new technologies that use four personality traits that are optimism, innovativeness, discomfort, and insecurity.

Technology Acceptance Model (TAM) is a theory about information system that models the process of how users are willing to accept and use the technology. This model can explain that when users use information systems, several factors can influence their decision about how and when to use the information system.

UTAUT that integrated eight models of behavioral theory. Those theories are Theory of Reasoned Action (TRA), Adoption Model Technology (TAM), Motivational Model (MM), Theory of Planned Behavior (TPB), Combined TAM and TPB (C-TAM - TPB), PC Utilization Model, Innovation Theory Diffusion (DOI), and Social Cognitive Theory.

Delone argued that the success of IT implementation can be assessed or predicted from user acceptance and user satisfaction using IT, where external factors influence user acceptance and satisfaction.



4.1. Hypothesis

Based on research model framework above, it can be formulated as follows:

Table 1: Regression Hypothesis Test (Simple Regression)

#	Hypothesis
H <sub>1</sub>	Perceived Usefulness variable has effect on variable Attitude in Siaran application usage.
H <sub>2</sub>	Variabel Perceived Ease of Use Variable has effect on Attitude Variable in Siaran application usage.
H <sub>3</sub>	Navigation variable has effect on Attitude variable in Siaran application usage.
H <sub>4</sub>	Accessibility variable has effect on Attitude variable in Siaran application usage.
H <sub>5</sub>	Privacy variable has effect on Attitude variable dalam in Siaran application usage.
H <sub>6</sub>	Security variable has effect on Attitude variable in Siaran application usage.
H <sub>7</sub>	Design Information Architecture variable has effect on variabel Attitude in Siaran application usage.
H <sub>8</sub>	Design Aesthetic Values variable has effect on Attitude variable in Siaran application usage.
H <sub>9</sub>	Innovativeness variable has effect on Attitude variable dalam Siaran application usage.
H <sub>10</sub>	Precision variable has effect on variabel Attitude in Siaran application usage.
H <sub>11</sub>	System Integration variable has effect on Attitude variable in Siaran application usage.
H <sub>12</sub>	Trust in Government variable has effect on variabel Attitude in Siaran application usage.
H <sub>13</sub>	Participation Efficacy variable has effect on variabel Perceived Behavioral Control in Siaran application usage.
H <sub>14</sub>	Facilitating Conditions variable has effect on Perceived Behavioral Control variable in Siaran application usage.
H <sub>15</sub>	Attitude variable has effect on Behavior Intention variable in Siaran application usage.
H <sub>16</sub>	Perceived Behavioral Control variable has effect on Behavior Intention variable in Siaran application usage.

4.2. Reliability Test

Reliability test can be conducted by using Cronbach’s alpha which assumes that all indicators are equally reliable. In this case, these indicators which reflect a construct must have the same outer loading value. However, due to the value of Cronbach’s alpha tends to understate internal consistency reliability, therefore, another method is introduced which is composite reliability. Cronbach's alpha values and composite reliability of 0.6 to 0.7 can be accepted [22]. Based on Table 2. Cronbach's Alpha value of all statements above 0.6 can thus be concluded that

all statements have been reliable as research instruments. The following below are the results of the Cronbach's alpha test and composite reliability:

Table 2: Cronbach’s Alpha

Construct	Cronbach’s Alpha	Composite Reliability
Perceived Usefulness	0.722	0.843
Perceived Ease of Use	0.761	0.862
Navigation	0.823	0.893
Accessibility	0.764	0.863
Privacy	0.646	0.789
Security	0.639	0.766
Design Information Architecture	0.806	0.885
Design Aesthetic Values	0.771	0.845
Innovativeness	0.851	0.91
Precision	0.669	0.856
System Integration	0.676	0.822
Trust in Government	0.89	0.922
Participation Efficacy	0.866	0.934
Facilitating Conditions	0.622	0.841
Attitude	0.714	0.84
Perceived Behavioral Control	0.691	0.827
Behavior Intention	0.806	0.873

4.3. Hypothesis Test

Based on the results of research which has been discussed above, hypothesis test can be conducted. The test is to determine the hypothesis which proposed previously, whether it can be accepted or not. To determine the effect or coefficient of relationship between variables can be seen from the original sample value, if the value positive or greater than 0 then the relationship has a positive effect. Vice versa if the value is negative or below 0, then the relationship has a negative influence. To analyze whether the hypothesis can be accepted or rejected, the standard used as a reference is the path coefficient value above 0.1. In addition to determining the coefficient, also checking whether the relationship is significant or not. According to [23] if coefficient lane is value above 0.1 and T-statistics value is greater than 1.96, it indicates significance in hypothesis testing.

Table 3: Hypothesis Test Results

Hypothesis	Lane Connection	Original Sample	T-Statistic	Inference
H1	PU > AT	-0.027	0.369	Not significant
H2	PEU > AT	0.195	2.701	Significant
H3	N > AT	-0.075	0.957	Not significant
H4	A > AT	0.081	0.848	Not significant
H5	PV > AT	0.016	0.226	Not significant
H6	SC > AT	0.257	4.495	Significant
H7	DIA > AT	0.046	0.559	Not significant
H8	DAV > AT	0.118	1.964	Significant
H9	IV > AT	0.039	0.46	Not significant

<b>H10</b>	PR > AT	0.228	2.333	<b>Significant</b>
<b>H11</b>	SI > AT	0.115	1.954	<b>Significant</b>
<b>H12</b>	TIG > AT	0.145	1.965	<b>Significant</b>
<b>H13</b>	PE > PBC	0.168	2.318	<b>Significant</b>
<b>H14</b>	FC > PBC	0.4	6.654	<b>Significant</b>
<b>H15</b>	AT > IT	0.594	9.653	<b>Significant</b>
<b>H16</b>	PBC > IT	0.254	3.258	<b>Significant</b>

## 5. Discussion

The following is a discussion of the test results based on the researcher's hypothesis:

- Perceived Usefulness towards Attitude. The results of data process on this research indicates that Perceived Usefulness (PU) factors does not have effect on Attitude (AT) significantly. The result is in line with the research which was conducted by [24]. It states that perception of usability has a poor relationship with user attitudes. The results of research that have been done, the users make complaints more concerned with the final results rather than increasing the speed of complaints.
- Perceived Ease of Use towards Attitude. The results of this data process in this research indicates that Perceived of Use (PEU) factor has a significant effect towards Attitude (AT). The result is in line with the research which was conducted by [25]. The research which has been conducted states that ease of use affects user behavior.
- Navigation towards Attitude. The Result of this data process indicate that Navigation (N) factor does not have significant effect towards Attitude (AT). The result is in line with the research which was conducted by [26].
- Accessibility towards Attitude. The results of data process in this research indicate that Accessibility (A) factor does not have significant effect towards Attitude (AT). The result is in line with the research which was conducted by [27]. The research that has been conducted states that the users are more concerned with the final results compared to the access speed. They prefer to make another complaint via email.
- Privacy towards Attitude. The results of data process in this research indicate that Privacy (PV) factor does not have significant effect towards Attitude (AT). The result is in line with the research which was conducted by [28]. In this research, the users feel privacy is not an important thing, due to the government has guaranteed to protect their privacy (trust the government).
- Security towards Attitude. The results of this data process in this research indicates that Security (SC) factor has a significant effect towards Attitude (AT). The result is in line with [29], [30] that state security factor has big effect towards attitude.
- Design Information Architecture towards Attitude. The results of data process in this research indicate that Design Information Architecture (DIA) factor does not have significant effect towards Attitude (AT). In this research the citizens tend to prioritize the functional system rather than the organizational structure of the complaint application interface.

- Design Aesthetic Values towards Attitude. The results of this data process in this research indicates that Design Aesthetic Values (DAV) factor has a significant effect towards Attitude (AT). The result is in line with [31].
- Innovativeness towards Attitude. The results of data process in this research indicate that Innovativeness (IV) factor does not have significant effect towards Attitude (AT). The result is in line with the research which conducted by [32]. The research which has been conducted stated that the users are more concern with the final result rather than application innovation. They feel that the complaint application is sufficient for their needs.
- Precision towards Attitude. The results of this data process in this research indicates that Precision (PV) factor has a significant effect towards Attitude (AT). The result is in line with [33].
- System Integration towards Attitude. The results of this data process in this research indicates that System Integration (SI) factor has a significant effect towards Attitude (AT). The result is in line with [34].
- Trust in Government towards Attitude. The results of this data process in this research indicates that Trust in Government (TG) factor has a significant effect towards Attitude (AT). The result is in line with the research which has been conducted by [35].
- Participation Efficacy towards Perceived Behavior Control. The results of this data process in this research indicates that Participation Efficacy (PE) factor has a significant effect towards Perceived Behavior Control (PBC). The result is in line with [25], [36] which stated that the level of participation is one of the determinants of perceived use behavior
- Facilitating Conditions towards Perceived Behaviour Control. The results of this data process in this research indicates that Facilitating Conditions (FC) factor has a significant effect towards Perceived Behavior Control (PBC). The result is in line with [25], [30] which stated that the condition of the facility is one of the determinants of perception of usage behavior
- Attitude towards Behavior Intention. The results of this data process in this research indicates that Attitude (AT) factor has a significant effect towards Behavior Intention (BI). The result is in line with [25] which stated that attitude factor is one of the main determinants and has a positive relationship to intention.
- Perceived Behavior Control towards Behavior Intention. The results of this data process in this research indicates that Perceived Behavior Control (PBC) factor has a significant effect towards Behavior Intention (BI). The result is in line with [25] which stated that the perception of usage behavior has a positive relationship with intention.

## 6. Conclusion

This Research aims to find out the factors which can affect the intention to use SIARAN e-government service by using Theory of Planned Behavior which is added by several variables. Several conclusions are obtained based on the results of testing between variables using the Structural Equation Modeling (SEM) method.

Therefore, the government is expected to improve services by increasing the interaction and integration, and communication between the community and the government through the exchange of ideas. This study has identified 17 main variables that will increase public confidence in e-government services. The results of the hypothesis test are:

- Accessibility factor has no significant effect on Attitude
- Attitude factor has a significant positive effect on Intention
- Aesthetic Design factor has a significant effect on Attitude
- Design Information Architecture has no significant effect on Attitude
- Facilitating Condition factor has a significant positive effect on Perceived Behavioral Control
- Innovativeness factor has no significant effect on Attitude
- Navigation factor has no significant effect on Attitude
- Participation Efficacy factor has a significant positive effect on Perceived Behavioral Control
- Perceived Behavioral Control factor has a significant positive effect on Intention
- Perceived Ease of Use factor has a significant effect on Attitude
- Perceived Usefulness factor has no significant effect on Attitude
- Precision factor has a significant positive effect on Attitude
- Privacy factor has not significant effect on Attitude
- Security factor has a significant positive effect on Attitude
- System Integration factor has a significant positive effect on Attitude
- Trust in Government factor has a significant positive effect on Attitude

In the end, the government must consider the unmet user needs and use them in e-government strategies. This aims to increase public trust in the government.

### Conflict of Interest

The authors declare no conflict of interest.

### Acknowledgment

This work is supported by Directorate General of Research and Development Strengthening, Indonesian Ministry of Research, Technology, and Higher Education, as a part of Penelitian Terapan Unggulan Perguruan Tinggi Research Grant to Binus University titled "Pengembangan Aplikasi e-Government Berbasis Media Sosial di Indonesia" or "Social Media Based Government Application Development in Indonesia" with contract number: 12/AKM/PNT/2019 and contract date: 27 March 2019.

### References

- [1] Kurnaedi, Penerapan "Live Smart City Kota Tangerang Selatan", 2017.
- [2] Legrisa, Inghamb, & Collette, "Why do people use information technology? A critical review of the technology acceptance model", *Information & Management* **40** 191–204, 2003.
- [3] Bharadwaj, Sawy, Pavlou dan Venkatraman, "Digital Business Strategy: Toward a Next Generation of Insights" *MIS Quarterly*, **37** (2), 471-482, 2013.
- [4] Hansen, Saridakis, & Benson, "Risk, Trust, and the Interaction of Perceived Ease of Use and Behavioral Control in Predicting Consumers", *Use of Social*

- Media for Transactions, PII: S0747-5632(17)30640-4, 2017. <https://doi.org/10.1016/j.chb.2017.11.010>
- [5] Harris, McKenzie, & Rentfro, "Building Trust or Mistrust: The Search for Performance Reporting from a Citizen's Perspective", 2009.
- [6] M. Lee, G. Neeley, & K. Stewart, "The practice of government public relations", FL: CRC Press, 2012.
- [7] E.M. Searson, & M.A. Johnson, "Transparency laws and interactive public relations: An analysis of Latin American government Web sites", *Public Relations Review*, **36**(2), 120–126, 2010. <https://doi.org/10.1016/j.pubrev.2010.03.003>
- [8] E. Welch, & C.C. Hinnant, "Internet use, transparency, and interactivity effects on trust in government", *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, HICSS 2003. <https://doi.org/10.1109/HICSS.2003.1174323>
- [9] H. Hong, "Government websites and social media's influence on government- public relationships", *Public Relations Review*, **39**(4), 346–356, 2013. <https://doi.org/10.1016/j.pubrev.2013.07.007>
- [10] Olson and Boyer, "Factors influencing the utilization of Internet purchasing in small organizations", *Technical Note. Information & Management* **40** 191–204, 2003. [https://doi.org/10.1016/S0272-6963\(02\)00089-X](https://doi.org/10.1016/S0272-6963(02)00089-X)
- [11] Pijpers, Bemelmans, Heemstra, "Senior executives' use of information technology", *Information and Software Technology* **43** 959-971, 2001.
- [12] Washburn, Sindhu, "Helping CIOs Understand Smart City Initiatives", 2010.
- [13] H. Partridge, "Developing a human perspective to the digital divide in the smart city", In *Proceedings of the Biennial Conference of Australian Library and information Association (Queensland, Australia, Sep 21-24)*, 2004. Available at <http://eprints.qut.edu.au/1299/1/partridge.h.2.paper.pdf>.
- [14] V.N. Keshamoni, D.D. Kang, M. Inukollu, "Factors Influencing Quality of Mobile Apps: Role of Mobile App Development Life Cycle", *International Journal of Software Engineering & Applications (IJSEA)*, **5**(5), 15-34, 2014. doi: <https://doi.org/10.5121/ijsea.2014.5502>
- [15] R.P. Dameri, "Searching for smart city definition: a comprehensive proposal", *International Journal of Computers & Technology*, **11**(5), 2544–2551, 2013.
- [16] R.T. Khasawneh, & E.A. Abu-shanab, "E-Government and Social Media Sites: The Role and Impact", *World Journal of Computer Application and Terchnology*, **1**(1), 10–17, 2013. <https://doi.org/10.13189/wjcat.2013.010103>
- [17] S. Atari, "Web 2.0 in Next-Generation Government and Governance A Middle East Point of View", February, 2011.
- [18] L. Hardjaloka, Studi Penerapan, "E-Government di Indonesia dan Negara Lainnya Sebagai Solusi Pemberantasan Korupsi di Sektor Publik", *Rechtsvinding*, **3**(3), 435–452, 2014.
- [19] Mustafid, & Sugiharto, A. Analisa, "Implementasi e-Government Untuk Pelayanan Publik di Kota Semarang", *Riptek*, **8**(2), 35–50, 2014.
- [20] R.S. Valle-Cruz, "Social Media in Local Governments in Mexico: A Diffusion Innovation Trend and Lessons", 2016. <https://doi.org/10.1007/978-3-319-17722-9>.
- [21] I. Ajzen, "The Theory of Planned Behavior. Organizational Behavior And Human Decision Prodeses", 179–211, 1991. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- [22] J.F. Hair, G.T. Hult, C.M. Ringle, & M. Sarstedt, "A Primer on Partial Least Squares Equation Modeling (PLS-SEM)", Thousand Oaks, CA: SAGE Publications, 2014.
- [23] L. dan-Ghozali, "Partial least squares: konsep, teknik dan aplikasi SmartPLS 2.0 M3 untuk penelitian empiris", 2012.
- [24] K. Nischay C. Gulati, K. Sharma & R. Jain, "To Assess the Effect of Perceived Usefulness, Ease of Use, Security & Privacy on Customer Attitude and Customer Adaptation: A Study of E-banking in India", 2017.
- [25] S.Y. Hung, C.M. Chang, & S.R. Kuo, "User acceptance of mobile e-government services: An empirical study", *Government Information Quarterly*, **30**(1), 33–44, 2013. <https://doi.org/10.1016/j.giq.2012.07.008>
- [26] Y. Kang, Y. JinKim, "Do visitors' interest level and perceived quantity of web page content matter in shaping the attitude toward a web site?", *Decision Support Systems*, **42**(2), 1187-1202, 2006. <https://doi.org/10.1016/j.dss.2005.10.004>
- [27] A. Irawati, D. Putra, "Analisis Technology Acceptance Model Dalam Memahami Niat Perilaku Mahasiswa Untuk Menggunakan E-Learning", *Prosiding Seminar Bisnis & Teknologi*, pp 177-189, 2014.
- [28] M. Jamalzadeh, N. Behravan, R. Masoudi, "An Empirical Study of Email-Based Advertisement and its Influence on Consumers' Attitude", *International Review of Management and Marketing* **2**(3), pp.130-138, 2012.
- [29] N. Jahangir & N. Begum, "The role of perceived usefulness, perceived ease of use, security and privacy, and customer attitude to engender customer

adaptation in the context of electronic banking”, *African Journal of Business Management* **2**(1), pp. 032-040, 2008.

- [30] Chiang, “Trust and security in the e-voting system”, *Electronic Government, An International Journal*, **6**(4), 2009.
- [31] Ivan Wen, “Factors affecting the online travel buying decision: a review”, *International Journal of Contemporary Hospitality Management*, **21**(6), pp. 752-765, 2009. <https://doi.org/10.1108/09596110910975990>
- [32] Hung, Chang, & Yu, “Determinants of user acceptance of the e-Government services: The case of online tax filing and payment system”, *Government Information Quarterly* **23**(1), Pages 97-122, 2006. <https://doi.org/10.1016/j.giq.2005.11.005>.
- [33] D. Rosna Y. Toin, “Faktor-faktor yang Mempengaruhi Kepuasan Penggunaan E-Procurement oleh Penyedia Barang dan Jasa (Studi Kasus di Pemerintahan Kota Surakarta)”, *Jurnal EKA CIDA* **1**(1) pp 16-29, 2016.
- [34] Mostafa, Seyedshahab, Narges, Maryam, Farokh, “Physicians' attitudes toward integrating consultation-liaison psychiatric services in four major general hospitals in Tehran”, *General Hospital Psychiatry* **37**(5), Pages 456-458, 2015. <https://doi.org/10.1016/j.genhosppsych.2015.05.006>.



## Ontologic Design of Software Engineering Knowledge Area Knowledge Components

Bulat Kubekov<sup>1,2</sup>, Anar Utegenova<sup>\*,1,3</sup>, Leonid Bobrov<sup>2</sup>, Vitaliy Naumenko<sup>1,2</sup>, Aibek Ibraimkulov<sup>1</sup>

<sup>1</sup>Institute Information and Computational Technologies, CS MES RK, 050056, Kazakhstan

<sup>2</sup>Faculty of Information Technology Novosibirsk state university of Economics and Management, Novosibirsk, 630000, Russia

<sup>3</sup>Department of Computer and program engineering, Turan University, Turan, 050056, Kazakhstan

### ARTICLE INFO

Article history:

Received: 25 December, 2019

Accepted: 23 May, 2020

Online: 06 July, 2020

Keywords:

Knowledge area engineering

Support concept

Support concept ontology

Knowledge trend

Knowledge content

Educational resources semantic context

Semantic interoperability

### ABSTRACT

The article sets forth the solution of the educational resources semantic context knowledge components development task, based on the learning technology project-oriented concepts, graduate's competency model and ontological; engineering. The being considered ontology model and knowledge display formalisms allow, firstly, relevantly image the educational resources semantic context in the support concepts ontology format, and their specifications in the form of knowledge expressions and knowledge components, and secondly, secure the knowledge components semantic interoperability withing the frame of their usage in educational environment and systems. The design process of disciplines knowledge content degree programs and individual learning guidelines is connected with specifying the knowledge content frame, the configuration of which is defined with competencies signatures and smart-contract parameters.

## 1. Introduction

Knowledge area engineering is the technique, which allows fulfill the educational resources and degree programs working out, based on the knowledge components collection.

Knowledge area engineering gives a possibility to define the knowledge components family members general features and distinct characteristics of each of them [1], [2]. In our researches, applying the knowledge area engineering is connected with developing the knowledge components and their repeated usage upon generating the degree programs new disciplines, taking into account new technological and instrumental trends of software.

In our creations the knowledge area engineering is understood as the activity on the analysis and representation of the semantic context of educational resources with necessary and sufficient support concepts set, their formalization and specification in the form of ontology model and knowledge expressions.

Being considered methodology of the SWEBOK (Software Engineering Body of Knowledge) knowledge domain knowledge

\*Corresponding Author: Utegenova Anar, Institute Information and Computational Technologies CS MES RK, [utegenova77@mail.ru](mailto:utegenova77@mail.ru)

components ontological design is based on the educational resources engineering and it is targeted at fulfilling the implementation of knowledge trend and knowledge content of disciplines working program syllabuses and software engineering degree programs, based on the knowledge components family. Requirements generality to separate components, in one knowledge area framework provides such components characteristics similarity. Therefore, the knowledge, having been worked out within SWEBOK knowledge domains, might be used in future upon preparing the new courses, including adjacent directions, as well, upon formulating individual degree programs and corresponding curricula [3].

Software engineering component is some working mechanism, the development of which is executed in reliance on the repeated usage. Such a working means in our researches is a knowledge component, representing a composition from the knowledge expressions and having the interface in the smart-contract form.

Models, having been introduced, knowledge representation arrangement formalisms, thus, allow maintain image pertinency of SWEBOK knowledge domains educational resources semantic context in the form of support concepts ontology, as well, their

semantic interoperability, upon designing the knowledge content of degree programs disciplines and individual learning guidelines.

Semantic interoperability is the possibility to interpret the educational resources semantic context by means of support concepts ontologies and repeatedly used knowledge components and to use them further in the new contexts, linked with design of software engineering basic and profile disciplines knowledge content.

Outlined knowledge components configuration conceptions are taken as a basis of being developed educational milieu, in the form of the Web-application with an adaptive cross-browser makeup, database and possibility of integrating into other informational-educational systems, in PDF, OWL and RDF formats.

## **2. Software engineering knowledge areas**

The world leading universities develop for their students the courses on the software engineering, which supplement existing programs on informatics and computer engineering. In order to formulate an efficient curriculum compiling principles, IEEE Computer Society and ACM have elaborated, in the frame of a larger scale project Computing Curriculum, the recommendations set Software Engineering 2004 (SE2004) [4].

SE2004 describes in detail the knowledge (Software Engineering Education Knowledge, SEEK), which in the academic process shall be rendered by the universities and colleges within the course on software engineering.

In the version (2004) SWEBOK breaks down the knowledge on software engineering into fifteen knowledge areas.

It should be noted, that SWEBOK, as distinguished from SEEK, envisages deeper knowledge, as it assumes the software engineering knowledge of a specialist, having several years' experience, while SEEK supposes the knowledge, necessary to render for the trainees. On the other hand, SEEK stipulates wider range of topics, which SWEBOK deliberately has left out of view, in order to concentrate at software engineering basics.

Despite the noted difference in interpreting and applying the knowledge areas, the methodology, used in our researches, anticipates usage of such conceptions and mechanisms, which allow on a consistent basis display and formulate the knowledge models with subsequent design of the knowledge components.

The article herein shows forth the methodology of arranging and representing the SWEBOK knowledge domain semantic context knowledge with the support concepts system, support concepts representation with the ontologies from identifying and concrete concepts and support concepts ontology specification in the knowledge expressions format. The methodology's final aim consists in SWEBOK knowledge areas representation in the form of knowledge components repositories and with the aim of their further usage for designing the degree programs disciplines knowledge content and individual learning guidelines.

For configuring the knowledge trajectory there is used a smart-contract, preconditions and postconditions of which are competencies signatures.

## **3. Engineering education project-oriented technology**

One of the important principles of programs upgrade for preparing the engineering staff to engineering activity is usage of CDIO – worldwide initiative approach, MIT [5,6]. CDIO approach gives possibility to form the degree programs disciplines optimal structures, and jointly with competency approach, to optimize the knowledge content and engineering education level programs quality[7]-[10].

CDIO declared aim is in the fact, that a higher school graduate shall be able to come up with a new product or new technical idea, fulfill all design works on their implementation, to give useful instructions to those, who will deal with them, introduce into production and apply the result. CDIO ideology and standards are applicable to training any engineering profile specialists as they are project-oriented learning technologies, focused at a student and integrated into the problems and experience of real production.

To reflect vocational specifics, at the level of graduates skills, abilities and capabilities in software engineering, the degree program shall be governed, firstly, with fifteen SWEBOK knowledge areas, defining basic engineering knowledge, secondly, with CDIO stages and standards, in which definite learning outcomes are represented as a detailed list of competences or qualification requirements.

Knowledge, skills and personal features are determined and codified in the list of planned learning outcomes in the so-called plan of CDIO Syllabus, which specifies, what the trainees shall know and be able upon completing their degree program.

As the educational needs and educational resources determining complex technology is the base for learning planning outcomes, in our researches the CDIO Syllabus is a plan, in which the main learning results are given as professional, basic and additional competencies signatures, used for defining the disciplines knowledge trend or learning trajectories.

As it has been noted, the CDIO Syllabus plan includes the list of employers requirements to engineering education, which, in respect to, learning project-oriented technology is interpreted in the form of competencies expressions of the CDIO project execution stages.

On the one hand, a higher school teacher's task consists in transferring the big subject knowledge volume to the students. On the other hand, engineers need necessary various personal and interpersonal competencies, as well, skills for objects, processes and systems creation, which will allow them work in a real team of engineers and benefit the society. CDIO approach has been developed with the aim to eliminate the conflict thereof and satisfy all needs of the trainees.

CDIO Syllabus is, as well, the main document for engineering educational programs reforming. From vocational engineers' point of view, it represents the employers list of requirements to engineering education level. With that, higher schools' teachers can consider the CDIO Syllabus plan to be a system of the main learning and competencies results, which shall own the graduates. List of students' knowledge, skills and competencies shall consist of sufficient general formulations, in order to apply it to many engineering areas. In the meantime, it shall be worked out thoroughly enough to be useful upon degree programs disciplines curricula planning and concrete specialty educational process assessing.

In our researches for the knowledge defining there is used the ontological engineering, the result of which is SWEBOK knowledge domain semantic imaging by means of support concepts ontologies, on the base of which there formed repeatedly used knowledge components, further applied for designing the degree program disciplines knowledge content and learning trajectory on the software engineering.

As the properties, endowed to the CDIO Syllabus plan are quite critical for their certain use upon planning the learning outcomes, the properties thereof are considered to be an integral part of degree programs learning trajectories design process. In our researches, under the properties we understand the smart-contract prerequisites and postrequisites, where prerequisites are statements on the key competencies, indispensable for trainee’s cognitive activity, and postrequisites are key competencies if the CDIO Syllabus plan has been acquired by the trainees upon successful mastering the knowledge components.

Smart-contract of the CDIO Syllabus plan initial level, has the following format:

$$\{P\} <CDIO Syllabus > \{Q\} \quad (1)$$

where {P} – prerequisites, connected with the statements on competencies, necessary for successful learning, and {Q} - postrequisites, which define the statements on competencies, which determined with success and quality of learning planned outcomes achievement.

Software engineering postrequisites are the statements, connected with trainee’s fundamental skills and knowledge, having been formulated as following competencies:

- Possessing the software engineering knowledge and skills, necessary to start the practical work.
- Individual or team work on qualitative software creation.
- Designing on one or several subject areas, using software engineering approaches, uniting “ethic, social, legal and economic interests”.
- Demonstrating such skills, as interpersonal communication, work efficient methods, leadership and intercourse.
- Studying the new models, methods and technologies as far as they appear.

Further decomposition of the CDIO Syllabus plan and smart-contract properties is linked with the structure of the educational process within the learning courses framework or individual learning trajectories. For example, for individual learning trajectories case the smart-contract will include the set of knowledge components (KC, Knowledge Components) and learning scenario configuration parameters:

$$\{P\} < \{KC\}, Col, Lev > \{Q\} \quad (2)$$

where, a parameter Lev assigns required for a trainee level of professional, basic or additional knowledge components competencies, and a parameter Col – references to those knowledge components, which are indispensable for a trainee to master successfully the knowledge component from the set {KC}.

#### 4. Formalisms and specifications of being used conceptions and mechanisms of their implementation

Ontology model is defined as multitudes trinary:  $Om = <C, R, F >$ , where C – educational space concepts (terms) multitude; R –

multitude of ratios between the concepts; F - multitude of interpretation functions, the definitions of which are assigned at ratios between concepts in ontology.

Knowledge specifications language formalisms: dyadic relations: “composition - symbol ‘\*’ ”- aggregation form with clearly-cut expressed relations of possession and coincidence of the life cycle between the concepts in ontology; “aggregation – symbol ‘+’ ” and "alternative selection– symbol ‘~’ ” - relations, by means of which there is imaged the semantic identity between concepts in the ontology [11-15].

Support concept is the basic abstraction of the educational milieu, reflecting the semantic generality and typical individual properties of the concepts of the given environment.

Identifying concepts are ontology concepts, by means of which the support concepts semantic and distinctive features, in reference to other educational space support concepts are defined.

Concrete concept is an ontology notion, which in the own identifying concept context configurate the support concept by means of typical, or different combination of child concepts, possessing clear and explicit description of semantic and distinctive properties of ontology support concept.

Knowledge expression is the specification of the support concept ontology by means of identifying and concrete concepts sequence with definite relations over the concepts: composition, aggregation and alternative selection. Knowledge expression implies the truth of conclusion on the support concept, that is, provided “parcel is the knowledge expression” is true, then the same is true for “conclusion is a support concept”.

For instance, specification of support concept ontology  $C_i$  is represented in the form of the following knowledge expression:

$$C_i \leq *C_{i,1} * C_{i,2} (*C_1 \sim C_2) * C_{i,3} + C_{i,4}; \quad (3)$$

From expression (3) it is seen, that the support concept ontology has a hierarchical structure, consisting of two levels, in which the first level concepts are identifying concepts -  $C_{i,1}$ ,  $C_{i,2}$ ,  $C_{i,3}$  and  $C_{i,4}$ , each of which might have the level from identifying concepts. For example, an identifying concept  $C_{i,2}$  has two concrete concepts from an obligatory concept  $C_1$  and alternative non-obligatory concept  $C_2$ .

Developed software editor allows forming the support concepts ontology in the form of knowledge expressions and visualizing them in the form of relational graph of support concept ontology  $C_i$ , as it is shown on the Figure 1.

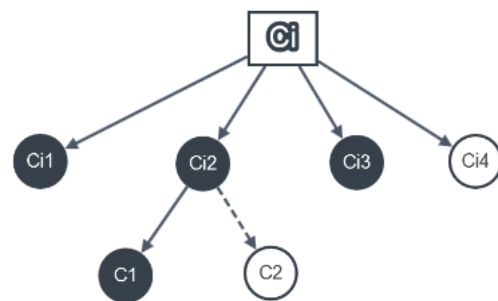


Figure 1: Relational graph of support concept ontology  $C_i$

As it has been noted above, the CDIO Syllabus plan includes the employers' list of requirements to engineering education, which, in reference to the project-oriented teaching technology are interpreted in the form of competencies expression of the CDIO project execution stages.

Competencies expression (ec - expression of competences) is the specification of CDIO stage competency model, in the form of professional, basic and additional competencies signatures.

Let's give an example of the CDIO first stage competencies expressions – Conceive stage, connected with an analytical type of professional activity:

ec.PrC.Conceive  $\leq$  \*PrC<sub>1</sub>(\*P<sub>1</sub>\*P<sub>2</sub>)\*PrC<sub>2</sub>\*PrC<sub>3</sub>; - professional competencies

ec.BaC.Conceive  $\leq$  \*BaC<sub>1</sub>~+BaC<sub>2</sub>\*BaC<sub>3</sub>(\*B<sub>1</sub>~+B<sub>2</sub>)+BaC<sub>4</sub>; - basic competencies

ec.AdC.Conceive  $\leq$  \*AdC<sub>1</sub>+AdC<sub>2</sub>\*AdC<sub>3</sub>~+AdC<sub>4</sub>; - additional competencies.

In the above examples under competencies we understand the dynamic knowledge aggregate, reflected in the support concepts, application of which is necessary to implement the project solutions in compliance with professional activity types of the CDIO stages.

Knowledge component (KC, Knowledge Component) is a composition of the knowledge expressions. Knowledge component possesses a clearly-cut definite smart-contract, by means of which the rules of its implementation and configuration management are represented upon the learning scenario designing. Absolutely all its dependencies on the environment have been described within the smart-contract framework.

Knowledge component is the one, from which there assembled the learning scenario family members. Constructively, the knowledge component is oriented to the knowledge duplexing minimization, raising the compatibility and, consequently, to the highest favor of repeated usage.

Knowledge trend is the knowledge model, being the composition of professional, basic and additional competencies of the CDIO stage competency model. Mastering each of the competency models assumes studying the support concepts set. Signatures of professional, basic and additional knowledge trend competencies represent learning planned outcomes - the CDIO Syllabus plan.

Knowledge component is the specification of the support concepts ontology of professional, basic and additional competencies in the knowledge expressions form. The knowledge content, therefore, defines the knowledge, the degree program discipline's working program or learning scenario shall possess [12-19].

### 5. Example of ontological engineering knowledge area “Software Configuration Management”

Proceeding from ontological engineering knowledge area “Software Configuration Management”, there have been selected six support concepts, representing the given knowledge area semantic context SWEBOK. As an example, we describe only one of the six support concepts of the field of knowledge «Software Configuration Management».

Support concept C<sub>1</sub> - SCM-process management.

SCM-activity, connected with product professional and integrity control via its elements identification, management and control over the changes, as well, checking and accountability according to configuration information.

Identifying concepts of the given support concept are:

C<sub>1.1</sub> - Organizational context SCM.

C<sub>1.2</sub> - SCM constraints and rules.

C<sub>1.3</sub> – Planning in SCM. The given identifying concept is the second level ontology semantic context, including the following concrete concepts: C<sub>1</sub>-Organization and duties; C<sub>2</sub>-Resources and schedule; C<sub>3</sub> – Tools and implementation; C<sub>4</sub> - Suppliers/contractors control; C<sub>5</sub> – Interfaces control.

C<sub>1.4</sub> - Configuration management plan.

C<sub>1.5</sub> - SCM-process execution control. The given identifying concept is the second level ontology semantic context, including the following concrete concepts: C<sub>1</sub> - Metrix and quantitative assessment process in SCM; C<sub>2</sub>- Audit in SCM framework.

Knowledge expression of the support concept “SCM-process management” is:

$$C_1 \leq *C_{1.1} * C_{1.2} * C_{1.3} (*C_1 * C_2 * C_3 + C_4 * C_5) * C_{1.4} * C_{1.5} (*C_1 * C_2); (4)$$

Thus, ontological engineering consists in displaying the SWEBOK knowledge areas semantic context, with necessary and enough support concepts set and with modeling the support concepts in ontology form-hierarchical structures from identifying and concrete concepts of the knowledge area herein.

To visualize the support concept ontology there has been used the software editor, allowing conduct the knowledge expression editing and fulfill the ontology assembly in relational graph view. For instance, a relational graph of the support concept ontology C<sub>1</sub> - SCM-process management, will have the view, as shown on the Figure 2.

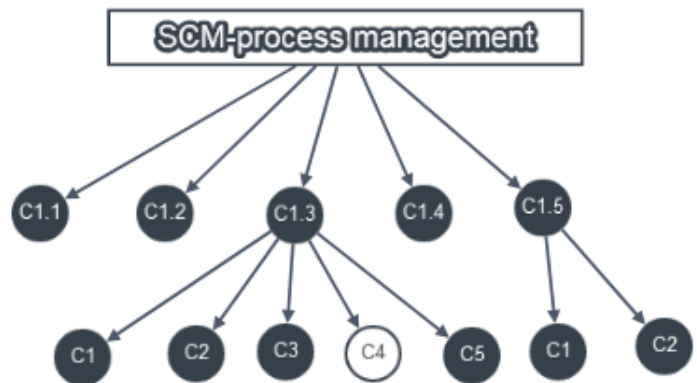


Figure 2: Relational graph of support concept ontology C<sub>1</sub> - SCM-process management

Let's consider the knowledge area – “Configuration management”, it is one of fifteen areas of SWEBOK, the semantic context of which is displayed with six support concepts, and



mastering of which secures vocational competencies, linked with configuration management processes.

## 6. Conclusion

The methodology, used concepts practical relevance is in: raising the training efficiency, based on schematic and symbolic teaching materials models; innovative solution on transfer to a new training information base, based on the project-competency model of knowledge organization and educational resources semantic context ontological engineering; development, based on the knowledge display ontological model, methodology of formulating and implementation of knowledge components, which might be recognized as an educational technology; consistency of representation and formalization of the knowledge components for designing the disciplines working programs knowledge content training guidelines, based on the smart-contract; providing the possibility for the knowledge components configuration in compliance with the demanded scenario and learning outcomes; habituation of students skills and rational techniques in structuring, compressing and visual representing the knowledge and, as a consequence, cognitive thinking development; possibility to form repeatedly used knowledge components, their design and adapting to a wide spectrum of degree programs and services.

## Acknowledgment

The work has been executed with the support of the Ministry of education and science of the Republic of Kazakhstan, grant № AP05134973.

## References

- [1] B. Kubekov, V. Naumenko, A. Utegenova, N. Zhaksybaeva, "Methodology of formation educational resources on basis of ontology", 12th IEEE International Conference "Application of Information and Communication Technologies – AICT2018", October 17-19, Almaty, Kazakhstan, 408-413, 2018.
- [2] K. Czarnecki, U. W. Eisenecker *Generative Programming. Methods, Tools, and Applications*/ Addison-Wesley, 731p, 2000.
- [3] E. Blomqvist, "Experiments on Pattern-Based Ontology Design/E. Blomqvist", In: Proc. of the 5th Int. Conf. on Knowledge Capture (K-CAP 2009) (September 1-4, Redondo Beach, California, USA). ACM, 41-48, 2009.
- [4] Software Engineering, Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering. A Volume of the Computing Curricula Series. IEEE CS and ACM, 134 p, 2014.
- [5] F.Y. Crawlye, "CDIO program: Description of goals and tasks of bachelor degree engineering education", CDIO report. Ed. #1. MIT, 2001. <http://www.cdio.org>
- [6] E.F. Crawley, D.R. Brodeur, D.H. Soderholm, "The education of future aeronautical engineers: Conceiving, designing, implementing and operating", *Journal of Science Education and Technology*. 17(2). 138-151, 2008.
- [7] A.I. Chuchalin, M.G. Minin, S.Y. Kulyukina, "Experience of forming vocational and universal competencies of engineering programs graduates at foreign higher schools", *Higher education in Russia*. #10, p.p.105-116, 2010.
- [8] A.I. Chuchalin, "Bachelor program upgrade in methods and technology area, with account of engineering education international standards", *Higher education in Russia*. #10, 2011.
- [9] A.I. Chuchalin, A.V. Yepikhin, Y.A. Muratova, "Learning outcomes assessment planning upon designing degree programs", *Higher education in Russia*, #1, – 13-20, 2013.
- [10] A.I. Chuchalin, "Concerning application of CDIO approach for designing engineering education level programs", *Higher education in Russia*. #4(200), 2016.
- [11] U. Indira, K. Saule & Z. Gulnaz, "Developing a Model for Managing the Curriculum Model Competency Form", 46-51, 2019. DOI: 10.1145/3343485.3343503.
- [12] B. Kubekov, I. Utebergenov, "The use of multiparadigm approach in knowledge modeling", 7th International Conference on Education and New Learning Technologies". - Barcelona (Spain), 6th-8th of July. - 136-141, 2015.
- [13] B. Kubekov, J. Kuandykova, I. Utebergenov, A. Utegenova, "Application of the conceptual model of knowledge to formalization of concepts of educational content", 9th International Conference on Application of Information and Communication Technologies AICT2015.- Rostov-on-Don, 14-16 October. - 294-306, 2015.
- [14] B. Kubekov, "Educational components formation technology for the planned CDIO SYLLABUS education", 9th annual International Conference of Education, Research and Innovation - ICERI2016, Seville (Spain), 14-16 November, 6139-6145, 2016.
- [15] B. Kubekov, Beyer, Ditmar, A. Utegenova, N. Zhaksybaeva, "Innovative paradigm of education of knowledge - competency form based on ontology", *Journal of theoretical and applied information technology* 15th November 2017, 95 (21), -ongoing JATIT@LLS (E-ISSN 1817-3195 / ISSN 1992-8645). 5859-5868, 2005.
- [16] W. Wojcik, B. Kubekov, V. Naumenko, S. Narynov, S. Toibayeva, A. Utegenova, "Project-competency based approach and the ontological model of knowledge representation of the planned learning", - *INTL Journal of electronics and telecommunications*, 65(1), 75-80. 2018. doi:10.2478
- [17] B. Kubekov, L.K. Bobrov, E.A. Savelyeva, A. Utegenova, V. Naumenko, "Project-competent paradigm of Knowledge representation of the three-level engineering education System", 10th International scientific-practical conference on software engineering and computer systems, section "Ontology Models and Semantic Web", MICSECS 2018. <https://www.scopus.com/authid/detail.uri?authorId=56826094200>
- [18] B. Kubekov, L. Bobrov, U. Anar, N. Vitaly & A. Raigul, "Model of engineering education with the use of the competence-project approach ontological engineering and smart contracts of knowledge components", *Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska*. 3, pp 14-17, 2019. DOI: 10.35784/iapgos.243.
- [19] B.S. Kubekov, "Organization and presentation of the planned training knowledge, based on the ontology", *Monograph. – Almaty: "LP-Zhasulan"*– 336p, 2019.

## Adapting to Individual Differences (ATID) For Inductive Thinking and Learning Purpose

Nia Rohayati<sup>1</sup>, Deni Darmawan<sup>\*2</sup>

<sup>1</sup>Language Education Indonesian Literacy and Regions Department, Galuh University, 46213, Indonesia

<sup>2</sup>Educational Technology and Communication Science Department, Indonesia University of Education, 40154, Indonesia

### ARTICLE INFO

Article history:

Received: 02 March, 2020

Accepted: 03 June, 2020

Online: 06 July, 2020

Keywords:

Adapting to individual differences

Inductive thinking

Learning purpose

### ABSTRACT

The research in general aims to develop students' writing skills by applying a model with reader response strategy and visual symbols to promote students' morals. Specifically, this inquiry is expected to describe the acceptability, comparison, impact, strength, and shortcomings, and the model Adapting To Individual Differences (ATID) of Indonesian literature learning to promote writing skills. This research used a quantitative approach to a quasi-experiment design. Nevertheless, a qualitative approach was also used to analyze students' writing. It was concluded that (1) the model of Adapting to Individual Differences (ATID) in Indonesian literature learning was able to promote students' writing skills, (2) the model of Adapting to Individual Differences (ATID) possessed the strength comparatively in quasi-experiment class. (3) the shortcomings of the model where it spent much time as students responded both nonverbally and verbally and guided questions facilitated students to express their responses. (4) the model of Adapting to Individual Differences (ATID) indicated the significant difference to the increasing of students' write skills in the quasi-experiment class And (5) the model of Adapting to Individual Differences (ATID) showed good qualification where the aspect of cognitive, affective, and psychomotor was developed.

### 1. Introduction

The teaching of literacy can not be separated from the teaching of the language even though it seems separated based on the basic competencies. In the education of the Indonesian language, the term "language" and "literary" tends to be interpreted separately [1]. The aforementioned tendencies are often led to discordant proposals that the language and literacy should be taught by different teachers. Based on the findings of [2] there are 91.6% of respondents who gave responses "agreed" that the literary should be separated from a language concerning the autonomy of literary teaching. Several reasons are supporting such separation, including 1) the position of the teaching of literary will be more stable and more focused because it has clear goals and the allocation of sufficient time, 2) the teaching of literacy will be taught by teachers who have a love and a great commitment in literary, 3) The government is expected to provide some special literary's textbooks and its support, 4) the teaching of literacy in schools will be able to achieve the expected goals

and objectives. The tendency on the separation or the unification of the language and literacy motivates writers to conduct research that does not make sense to both of these competencies as a dichotomy. This study put the languages and literacies on an equal footing. Equality is supported by the opinion of [3] who stated that the students' language skills can be developed systematically if the literacy is taught side by side with the language. So both of these competencies become the basis for conducting research. Furthermore, the literature can enhance the students' understanding of the Indonesian language because of the variety of sentences and vocabulary that learned. According to [4] Students will understand the language well through the learning of literacy. However, currently, the portion of basic competence of literacy is very small compared with the portion of teaching the languages; let alone four basic competencies are closely related to language. This is caused by curriculum guidelines applicable at the time (the communicative approach/meaningfulness). Even pointed out the literacy teaching at elementary school is very far from satisfactory because the basic competence can not develop the There are some problems concerning the low quality of learning the language skills, [5] explains that according to

\*Deni Darmawan, Indonesia University of Education, Street of Dr. Setiabudi, 229 Bandung, 40154, Phone (022)2013163, Email: deni\_darmawan@upi.edu



research conducted by Taufik Ismail, turns out that the writing skills of Indonesian students are the lowest in Asia. Similarly, according to a[6], about the test results of the grade IV elementary school students, Indonesia was ranked as the lowest in East Asia. The average reading test results in some states are as follows: Hong Kong 75.5%, Singapore 74%, Thailand 65.1%, Philippines 52.6%, and Indonesia 51.7% [7]. Furthermore, Semiawan also explained that the results of the study showed that the students in Indonesia was only able to understand 30% of the reading material and had difficulty answering questions that require reasoning description. The low level of literacy in Indonesia, because there are too many students and scattered in a number of remote islands. In addition, the area is difficult to reach, so the literacy program requires access and equity, which requires time. Besides the culture of each region is different so that literacy learning requires variation and creativity from its educators.

Difficulty in writing experienced by students is caused by the fact that the students are not getting used to being trained to write from the beginning. In a writing exercise, students experienced difficulties because of the difficulty to prepare the first sentence. They are confused about where to start writing and how to open the first sentence in writing. Determining the points of writing is a difficult thing for students. Student's utterances like "I'm confused do not know what I would write." " I have a lot of material/things I want to write, but I do not know how to vote". "Several times I change my mind about the subject I wanted to write but do not get a steady". The remarks indicate that the student is difficult to start writing.

In connection with the statement above, the teaching of writing need to switch from a conventional learning model (which is based on the assumption that "knowledge can be transferred intact from that of the teacher to the student's mind") to the learning modern learning model, namely the model of Adapting To Individual Differences (ATID). This model is based on the assumption that "knowledge is constructed in the minds of students". Following the scope of the problem as has been described above, the key problem of research is formulated as follows: "Is the implementation of the ATID model in learning of Indonesian literacy can improve the writing skills of Indonesian literacy in level VIII junior high school students?". This study aimed to describe the increase in the ability to write Indonesian literacy in the level of VIII junior high school by using the ATID model.

**2. Theoretical Review**

*2.1. Adapting to Individual Differences (ATID) in Language and Literacy Learning*

ATID is a theory of learning to improve Indonesian literacy writing skills. Learning theory developed by [8], explains that strategies and techniques that will be developed in the form of a media print to appreciate the literacy that use non-print media to express the results of the literacy. As noted by [9] who formulated four visual dimensions that can be used to respond to a literacy works, namely (a) graph, consisting of sociogram, maps, cards, diagrams, and calligraphy, (b) Illustration, consisting of poster, photographs, collections, and drawings, (c) film/video, consisting of narrative texts, animations, movies, and special effects, and (d) performing arts, consisting of the tableau, dance, music, and pantomime. The use of visual symbols is an attempt to expand and enrich the interpretation and response toward literacy. Following

the research subject which are the elementary school students, for this research, the visual symbols used is limited to sociogram and images.

The ATID model is a personal development learning model that influenced by congenital and external factors. Thus, in lesson planning, it must consider two factors. In this study, the model is applied concerning a verbal response which is owned by the students.

Verbal responses are an important factor in the learning process that can be applied to appreciate the literacy by using words that expressed orally and in writing. In this study, the verbal response in the form of writing is the reader response strategy proposed by [10]. Reader response is a theory or a strategy of contemporary literacy that oriented to the role of readers who interact with literacy at the time the work was assessed. Reader response strategies consists of seven factors, namely: (a) engaging, (b) describing, (c) conceiving, (d) explaining, (e) connecting, (f) interpreting, and (g) judging.

*2.2. Literacy Learning*

Literacy learning is oriented to the visual dimensions which relevant to the reader response, i.e. with the literacy learning theory, basic operations theory, and steps of the learning model of inductive thinking (basic inductive model). The learning model of inductive thinking is put forward by [8] as follows: (a) the formation of concepts (identifying, investigating, and conclusion), (b) the interpretation of data (predict, explain, and testing), and (c) application of the principle (predict, explain, and test).

Literacy learning in this study will be carried out by presenting the subject material about: "The insight into the knowledge of who he is in the family", which described in figure 1.

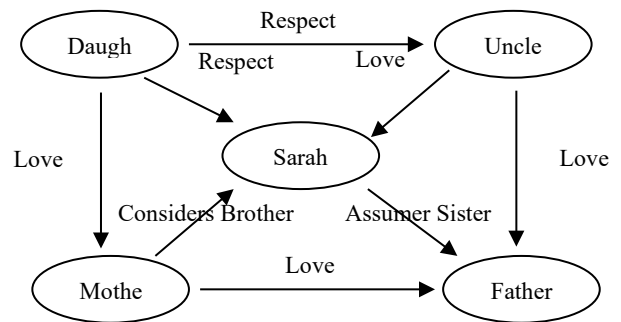


Figure 1: Map material-1 which will be experimented to guide students for develop writing skills of Indonesia literacy using the model of ATID

Based on Figure 1, students are expected to be able to develop the literacy work with the ideas of attitudes and behavior of himself in his family environment. The family map above educates in developing literacy skills with ATID. Thus, all students throughout the archipelago and inland areas in Indonesia certainly have a family. Thus, in improving literacy ability through the ATID model it can be done by all educators in all regions in Indonesia equally.

**3. Research Methodology**

The research methodology for ATID uses a quantitative approach in quasi-experiment [9] to see differences in the ability of students to write Indonesian Literacy. The population of the research is students of junior high school in Tasikmalaya and

Singaparna town. There are two class sample: one class as a class experiment that learning using ATID, and the other class as a class control that learning using a traditional model.

The technique of data collection is performed using a learning test in the form of questions about the ability to write a literacy in the Indonesian Language. On the other hand, the technique of data analysis uses a t-test to evaluate the average differences outcomes of student learning in writing literacy of the Indonesian Language between experimental class and control class. To get students who have similar achievements and prior learning abilities, the homogeneity of the experimental class and the control class are used in the sampling technique.

**4. Result and Discussion**

The advantage of learning models with ATID can facilitate students in acquiring the ability to write literacy about "Insight knowledge of who he or she was in the family." Table 1 shows the result of the research in the experimental class and control class.

Table 1: The Score of Writing Skills of Indonesian Literacy Pre-Test and Post-Test Class Quasi Experiment

The Average Results of Two Evaluator in Class Quasi-Experiments			
No.	Pre-Test	Post-Test	Gain
1	75	90	15
2	72	85	13
3	75	90	15
4	72	80	8
5	75	90	15
6	68	85	17
7	65	75	10
8	65	75	10
9	65	79	14
10	63	74	11
11	59	67	8
12	61	72	11
13	59	65	6
14	65	73	8
15	55	69	14
16	55	72	17
17	65	65	0
18	49	61	12
19	50	59	9
20	50	59	9
Σ	1263	1485	222
$\bar{X}$	63.15	74.25	11.1

Tabel.1 indicates that the average score of pretest about writing literacy in the experimental class was 63.15. To realize the student's competence, the teacher has to explain how to bring the students feeling towards the story action that read by students [10]. Besides, teachers should explain how to connect all the elements

in the story with the social and cultural life. Competence in identifying the literary works presented by [11] while the religious perspective by [7]. Meanwhile, listening and speaking activities can be conducted in the discussion forum, which discussed the results of reading by all students.

After learning literature by applying the model Atid implemented, the average value of responding turned into 74.25. The average gain between the pre-test and post-test scores was 11.1. It shows that a significant increase in the value of post-test students after learning literature by applying the model ATID does. The following table clearly shows the difference between the pre-test to post-test values obtained in a quasi-experimental class of the capability of writing an Indonesian literacy [12].

Results of further testing in the form table below which shows the results of pre-test and post-test grade control, the ability to write an Indonesia literacy.

Table 2: Score of Writing Skills of Indonesian Literacy

The Average Results of Two Evaluator in Class Control			
No.	Pre-test	Post-test	Gain
1	80	82	2
2	79	80	1
3	75	77	2
4	72	75	3
5	72	73	1
6	72	75	3
7	75	79	4
8	69	72	3
9	69	70	1
10	65	66	1
11	61	65	4
12	69	70	1
13	55	60	5
14	55	58	3
15	59	64	5
16	59	61	2
17	61	61	0
18	59	61	2
19	59	61	2
20	50	58	8
Σ	1315	1368	53
$\bar{X}$	65.75	68.4	2.65

Table 2 above indicates that the average value of pretest responded to stories in the control class is 65.75. The value increased to 68.4 after students take the post-test with a value gain of 2.65. The average gain between the pre-test and post-test scores was 2.65. It shows that an increase of the value pretest to the value of post-test although the increase is much lower than grade

students quasi-experimental experience learning by applying the model ATID treatment with students who are not being treated. Because the value signification post-test capability wrote literature students is 0,000 less than the real value of 0.05, then Ho is rejected. This fact can be seen in the table on the average difference test writing skills literary works Indonesia between the two groups of students are students of a quasi-experimental class and control class as follows [13]. This finding shows the ATID model is very effective in helping improve language literacy in students.

Table 3: Writing Skills Test Results Literacy of Indonesian Students between pre-test and post-test Class Quasi-Experiment

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Dev	Std. Error Mean	95% Confiden Interval of the Difference				
					Lower	Upper			
Pair 1	ost test Experiment Post test Control	5,850	5,53719	1,2382	3,2585	8,4415	4,725	19	.000

Based on the results of testing the application of learning models ATID above either through an experimental class and control, as well as mapping the topics that have experimented then be analyzed as follows. From the aspect of materials developed maps, classes that are subjected to ATID models further demonstrate the superiority that is, when compared with the acquisition of control class capabilities. This is evident from the difference in mean pre-test / post-test experimental and control classes. The next important thing is inferred ATID models that had been developed before students write short stories. Based on the results of analyzing aspects of clarity, accuracy, interest, inclusiveness, and completeness, all the students' concept is based on individual differences in which includes seven elements of the set (determination of the characterization of short stories that became the focus of the story; the laying of the characters; creativity to color each character or the gist of the story; the use of symbols, and the nature of the schematic in the concept has shown its quality is right, because the average concepts students have demonstrated the clarity, precision, interest, inclusiveness, and completeness well, imagination, and schematic. These findings have implemented the opinion of [11], which suggests that teachers should explain how to engage the students to feel the characters were read by the students and how to connect all the elements in the story with the social and cultural life in identifying literature [14]. Through the process of and findings in his research, particularly efforts to enable the students for learning Indonesian, it is expected to answer the concerns of the opinion from [15] which states that students are still less skilled Indonesian language.

During the process of revision of the test in the form of experimental activities carried out during the study, indirect stages in this study require students to keep repeating about the ability of ideas, story ideas, and contents of the map material have experimented. This finding is consistent with the policy of the [16], one of the specific goals of learning the use of language is that students should be able to express ideas, opinions, knowledge in writing, and has a penchant for writing. To express ideas,

opinions, and knowledge, students need to be given training on an ongoing basis.

Students' attitudes towards writing activities positively, because students perceive, that after reading activities should be followed by a short story writing activity is an activity that first time they have experienced, and generally students always scored less satisfactory as long as they do learning activities during this writing. Students' interest in learning writing looked on: students are always trying to do the tasks write well; students have never submitted responses to literature critically; students feel impressed, when analyzing the literature Indonesia / stories should be linked to the values of life in the real world: students feel happy if the reading of literary works continued with analyzing the literature, and, students feel happy if his writings published in the magazine wall. From these findings, it has put role models ATID Indonesian teachers who can make skillful use of Indonesian students in all its functions, especially the function of communication. This finding is consistent with the opinion of [17] states that "To teach literature Correctly is to emphasized the aesthetic stance and to emphasize the efferent." That is, students are not only able to identify what is contained in literary works such as background, character, and characterization, as well as the storyline, but they were also able to identify what is on the outside of the literary works such as the author's intention, symbolism, storytelling and so forth.

Besides, students should not only be trained in listening skills, speaking, reading, and writing, but are trained to think and reason in an orderly manner in the Indonesian language. [18] describes the relationship of writing to think as one aspect of language skills. Nickerson said that frozen speech or utterance is another name for the author. Writing a valuable contribution to civilization, namely an enduring record. Note perennial accelerate the accumulation of knowledge [19].

The externally within the meaning of the teacher's role when it illustrates the material into a story, among others, played by the students then students will be more motivated to understand it. This finding is supported by the opinion of [20] illustrates if there are students who are not able to express a new story read verbally, but he was unable to explain what happened in a story in creating the image, the teacher should receive a response so. Thus, students are motivated to learn to write short stories, because teachers create an atmosphere of fun teaching; because teachers provide a model of analysis of literary work that is attractive and easy [21].

## 5. Conclusion and Recommendation

### 5.1 Conclusion

Based on analyzing and testing through the research that has been done the level of ability to write short stories before and after the model Adapting to Individual Differences (ATID) eighth-grade students at two public junior high school in Tasikmalaya, shows that there increasing ability Indonesia wrote literature student after the student gets treatment the ATID models [22].

From the findings and discussion, the conclusions are real that the model of ATID in Indonesian literature learning was able to promote students' writing skills. The model of ATID possessed strength comparatively in the quasi-experiment class. Then the shortcomings of the model were it spent much time as students responded both non verbally and verbally and guided questions

facilitated students to express their responses. The model of ATID indicated the significant difference to the increasing of students' writing skills in the quasi-experiment class and this model showed good qualification where the aspect of cognitive, affective, and psychomotor was developed [23].

### 5.2 Recommendation

Based on these results it can be recommended that with the application of this ATID very impressed with the students' learning model with a model to write short stories with the ATID model [12]. For teachers should be able to pay attention to aspects of the utilization of the model both in terms of ease of understanding of learning, in terms of function chart and coloring creativity and depiction of the emblem in sociogram, in terms of usability order of ideas in concept, in terms of the psychology of learning [24], as well as in terms of the sustainability of the use of models ATID. Another recommendation is expected that teachers should continue to support the students' ability to write literary works to cultivate the application of this model for future learning.

### Conflict of Interest

The authors declare no conflict of interest.

### Acknowledgment

The author would like to express gratitude to the Ministry of Research and Technology of Indonesia that has provided research grants, and to the leadership of the Galuh University who has permitted to complete this research to provide benefits in the world of education.

### References

- [1] H. Yu, J. An, J. Yoon, H. Kim, and Y. Ko, "Simple methods to overcome the limitations of general word representations in natural language processing tasks I," *Comput. Speech Lang.*, **59**, 91–113, 2020, DOI: 10.1016/j.csl.2019.04.009.
- [2] K. A. Harras, "Problems of Literature Teaching Problems: Language and Literature," *J. Lang. Educ. Lit. Teaching.*, **3** Teaching Language, 2003.
- [3] R. dan M. N. L. Carter, "Teaching Literature," New York Longman, Inc., 1991.
- [4] M. R. Zughoul, "English Department in the Third World Universities: Language, Linguistic, or Literature," *English Teach. Forum*, **4**, 2006.
- [5] A. Imran, "The Lowest Writing Skills of Indonesian Students in Asia," *People's Mind*, **26**, October 2000.
- [6] W. B. Report, "Writing skill in Indonesia," *Annu. Report Lit. The Asia Pacific*, 1998.
- [7] Semiawan, "Rhetoric in Writing," Jakarta: Depdikbud Dikti PPLPTK, 2003.
- [8] David Hunt and Schroeder, "Writing Skill in Language Literate," *Lang. Literature*, 7 Oct 2019.
- [9] W.R.B. Gall, Meredith D., J.P. Gall, "Educational Research: An Introduction," NY Pearson Educ. Inc., 7, 2003.
- [10] C. Liu, R. Wang, L. Li, G. Ding, J. Yang, and P. Li, "Effects of encoding modes on the memory of naturalistic events," **53**, August 2019, 2020, DOI: 10.1016/j.jneuroling.2019.100863.
- [11] R. W. & J. D. M. Beach, "Teaching Literature in the Secondary School," New York Harcourt Brace Jovanovich, Inc., 1991.
- [12] M. Michel, J. Kormos, T. Brunfaut, and M. Ratajczak, "Journal of Second Language Writing The role of working memory in young second language learners' written performances," *J. Second Lang. Writ.*, **45**, May 2018, 31–45, 2019, DOI: 10.1016/j.jslw.2019.03.002.
- [13] M. K. Tåqvist, "Journal of Second Language Writing 'A wise decision': Pre-Modification of discourse-organizing nouns in L2 writing," *J. Second Lang. Writ.*, **41**, June, 14–26, 2018, DOI: 10.1016/j.jslw.2018.05.003.
- [14] M. Yoyo, "The Effectiveness of the Teaching Model on Reader Responses in Teaching Poetry Study; Experimental Study on Students of Indonesian Language and Literature Education Department FPBS Bandung," Diss. Bandung UPI PPS., **1998/1999**, 2000.
- [15] J. Badudu, "This is the correct Indonesian IV," Jakarta: Gramedia Pustaka Utama., 2015.
- [16] D. of E. and Culture., "Socialization of Indonesian Language Learning," Jakarta: Balai Pustaka, 2018.
- [17] L.M. Rosenblatt, "The Reader, the Text, the Poem: The Transactional Theory of the Literary Work," Illinois South. Illinois Univ. Press., 1978.
- [18] B.K. Owusu and H. Burianová, "Transcranial direct current stimulation improves novel word recall in healthy adults," *J. Neurolinguistics*, **53** no. August 2019, p. 100862, 2020, DOI: 10.1016/j.jneuroling.2019.100862.
- [19] D. Darmawan, Y. Ruyadi, W. J. Abdu, and A. Hufad, "Efforts to know the rate at which students analyze and synthesize information in science and social science disciplines: A multidisciplinary bio-communication study," *Online J. Biol. Sci.*, **17**(3), 2017, DOI: 10.3844/ojbsci.2017.226.231.
- [20] A.C. Purves, "How Porcupines Make Love II: Teaching a Response-Centered Literature Curriculum," New York Longman Group, Ltd., 1990.
- [21] K.D. Njikeh, "The Mentally Offset Human and the Society," *Open J. Psychiatry*, **06**(02), 188–190, 2016, DOI: 10.4236/ojpsych.2016.62022.
- [22] R. Tanaka, A. Ozeki, S. Kato, and A. Lee, "Computer Speech & Language Context and knowledge aware conversational model and system combination for grounded response generation," *Comput. Speech Lang.*, **62**, p. 101070, 2020, DOI: 10.1016/j.csl.2020.101070.
- [23] L. Nuryanti, D. Darmawan, and A. Hufad, "The Implementation Model of Character Building in Curriculum 2013 Through The Scout Movement as A Required Extracurricular in Primary And Secondary Education," **5** (1), 361–367, 2020.
- [24] S. Gosselke, M. Horne, K. J. Br, Y. Shtyrov, and M. Roll, "Neural processing of morphosyntactic tonal cues in second-language learners," **45**, 2018, doi: 10.1016/j.jneuroling.2017.09.001.



## A Mobile Application Design to Prevent Criminal Acts in Lima, Peru

Alexi Delgado<sup>\*1</sup>, Enrique Lee Huamani<sup>2</sup>, Elizabeth Jenny Cortez-De La Peña<sup>2</sup>

<sup>1</sup>Mining Engineering Section, Pontificia Universidad Católica del Perú, 15088, Peru

<sup>2</sup>Systems Engineering Program, Universidad de Ciencias y Humanidades, 15307, Peru

---

### ARTICLE INFO

Article history:

Received: 25 March, 2020

Accepted: 14 June, 2020

Online: 06 July, 2020

---

Keywords:

Balsamiq Wireframes

Design

Mobile Application

---

---

### ABSTRACT

*In the province of Lima, Peru, criminal acts are increasing every day; it is a social problem that affects many people, who often try with the life of the indignant person, becoming the greatest fear of the community. For this reason, a mobile application was designed that allows the inhabitants to report and share the criminal acts that occur in their environment in real time, allowing them to prevent such acts. This application was designed using the Balsamiq tool, which allows each sketch to be developed in an organized manner. The result was the development of each module that addresses each of the functionalities of the application design for its correct structuring with the Balsamiq tool. These results will help the inhabitants of Lima, Peru, to expose the facts of which they are victims, without the immediate need to go to the police station, and which will be proposed for development to different competition funds.*

---

### 1. Introduction

Today, the area of criminality contains many criminal problems that constantly affect people's lives. Globally, the list of crimes is expanding dramatically every day [1]. In Peru, it has become a problem that, along with corruption, affects the lives of the population in about 40.4% of cases, making it the country's main problem, followed by poverty with 38.9% [2]. Thus, a safe environment has become a major public and social concern. Therefore, and taking into consideration that the purpose of engineering is to reduce human efforts and make life more comfortable in order to overcome the obstacles that arise [3], a contribution to this social problem will be sought by designing an application that is useful for the citizens of the province of Lima, Perú.

Several methodologies were found for the design of a mobile application, the first of them is the tool Justin mind platform for sharing and evaluating prototypes of web applications, iOS and Android [4]. It contains an intuitive and dynamic interface, has access to indicate actions on each element of the design and has both free and paid versions. Additionally, there is Proto.io that allows applying screen transitions and defining touch events for each design element on the screen. That creates a high-quality prototype. The prototypes can be viewed and experienced interactively within the browser, but most importantly they can be

launched on the actual mobile device [5]. Finally, there is the Balsamiq tool, the most popular being helpful for fast model building [6], it develops easily interactive models and quickly, they can be used to verify the key ideas behind an application fulfilling the user's objective [7]. Therefore, Balsamiq was chosen for the development of the design since it is useful to communicate ideas [8], speeding up the creation of sketches and schemes in a simple way. The reason for using the Balsamiq tool has been due to the experience gained in its use with respect to the prototyping of projects related to the development of mobile and web applications in addition to having the paid version that provides a wide selection of basic elements along with navigation elements, video and image areas, forms and many other functions, this tool allows you to move around the design surface simply by dragging and dropping, thus creating the wireframe.

This article uses the design of a mobile application to analyze crime in the province of Lima, which is useful for evaluating the increase or decrease of criminal acts [9], information that could be made available police officers. The application allows the affected areas to be displayed on the map and is useful for alerting and preventing citizens. It also includes a search and database of reports and data from the nearest police stations.

The objective of this investigation is to design a mobile application that allows citizens to report and share criminal acts that have occurred and, in this way, contribute to the authorities

---

\*Alexi Delgado, Email: [kdelgadov@pucp.edu.pe](mailto:kdelgadov@pucp.edu.pe)

taking precise actions, as well as alerting citizens who may be in the area.

The present work is constituted in the following way: section II explains the methodology by means of stages, detailing the tool used. Section III introduces the application developed sequentially showing the interfaces. Section IV shows the results obtained and the development of the discussion. Finally, section V presents the conclusions.

## 2. Methods

This research will present the design of a mobile application, to structure, simulate and make the visually friendly prototypes for the planning of the mobile development environment, to develop the application work is deployed in phases corresponding to the creation of the design.

### 2.1. Creation of Users

The mobile application has the well-known “User Stories”, which allow a set of tasks to be performed, and have different accesses, permissions and privileges according to the user's functions.

### 2.2. Initial Screen Design

When the design phase of the project starts, the first and fundamental step is to represent the layout development of the mobile application, the main view; in this case a suitable tool was selected for its design: Balsamiq Mockups, which is a medium fidelity prototyping software, which allows the elaboration of functional interfaces and gives the user an initial impression of the application's operation [10]. Following this, the creation of schematics was outlined in order to speed up the design phase with functionalities represented in static form. It is worth highlighting the usefulness of this tool in this stage, since it facilitates the creation of instantaneous, polished, precise sketches to design accessible and organized tests; in other words, it has the purpose of showing the idea of the final project.

### 2.3. Elaboration of the modules

Essentially, in this phase the modeling of diverse designs within each module is produced. The general view includes the modules that provide a global appreciation that should captivate the user at first sight, contains all the functionalities by means of buttons and is intended to be direct. It is crucial to define the modules that the application will contain, because it allows the user to perform several functions that can be adapted to the development.

The application is composed of three parts as shown in Figure 1, where the user can report and/or search for types of criminal acts, as well as communication management, where the user can observe the news section and Home module, where the user has access to his personal account.

Eventually in every mobile application modifications and variations can be presented, it is possible to improve the structure that was thought, being able to verify certain deficiencies quickly and to add interaction to the draft.

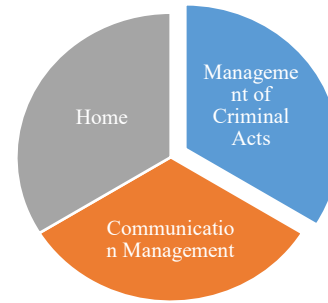


Figure 1: Mobile Application Screen Diagrams

### 2.4. Implementing Mockups

In the last phase, the project will use the Android operating system, which has developed in a major role to this eco system known as Consumer [11], Android Studio is the official environment for the development of applications in Android, written in Java language. Application Programming Interface (API) can be provided through the help of GPS technology and map manipulation using Google Maps API tools in Android [12]. Some features used are camera, microphone, map and Internet access. SQLite directly calls the corresponding API functions to perform data access operations [13].

Table 1: Functions of each user

User	Function
Administrator	Process of managing and controlling all the functions and confidential information of users, observe the reports filtered by police station, type of crime, map of affected areas, and can add modules with new functions if needed.
Citizen	Access the mobile application that will be validated by the administrator user for registration in order to have a correct user identification and avoid fraud when reporting an incident. In the application you can send and share criminal events that occur or happen to around you and even broadcast live, view a search engine by incident type, view the news section and map the affected areas.
Police officers	Receives all notifications from citizens with their corresponding data (type of incident, place of the event, description, date, evidence, etc.) with the aim of taking the corresponding measures in the shortest possible time. When the police accept the confirmation, a notification will be sent to the citizen indicating that they are going to the place of the incident.



### 3. Case Study

#### 3.1. Creation of Users

It has been considered pertinent to take into account three basic specific users, of which the functions of each user are shown in Table 1.

#### 3.2. Initial Screen Design

In this phase of the design, we proceed to show the main view, the initial sketches to be taken into account for the mobile application, made with the Balsamiq tool.

When entering the application a screen is accessed directly, if the user is not in an active session, the application asks for a login either by Facebook or Google or by a user already created, as shown in Figure 2(a), likewise the option to register by filling in your personal information in the system, in case you do not have a user created, as shown in Figure 2(b).



Figure 2: (a) Login Prototype, (b) Registration Prototype

Table 2: Composition of the Modules

General Module	Second Level Module
Management of Criminal Acts	Searching for Criminal Acts Types of Criminal Events Map of concerned areas Report Criminal Activity
Communication Management	News Live broadcast Contacts Commissaries
Home	Profile History of Reports

#### 3.3. Elaboration of the modules

The general module is shown in Table 2, with ten buttons that direct to each of the functionalities or instances of the mobile application.

##### 3.3.1. Criminal Management Module:

This module consists of four instances.

- a) *Search for Criminal Acts*: The user can filter by type of criminal act, police station, date and area to find out about the reports presented, as shown in Figure 3.



Figure 3: Search for Criminal Acts

- b) *Types of criminal events*: The user can observe a classification of the different criminal acts that exist, this is represented in Figure 4.

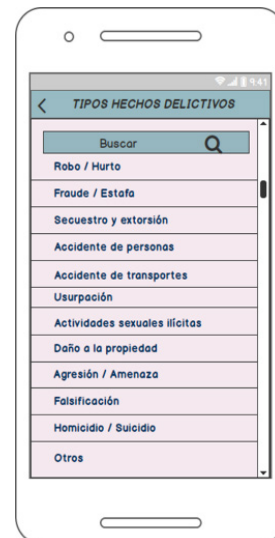


Figure 4: Types of Criminal Events

- c) *Map of concerned areas*: This is a Google Maps map which indicates the offended areas.
- d) *Report of criminal activity*: The contribution of the user is presented on the report (type, place, description, date, time) where it is possible to add text,

attach photo, video, audio and location that objectively support the affected incident.

- 3.3.2. *Communication Management Module:* The module contains four instances.
- News: This section shows the publications made by the users; it has the option to comment, share and view the details.
  - Live transmission: It allows to transmit a live video sharing some criminal event in real time.
  - Contacts: The user can determine which will be the emergency contacts and have direct access to them if needed.
  - Commissaries: In this section you can see all the commissaries with their respective information as an addition, it contains a search engine for greater access.
- 3.3.3. *Start Module:* The module is contained by two instances.
- Profile: In this section the information of the user is shown giving the opportunity to access to the edition of their information if needed.
  - History of reports: Screen in which the user visualizes his reports with its respective information.

3.4. *Implementation of Mockups*

To complete the design development, it will be implemented, in the future, in the Android Studio program which contains a wide range of views that can be used to develop an application containing lists, grids, text boxes, buttons and also an embedded web browser [14]. This mobile application is set up to publish, collaborate and report crime information, building a crime information map, as well as providing the opportunity for immediate assistance in emergency situations and pre-existing events, using Google libraries. The user functions are shown in Table 3.

Table 3: Characteristics

Function	Description
Localization	The user allows the interface to recognize their location when making a report. This option is also imperative when searching for affected areas and when sending alerts to different police stations and users.
Photo and Video	The user can attach any photo or video of the act that serves as evidence for the complaint, as well as so that other users can recognize those involved in a more optimal way.
Audio	The user can add voice audio as a sample if required.

Figure 5(a) shows the prototype of how the interface that shows the location of the user will display; Figure 5(b) will show the areas where recent complaints have been made so that users in the environment have more information to take appropriate action.

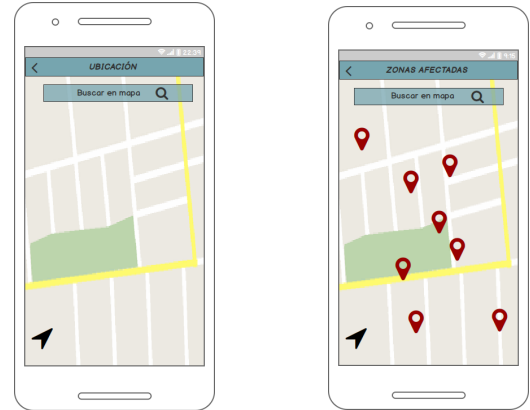


Figure 5: (a) Localization, (b) Concerned areas

The Google APIs will be used to obtain the location with the GPS network and another one to show the affected areas on a map to the user. These services are used to obtain the longitude and latitude of the reported location [15]. Records in the SQLite database are updated and uploaded to the cloud each time the user modifies any corresponding data [16]. Figure 6 shows the design that was developed to report criminal acts in order to provide citizens with a tool to validate their report, which records the type of incident, the place where it occurred, the date and time, also allows the user to add more details if desired, allowing them to publish anonymously for greater security if deemed appropriate, this anonymous information can be seen by police officers and the administrator with the limitation of not knowing which user has registered the criminal acts. Finally, in the application there will be previously detailed options to attach photo, video or share live video which will make the mobile application interactive and easy to use.

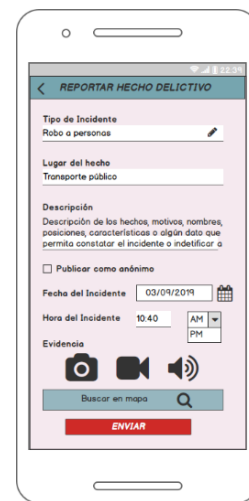


Figure 6: Criminal Act Report

3.5. *Transfer of prototypes to Android Studio*

In order to transfer all the prototypes to Android Studio for the development of the application will have to follow in the

structure of the layout where we must open the Android Studio and create a project with the name of the application in this case we are calling “Mobile application to prevent criminal acts” and hit the Next button as shown in Figure 7.

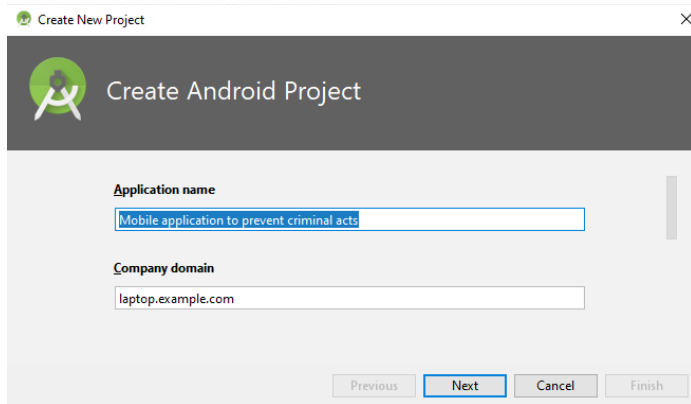


Figure 7: Project Creation in Android Studio

Then select a version of our application is recommended to use Nougat because it is currently the most stable version and give the next button as shown in Figure 8.

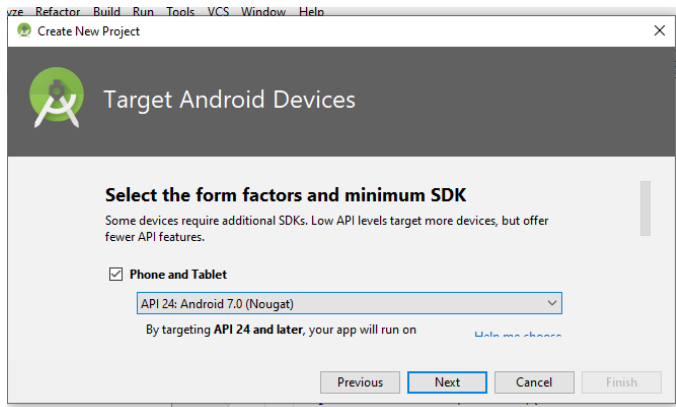


Figure 8: SDK version selection

To make it a worked from scratch application a blank template will be used as shown in Figure 9.

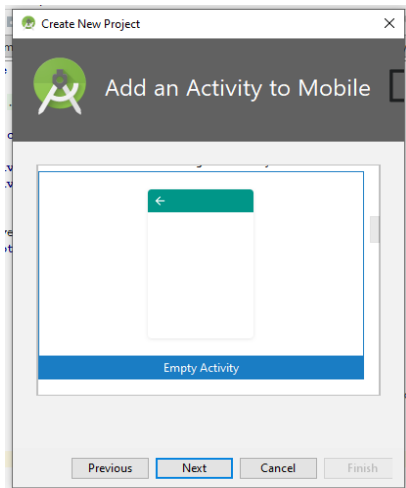


Figure 9: Empty Activity selection

Then we must create an initial project with the name MainActivity to finally give the Finish button as shown in Figure 10.

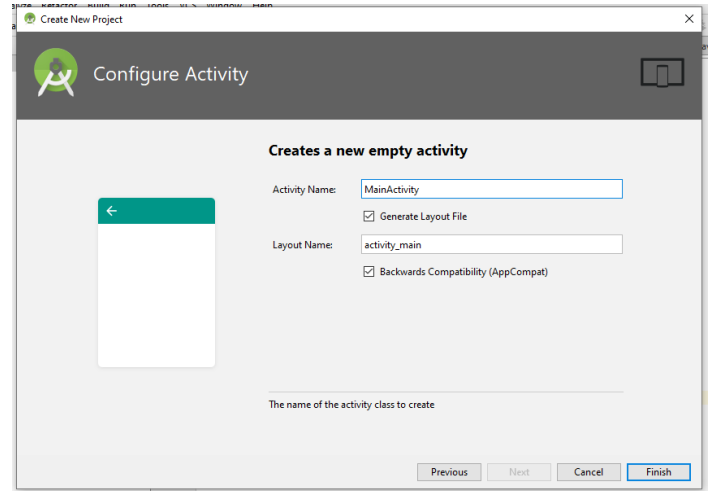


Figure 10: Completion of project creation

For the graphic development we must access the MainActivity.xml where we will have to structure the forms with respect to the model, as shown in figure 11 that was developed following the structure of the model in figure 6.



Figure 11: Transfer of prototype to Android Studio

#### 4. Results and Discussion

As a result of this research we can indicate that it has contributed to propose a solution for one of the cities with the highest percentage of crime as it is the city of Lima, the proposal of this research is very important because currently there are funds

available to compete for the implementation of mobile applications that improve the security of citizens, in Table 4 with respect to the research [2] that was mentioned in the introduction section we can validate that there is a higher percentage of crimes in Lima itself being necessary these types of proposals that can help reduce these types of incidence.

Table 4: Regions with the highest crime rate

Region	Percentage
Puno	37.4%
Huancavelica	37.8%
Callao	38.9%
Ayacucho	26.9%
Loreto	30.2%
Lima Metropolitana	40.4%

Since in Peru we have a large number of regions, the 6 with the highest crime rate were placed with respect to Table 4 with respect to 2018, where the percentage is considered by the number of people in their region, therefore contributing with a proposal for Metropolitan Lima it could help reduce the percentage.

#### 4.1. About the Case Study

As a case study, the design of the mobile application has been carried out with special dedication and in the most appropriate way, achieving a series of stages and instantaneous functionalities for its creation and correct structuring by means of the Balsamiq Mockups tool. This research tries to propose an application for the benefit of society as proposed by the research of [17] that makes a prototype proposal for a mobile application in order to reduce pollution in Peru. The prototype development serves as a basis to have a general concept of what is proposed to be done, as we can also see in the Figures with respect to the prototype a user-friendly environment, this can be modified when it is taken to the development of the application depending on the requirements that are requested.

#### 4.2. About the Methodology

In Balsamiq Mockups, mobile applications are designed in a more intelligent and easy way [18], they are also used to design the structure, hierarchy and relationship between the elements that compose the mobile application [19]. Future integration will be done with Android Studio IDE, and there are several methods to make a mobile application. For this purpose, the prototype work of the Mobile Application for pneumonia management obtained in the community [20] has been observed, which had a system written in the Objective-C language using a development approach called Model View Controller (MVC); which turns out to be an interesting proposal for the proposed future objective, to also be coupled in the iOS system, as it has been evaluated in the mentioned research.

### 5. Conclusion

In this paper we managed to design the prototype of the mobile application with the Balsamiq tool showing the digital outlines of all the application modules, helping to reach a better solution to the situations of the role of the user, a design that allows the

community to report any crime by attaching different evidence (such as audio, photo and/or video) and even broadcast live, also accessing the location (after activating the GPS).

The method applied was the most deductive, seeking to encompass the characteristics of the design and for this purpose the Balsamiq tool was used to contribute to the design of the application and its components defined in their respective sketches, giving a series of special features to generate interactive drawings to the benefit of end users, ensuring that the development is not compromised or slowed down at any stage.

It is recommended in future investigations to validate this design to a concrete application integrated with the National Registry of Identification and Marital Status (RENIEC by its Spanish acronym) to authenticate the data of the citizen, complementing it with topics of satellite tracking or tracing, interconnected in real time with the police stations in order to enable the immediate arrival of police and security personnel to the place of the event, and to contemplate its operation with Android Studio.

With respect to the design of the mobile application explained in this research, this proposal will be submitted to competitive seed funds such as Fondo Nacional de Desarrollo Científico y Tecnológico (Fondecyt) and the National Innovation Program for Competitiveness and Productivity (Innovate Perú) so that it can be developed and used by citizens. This research can be improved in the visual and functional part thanks to the direct communication that would be had with the citizens when this proposal is launched.

### References

- [1] C. Guevara, D. Bonilla, J. Pozo, R. Pérez, H. Arias, and L. Martínez, "Mobile geographic information system for citizen security", in Iberian Conference on Information Systems and Technologies, CISTI, 2019-June, 2019. <https://doi.org/10.23919/CISTI.2019.8760713>
- [2] C. M. Soto et al., "PLAN NACIONAL DE SEGURIDAD CIUDADANA 2019-2023" (report in Spanish), 2018.
- [3] S. Bhattacharjee and G. C. Somashekhar, "Artificial intelligence to impart surveillance, tracking, & actuation on suspicious activities", in Proceedings - 7th IEEE International Advanced Computing Conference, IACC 2017, 1-5, 2017. <https://doi.org/10.1109/IACC.2017.0016>
- [4] W. T. Mok, R. Sing, X. Jiang, and S. L. See, "Proposal of a depression detector", in 2014 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA 2014, 2014. <https://doi.org/10.1109/APSIPA.2014.7041742>
- [5] "Proto.io: precios, funciones y opiniones | GetApp". [Online]. Available in: <https://www.getaes/software/90497/proto-dot-io>. [Accessed: 03-jan-2020].
- [6] J. M. Rivero, G. Rossi, J. Grigera, E. Robles Luna, and A. Navarro, "From interface mockups to Web application models", in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6997 LNCS, 257-264, 2011. [https://doi.org/10.1007/978-3-642-24434-6\\_20](https://doi.org/10.1007/978-3-642-24434-6_20)
- [7] S. N. Fazelah, M. Noor, Z. A. Kadir, B. Pahat, and J. Malaysia, "Students' Learning Preferences of English for Academic Purposes-A KUiTTHO Affair". Available in: [https://www.academia.edu/3608107/Students\\_Learning\\_Preferences\\_of\\_English\\_for\\_Academic\\_Purposes\\_-\\_A\\_KUiTTHO\\_Affair](https://www.academia.edu/3608107/Students_Learning_Preferences_of_English_for_Academic_Purposes_-_A_KUiTTHO_Affair)
- [8] I. Bouchrika, L. Ait-Oubelli, A. Rabir, and N. Harrathi, "Mockup-based navigational diagram for the development of interactive web applications", in ACM International Conference Proceeding Series, 27-32, 2013. <https://doi.org/10.1145/2503859.2503864>
- [9] W. Chi and X. Wang, "Design of criminal analysis system based on OLAP", in ICCSE 2012 - Proceedings of 2012 7th International Conference on Computer Science and Education, 838-841, 2012. <https://doi.org/10.1109/ICCSE.2012.6295200>
- [10] D. S. Alves, E. M. A. Da Silva, M. B. Honorato, and M. De Araújo Novaesa, "Prototype of care application for obstetric telemonitoring of hypertensive syndromes in high risk pregnancy", in Studies in Health Technology and Informatics, 264, 1769-1770, 2019. <https://doi.org/10.3233/SHTI190639>

- [11] S. Patel, "Behavioural study of memory allocators for Android platform", in 2017 IEEE International Conference on Consumer Electronics-Asia, ICCE-Asia 2017, **2018**-January, 52-55, 2018. <https://doi.org/10.1109/ICCE-ASIA.2017.8309320>
- [12] I. P. G. A. A. Putra, E. Sedyono, and A. Setiawan, "E-land design of mobile application for land information system using Android-based Google Maps API V2", in Proceedings - 2017 International Conference on Innovative and Creative Information Technology: Computational Intelligence and IoT, ICITech, **2018**-January, 1-5, 2018. <https://doi.org/10.1109/INNOCIT.2017.8319145>
- [13] K. Yue, L. Jiang, L. Yang, and H. Pang, "Research of embedded database SQLite application in intelligent remote monitoring system", in Proceedings - 2010 International Forum on Information Technology and Applications, IFITA, **2**, 96-100, 2010. <https://doi.org/10.1109/IFITA.2010.241>
- [14] R. V. Golhar, P. A. Vyawahare, P. H. Borghare, and A. Manumare, "Design and implementation of android base mobile app for an institute", in International Conference on Electrical, Electronics, and Optimization Techniques, ICEEOT 2016, 3660-3663, 2016. <https://doi.org/10.1109/ICEEOT.2016.7755391>
- [15] Z. Jiantao and P. Xuan, "Design and implementation of field staff management system on android mobile", in Proceedings - 3rd International Conference on Instrumentation and Measurement, Computer, Communication and Control, IMCCC 2013, 791-794, 2013. <https://doi.org/10.1109/IMCCC.2013.176>
- [16] C. Doukas, T. Pliakas, and I. Maglogiannis, "Mobile healthcare information management utilizing Cloud Computing and Android OS", in 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10, 1037-1040, 2010. <https://doi.org/10.1109/IEMBS.2010.5628061>
- [17] A. Delgado, D. Vriclizar and E. Medina, "Artificial intelligence model based on grey systems to assess water quality from Santa river watershed," Proceedings of the 2017 Electronic Congress, E-CON UNI 2017, January, 1-4, 2018.
- [18] K. Kalimullah and D. Sushmitha, "Influence of Design Elements in Mobile Applications on User Experience of Elderly People", in Procedia Computer Science, **113**, 352-359, 2017. <https://doi.org/10.1016/j.procs.2017.08.344>
- [19] E. Canbazoglu, Y. B. Salman, M. E. Yildirim, B. Merdenyan, and I. F. Ince, "Developing a mobile application to better inform patients and enable effective consultation in implant dentistry", Comput. Struct. Biotechnol. J., **14**, 252-261, 2016. <https://doi.org/10.1016/j.csbj.2016.06.006>
- [20] A. Delgado, A. Aguirre, E. Palomino and G. Salazar, "Applying triangular whitenization weight functions to assess water quality of main affluents of Rimac river", Proceedings of the 2017 Electronic Congress, E-CON UNI 2017, January, 1-4, 2018.



## Energy Recovery Equipment and Control Strategies in Various Climate Regions

Rand Talib, Alexander Rodrigues, Nabil Nassif\*

Department of Civil and Architectural Engineering and Construction Management, University of Cincinnati, Cincinnati, OH 45221, USA.

### ARTICLE INFO

Article history:

Received: 27 March, 2020

Accepted: 26 May, 2020

Online: 06 July, 2020

Keywords:

Building energy consumption

Energy recovery systems

Physics based simulation

Bin method energy model

Climate zone design optimizations

### ABSTRACT

Different types of air-to-air energy recovery technologies such as coil loops, heat pipes, sensible wheels, and total energy wheels are frequently incorporated in HVAC systems in an attempt to reduce energy consumption. This study examines the impact of various types of energy recovery technologies and capacity control strategies on a building's cooling, heating, and fan energy consumption across different climate zones, including Fargo, ND; Cincinnati, OH; Miami, FL; San Francisco, CA; and Phoenix, AZ. A self-developed analysis model simulates a typical HVAC system and compares data that will aid in evaluating different energy recovery equipment and control strategies to achieve maximum energy conservation. Conversely, the results of the study show that the improper operation and incorrect selection of energy recovery technologies could lead to increased energy consumption, further emphasizing the need for proper implementation of controls in energy recovery technologies.

### 1. Introduction

The building stock is the largest consumer of energy in the United States, estimated to consume 40.3 quadrillion Btus or about 39% of total U.S energy consumption, making the reduction of energy consumed in buildings of paramount importance [1]. The building's heating, cooling, and ventilation systems are the single largest energy consumers of any other system, totaling about 13.5 quadrillion Btus or 35% of total energy use [2]. With humans spending 90% of their time indoors and the increased reliance on buildings in everyday life, creating an indoor environment that is healthy, comfortable, and productive becomes critical [3]. ASHRAE Standard 62.1-2016 addresses the ventilation requirements to ensure these indoor spaces are comfortable to the occupants [4]. Hence, the area of overlap between human comfort and energy efficiency has warranted further research and enhancement.

With the arising issues of global warming and the cost of energy the need for more efficient building systems became more crucial [5]. Thus, system designers and engineers always exploring new ways of implementing new technologies for a better design and operation of HVAC systems. One of the methods is implementing the heat/energy recovery devices [6]. An air to air energy recovery heat exchanger includes but is not limited to, heat recovery ventilators and energy recovery ventilators. Those devices are placed in the units where the exchange between the outdoor air and return air is happening. The exhaust air of the

building passes through one side of the exchanger while the outside makeup air passes through the other side of the exchanger. In the cooling season the return cooled indoor air is used to precool the outside incoming air. While, this process is reversed during the heating season when the exchanger is using the exhaust warm air to preheat the outside incoming air. However, heat exchanger might transfer sensible heat only or sensible and latent heat depending on the use. Various types of heat exchangers are available nowadays with efficiencies ranging from 55% to almost 80% [7]. Moreover, various types of air to air energy recovery devices are available such as, the heat pipe exchanger, the open and closed runaround systems, the thermal wheel, and the plate exchanger [8].

Therefore, system designers are more frequently considering the use of air to air energy recovery systems to reduce the amount of energy HVAC systems consume. Local building jurisdictions and national design standards have also evolved to require energy recovery in more applications. Although established standards indicate when energy recovery systems are required and their minimum technical requirements, little guidance is available in the optimal selection of the type of system, and the control strategies that will produce maximum energy reduction [9]. Moreover, with the many benefits that comes with using a heat exchanger device some disadvantages may occur. One of the problems is the frosting inside the exchangers especially when used in cold regions as Canada and Northern Europe. This issue might result in decreasing the performance of the exchangers. Therefore, it should be taken into consideration when selecting the exchangers [10]. However,

\*Nabil Nassif, Email: [nassifnl@ucmail.uc.edu](mailto:nassifnl@ucmail.uc.edu)



with the growing industry and the new technologies that are being discovered every and the many heat recovery devices that are available nowadays. Many engineers wonder about the type of energy recovery device that should be used, what size, what type, what location and when? Also, what is the algorithm of integrating it into the HVAC system? What is the required maintenance? What is the payback period of that installation? And most importantly what can I expect the performance to be like? Since the answer might require an exact and detailed calculations. Many answers were discussed in few articles, but the search is still ongoing.

Previous studies have compared between multiple types of heat recovery systems. However, the studies have been done on the same zone, same AHU, and or same building but different AHU. On that note, a study was conducted on southern Illinois university laboratory building comparing between two types of energy recovery systems on two different zones. Compare to a 100% fresh air base case. It was found that the glycol runaround loop heat recovery system gave an annual energy saving of 17.8%. while, implementing a return air heat recovery system resulted in a 17.9% savings of the annual energy use [11]. The downside of those studies is that the comparison was done based on the same climate zone or location. For better estimation of the energy savings of implementing the energy recovery systems, different climate zones should be assessed. Since climate conditions of each location has tremendous effect on the performance of the energy recovery system.

This paper’s objective is to study the differences in energy consumption among various energy recovery systems located in various climate zones and provide useful guidance in selecting an efficient system that yields maximum energy savings. The proposed modelling technique is tested by performing experiments on a 3-ton DX split-system heat pump. The experiment is prepared utilizing a specially integrated unit in a controlled setting. The heat pump runs under various outside conditions and in heating and cooling modes. All the various types of energy recovery systems will be implemented in all the selected zones. And, the results will be fully discussed.

## 2. Energy Recovery Systems

The concept of air-to-air energy recovery is to use the energy of the exhaust air stream to precondition the ventilation air being brought in from outside. According to ASHRAE 90.1-2016, systems with good potential for energy recovery, dictated by location, operational hours, flow rates, and outside air percentages, shall implement energy recovery systems with an Energy Recovery Effectiveness (ERE) of  $\geq 50\%$ .

$$ERE = \frac{\dot{m}_{oa} (h_o - h_{o'})}{\dot{m}_{oa} (h_o - h_{ea})} \quad (1)$$

Heat recovery installation can considerably reduce the total energy consumption of the system as well as the capital cost of the secondary heating and cooling equipment. Air to air heat recovery is essential for HVAC design and operation for multiple reasons like:

- 1) Reducing the peak energy rates as well as operation cost
- 2) Allows for a higher ventilation rates that will help in creating a better indoor air quality rates at a minimum auxiliary energy consumption that will help not in only provide a better environment for the inhabitants but will also reduce the cost of operation [12].

Also, a study examined a CFD model implemented to optimize the energy performance while maximizing the indoor air quality of a building that is equipped with a heat recovery wheel. It was found that 43% of savings in the annual energy consumption was achieved when using the energy recovery wheel comparing to a 100% fresh air case (no energy recovery wheel). While maintain an acceptable IAQ level without violating the ASHRAE standard for acceptable CO2 concentration in breathing spaces [13].

Depending on various project factors, a design engineer may choose to implement different forms of air-to-air energy recovery systems for various reasons. For the scope of this study, the performance of a coil loop, heat pipe, sensible wheel, and total energy wheel will be evaluated against and optimized for various locations with various weather conditions.

Coil loops consist of two hydronic coils with a glycol solution as the energy transfer medium. Sensible heat is transferred between the exhaust air and the outside air, cooling the outside air in the summer and heating it in the winter. Coil loops have the advantage of not having any cross contamination between exhaust air and outside air, and not requiring air intakes and exhaust locations to be located near one another [14].

Heat pipes are similar to coil loops as they both recover only sensible heat through hydronic piping. Heat pipes however require the air streams to be side by side. The coils are continuous from one air stream to the other and are filled with refrigerant. The refrigerant is pressurized, and when exposed to warm air on one side and cold air on the other, convection is induced, allowing the system to operate without a pump. Though slightly more efficient than a coil loop, the continuous piping between air streams makes it hard to completely prevent cross contamination [14].

Sensible wheels are rotating cylindrical wheels positioned in the duct allowing air to pass through it. In the duct, the wheel is divided into two equal sections, allowing for both the outside air and the exhaust air to pass through. As the wheel rotates, it indirectly exchanges sensible heat between the outside air and the exhaust air streams. In winter, it exchanges heat from the exhaust air to the outside air and in summer, it exchanges heat from the outside air to the exhaust air. Though the heat transfer is indirect, there are areas of carry over near the divide of the wheel, where cross contamination of the air streams is possible. In order to make up for the cross leakage of airstreams, additional outside air is brought in based on the outside air correction factor [15].

Total energy wheels are similar to sensible wheels in their installation and rotating nature, however total energy wheels have the added benefit of also transferring latent energy between the air streams. The majority of wheels are made with a honeycomb type aluminum matrix coated with desiccant. This desiccant allows the wheel to run at lower outdoor temperatures in the winter without ice buildup. Two main media types used in wheels are synthetic fiber and polymers, both with similar effectiveness; they do, however, vary in depth, maintenance, and weight. In order to rotate the wheel a motor ranging from 1/3 to 5 HP is required depending on the media type [15].

In all these forms of energy recovery systems, there is additional fan power needed to overcome the added pressure drop caused by the added recovery systems. The fan energy consumed by a system is a function of the volumetric flow rate, differential pressure, and the efficiency of the fan. Listed in Table 1 is a breakdown of the various energy recovery systems in the study, and their associated technical information.

Table 1: Energy recovery systems and technical information

System	Transfer Medium	Sensible Eff.	Latent Eff.	Fan Pressure Drop
Coil Loop	30% glycol	45%	--	0.5 inwg
Heat Pipe	R134A	50%	--	0.7 inwg
Sensible Wheel	Wheel	70%	--	0.8 inwg
Total Energy Wheel	Wheel	70%	70%	1 inwg

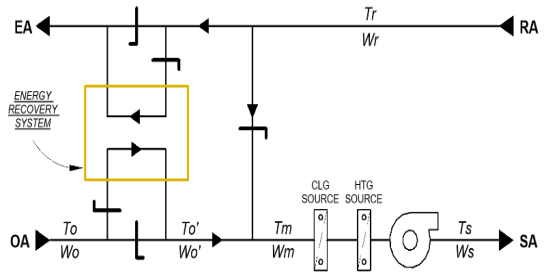


Figure 2: System schematic diagram with noted air stage points

### 3. Building Description

For the model, a 72,000 square-foot office building was assumed. Using ASHRAE Standard 62.1-2016, the minimum outside air flow rate was computed to be 7,000 CFM with 6,300 CFM of exhaust air. The building was assumed to be occupied on business days during business hours, 8am-4pm, with 1,000 people in the space. Building was modelled using eQuest software [16].

Table 2: Assumed building characteristics

Building Assumption	Value
Building Type	Office Space
Building Size	72,000 square-feet
OA to Building	7,000 CFM
EA out of Building	6,300 CFM
Building Schedule	Operational during business days; 8am – 4pm; 1000 building occupants

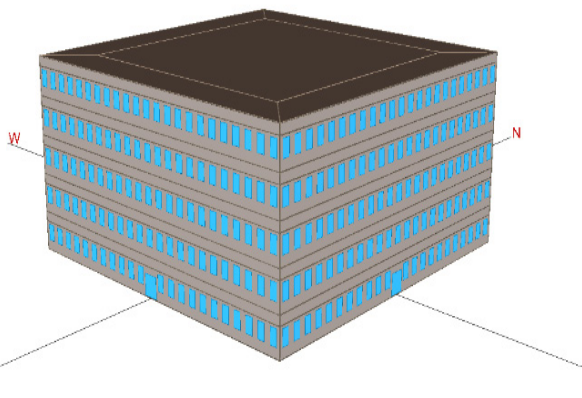


Figure 1: Building model in eQuest [16].

The model’s intended goal is to replicate the mechanical conditioning process that occurs in the building’s air handling units. In order to accomplish this, the process was broken down into four major steps. The stage points mentioned in the modelling section are illustrated in Figure 2 for ease of understanding the whole aspect.

### 4. Modeling

To study the building energy consumption with various energy recovery systems in different climate zones, a bin energy model was created [17]. Bin energy models refer to a procedure where annual weather data is sorted into discrete bins of weather conditions. As this study focuses on the comparison of energy consumption and associated energy reduction among different energy recovery systems in various locations, a bin model was deemed appropriate. Each bin contains the average number of hours of occurrence for a specific dry bulb temperature set during a year. The 8,760 hours in the year are divided into three scheduled categories based on occurrence: 12am-8am, 8am-4pm, and 4pm-12am. Along with each dry bulb temperature bin, the mean coincident wet bulb temperature is given. For this model, weather data was acquired from the National Climatic Data Centre [18]. The locations being tested are displayed in Table 3.

Table 3: Modeled locations and ASHRAE Climate Zone

Location	ASHRAE Climate Zones
Cincinnati, Ohio	4A
Fargo, North Dakota	7A
Miami, Florida	1A
San Francisco, California	3C
Phoenix, Arizona	2B

Moreover, the specification of each location as the environmental parameters and the annual heating, cooling and dehumidification design conditions are shown in table 4. Those data are given to illustrate the weather tendency in each location. The climate design information is obtained from the ASHRAE handbook of fundamentals [19]

Table 4: Climate design information for the modeled locations (Latitude and longitude: ° elevation: ft/ DB: °F/ HDD and CDD 65: Annual heating and cooling degree-days, base 65°F, °F-day)

Location	Cincinnati, Ohio	Fargo, North Dakota	Miami, Florida	San Francisco, California	Phoenix, Arizona
latitude	39.1	46.93	25.82	37.62	33.44
Longitude	81.44	96.81	80.3	122.4	111.99
elevation	1237	899	30	20	1106

Heating DB (F°)	8.1	-19.3	47.6	39.1	38.7
Cooling DB (F°)	92.8	90.7	91.8	82.8	110.3
Enthalpy	88.1	84.5	90	74.2	106.4
HDD	4744	8729	126	2689	923
CDD	1155	555	4537	144	4626

#### 4.1. Outside-Air Condition Calculation

Calculate the outside air temperature and humidity ratio after passing the energy recovery system based on the efficiency of the specific system. The formula used was a derivation based on the efficiency formula provided by ASHRAE Standard 90.1-2016.

$$T'_o = -\eta_s * \left(\frac{\dot{m}_{ea}}{\dot{m}_{oa}}\right) * (T_o - T_r) + T_o \quad (2)$$

$$W'_o = -\eta * \left(\frac{\dot{m}_{ea}}{\dot{m}_{oa}}\right) * (W_o - W_r) + W_o \quad (3)$$

#### 4.2. Mixed-Air Condition Calculation

Calculate the mixed air temperatures between the outside air stream and return air stream. In this energy model, the mixed air is used as the control point for the energy recovery system. For two different cases, (1) with energy recovery and (2) without energy recovery, the required outdoor air fraction to meet the supply air temperature is calculated.

$$\lambda_{req} = \frac{T_s - T_r}{T'_o - T_r} \quad (4)$$

$$\lambda = \frac{\dot{m}_{oa}}{\dot{m}_{total}} \quad (5)$$

For both cases, the mixed air temperature and humidity ratio is calculated with the subsequent outside air ratio.

$$T_m = \lambda_{req} * T'_o + (1 - \lambda)T_r \quad (6)$$

$$W_m = \lambda_{req} * W'_o + (1 - \lambda)W_r \quad (7)$$

The difference between the supply air temperature and the mixed air temperature with energy recovery is calculated. Next, the difference between the supply air temperature and the mixed air temperature without energy recovery is calculated. After evaluating the two results, the energy recovery system is signaled to run or not to run [20]. If the energy recovery system was seen to remove or add too much heat, the energy recovery system would be signaled to modulate in order to meet supply air conditions. The supply air temperature is calculated based on a conditional equation, where in full cooling the supply air temperature is 55°F, and in full heating the supply air temperature is 65°F. Between the two temperatures, the supply air varies linearly with a slope of 0.5°F. The formula can be seen in Eq. 8.

$$T_s = 55 + 0.5(65 - T_{oa})$$

Lower limit = 55 °F  
Upper limit = 65 °F

(8)

#### 4.3. Load Calculation

The load associated with cooling and heating the air on a system level is then calculated for each temperature bin using the following equations.

$$q_s = \dot{m}_{total} * C_p * \Delta t \quad (9)$$

$$q_l = \dot{m}_{total} * h_g * (W_m - W_s) \quad (10)$$

$$q_t = \dot{m}_{total} * \Delta h = q_s + q_l \quad (11)$$

To find the annual load energy, the load is multiplied by the temperature hourly bin for the occupied schedule in the specific weather file.

#### 4.4. Energy Consumption Calculation

In order to calculate the building energy consumption, further assumptions were made about the systems within the building [21]. The building is equipped with an air-cooled chiller, with the efficiency varying as a function of the outside air temperature, between 1.2 kw/ton and 1.0 kw/ton [22]. A natural gas, non-condensing boiler with a COP of 0.85, meets the building's heating demand. In order to get comparative results for the whole building's energy consumption, eQuest was used to determine the zone level heating requirements for each temperature bin. The building was modeled in each tested city using ASHRAE Standard 90.1-2016 minimum values.

Mentioned in the 'Energy Recovery Systems' section are listed parasitic losses associated with implementing energy recovery systems. For the purpose of this study, these losses are assumed to be limited to the additional fan energy required to overcome the additional pressure drop over the energy recovery system. To calculate the fan energy associated with the air handling unit, it was assumed that the building level static pressure drop is 7 in.w.g. The additional static pressure added to this assumption by the energy recovery system is given in Table 1. The efficiency of the fan is given by the minimum ASHRAE 90.1-2016 value.

$$Fan\ Break\ Horse\ Power = \frac{(CFM_{total} * \Delta p)}{6356 * \eta_{fan}} \quad (12)$$

### 5. Results

For each location, four energy recovery systems were compared to the baseline case with no energy recovery system. The data is broken down by cooling, heating, and fan energy. The total system energy consumed is the summation of the energies. Note, that only energy within the variable air volume system is considered; lighting, plug loads, space equipment, etc. are not being accounted for in the total energy consumption calculation.

In Cincinnati, OH, it was found that as the efficiency of the energy recovery system increases, the total energy consumption decreases. With no energy recovery the total energy consumed by the system was 949,947 kbtu/year. With the coil loop, heat pipe, sensible wheel, and total energy wheel, the total energy consumption was 855,458 kbtu/year; 834,800 kbtu/year; 794,719 kbtu/year; and 784,438 kbtu/year respectively. The energy consumed by the fan increases marginally as the static pressure over the recovery system increases, but not enough to outweigh the energy savings.

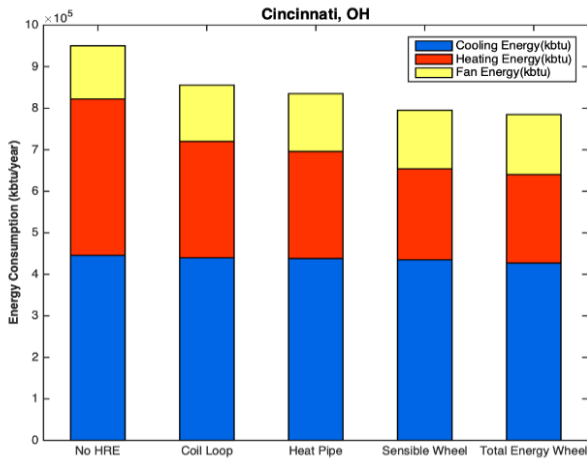


Figure 3: System energy consumption for various energy recovery systems; Cincinnati, OH.

When analysing the differences between energy recovery systems in Fargo, ND, the total energy consumption is noticeably reduced as the efficiency of the recovery system increases. The base case with no recovery system consumed 1,180,362 kbtu/year. Implementing a coil loop, heat pipe, sensible wheel, and total energy wheel yielded energy consumptions of 978,127 kbtu/year, 930,013 kbtu/year, 833,789 kbtu/year and 827,387 kbtu/year respectively. Comparing the sensible energy wheel and the total energy wheel, the sensible wheel saves 29.36% compared to the baseline, and the total energy wheel saves 29.90%, a very minute difference. Further analysing the breakdown of the total energy consumption, the required cooling energy decreases as the efficiency increases, with the noticeable difference of the total energy wheel. Cooling energy for the total energy wheel was reduced by 1.04% compared to the baseline, whereas the sensible wheel reduced cooling energy by 2.33%. When reviewing the energy model, the higher consumption in cooling energy seen in the total energy wheel compared to the sensible wheel can be attributed to the control sequences using dry bulb temperature. Because the system is being controlled by dry bulb temperature, additional moisture is transferred into the mixed air which now must be removed through mechanical cooling. Due to Fargo's cold temperature, the large amount of heating energy saved by the total energy wheel outweighs the added fan and cooling energy, resulting in the most energy saved.

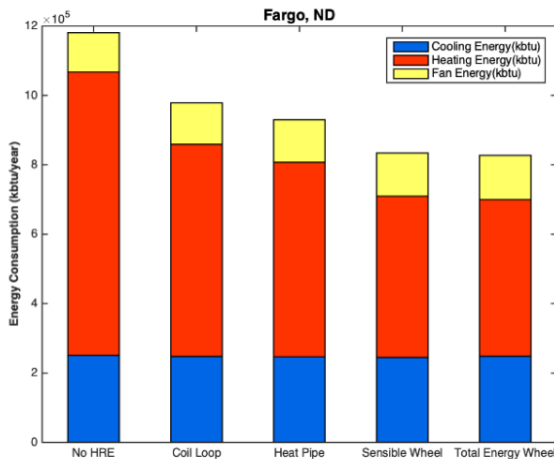


Figure 4: System energy consumption for various energy recovery systems; Fargo, ND.

In Phoenix, AZ, the baseline with no recovery system consumed 844,231 kbtu/year. The coil loop, heat pipe, sensible wheel, and total energy wheel consumed 824,696 kbtu/year, 822,316 kbtu/year, 811,621 kbtu/year, 843,835 kbtu/year respectively. The system with the largest reduction in energy is the sensible wheel, producing a savings of 3.86% compared to a reduction of only 0.05% by the total energy wheel. The sensible wheel reduced cooling energy by 6.60% while the energy wheel reduced it by only 2.35%. Similar to Fargo, the lower reduction in cooling energy of the total energy wheel can be attributed to the programmed system being controlled by dry bulb temperature.

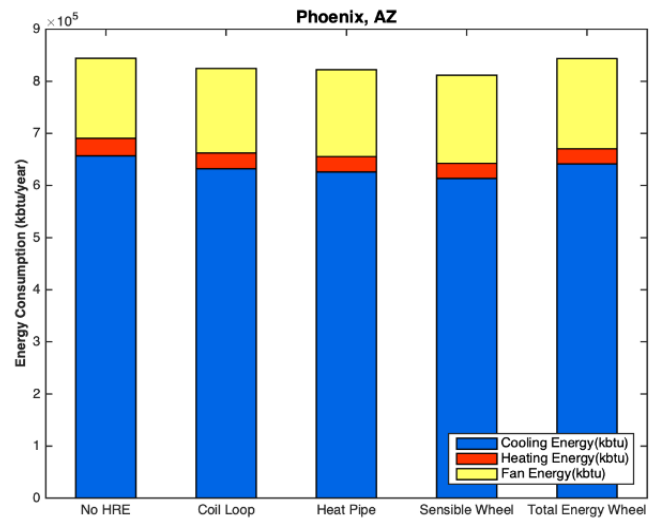


Figure 5: System energy consumption for various energy recovery systems; Phoenix, AZ.

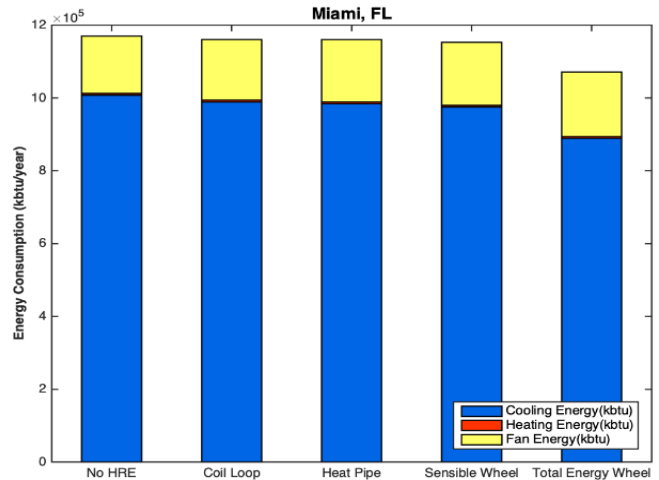


Figure 6: System energy consumption for various energy recovery systems; Miami, FL

In Miami, FL all systems show a savings in total energy consumption as the efficiency of the recovery systems increased. The baseline case consumed 1,170,108 kbtu/year, whereas the coil loop, heat pipe, sensible wheel, and total energy wheel consumed 1,160,188 kbtu/year, 1,159,955 kbtu/year, 1,152,901 kbtu/year, and 1,071,395 kbtu/year respectively. The sensible recovery technologies, the coil loop, heat pipe, and sensible wheel yielded savings from the baseline of 0.85%, 0.87%, and 1.47% respectively. Total energy wheels, which transfer both sensible and latent energy, reduces total energy consumption by 8.44%, almost a 7% gain in efficiency compared to the sensible wheel. The



sensible wheel and total energy wheel have the same heating consumption, 4,774 kbtu/year, because the heating load in the air handling unit is entirely accomplished through energy recovery. This remaining heat is added at the zone level.

In San Francisco, CA the energy consumption for all energy recovery systems increase compared to the baseline. The baseline model with no recovery system consumed 453,100 kbtu/year, and the coil loop, heat pipe, sensible wheel, and total energy wheel consumed 453,395 kbtu/year, 456,159 kbtu/year, 457,069 kbtu/year, 462,986 kbtu/year respectively. For sensible recovery systems (coil loop, heat pipe, and sensible wheel), the heating and cooling energy is minimally reduced. The increase in parasitic energy used by the fan outweighs the savings from heating and cooling. In contrast the total energy wheels usage of cooling energy is greater than that of the baseline due to the control sequences based on dry bulb temperature. The implementation of the total energy wheel increased energy consumption by 2.18% compared to the baseline.

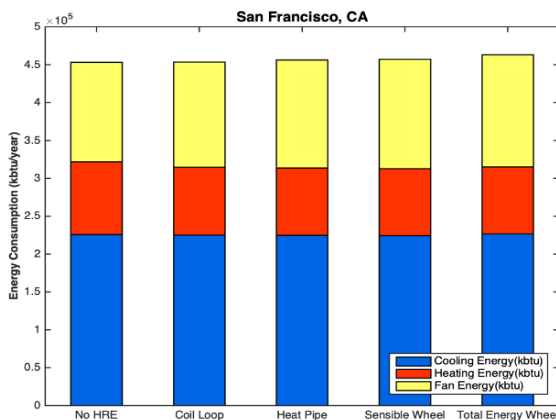


Figure 7: System energy consumption for various energy recovery systems; San Francisco, CA.

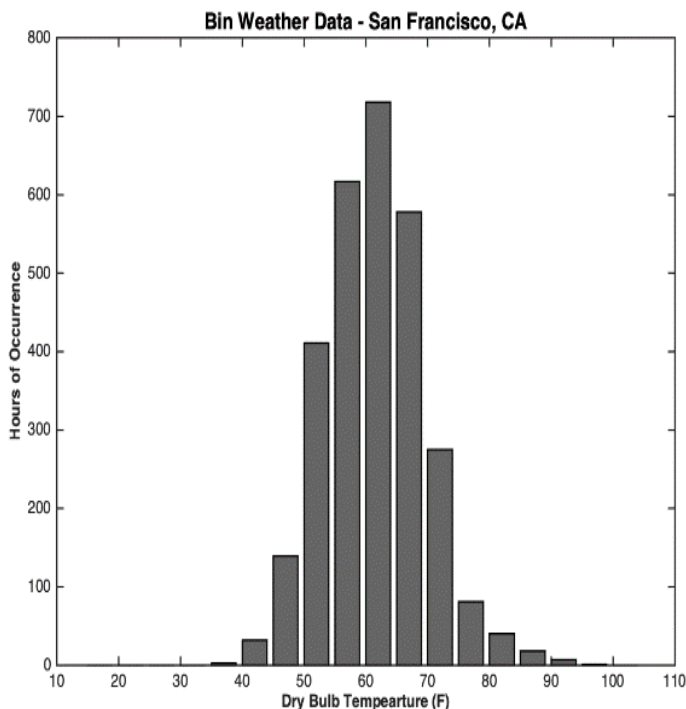


Figure 8: Hourly temperature occurrence; San Francisco, CA.

San Francisco has a noticeably mid climate, with 70% of the hours in a year falling within the 52F-62F range. This climate requires little to no air conditioning. When looking at the energy model, it became clear that for most temperature bins, the control sequencing was able to control the dampers in order to meet the supply air requirements without mechanical conditioning.

## 6. Conclusion and Discussion

Air to air energy recovery devices are reliable and effective methods that are implemented to reduce the total auxiliary energy use as well as improving the indoor air quality through higher ventilation rates. Systems that are equipped with energy recovery devices were proven to be more effective than the ones who don't. With the availability of several types of energy recovery devices choosing the most suitable one for the application is crucial to achieve comfortable and cost-effective buildings [23].

In this paper, four types of energy recovery systems were studied in five different climate zones, and their system's energy consumption were analysed. The energy recovery systems are:

1. Coil loop,
2. Heat pipe,
3. Sensible wheel,
4. Total energy wheel.

The locations are Cincinnati, OH, Fargo, ND, Miami, FL, San Francisco California, and Phoenix Arizona. Figure 9 illustrates the reduction in energy for each system in each location.

From the results of this analysis, an underlying takeaway is that energy recovery systems produce different results depending on the conditions and locations in which they operate. The largest savings in energy consumption occurs in climates that are heating dominated, as seen in the results from the simulation in Fargo, ND (ASHRAE Climate Zone 7). It is also notable that in areas with dry climates and high sensible heat ratios, the use of a system that transfers both sensible and latent energy must be studied with care in regard to the control sequencing.

Moreover, in this study the sole control point of the system was based on dry bulb temperature, limiting the optimization of the systems. Expanding the possible control sequences to include the air's humidity ratio or enthalpy will create a more perspective aid in selecting energy recovery systems and control strategies. This study is the basis for the development of a design tool that will pair with known energy simulation software such as eQuest to provide project specific energy reduction values based on various energy recovery systems.

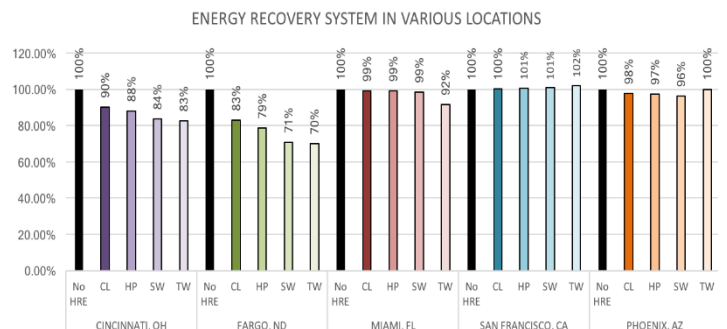


Figure 9: Comparison of all energy recovery systems in all tested locations values are presented as a percentage of the baseline with no energy recovery system.

**Conflict of Interest**

The authors declare no conflict of interest.

**References**

[1] U.S. Energy Information Administration, “How Much Energy is Consumed in U.S. Residential and Commercial Buildings?”, 3 May, 2018. Available: <https://www.eia.gov/tools/faqs/faq.php?id=86&t=1>

[2] U.S. Department of Energy, “Chapter 5: An Assessment of Energy Technologies and Research Opportunities”, September, 2015. Available: <https://www.energy.gov/sites/prod/files/2017/03/f34/qtr-2015-chapter5.pdf>

[3] International Well Building Institute, “The WELL Building Standard”, 20 October, 2014.

[4] American Society of Heating Refrigerating and Air-conditioning Engineers, Inc. (ASHRAE), Standard 62.1-2016, “Ventilation for acceptable Indoor Air Quality”, 2016.

[5] R. Talib, N. Nassif, M. Arida, & T. Abu-Lebdeh, “Chilled water VAV system optimization and modeling using artificial neural networks”, American Journal of Engineering and Applied Sciences, **11** (4), 1188-1198, 2018. doi:10.3844/ajeassp.2018.1188.1198

[6] W. Wu, X. Li, T. You, & Ohio Library and Information Network, “Absorption heating technologies: Efficient heating, heat recovery and renewable energy”, Singapore: Springer, 2018. doi:10.1007/978-981-15-0470-9

[7] Dieckmann, John, Kurt W. Roth, and James Brodrick. "Air-to-air energy recovery heat exchangers." ASHRAE Journal **45**(8), 57, 2003.

[8] Sauer Jr, H. J., and R. H. Howell. "Promise and potential of air-to-air energy recovery systems." International Journal of Refrigeration **4**: 182-194. 1981.

[9] American Society of Heating Refrigerating and Air-conditioning Engineers, Inc. (ASHRAE), Standard 90.1-2016, Energy Standard for Buildings Except Low-Rise Residential Buildings, Atlanta, GA, USA, 2016.

[10] Nasr, Mohammad Rafati, et al. "A review of frosting in air-to-air energy exchangers." Renewable and Sustainable Energy Reviews **30**: 538-554, 2014.

[11] N.T. Madineedi, J.M. Harrell, & J.A. Mathias, “Redesigning the HVAC system of a university laboratory building”, ASHRAE Transactions, **123**(2), 217, 2017.

[12] Besant, Robert W., and Carey J. Simonson. Air-to-air energy recovery. ASHRAE journal **42**(5), 31-43, 2000.

[13] M. Kanaan, “CFD optimization of return air ratio and use of upper room UVGI in combined HVAC and heat recovery system”, AMSTERDAM: Elsevier Ltd, 2019. doi:10.1016/j.csite.2019.100535.

[14] Trane Engineers Newsletter Live – Ronnie Moffitt, Dennis Stanke, John Murphy, Jeanne Harshaw, Air-to-Air Energy Recovery, 10 August 2012.

[15] Freund, Sebastian, “Simulation of Air-to-Air Energy Recovery Systems for HVAC Energy Conservation in an Animal Housing Facility,” Master of Science thesis, Solar Energy Laboratory University of Wisconsin-Madison, 2003.

[16] eQuest. Quick Energy Simulation Tool, eQuest Version 3.65. Available: <http://doe2.com/equest/>

[17] M.N. Hydeman, P. Sreedharan, and S. Blanc, “Development and Testing of a Reformulated Regression-Based Electric Chiller Model, ASHRAE Transactions **105**(2), 1118-27, 2012.

[18] National Climatic Data Center - National Oceanic and Atmospheric Administration: Cincinnati, Ohio; Fargo North Dakota; Miami, Florida; San Francisco, California; Phoenix, Arizona. Available: <http://www.ncdc.noaa.gov>

[19] Ashrae handbook, “Fundamentals by American Society of Heating, Refrigerating and Air-Conditioning Engineers”, 2013. eBook: Full Text Online, TH7011.A4222 2013eb

[20] N. Nassif, S. Kaji, and R. Sabourin, “Optimization of HVAC Control System Strategy Using Two-Objective Genetic Algorithm”, HVAC&R Research. **11**(3), 459-486, 2005. DOI: 10.1080/10789669.2005.10391148

[21] W. Wu, H.M. Skye, H. M., & P.A. Domanski, “Selecting HVAC systems to achieve comfortable and cost-effective residential net-zero energy buildings. Applied Energy, **212**, 577-591, 2018. doi:10.1016/j.apenergy.2017.12.046.

[22] Ali, M. Tauha – MathWorks MATLAB, SI Psychrometric Chart, 2015. Available: <https://www.mathworks.com/matlabcentral/fileexchange/49154-si-psychrometric-chart>

[23] X. Pei, “Application of exhaust heat recovery in energy saving of HVAC”, IOP Conference Series: Earth and Environmental Science, **295**, 52009, 2019. doi:10.1088/1755-1315/295/5/052009

**Nomenclature**

$ERE$	energy recovery effectiveness
$\dot{m}_{oa}$	mass flow rate – outside air (lb/hr)
$\dot{m}_{ea}$	mass flow rate – exhaust air (lb/hr)
$\dot{m}_{total}$	mass flow rate – total system air (lb/hr)
$CFM_{oa}$	outside air volumetric flow rate (ft <sup>3</sup> /min)
$CFM_{ea}$	exhaust air volumetric flow rate (ft <sup>3</sup> /min)
$CFM_{total}$	system total air volumetric flow rate (ft <sup>3</sup> /min)
$CDD$	Cooling degree days
$HDD$	Heating degree days
$DB$	Dry bulb temperature (F°)
$h_o$	outside air enthalpy (btu/lb)
$h_{o'}$	outside air enthalpy after energy recovery equipment (btu/lb)
$h_{ea}$	exhaust air enthalpy (btu/lb)
$\eta_s$	sensible efficiency of recovery system (%)
$T_o$	outside air temperature (°F)
$T_o'$	outside air temperature after energy recovery system(°F)
$T_r$	return air temperature (°F)
$h_g$	latent heat of vaporization (btu/lb)
$\eta_l$	latent efficiency of energy recovery system (%)
$W_o$	outside air humidity ratio (lb/lb)
$W_o'$	outside air humidity ratio after energy recovery system (lb/lb)
$W_r$	return air humidity ratio (lb/lb)
$\lambda_{req}$	outside air ratio to meet T <sub>s</sub> (%)
$\lambda$	Outside air ratio (%)
$T_m$	mixed air temperature (°F)
$W_m$	mixed air humidity ratio (lb/lb)
$q_s$	sensible load (btu/hr)
$q_l$	latent load (btu/hr)
$q_t$	total load (btu/hr)
$\Delta p$	differential pressure (in w.g.)
$\eta_{fan}$	fan efficiency (%)
$OSCF$	outside air correction factor



## Fraud Detection Call Detail Record Using Machine Learning in Telecommunications Company

Ma'shum Abdul Jabbar<sup>1,\*</sup>, Suharjito<sup>2</sup>

<sup>1</sup>Computer Science Department, Binus Graduate Program - Master of Computer Science, Bina Nusantara University, Jakarta 11480, Indonesia

<sup>2</sup>Computer Science Department, Binus Online Learning, Bina Nusantara University, Jakarta 11480, Indonesia

### ARTICLE INFO

Article history:

Received: 13 May, 2020

Accepted: 19 June, 2020

Online: 06 July, 2020

Keywords:

Fraud

Call Detail Record

K-Means

DBSCAN

### ABSTRACT

Fraud calls have a serious impact on telecommunications operator revenues. Fraud detection is very important because service providers can feel a significant loss of income. We conducted a fraud research case study on one of the operators that experienced fraud in 2009 and 2018. Call Detail Record (CDR) containing records of customer conversations such as source and destination number, call start time, duration of calls at the operator can be a source of information to use in fraud detection. The method used in this study uses machine learning with unsupervised learning techniques which are quite popular methods used in fraud detection. The purpose of this study is to propose an effective method that can be applied to detect fraud on the CDR. Variables used include caller number, number dialled, duration, fee and destination city of the dataset totalling 11,418 rows from record periods 01 to 31 May 2018. In analyzing our CDR using the K-Means and DBSCAN algorithms, we then evaluate the results to calculate accuracy by comparing to actual fraud data. Based on evaluations using confusion matrix on actual CDR fraud, we obtained the K-Means algorithm to show a better accuracy value to model fraud on telecommunications CDR compared to DBSCAN.

### 1. Introduction

Fraud in the telecommunications industry has a serious impact on revenue and customer relationships, especially when losses continue to increase over time [1]. As the telecommunications industry advances, the problem of telecommunications fraud has also grown in recent years [2]. The telecommunications company, which is the case study in this research has adopted VoIP (Voice over Internet Protocol) since 2015, which previously applied EWSD (Electronic Worldwide Digital Switch) technology. The company experienced a fraud incident in 2009 with a total loss of up to Rp 1 billion, fully borne by the company. The case was caused by traffic from customers to use telephone numbers to international numbers with quite expensive rates and frequent calls. As we know IP technology allows connections that are more flexible than analogue but also open more gaps for attackers to abuse the network and use unauthorized services. Despite using VoIP technology much better than EWSD, it still happens in 2018 with a total loss of up to Rp 28 million in which the company bore 15%, the customer bore 85%. This number has dropped dramatically compared to cases that have been experienced before,

but this is still an important thing to follow up for the company due to the unavailability of anomalous detection that handles this event. Instead of reducing similar risks, precautions are taken by closing international call gates for call traffic to the destination country that was once indicated as an anomaly. This problem motivates us to provide a solution for fraud modelling to find out the call history of potential customers as an anomaly, by grouping potential calls as a fraud. With this effort, it is expected that the customer will fully bear the final use burden that is billed to the customer because it does not include improper use.

Fraud according to the Black's Law Dictionary is intentional fraud or dishonesty committed by one or more individuals, generally intended for financial gain [3]. Telecommunications fraud is a combined illegal activity of subscription identity theft, unauthorized access and fraudulent sharing of international traffic revenue [4]. Fraud detection is a popular application used by the Revenue-assurance division to detect and stop all types of fraud on wired and wireless networks [5]. Fraud detection is very important because providers can feel a significant loss of income due to anomalies [6]. Telecommunications companies often experience huge financial losses due to fraud incidents caused by their services, and this makes the importance of fraud detection to

\*Corresponding Author: Ma'shum Abdul Jabbar, [mashum.jabbar@binus.ac.id](mailto:mashum.jabbar@binus.ac.id)

reduce the impact of this risk [7]. Fraud on telecommunications can be divided into several types, and the most typical is accessing calls using the original customer account to make fraudulent calls [8]. Imagine an unknown call from a local number, and that from a friend or family who lives abroad, it's really strange to receive international calls from a local number, basically this also fraud [9]. Current technology has provided protection against various attacks, but not many are designed to detect complex fraud operations [10]. When using VoIP, several problems in security and quality of service will arise, VoIP infrastructure must be equipped with a security shield to protect itself from various forms of security threats [11]. The traditional approach to detecting fraud in telecom is usually only by blacklisting the numbers that commit fraud. It can be tricked by changing their numbers which is very easy to do on VoIP technology [12].

The pattern of call traffic is expected to be understood by telecommunications provider with the availability of CDR (Call Detail Record) [13]. The CDR contains spatio-temporal data sets that researchers can use in various fields, and also by investigators [14]. CDR have contained metadata such as source and destination number, call start time, call duration, associated with each telephone call [15]. When the CDR is combined, we can pair accounting and historical data from each call so that it is possible to get a complete picture of service and customer billing [16]. The CDR contains billions of detailed call records with descriptive information available for mining, which is very important for marketing and can also be used to identify fraud activities [17]. Analysis of the CDR can determine the specific activities of users in a particular region at a certain date or time, abnormal behavior, or when there are changes in the pattern of user activity in the region can be considered an anomaly [18]. Massive data sources such as CDR to effectively analyze large amounts of data are the main challenges faced by telecommunications operators, especially when new techniques and machine learning algorithms that require large enough memory [19].

In large data sets in many cases, the original tags of the data are not always available, so it is not possible for researchers to use supervised learning methods. This is the main reason why unsupervised learnings such as clustering are widely used in data mining for CDR. K-means algorithm is the most common partitioning method [20]. Another technique is to classify data into groups using DBSCAN (Density-Based Spatial Clustering of Applications with Noise), which is a density-based clustering algorithm [21]. Some existing clustering methods can balance accuracy and efficiency, but cluster quality is not so good that researchers usually combine other algorithms or use optimization strategies to improve the clustering method [20].

Based on the fraud problem above, we propose a fraud detection mechanism that can be applied effectively in telecommunications companies. Our approach is to analyze and clustering the CDR to find anomalies traffic using machine learning. Through machine learning, we can get a bunch of suspicious calls and help companies find fraud in the cluster. In this study, we made the following contributions:

- Propose an effective method implemented to detect fraud within call detail record analysis, utilizing variables that are quite prominent on traffic patterns in the telecommunications company case study.

- Using machine learning with unsupervised learning techniques to detect anomaly traffic, using the K-means algorithm and DBSCAN to find the best performing techniques for detecting fraud.
- In this study, we used a CDR which contained a fraud label to evaluate the performance of the method used. The results obtained show that our approach can detect fraud quite accurately.

## 2. Related Works

Related works are based on fraud detection using the methodology of unsupervised learning and semi-supervised. Presentation of the concept of identification and verification of anomalies using clustering methods to process the partitioning of data groups into a small number of groups and subgroups using the K-means algorithm. Information is obtained from the cellular network CDR which is exploited to determine the user's anomalous behavior at a certain date and time. The results obtained are a number of clusters with different data points, where groups that have fewer objects or data points are groups that contain anomalous activity. Evaluation is done by training the neural network model with anomaly and anomaly free data and observing the mean squared error. The conclusion is to use the k-mean clustering technique to show unusual user activity or referred to as an anomaly [22].

Another approach taken is to analyze the difference between fraud and normal call characteristics using the K-means algorithm for a series of statistical features, including the number of calls per number, average talk time, variant talk, time, number of roaming calls, and number of numbers called, etc. The results show that elements in the cluster with high call counts, low talk ratios, low call success rates, call durations have high peak values and large standard deviations, but the small average value in this cluster is most likely fraudulent activity. Evaluate using Hopkins statistics and choose the appropriate grouping parameters through network search. Hopkins statistics are calculated on a data set obtained after principal component analysis, and the results are 0.998 with data having a high tendency to the group. The conclusion in this paper using CDR data to train the K-Means model, potential category information and some feature patterns can be found through various cluster analysis. The method in this article does not require class annotations, which reduces data acquisition difficulties. When the amount of data increases, the accuracy in fraud detection increases [23].

There is further research that combines K-means with Dynamic Time Warping (DTW) to determine the distance between sample points and produce a partition basis on time series data with each individual feature. Furthermore, the results of grouping individual features are included in the final result. The results show that experiments in real-time data sets show the effectiveness of this method. The results of the grouping were evaluated using internal (Number of Quadratic Errors) and external indicators (Purity of each group was calculated from primitive data labels). On this basis, several features were chosen to do a combinational grouping based on K-means. The conclusion in this study, most of the features built can describe user behavior and identify groups with a higher proportion of fraud. Experiments show that the grouping method that combines better than traditional K-means. In addition, combinational

grouping methods based on grouping single features have helpful results, which can help identify fraud groups that have certain behavioral models [20].

The difference between this study and the previous study is that practiced two algorithms to analyze CDR to provide results in fraud detection, which are K-Means and DBSCAN. The minPts value for calculations on DBSCAN is same with k value in K-Means clustering, which is obtained through the elbow method. While the eps parameter values on DBSCAN, we propose various from the default is 0.5, 1 and 1.5 therefore to obtain the appropriate eps value used that is 1.5, and the variables we use in the experiment are only those that we think need to be analyzed and provide enough information to analyze fraud. For the results of this study, we conducted an evaluation using accuracy, recall, and precision from the CDR which contained actual fraud information. Accordingly, we will immediately receive result the methods we use with machine learning.

### 3. Research Methodology

The research framework can be the basis for determining the steps to be taken in research. Next will be followed by a discussion of methods that will be used to realize solutions to research problems.

#### 3.1. Proposed Methods

Based on the latest methods and trends methods for detection of anomalies in the field of telecommunications. In this paper we propose that DBSCAN and K-Means use research variables that influence fraud activities, namely source number, destination number, duration, cost, and destination country. With this combination it is hoped that CDR can better analyze the identification of anomalies in the call log. In addition, the CDR used contains actual fraud in accordance with identification from the company. So, it will be easier to evaluate the results of the algorithm used to measure the accuracy of the results obtained. Research framework can be seen in Figure 1.

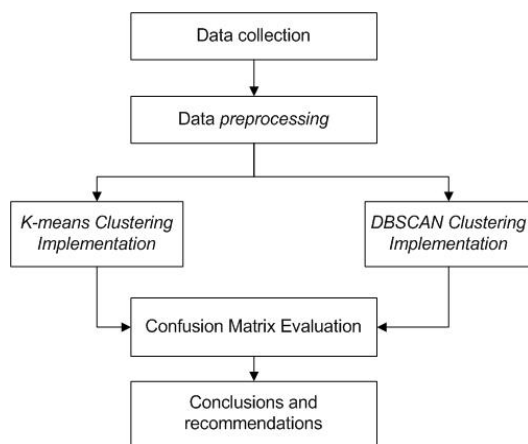


Figure 1: Research Framework

#### 3.2. K-Means

K-means is one of the algorithms that enters the application of data mining with a non-hierarchical clustering method, while the required clustering is data that has no label or class [24]. K-means

method is quite simple, starting with the selection of the number of clusters as many K pieces and then K pieces of data are taken randomly from the dataset as centroids to represent a cluster. All data is then calculated the distance from the centroid and each data will be a member of a cluster represented by a centroid that has the closest distance to the data. Finally, the recalculation of the centroid value obtained from the average value of each cluster [25]. The steps in the K-Means algorithm are as follows [24]:

1. Determine the value of K.
2. Take K from the dataset randomly as centroid.
3. Calculate the distance of each data with each centroid

**Definition 1:** Calculate the Euclidean distance between two points in space [26] :

$$d(x_i, y_j) = \sqrt{(x_{i1} - y_{j1})^2 + (x_{i2} - y_{j2})^2 + \dots + (x_{in} - y_{jm})^2} \quad (1)$$

Where  $d(x_i, y_j)$  : distance between data x to data y,  $x_i$  : data x, and  $y_j$  : data y.

**Definition 2:** The average distance between data objects:

$$MeanDist = \frac{1}{C_n^2} \sum d(x_i, x_j) \quad (2)$$

Where  $n$  is the number of data objects in the data set,  $C_n^2$  is the number of couples taken from  $n$  data points.

4. Determine the cluster membership of each data. Data becomes a member of the cluster with the closest centroid distance.
5. Update the centroid point value.

$$\mu_k = \frac{1}{N_k} \sum_{i=1}^{N_k} x_i \quad (3)$$

$\mu_k$  : centroid point of the K-cluster  
 $N_k$  : the amount of data in the K-cluster  
 $x_i$  : i data in K-cluster

6. Repeat steps two through five until there are no more cluster member changes.

The K-means Flowchart can be seen in Figure 2

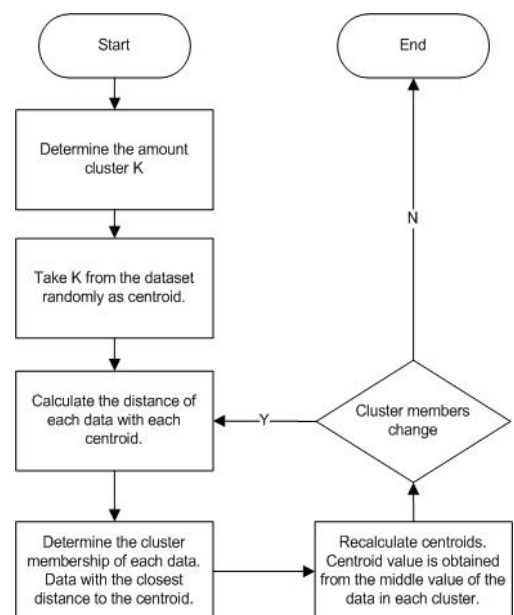


Figure 2: K-Means Flowchart [24]

### 3.3. DBSCAN

Density-Based Spatial Clustering of Application with Noise (DBSCAN) is a clustering algorithm based on data density. The concept of density is the amount of data (Minpts) within the eps radius ( $\epsilon$ ) of each data [27]. The DBSCAN flowchart algorithm can be seen in Figure 3.

$$N_{Eps(p)} = \{ q \in D \mid dist(p, q) < eps \} \tag{4}$$

$N_{Eps(p)}$  : Minimum number of points for a cluster  
 $eps$  : Maximum distance from spatial data  
 $dist(p, q)$  : The distance between point p and point q

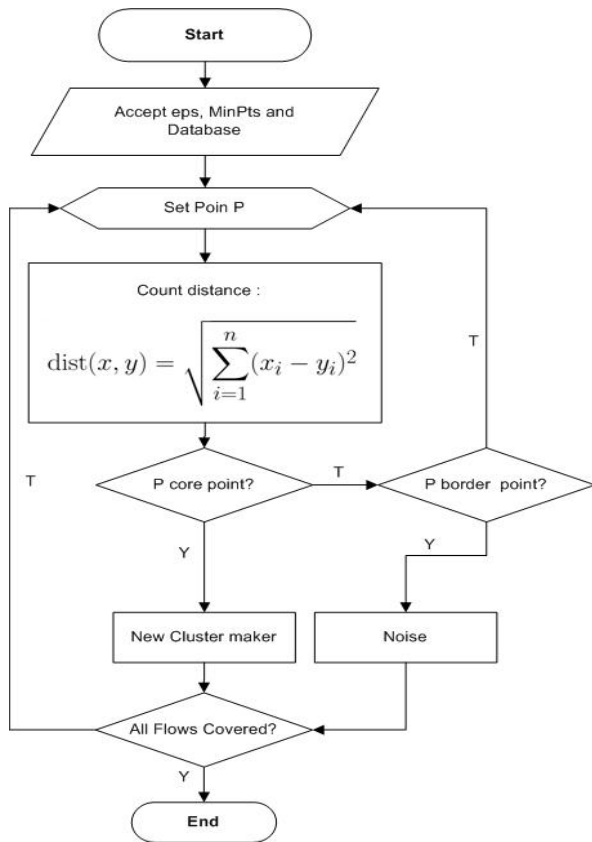


Figure 3: DBSCAN Flowchart [27]

The steps in the DBSCAN algorithm on the flowchart are as follows :

1. Determine the eps and minpts values.  
 DBSCAN algorithm requires two parameters : Eps, which is used to determine the neighbouring area of an object (or point) and MinPts, which is the minimum number of points in the Eps radius [28]. In this study, it is proposed to determining the value of Eps and MinPts with trial and error parameter, meaning that it determines the value of a parameter must be doing several times to get the expected number of clusters.
2. Specify random P points to start clustering.
3. Calculate all distances between points that have a density-reachable to the value of P.
4. If the number of points is greater or equal to MinPts then P is a core point and a cluster will be formed.
5. If p is a border point, it will proceed to another point.

DBSCAN can only be used on 2-dimensional spatial data. In another study, the DBSCAN algorithm was modified so that it could be used on spatial objects in 3-dimensional space [29]. This equation will be used to calculate distances in 3D space :

$$d(x_1, y_1, z_1), d(x_2, y_2, z_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \tag{5}$$

### 3.4. Confusion Matrix

Accuracy singly takes not tell the full story when working with unbalanced class datasets, where there is a significant discrepancy between the number of positive (fraud) and negative (non-fraud) labels [1]. Precision, Recall [30] and accuracy [31] can be used as criteria for classifier evaluations. This parameter is related to True Positive and False Positive (TP / FP) which refers to the number of positive predicted true / false, and also True Negative and False Negative for the number of negative predicted true / false (TN / FN). Confusion Matrix can be used to measure machine learning performance in classification, it can be seen in Figure 4.

		Actual Value	
		Positive (1)	Negative (0)
Predicted Value	Positive (1)	<b>TP</b>	<b>FP</b>
	Negative (0)	<b>FN</b>	<b>TN</b>

Figure 4: Confusion Matrix [31]

$$Recall = \frac{TP}{TP+FN} \tag{6}$$

$$Precision = \frac{TP}{TP+FP} \tag{7}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

## 4. Results & Discussion

Data preprocessing and evaluation of call fraud detection proposed using machine learning by utilizing Jupyter Notebook (Anaconda3) as a Python programming language tool.

### 4.1. Data collection

We use CDRs from customer billing records for the usage period from 01 to 31 May 2018. The data sources obtained were 11,418 rows and had many atribut and we choose 6 attributes to be used as research variables like anumber (Caller ID), bnumber (Called ID), duration, fee (cost), destination, and type. Dataset can be seen in Table 1.

Table 1. Dataset

Index	anumber	bnumber	duration	fee	destination	type
0	6277XXXX	15002XX	434	6000	callcenter	Injapati
1	6277XXXX	15002XX	172	2250	callcenter	Injapati
2	6277XXXX	15002XX	183	3000	callcenter	Injapati
...	...	...	...	...	...	...
...	...	...	...	...	..	...
11415	6277XXXX	0101760X	242	18000	Malaysia	VoIP 01017
11416	6277XXXX	0101760X	3	3600	Malaysia	VoIP 01017
11417	6277XXXX	0101760X	7	3600	Malaysia	VoIP 01017



4.2. Data Preprocessing

In the data preprocessing phase, the data will be cleaned to correct or delete data that is incorrect, incomplete, inaccurate, or has the wrong format to produce high-quality data. We also drop data with Local, Injapati and SLJJ call types. The remaining types of calls are international calls, among others SLI 001, SLI 007, and VoIP 01017. So, the current amount of data used for the calculation is 4,158. New dataset can be seen in Table 2.

Table 2. New dataset

Index	anumber	bnumber	duration	fee	destination	type
167	6277XXXX	00176XXX	14	2864	Russia Fed	SLI 001
168	6277XXXX	00196XXX	7	3818	Uni Emirat Arab	SLI 001
169	6277XXXX	00149XXX	20	4295	Germany	SLI 001
...	...	...	...	...	...	...
...	...	...	...	...	...	...
11415	6277XXXX	0101760X	242	18000	Malaysia	VoIP 01017
11416	6277XXXX	0101760X	3	3600	Malaysia	VoIP 01017
11417	6277XXXX	0101760X	7	3600	Malaysia	VoIP 01017

After that, we change the value of the data of strings by encoding it into numbers so that it can be used as a variable for the clustering method. Encoding dataset can be seen in Table 3.

Table 3. Encoding dataset

Index	anumber	bnumber	duration	fee	destination	type
167	263	952	12	3	33	0
168	256	1167	5	6	46	0
169	169	126	18	8	15	0
...	...	...	...	...	...	...
...	...	...	...	...	...	...
11415	262	1210	240	37	52	2
11416	262	1211	1	5	52	2
11417	262	1211	5	5	52	2

Finally, we create a scaler for change the data will have an average value of 0 and a standard deviation of 1 with Standard Scaler. The standard scaler technique is useful for compiling data according to different units from all dataset. Scaler dataset can be seen in Figure 6.

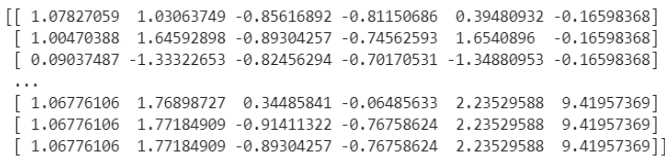


Figure 5: Scaler dataset

4.3. K-Means Clustering Implementation

In this research, the K value is defined by using the Elbow Method as a method that can be used to determine the best number of clusters by seeing at the percentage of comparison results within the number of clusters that will form an elbow at a point. While determining the value of k cluster on K-means the author utilises machine learning to find the Elbow Method to determine the number of clusters from the data set.

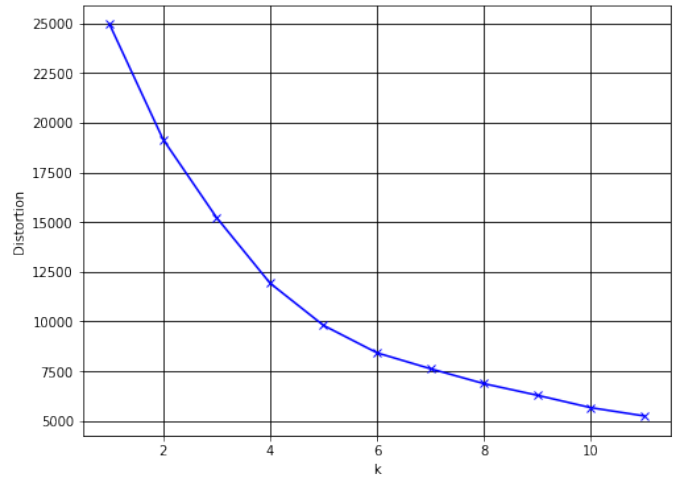


Figure 6: Elbow Method

Figure 7 shows the curve value from Elbow method. Based on the graphic, it can be seen that the number of the maximum number of clusters is 6.

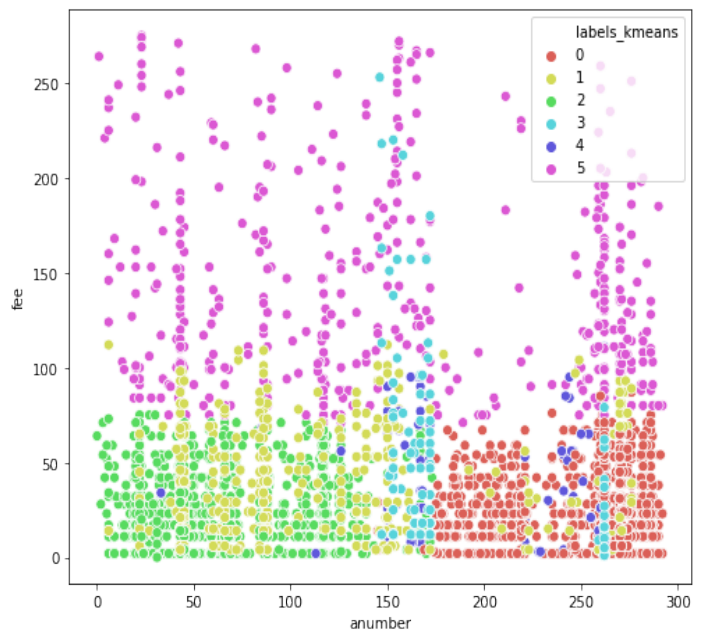


Figure 7: K-Means Clustering

The above plot in figure 8 is useful for interpretation of normal traffic and anomalies in data groups. This graphical representation provides how well each object is located in each cluster. Axis x describes Anumber and Axis y describes Fee. From the graph, we clearly observe that some points originate from cluster 0 to cluster 5, where cluster 4 we consider to be an anomaly data and other clusters are normal data.

4.4. DBSCAN Clustering Implementation

When determining the eps value in DBSCAN clustering, the author tries a suitable value to use that is 1.1 and for the minPts parameter uses the same k value as K-Means which is 6.

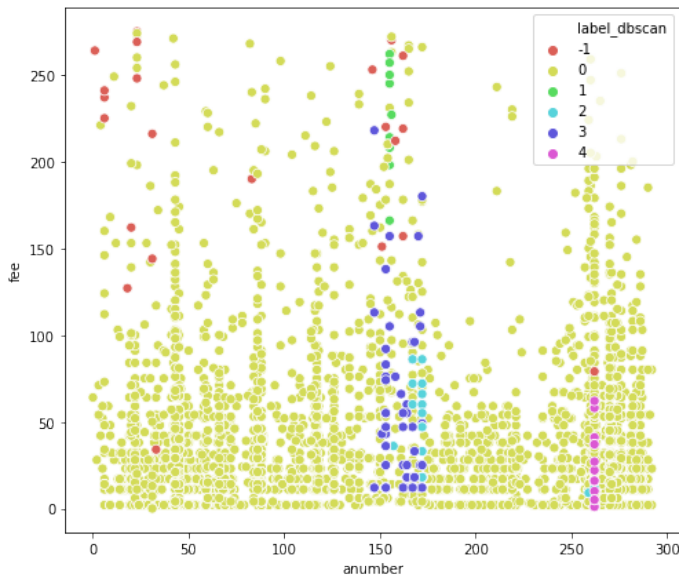


Figure 8: DBScan Clustering

The above plot in figure 9 is useful for interpretation of normal traffic and anomalies in data groups. Axis x describes Anumber and Axis y describes Fee. From the graph, we clearly observe that some points are from cluster -1 to cluster 4 shows that the point in cluster -1 is an anomaly data.

4.5. Confusion Matrix Evaluation

Based on the proposed method, we evaluate the results of fraud discovered from K-Means and DBSCAN clustering with actual fraud data. The evaluation practices Confusion Matrix to calculate accuracy, precision, and recall. Figure 10 shows the results of the confusion matrix for K-Means, the number of True Positives (TP) obtained is 3690, False Positive (FP) is 71, False Negative (FN) is 105 and True Negative (TN) is 292.

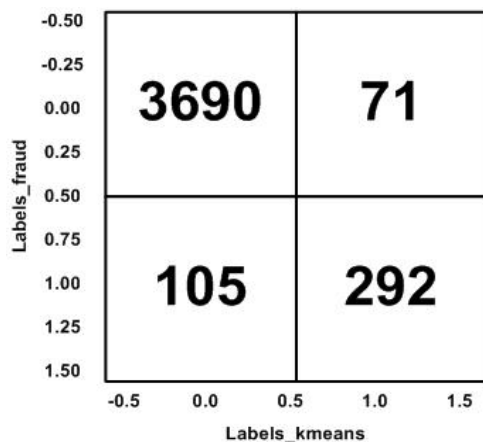


Figure 9: Confusion Matrix on K-Means

The result of the confusion matrix for DBSCAN can be shown on Figure 11, for the number of True Positives (TP) obtained is 3738, False Positive (FP) is 23, False Negative (FN) is 397 and True Negative (TN) is 0.

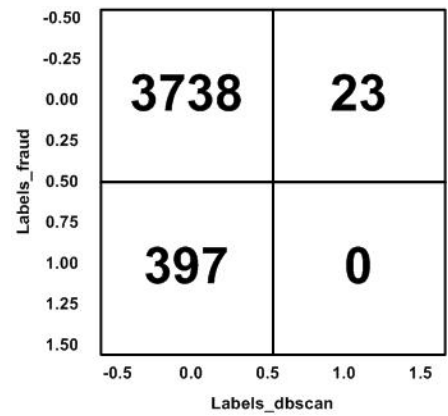


Figure 10: Confusion Matrix on DBSCAN

Table 4 below shows the performance comparison from the evaluation using the confusion matrix for the accuracy, precision, and recall of each algorithm.

Table 4. Performance comparison of the K-Means and DBSCAN algorithms

Algorithms	Accuracy	Precision	Recall
K-Means	0.972332016	0.981122042	0.957671958
DBSCAN	0.903990326	0.993884605	0.898989899

5. Conclusion

We have detected anomalies in telecommunications CDR using machine learning methods with the K-Means and DBSCAN algorithms. Our results found that this algorithm was quite good to be used in detecting fraud on telecommunications CDRs. This is demonstrated through the results of accuracy based on evaluations using actual fraud data. The problem of quality in clustering can be tested by comparing the results with other clustering algorithms. In our study, the accuracy of the algorithm used was not much different. K-Means algorithm shows better accuracy value compared to DBSCAN. So that we can conclude an effective method that can be applied to detect fraud in the CDR is the machine learning method using the K-Means algorithm.

The results of this study also address the problems that occur in case studies of telecommunications operators that do not yet have alternative approaches in detecting fraud in their services. Other problems such as fraud detection in large datasets can also be solved using this method because K-Means as one of the techniques of unsupervised learning can properly handle large amounts of data. Future research includes the application of other clustering methods, or other unsupervised learning techniques such as hierarchical grouping and other clustering algorithms. It is expected that further research can provide better results in telecommunications fraud detection includes more techniques for fraud detection then compared to each other.

Based on the results of our studies, we realize that fraud problems will always develop over time and their handling will also be different in each case. This causes the results of our research will not necessarily give good results if applied to other telecommunications operators and or indifferent fraud cases.

References

[1] M. Arafat, A. Qusef and G. Sammour, "Detection of Wangiri Telecommunication fraud using ensemble learning" in 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), 2019. <https://doi.org/10.1109/JEEIT.2019.8717528>



- [2] M. Liu, J. Liao, J. Wang and Q. Qi, "AGRM: Attention-based graph representation model for Telecom fraud detection" in ICC 2019-2019 IEEE International Conference on Communications (ICC), 2019. <https://doi.org/10.1109/ICC.2019.8761665>
- [3] S. Gee, *Fraud and Fraud Detection: A Data Analytics Approach*, Hoboken: John Wiley & Sons, Inc., 2015.
- [4] V. Jain, "Perspective analysis of telecommunication fraud detection using data stream analytics and neural network classification based data mining" *International Journal of Information Technology*, **9**(3), 303-310, 2017. <https://doi.org/10.1007/s41870-017-0036-5>
- [5] T. Russell, *Signaling system# 7 (Vol. 2)*, New York: McGraw-Hill, 2002.
- [6] C. S. Hilaras, P. A. Mastorocostas and I. T. Rekanos, "Clustering of telecommunications user profiles for fraud detection and security enhancement in large corporate networks: a case study" *Applied Mathematics & Information Sciences*, **9**(4), 1709, 2015. <https://doi.org/10.12785/amis/090407>
- [7] A. S. Yesuf, L. Wolos and K. Rannenber, "Fraud risk modelling: requirements elicitation in the case of telecom services" in *International Conference on Exploring Services Science*, 2017. [https://doi.org/10.1007/978-3-319-56925-3\\_26](https://doi.org/10.1007/978-3-319-56925-3_26)
- [8] S. Subudhi and S. Panigrahi, "A hybrid mobile call fraud detection model using optimized fuzzy C-means clustering and group method of data handling-based network" *Vietnam Journal of Computer Science*, **5**(3-4), 205-217, 2018. <https://doi.org/10.1007/s40595-018-0116-x>
- [9] I. Ighneiwa and H. Mohamed, "Bypass fraud detection: Artificial intelligence approach," *arXiv preprint arXiv:1711.04627*, 2017.
- [10] E. Eifrem, "Graph databases: the key to foolproof fraud detection?" *Computer Fraud & Security*, **2016**(3), 5-8, 2016. [https://doi.org/10.1016/S1361-3723\(16\)30024-0](https://doi.org/10.1016/S1361-3723(16)30024-0)
- [11] M. Kolhar, A. Alameen and M. Gulam, "Performance evaluation of framework of VoIP/SIP server under virtualization environment along with the most common security threats" *Neural Computing and Applications*, **30**(9), 2873-2881, 2018. <https://doi.org/10.1007/s00521-017-2886-y>
- [12] Q. Zhao, K. Chen, T. Li, Y. Yang and X. Wang, "Detecting telecommunication fraud by understanding the contents of a call" *Cybersecurity*, **1**(1), 8, 2018. <http://doi.org/10.1186/s42400-018-0008-5>
- [13] S. Zoldi, "Using anti-fraud technology to improve the customer experience" *Computer Fraud & Security*, **2015**(7), 18-20, 2015. [https://doi.org/10.1016/S1361-3723\(15\)30067-1](https://doi.org/10.1016/S1361-3723(15)30067-1)
- [14] E. Abba, A. M. Aibinu and J. K. Alhassan, "Development of multiple mobile networks call detailed records and its forensic analysis" *Digital Communications and Networks*, **5**(4), 256-265, 2019. <https://doi.org/10.1016/j.dcan.2019.10.005>
- [15] J. Liu, B. Rahbarinia, R. Perdisci, H. Du and L. Su, "Augmenting telephone spam blacklists by mining large CDR datasets" in *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*, 2018. <https://doi.org/10.1145/3196494.3196553>
- [16] A. Minessale II and G. Maruzzelli, *Mastering FreeSWITCH*, Packt Publishing Ltd, 2016.
- [17] K. C. Mondal and H. B. Barua, "Fault analysis and trend prediction in telecommunication using pattern detection: Architecture, Case Study and Experimentation" in *International Conference on Computational Intelligence, Communications, and Business Analytics*, 2018. [https://doi.org/10.1007/978-981-13-8578-0\\_24](https://doi.org/10.1007/978-981-13-8578-0_24)
- [18] C. Gunavathi, R. S. Priya and S. L. Aarthy, "Big data analysis for anomaly detection in telecommunication using clustering techniques" in *Information Systems Design and Intelligent Applications*, 2019. [https://doi.org/10.1007/978-981-13-3329-3\\_11](https://doi.org/10.1007/978-981-13-3329-3_11)
- [19] N. R. Al-Molhem, Y. Rahal and M. Dakkak, "Social network analysis in Telecom data" *Journal of Big Data*, **6**(1), 99, 2019. <https://doi.org/10.1186/s40537-019-0264-6>
- [20] Y. Yu, X. Wan, G. Liu, H. Li, P. Li and H. Lin, "A combinatorial clustering method for sequential fraud detection" in *2017 International Conference on Service Systems and Service Management*, 2017. <https://doi.org/10.1109/ICSSSM.2017.7996302>
- [21] R. Hong, W. Rao, D. Zhou, C. An, Z. Lu and J. Xia, "Commuting Pattern Recognition Using a Systematic Cluster Framework" *Sustainability*, **12**(5), 1764, 2020. <https://doi.org/10.3390/su12051764>
- [22] K. Sultan, H. Ali and Z. Zhang, "Call detail records driven anomaly detection and traffic prediction in mobile cellular networks" *IEEE Access*, **6**, 41728-41737, 2018. <https://doi.org/10.1109/access.2018.2859756>
- [23] X. Min and R. Lin, "K-means algorithm: Fraud detection based on signaling data" in *2018 IEEE World Congress on Services (SERVICES)*, 2018. <https://doi.org/10.1109/services.2018.00024>
- [24] J. Suntoro, *Data Mining: Algoritma dan Implementasi dengan Pemrograman PHP*, Elex Media Komputindo, 2019.
- [25] S. Adinugroho and Y. A. Sari, *Implementasi Data Mining Menggunakan Weka*, Universitas Brawijaya Press, 2018.
- [26] C. Xiong, Z. Hua, K. Lv and X. Li, "An improved K-means text clustering algorithm by optimizing initial cluster centers" in *2016 7th International Conference on Cloud Computing and Big Data (CCBD)*, 2016. <https://doi.org/10.1109/CCBD.2016.059>
- [27] G. C. Ngo and E. Q. B. Macabebe, "Image segmentation using K-means color quantization and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) for Hotspot Detection in Photovoltaic modules" in *2016 IEEE region 10 conference (TENCON)*, 2016. <https://doi.org/10.1109/tencon.2016.7848290>
- [28] F. O. Ozkok and M. Celik, "A new approach to determine Eps parameter of DBSCAN algorithm" *International Journal of Intelligent Systems and Applications in Engineering*, **5**(4), 247-251, 2017. <https://doi.org/10.18201/ijisae.201753899>
- [29] I. D. Id and E. Mahdiyah, "Modifikasi DBSCAN (Density-Based Spatial Clustering With Noise) pada Objek 3 Dimensi," *Jurnal Komputer Terapan*, **3**(1), 41-52, 2017. <https://doi.org/10.13140/RG.2.2.22346.67529>
- [30] M. Lenning, J. Fortunato, T. Le, I. Clark, A. Sherpa, S. Yi, P. Hofsteen, G. Thamilarasu, J. Yang, X. Xu, T. K. Hsiai, H. Cao and H. D. Han, "Real-time monitoring and analysis of zebrafish electrocardiogram with anomaly detection" *Sensors*, **18**(1), 61, 2018. <https://doi.org/10.3390/s18010061>
- [31] E. Alpaydin, *Introduction to machine learning*, MIT press, 2020.

# Robust Static Output-Feedback Fault Tolerant Control for a Class of T-S Fuzzy Systems using Adaptive Sliding Mode Observer Approach

Slim Dhahri<sup>\*1,2</sup>, Essia Ben Alaia<sup>1,3</sup>

<sup>1</sup>Department of Computer Engineering and Networks, College of Computer and Information Sciences, Jouf University, 72388, Sakaka, Saudi Arabia

<sup>2</sup>Engineering Laboratory of Industrial Systems and Renewable Energies, National Higher School of Engineers of Tunis, Tunis University, 1008, Tunis, Tunisia

<sup>3</sup>Laboratory of Analysis, Conception and Control of Systems (LACCS), National School of Engineers of Tunis, Tunis El Manar University, 1002, Tunis, Tunisia

## ARTICLE INFO

### Article history:

Received: 9 April, 2020

Accepted: 1 June, 2020

Online: 6 July, 2020

### Keywords:

TS fuzzy systems

Sliding mode observer

Fault estimation

Fault-tolerant control

## ABSTRACT

*In this paper, the problems of actuator and sensor fault estimation (FE) and fault-tolerant control (FTC) for uncertain nonlinear systems represented by Takagi-Sugeno (T-S) fuzzy models are investigated. First, a robust fuzzy adaptive sliding mode observer (SMO) is designed to simultaneously estimate system states and both actuator and sensor faults. Then, using the obtained on-line FE information, a static output-feedback fault-tolerant control (SOFFTC) is developed to compensate the fault effects and stabilize the closed-loop system. Moreover, sufficient conditions for the existence of the proposed observer and controller are given in terms of linear matrix inequalities (LMIs). The robustness against uncertainties is treated using the  $H_\infty$  optimization technique to attenuate its effect on the estimation error. Finally, the simulation results of nonlinear inverted pendulum with cart system validate the efficiency of the proposed method.*

## 1 Introduction

Modern industrial systems are affected usually by various event of faults such as, loss of actuator effectiveness, failures or offsets of actuators/sensors, deviations of output measurement, etc. Indeed, the presence of fault causes an unacceptable performances of design controllers, thus deteriorating the overall system execution, and so leading to wrong dangerous situations.

Thus, it is important to encourage the development of research on fault tolerant control (FTC), which is divided on two types. The first one, the so-called passive FTC, is focused on to conceive a robust controller against disturbances and uncertainties. A key limitation is that the system stability can't be guaranteed in the presence of faults.

Nevertheless, based on online fault estimation (FE), such as the size and the shape, active FTC can develop robust controller such that the fault effects are eliminated and the system stability is achieved. In the literature, several research results on the FTC techniques are documented, see for example [1–9], and the references

therein.

In industrial processes, most of systems are described by nonlinear mathematical models. Takagi-Sugeno (TS) fuzzy systems [10] provide a powerful tool to approximate nonlinear characteristics. T-S fuzzy systems are nonlinear models represented by a set of local linear models. By fuzzy blending of linear representations with appropriate membership functions, the overall fuzzy model of the system is achieved, which greatly simplifies the analysis and control for complex nonlinear systems. Therefore, excellent results in FE and FTC problems of T-S fuzzy systems are developed in [11–17]. In [18], a FTC is designed for TS fuzzy systems subject to actuator faults. However, this result must verify the rank condition, which is really hard to fulfill for many practical systems. In [19], the problem of FE and FTC for a T-S fuzzy systems with uncertainties and actuator faults is investigated without the requirement of rank condition. It deals only with constant faults, however, the faults are time-varying in many real systems. In [20], a sliding mode observer (SMO) is designed to estimate sensor fault for nonlinear stochastic systems for FTC. However, the actuator fault is not con-

\*Corresponding Author: Slim Dhahri, Department of Computer Engineering and Networks, College of Computer and Information Sciences, Jouf University, Sakaka, Saudi Arabia, dhahri.slim@yahoo.fr

sidered. In [21], a FTC is developed for T-S fuzzy systems affected by actuator faults. Since in many practical systems, actuator and sensor faults may occur at the same time and uncertainties may exist. It is desirable to consider actuator and sensor faults under one unified framework. For example, in [22], a SMO-based FE and FTC is designed for a class of nonlinear systems subject to actuator and sensor faults. A major disadvantage of this approach is the appearance of chattering mode. In [23], FE and FTC problems are studied to simultaneously estimate sensor and actuator faults using two observers and filter. However, this method is more expensive since it requires a high additional computation cost. In [24], a robust adaptive observer is developed to simultaneously estimate state and both sensor and actuator faults for nonlinear systems despite the presence of disturbances. A FTC law is applied to stabilize the closed-loop system and compensate the fault effects. However, sufficient conditions of observer and controller gains are formulated in an unified optimization problem and computed by solving a set of LMIs only in single step. Unfortunately, these results need the knowledge of the upper bounds of faults. If the information of fault is unknown, the SMO cannot be obtained.

The aim of this work is to address fault estimation and fault tolerant control problems for T-S fuzzy systems subject to simultaneously actuator faults, sensor faults and uncertainties. First, a novel robust adaptive SMO is proposed to estimate the states and both actuator and sensor faults using equivalent output error injection approach. Then, based on online fault information a static output-feedback fault-tolerant control (SOFFTC) is designed to compensate the fault effects and stabilize the closed-loop system. All the design conditions are formulated in an optimization problem under LMIs constraints. Finally, the simulation result of an inverted pendulum with cart system is given to prove the effectiveness of the proposed method.

The main contributions of the present work are the following:

1. A novel fuzzy adaptive SMO is designed for the estimation of states and faults in a T-S fuzzy systems affected by simultaneously actuator faults, sensor faults and uncertainties. Robustness against uncertainties is analyzed using the  $H_\infty$  technique to reduce its effect.
2. Most existing SMO design methods such as those reported in [22–24] assume that the value of the upper bounds of actuator faults  $\rho_a$  and sensor faults  $\rho_s$  is known. If the information of fault is unknown or exceeds the admissible value, these methods cannot be feasible. To overcome this problem, a new adaptive law is constructed to estimate the upper bounds online.
3. The problem of both actuator and sensor FE under one unified framework for T-S fuzzy systems is investigated. Whereas, many researchers have considered only sensor faults [25–27] or actuator faults [28, 29].
4. Based on the FE, a SOFFTC is designed to effectively accommodate the influence of fault and ensure the stability of the resulting closed-loop system. The proposed method is easily be implemented in practice and is much simpler than dynamic output feedback fault tolerant controller.

5. Sufficient conditions of the observer and controller are formulated in an optimization problem under LMIs constraints which can be designed separately.

The rest of this paper is organized as follows: Section 2 presents the problem formulation and preliminaries. The design of the observer and the analysis of the stability of the error dynamics are given in Section 3. FE is studied in Section 4. Section 5 gives the SOFFTC scheme. Finally, simulation example in Section 6 validates the efficiency of the proposed algorithm.

## 2 Problem Formulation and Preliminaries

Consider a TS fuzzy model with actuator faults, sensor faults and uncertainties. The  $i$ th rule of the T-S fuzzy model is of the following form:

Plant Rule  $i$ : If  $\xi_1(t)$  is  $\mu_{1,i}$  and ...  $\xi_g(t)$  is  $\mu_{g,i}$ , Then

$$\begin{cases} \dot{x}(t) = A_i x(t) + B_i u(t) + M_i f_a(t) + E_i d(x, u, t) \\ y(t) = C_i x(t) + N f_s(t) \\ y_c(t) = C_{ci} x(t) \end{cases} \quad (1)$$

where  $x(t) \in R^n$  represents the state vector;  $u(t) \in R^m$  is the input;  $y(t) \in R^p$  is the output;  $y_c(t) \in R^{p1}$  is the controlled output;  $A_i, B_i, M_i, E_i, C_i, N$  and  $C_{ci}$  are real known constant matrices with appropriate dimensions;  $f_a(t) : R^+ \rightarrow R^q$  and  $f_s(t) : R^+ \rightarrow R^h$  represent additive actuator fault and sensor fault vector, respectively;  $d(x, u, t) \in R^l$  models the uncertainties, which is assumed to belong to  $\mathcal{L}_2 [0, \infty)$ ; the pairs  $(A_i, C_i)$  are observable, and the pairs  $(A_i, B_i)$  are controllable;  $\xi_j (j = 1, \dots, g)$  are the premise variables, and  $\mu_{j,i} (j = 1, \dots, g; i = 1, \dots, k)$  are fuzzy sets;  $g$  and  $k$  are the number of premise variables and IF-THEN rules, respectively. The fuzzy model is given by:

$$\begin{cases} \dot{x}(t) = \sum_{i=1}^k \mu_i(\xi(t)) (A_i x(t) + B_i u(t) + M_i f_a(t) + E_i d(x, u, t)) \\ y(t) = \sum_{i=1}^k \mu_i(\xi(t)) (C_i x(t) + N f_s(t)) \\ y_c(t) = \sum_{i=1}^k \mu_i(\xi(t)) (C_{ci} x(t)) \end{cases} \quad (2)$$

where  $\xi(t) = [\xi_1(t), \dots, \xi_g(t)]$ ,  $\mu_i(\xi(t)) = \frac{w_i(\xi(t))}{\sum_{i=1}^k w_i(\xi(t))}$ ,  $w_i(\xi(t)) = \prod_{j=1}^g \theta_{ij}(\mu_j(t))$  and here  $\theta_{ij}(\cdot)$  stands for the order of the membership function of  $\theta_{ij}$ . It is assumed that

$$w_i(\xi(t)) \geq 0, \quad i = 1, \dots, k, \quad \sum_{i=1}^k w_i(\xi(t)) > 0 \quad (3)$$

for any  $\xi(t)$ . Thus, for any  $\xi(t)$ ,  $\sum_{i=1}^k \mu_i(\xi(t))$  satisfies

$$\mu_i(\xi(t)) \geq 0, \quad i = 1, \dots, k, \quad \sum_{i=1}^k \mu_i(\xi(t)) = 1 \quad (4)$$

For simplicity, we will use  $\mu_i$  to represent  $\mu_i(\xi(t))$ .

**Assumption 1.**  $f_a(t)$  and  $f_s(t)$  are unknown but norm bounded

$$\|f_a(t)\| \leq \rho_a, \quad \|f_s(t)\| \leq \rho_s \quad (5)$$

where  $\rho_a$  and  $\rho_s$  are unknown positive scalars.

**Assumption 2 [30].** The actuator fault distribution matrices  $M_i$  in (2) satisfy:

$$\text{rank}(CM_i) = \text{rank}(M_i), \quad i = 1, \dots, k \quad (6)$$

**Assumption 3 [30].**

$$\text{rank} \begin{bmatrix} sI_n - A_i & M_i \\ C_i & 0 \end{bmatrix} = n + \text{rank}(M_i), \quad i = 1, \dots, k \quad (7)$$

**Lemma 1 [31].** For matrices  $A$  and  $B$  and any scalar  $\varepsilon > 0$ , we have

$$AB + (AB)^T \leq \varepsilon^{-1}AA^T + \varepsilon B^T B \quad (8)$$

**Lemma 2 [32].** If

$$S_{ii} < 0, \quad 1 \leq i \leq k \quad (9)$$

$$\frac{2}{r-1}S_{ii} + S_{ij} + S_{ji} < 0, \quad 1 \leq i \neq j \leq k \quad (10)$$

then, we have

$$\sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j S_{ij} < 0 \quad (11)$$

**Lemma 3 [33].** Under Assumption 2, there exists coordinate transformations

$$z(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = T_i x(t), \quad v(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = S_i y(t)$$

such that

$$T_i A_i T_i^{-1} = \begin{bmatrix} A_{11,i} & A_{12,i} \\ A_{21,i} & A_{22,i} \end{bmatrix}, T_i B_i = \begin{bmatrix} B_{1,i} \\ B_{2,i} \end{bmatrix}, T_i M_i = \begin{bmatrix} M_{1,i} \\ 0 \end{bmatrix}$$

$$T_i E_i = \begin{bmatrix} E_{1,i} \\ E_{2,i} \end{bmatrix}, S_i N = \begin{bmatrix} 0 \\ N_2 \end{bmatrix}, S_i C_i T_i^{-1} = \begin{bmatrix} C_{11,i} & 0 \\ 0 & C_{22,i} \end{bmatrix}$$

where  $A_{11,i} \in R^{q \times q}$ ,  $A_{22,i} \in R^{(n-q) \times (n-q)}$ ,  $B_{1,i} \in R^{q \times m}$ ,  $M_{1,i} \in R^{q \times q}$ ,  $E_{1,i} \in R^{q \times l}$ ,  $N_2 \in R^{(p-q) \times h}$ ,  $C_{11,i} \in R^{q \times q}$  and  $C_{22,i} \in R^{(p-q) \times (p-q)}$  is invertible,  $i = 1, \dots, k$ .

Through coordinate transformations, the system (2) is converted into the following two subsystems:

$$\begin{cases} \dot{z}_1(t) = \sum_{i=1}^k \mu_i (A_{11,i} z_1(t) + A_{12,i} z_2(t) + B_{1,i} u(t) + M_{1,i} f_a(t) \\ \quad + E_{1,i} d(x, u, t)) \\ v_1(t) = \sum_{i=1}^k \mu_i (C_{11,i} z_1(t)) \end{cases} \quad (12)$$

$$\begin{cases} \dot{z}_2(t) = \sum_{i=1}^k \mu_i (A_{21,i} z_1(t) + A_{22,i} z_2(t) + B_{2,i} u(t) \\ \quad + E_{2,i} d(x, u, t)) \\ v_2(t) = \sum_{i=1}^k \mu_i (C_{22,i} z_2(t) + N_2 f_s(t)) \end{cases} \quad (13)$$

In addition, partition the matrix  $S_i$  as:

$$S_i = \begin{bmatrix} S_{11,i} \\ S_{22,i} \end{bmatrix} \quad (14)$$

where  $S_{11,i} \in R^{(p-q) \times p}$  and  $S_{22,i} \in R^{q \times p}$ . The variable  $z_1(t)$  can be obtained by:

$$z_1(t) = \sum_{i=1}^k \mu_i (C_{11,i}^{-1} S_{11,i} y(t)) \quad (15)$$

We define a new state  $z_3(t) = \int_0^t v_2(\tau) d\tau$  where  $\dot{z}_3(t) = \sum_{i=1}^k \mu_i (C_{22,i} z_2(t) + N_2 f_s(t))$ . Then the augmented system with the new state  $z_3(t)$  is given as:

$$\begin{cases} \dot{z}_0(t) = \sum_{i=1}^k \mu_i (A_{0,i} z_0(t) + A_{3,i} z_2(t) + B_{0,i} u(t) + M_{0,i} f_s(t) \\ \quad + E_{0,i} d(x, u, t)) \\ v_3(t) = \sum_{i=1}^k \mu_i (C_{0,i} z_0(t)) \end{cases} \quad (16)$$

where  $\dot{z}_0(t) = \begin{bmatrix} z_2(t) \\ z_3(t) \end{bmatrix} \in R^{n+p-2q}$ ,  $v_3(t) \in R^{p-q}$ ,  $A_{0,i} = \begin{bmatrix} A_{22,i} & 0 \\ C_{22,i} & 0 \end{bmatrix} \in R^{(n+p-2q) \times (n+p-2q)}$ ,  $A_{3,i} = \begin{bmatrix} A_{21,i} \\ 0 \end{bmatrix} \in R^{(n+p-2q) \times q}$ ,  $B_{0,i} = \begin{bmatrix} B_{2,i} \\ 0 \end{bmatrix} \in R^{(n+p-2q) \times m}$ ,  $E_{0,i} = \begin{bmatrix} E_{2,i} \\ 0 \end{bmatrix} \in R^{(n+p-2q) \times l}$ ,  $M_{0,i} = \begin{bmatrix} 0 \\ N_2 \end{bmatrix} \in R^{(n+p-2q) \times h}$  and  $C_{0,i} = \begin{bmatrix} 0 & I_{p-q} \end{bmatrix} \in R^{(p-q) \times (n+p-2q)}$ .

**Lemma 4 [33].** The pair  $(A_{0,i}, C_{0,i})$  is observable, if the pair  $(A_{22,i}, C_{22,i})$  is detectable,  $i = 1, \dots, k$ . Then, there exists matrices  $L_i$ , having the special structure  $L_i = \begin{bmatrix} L_{1,i} & 0 \end{bmatrix}$ , such that  $A_{22,i} + L_i C_{22,i}$  is stable,  $i = 1, \dots, k$ .

Let the transformation of coordinates  $h(t) = \begin{bmatrix} h_1^T(t) & h_2^T(t) \end{bmatrix}^T = T_{L,i} z_0(t)$  with

$$T_{L,i} = \begin{bmatrix} I_{n-q} & L_i \\ 0 & I_{p-q} \end{bmatrix}, \quad i = 1, \dots, k \quad (17)$$

where  $h_1(t) \in R^{n-q}$  and  $h_2(t) \in R^{p-q}$ . Therefore, the system (16) is converted into the following system:

$$\begin{cases} \dot{h}(t) = \sum_{i=1}^k \mu_i (A_{h,i} h(t) + T_{L,i} A_{3,i} z_2(t) + B_{h,i} u(t) \\ \quad + E_{h,i} d(x, u, t) + M_{h,i} f_s(t)) \\ v_3(t) = \sum_{i=1}^k \mu_i (C_{h,i} h(t)) \end{cases} \quad (18)$$

where

$$A_{h,i} = \begin{bmatrix} A_{22,i} + L_i C_{22,i} & -(A_{22,i} + L_i C_{22,i}) L_i \\ C_{22,i} & -C_{22,i} L_i \end{bmatrix}, B_{h,i} = \begin{bmatrix} B_{1,i} \\ 0 \end{bmatrix}$$

$$E_{h,i} = \begin{bmatrix} E_{1,i} \\ 0 \end{bmatrix}, M_{h,i} = \begin{bmatrix} 0 \\ N_2 \end{bmatrix}, C_{h,i} = \begin{bmatrix} 0 & I_{p-q} \end{bmatrix}$$

Therefore, T-S fuzzy subsystems (12) and (13) can be rewritten respectively as:

$$\begin{cases} \dot{z}_1(t) = \sum_{i=1}^k \mu_i (A_{11,i} z_1(t) + A_{12,i} z_2(t) + B_{1,i} u(t) + M_{1,i} f_a(t) \\ \quad + E_{1,i} d(x, u, t)) \\ v_1(t) = \sum_{i=1}^k \mu_i (C_{11,i} z_1(t)) \end{cases} \quad (19)$$

$$\begin{cases} \dot{h}_1(t) = \sum_{i=1}^k \mu_i ((A_{22,i} + L_i C_{22,i})h_1(t) - (A_{22,i} + L_i C_{22,i})L_i h_2(t) \\ \quad + A_{12,i}z_1(t) + B_{1,i}u(t) + E_{1,i}d(x, u, t)) \\ \dot{h}_2(t) = \sum_{i=1}^k \mu_i (C_{22,i}h_1(t) - C_{22,i}L_i h_2(t) + N_2 f_s(t)) \\ v_3(t) = h_2(t) \end{cases} \quad (20)$$

### 3 Adaptive Sliding Mode Observers Design

For system (19), we construct the following adaptive SMO:

$$\begin{cases} \dot{\hat{z}}_1(t) = \sum_{i=1}^k \mu_i (A_{11,i}\hat{z}_1(t) + A_{12,i}\hat{h}_1(t) - A_{12,i}L_i v_3(t) \\ \quad + B_{1,i}u(t) + M_{1,i}v_{1,i}(t) + (A_{11,i} - A_{11,i}^s)C_{11,i}^{-1}(v_1(t) - \hat{v}_1(t))) \\ \hat{v}_1(t) = \sum_{i=1}^k \mu_i (C_{11,i}\hat{z}_1(t)) \end{cases} \quad (21) \text{ where}$$

where  $\hat{z}_1(t)$ ,  $\hat{h}_1(t)$  and  $\hat{v}_1(t)$  denote, respectively, the estimated  $z_1(t)$ ,  $h_1(t)$  and  $v_1(t)$ .  $A_{11,i}^s \in R^{q \times q}$  is a stable matrix and  $v_{1,i}(t)$  is defined by:

$$v_{1,i}(t) = \begin{cases} (\hat{\rho}_a + l_{a,i}) \frac{M_{1,i}^T P_1 (C_{11,i}^{-1} S_{11,i} v_1(t) - \hat{z}_1(t))}{\|M_{1,i}^T P_1 (C_{11,i}^{-1} S_{11,i} v_1(t) - \hat{z}_1(t))\|} & \text{if } C_{11,i}^{-1} S_{11,i} v_1 - \hat{z}_1 \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

where  $P_1 \in R^{q \times q} > 0$  is the Lyapunov matrix for  $A_{11,i}^s$ ,  $\hat{\rho}_a$  is adaptive parameter to estimate the unknown parameter  $\rho_a$ , and the scalar  $\hat{\rho}_a$  is introduced using an update law

$$\dot{\hat{\rho}}_a = \sigma_1 \|M_{1,i}^T P_1 (C_{11,i}^{-1} S_{11,i} v_1(t) - \hat{z}_1(t))\| \quad (22)$$

with constant  $\sigma_1 > 0$ .

For system (20), we design the following adaptive SMO:

$$\begin{cases} \dot{\hat{h}}_1(t) = \sum_{i=1}^k \mu_i ((A_{22,i} + L_i C_{22,i})\hat{h}_1(t) - (A_{22,i} + L_i C_{22,i}) \\ \quad \times L_i v_3(t) + B_{1,i}u(t) + A_{21,i}C_{11,i}^{-1}v_1(t)) \\ \dot{\hat{h}}_2(t) = \sum_{i=1}^k \mu_i (C_{22,i}\hat{h}_1(t) - C_{22,i}L_i \hat{h}_2(t) \\ \quad - (C_{22,i}L_i + K_i)(v_3(t) - \hat{v}_3(t)) + N_2 v_{2,i}(t)) \\ \hat{v}_3(t) = \hat{h}_2(t) \end{cases} \quad (23)$$

where  $\hat{h}_1(t)$  and  $\hat{v}_3(t)$  denote, respectively, the estimated of  $h_1(t)$  and  $v_3(t)$ ,  $K_i \in R^{(p-q) \times (p-q)}$  is the observer gains, and  $v_{2,i}(t)$  is defined by:

$$v_{2,i}(t) = \begin{cases} (\hat{\rho}_s + l_{s,i}) \frac{N_2^T P_{02}(v_3(t) - \hat{v}_3(t))}{\|N_2^T P_{02}(v_3(t) - \hat{v}_3(t))\|} & \text{if } v_3(t) - \hat{v}_3(t) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

where  $P_{02} \in R^{(p-q) \times (p-q)} > 0$ ,  $\hat{\rho}_s$  is adaptive parameter to estimate the unknown parameter  $\rho_s$ , and the scalar  $\hat{\rho}_s$  is introduced using an update law

$$\dot{\hat{\rho}}_s = \sigma_2 \|N_2^T P_{02}(v_3(t) - \hat{v}_3(t))\| \quad (24)$$

with constant  $\sigma_2 > 0$ .

Let us define  $e_1(t) = z_1(t) - \hat{z}_1(t)$ ,  $e_2(t) = h_1(t) - \hat{h}_1(t)$  and  $e_3(t) = h_2(t) - \hat{h}_2(t)$ , then the error dynamic system as follows:

$$\begin{cases} \dot{e}_1(t) = \sum_{i=1}^k \mu_i (A_{11,i}e_1(t) + A_{12,i}e_2(t) + E_{1,i}d(x, u, t) \\ \quad + M_{1,i}(f_a(t) - v_{1,i}(t))) \\ \dot{e}_2(t) = \sum_{i=1}^k \mu_i ((A_{22,i} + L_i C_{22,i})e_2(t) + E_{2,i}d(x, u, t)) \\ \dot{e}_3(t) = \sum_{i=1}^k \mu_i (C_{22,i}e_2(t) + K_i e_3(t) + N_2(f_s(t) - v_{2,i}(t))) \end{cases} \quad (25)$$

Define  $r(t)$  as

$$r(t) = He(t) = H \begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{bmatrix} \quad (26)$$

$$H := \begin{bmatrix} H_1 & 0 & 0 \\ 0 & H_2 & 0 \\ 0 & 0 & H_3 \end{bmatrix} \quad (27)$$

with  $H_1 \in R^{q \times q}$ ,  $H_2 \in R^{(n-q) \times (n-q)}$  and  $H_3 \in R^{(p-q) \times (p-q)}$ . The adaptive SMO design method under  $H_\infty$  performance to be addressed in this work is

- (i) The observer error dynamics system (25) with  $d(x, u, t) = 0$  is asymptotically stable, namely, there is no uncertainty;
- (ii) For a given  $\gamma_1 > 0$ . The following  $H_\infty$  performance is satisfied:

$$\int_0^T r^T(t)r(t)dt < \gamma_1 \int_0^T d^T(x, u, t)d(x, u, t)dt \quad (28)$$

for all  $T > 0$  and  $d(x, u, t) \in \mathcal{L}_2 [ 0 \quad \infty )$  under zero initial conditions.

#### 3.1 Stability analysis

**Theorem 1.** Consider T-S fuzzy system (2) under Assumptions 13. The observer error dynamics system (25) is asymptotically stable and satisfy (28) with attenuation level  $\gamma_1 > 0$ , if there exist matrices  $P_1 > 0$ ,  $P_{01} > 0$ ,  $P_{02} > 0$ ,  $X_i, Y_i, i = 1, \dots, k$ , such that:

Minimize  $\gamma_1$  subject to

$$\begin{bmatrix} \Gamma_{1,i} & P_1 A_{12,i} & 0 & P_1 E_{1,i} \\ * & \Gamma_{2,i} & C_{22,i}^T P_{02} & P_{01} E_{2,i} \\ * & * & \Gamma_{3,i} & 0 \\ * & * & * & -\gamma_1 I \end{bmatrix} < 0 \quad (29)$$

where

$$\begin{aligned} \Gamma_{1,i} &= (A_{11,i}^s)^T P_1 + P_1 A_{11,i}^s + H_1^T H_1 \\ \Gamma_{2,i} &= A_{22,i} P_{01} + P_{01} A_{22,i}^T + X_i C_{22,i} + C_{22,i}^T X_i^T + H_2^T H_2 \\ \Gamma_{3,i} &= Y_i + Y_i^T + H_3^T H_3 \end{aligned}$$

If the optimization problem is solved, then we can obtain the following observer gains

$$\begin{aligned} L_i &= P_{01}^{-1} X_i \\ K_i &= P_{02}^{-1} Y_i \end{aligned}$$



**Proof.** Let the following Lyapunov functional candidate:

$$V(t) = V_1(t) + V_2(t) + V_3(t) \tag{30}$$

where  $V_1(t) = e_1^T(t)P_1e_1(t) + \frac{1}{\sigma_1}\tilde{\rho}_a^2$ ,  $V_2(t) = e_2^T(t)P_{01}e_2(t)$ ,  $V_3(t) = e_3^T(t)P_{02}e_3(t) + \frac{1}{\sigma_2}\tilde{\rho}_s^2$ ,  $\tilde{\rho}_a = \rho_a - \hat{\rho}_a$  and  $\tilde{\rho}_s = \rho_s - \hat{\rho}_s$ . The derivative of  $V_1(t)$  satisfy:

$$\begin{aligned} \dot{V}_1(t) &= \sum_{i=1}^k \mu_i \left( e_1^T(t) \left( (A_{11,i}^s)^T P_1 + P_1 A_{11,i}^s \right) e_1(t) \right. \\ &\quad + 2e_1^T(t)P_1A_{12,i}e_2(t) + 2e_1^T(t)P_1E_{1,i}d(x, u, t) \\ &\quad \left. + 2e_1^T(t)P_1M_{1,i}(f_a(t) - v_{1,i}(t)) + \frac{2}{\sigma_1}\tilde{\rho}_a(-\dot{\hat{\rho}}_a) \right) \end{aligned} \tag{31}$$

Using the definition of  $v_{1,i}(t)$  and the bound of  $f_a(t)$ , we have

$$\begin{aligned} &e_1^T(t)P_1M_{1,i}(f_a(t) - v_{1,i}(t)) + \frac{1}{\sigma_1}\tilde{\rho}_a(-\dot{\hat{\rho}}_a) \\ &= e_1^T(t)P_1M_{1,i}f_a(t) - (\hat{\rho}_a + l_{a,i})e_1^T(t)P_1M_{1,i} \frac{M_{1,i}^T P_1 e_1(t)}{\|M_{1,i}^T P_1 e_1(t)\|} \\ &\quad + \frac{1}{\sigma_1}(\rho_a - \hat{\rho}_a)(-\sigma_1 \|M_{1,i}^T P_1 e_1(t)\|) \\ &= e_1^T(t)P_1M_{1,i}f_a(t) - (\rho_a + l_{a,i})\|M_{1,i}^T P_1 e_1(t)\| \\ &\leq \|M_{1,i}^T P_1 e_1(t)\|\rho_a - (\rho_a + l_{a,i})\|M_{1,i}^T P_1 e_1(t)\| \\ &= -l_{a,i}\|M_{1,i}^T P_1 e_1(t)\| < 0 \end{aligned} \tag{32}$$

Therefore

$$\begin{aligned} \dot{V}_1(t) &\leq \sum_{i=1}^k \mu_i \left( e_1^T(t) \left( (A_{11,i}^s)^T P_1 + P_1 A_{11,i}^s \right) e_1(t) \right. \\ &\quad \left. + 2e_1^T(t)P_1A_{12,i}e_2(t) + 2e_1^T(t)P_1E_{1,i}d(x, u, t) \right) \end{aligned} \tag{33}$$

Similarly, the derivatives of  $V_2(t)$  and  $V_3(t)$  can be obtained as:

$$\begin{aligned} \dot{V}_2(t) &= \sum_{i=1}^k \mu_i \left( e_2^T(t) \left( (A_{22,i} + L_i C_{22,i})^T P_{01} \right. \right. \\ &\quad \left. \left. + P_{01}(A_{22,i} + L_i C_{22,i}) \right) e_2(t) \right. \\ &\quad \left. + 2e_2^T(t)P_{01}E_{2,i}d(x, u, t) \right) \end{aligned} \tag{34}$$

$$\begin{aligned} \dot{V}_3(t) &= \sum_{i=1}^k \mu_i \left( e_3^T(t) \left( K_i^T P_{02} + P_{02}K_i \right) e_3(t) \right. \\ &\quad \left. + 2e_3^T(t)P_{02}C_{22,i}e_2(t) \right. \\ &\quad \left. + 2e_3^T(t)P_{02}N_2(f_s(t) - v_{2,i}(t)) \right) \end{aligned} \tag{35}$$

Similarly, we obtain

$$2e_3^T(t)P_{02}N_2(f_s(t) - v_{2,i}(t)) \leq -l_{s,i}\|N_2^T P_{02}e_3(t)\| < 0 \tag{36}$$

From (30), (33)(36), the time derivative of  $V(t)$  is

$$\begin{aligned} \dot{V}(t) &\leq \sum_{i=1}^k \mu_i \left( \xi(t) \left( e_1^T(t) \left( (A_{11,i}^s)^T P_1 + P_1 A_{11,i}^s \right) e_1(t) \right. \right. \\ &\quad \left. \left. + 2e_1^T(t)P_1A_{12,i}e_2(t) + 2e_1^T(t)P_1E_{1,i}d(x, u, t) \right. \right. \\ &\quad \left. \left. + e_2^T(t) \left( (A_{22,i} + L_i C_{22,i})^T P_{01} + P_{01}(A_{22,i} + L_i C_{22,i}) \right) e_2(t) \right. \right. \\ &\quad \left. \left. + 2e_2^T(t)P_{01}E_{2,i}d(x, u, t) + e_3^T(t) \left( K_i^T P_{02} + P_{02}K_i \right) e_3(t) \right. \right. \\ &\quad \left. \left. + 2e_3^T(t)P_{02}C_{22,i}e_2(t) \right) \right) \end{aligned} \tag{37}$$

When  $d(x, u, t) = 0$ , we have

$$\dot{V}(t) \leq \sum_{i=1}^k \mu_i \left( \begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{bmatrix}^T \Lambda_i \begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{bmatrix} \right) \tag{38}$$

where

$$\Lambda_i = \begin{bmatrix} Q_{1,i} & P_1A_{12,i} & 0 \\ A_{12,i}^T P_1 & Q_{2,i} & C_{22,i}^T P_{02} \\ 0 & P_{02}C_{22,i} & Q_{3,i} \end{bmatrix} \tag{39}$$

with

$$\begin{aligned} Q_{1,i} &= (A_{11,i}^s)^T P_1 + P_1 A_{11,i}^s \\ Q_{2,i} &= (A_{22,i} + L_i C_{22,i})^T P_{01} + P_{01}(A_{22,i} + L_i C_{22,i}) \\ Q_{3,i} &= K_i^T P_{02} + P_{02}K_i \end{aligned}$$

If  $\Lambda_i < 0$ , then  $\dot{V}(t) < 0$ , which implies that  $e \rightarrow 0$  as  $t \rightarrow \infty$ . Therefore, the error dynamics system is asymptotically stable.

When  $d(x, u, t) \neq 0$ , we define

$$J_1(t) = \dot{V}(t) + r^T(t)r(t) - \gamma_1 d^T(x, u, t)d(x, u, t) \tag{40}$$

Substituting (37) and (26) into (28) yields

$$\begin{aligned} J_1(t) &= \dot{V}(t) + r^T(t)r(t) - \gamma_1 d^T(x, u, t)d(x, u, t) \\ &= \sum_{i=1}^k \mu_i \left( e^T \left( \Lambda_i + H^T H \right) e + 2e_1^T(t)P_1E_{1,i}d(x, u, t) \right. \\ &\quad \left. + 2e_2^T(t)P_{01}E_{2,i}d(x, u, t) - \gamma_1 d^T(x, u, t)d(x, u, t) \right) \\ &= \sum_{i=1}^k \mu_i \left( \begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{bmatrix}^T \Lambda_i \begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{bmatrix} \right. \\ &\quad \left. + 2 \begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{bmatrix}^T \begin{bmatrix} P_1 & 0 & 0 \\ 0 & P_{01} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} E_{1,i} \\ E_{2,i} \\ 0 \end{bmatrix} d(x, u, t) \right. \\ &\quad \left. - \gamma_1 d^T(x, u, t)d(x, u, t) \right) \\ &= \sum_{i=1}^k \mu_i \left( \begin{bmatrix} e_1(t) \\ e_2(t) \\ e_2(t) \\ d(x, u, t) \end{bmatrix}^T \psi_i \begin{bmatrix} e_1(t) \\ e_2(t) \\ e_2(t) \\ d(x, u, t) \end{bmatrix} \right) \end{aligned} \tag{41}$$

with

$$\psi_i = \begin{bmatrix} Q_{1,i} + H_1^T H_1 & P_1A_{12,i} & 0 & P_1E_{1,i} \\ * & Q_{2,i} + H_2^T H_2 & C_{22,i}^T P_{02} & P_{01}E_{2,i} \\ * & * & Q_{3,i} + H_3^T H_3 & 0 \\ * & * & * & -\gamma_1 I \end{bmatrix} \tag{42}$$

The previous inequalities are nonlinear because of  $P_{01}L_i$  and  $P_{02}K_i$ . This problem can be solved using the variable change  $X_i = P_{01}L_i$  and  $Y_i = P_{02}K_i$ . Applying the Schur complement, we can obtain the LMI form (29). So if  $J_1(t) < 0$ , the error dynamics system (25) is stable satisfying the  $H_\infty$  performance (28).  $\square$

### 3.2 Sliding motion analysis

For system (25), Let

$$S = \{e_1(t), e_2(t), e_3(t) | e_1(t) = 0, e_3(t) = 0\} \quad (43)$$

**Theorem 2.** Consider system (2) satisfying Assumptions 13. The system error dynamics (25) can be driven to the sliding surface  $S$  in finite time and remain on it if the LMI (29) is solvable and the gains  $l_{a,i}$  and  $l_{s,i}$  satisfy:

$$l_{a,i} \geq \|M_{1,i}^{-T}\| \left( \|A_{12,i}\| \|e_2(t)\| + \|E_{1,i}\| \|d(x, u, t)\| \right) + \eta_{a,i} \quad (44)$$

$$l_{s,i} \geq \|N_2^{-T}\| \|C_{22,i}\| \|e_2(t)\| + \eta_{s,i} \quad (45)$$

where  $\eta_{a,i}$  and  $\eta_{s,i}$  are two positive scalars,  $i = 1, \dots, k$ .

**Proof:**

Consider  $V_1(t) = e_1^T(t)P_1e_1(t) + \frac{1}{\sigma_1^2}\tilde{\rho}_a^2$  and  $V_3(t) = e_3^T(t)P_{02}e_3(t) + \frac{1}{\sigma_2^2}\tilde{\rho}_s^2$ . The differentia of  $V_1(t)$  can be obtained as:

$$\begin{aligned} \dot{V}_1(t) &= \sum_{i=1}^k \mu_i \left( e_1^T(t) \left( (A_{11,i}^s)^T P_1 + P_1 A_{11,i}^s \right) e_1(t) \right. \\ &\quad + 2e_1^T(t)P_1A_{12,i}e_2(t) + 2e_1^T(t)P_1E_{1,i}d(x, u, t) \\ &\quad \left. + 2e_1^T(t)P_1M_{1,i}(f_a(t) - v_{1eq,i}(t)) \right) + \frac{2}{\sigma_1} \tilde{\rho}_a (-\dot{\tilde{\rho}}_a) \quad (46) \end{aligned}$$

Since  $(A_{11,i}^s)^T P_1 + P_1 A_{11,i}^s < 0$ , then we can get

$$\begin{aligned} \dot{V}_1(t) &\leq \sum_{i=1}^k \mu_i \left( 2e_1^T(t)P_1A_{12,i}e_2(t) + 2e_1^T(t)P_1E_{1,i}d(x, u, t) \right. \\ &\quad \left. - l_{a,i} \|M_{1,i}^T P_1 e_1(t)\| \right) \\ &\leq \sum_{i=1}^k \mu_i \left[ \|P_1 e_1(t)\| \left( \|A_{12,i}\| \|e_2(t)\| + \|E_{1,i}\| \|d(x, u, t)\| \right) \right. \\ &\quad \left. - l_{a,i} \|M_{1,i}^T P_1 e_1(t)\| \right] \\ &\leq \sum_{i=1}^k \mu_i \left[ \|M_{1,i}^T P_1 e_1(t)\| \left( \|M_{1,i}^{-T}\| \left( \|A_{12,i}\| \|e_2(t)\| \right. \right. \right. \\ &\quad \left. \left. \left. + \|E_{1,i}\| \|d(x, u, t)\| \right) \right) - l_{a,i} \right] \quad (47) \end{aligned}$$

From (46), one obtains

$$\dot{V}_1(t) \leq -2\eta_{a,i} \|M_{1,i}^T P_1 e_1(t)\| \leq -2\eta_{a,i} \|M_{1,i}\| \sqrt{\lambda_{\min}(P_1)} V_1^{1/2}(t) \quad (48)$$

If (45) is verified, then

$$\dot{V}_3(t) \leq -2\eta_{s,i} \|N_2^T P_{02} e_3(t)\| \leq -2\eta_{s,i} \|N_2\| \sqrt{\lambda_{\min}(P_{02})} V_3^{1/2}(t) \quad (49)$$

Then the reachability condition [34] is verified.

## 4 Fault Estimation

From Theorem 2, an ideal sliding mode take place on  $S$  and  $\dot{e}_1(t) = e_1(t) = 0$ . Consequently, the error dynamics of  $e_1(t)$  becomes:

$$0 = \sum_{i=1}^k \mu_i \left( A_{12,i}e_2(t) + E_{1,i}d(x, u, t) + M_{1,i}(f_a(t) - v_{1eq,i}(t)) \right) \quad (50)$$

where  $v_{1eq,i}(t)$  denotes the equivalent term [34] replaced by:

$$v_{1eq,i}(t) = (\hat{\rho}_a + l_{a,i}) \frac{M_{1,i}^T P_1 \left( C_{11,i}^{-1} S_{11,i} v_1(t) - \hat{z}_1(t) \right)}{\left\| M_{1,i}^T P_1 \left( C_{11,i}^{-1} S_{11,i} v_1(t) - \hat{z}_1(t) \right) \right\| + \delta_a} \quad (51)$$

where  $\delta_a > 0$ . Since  $M_{1,i}$  is invertible, (50) can be rewritten as:

$$v_{1eq,i}(t) - f_a(t) = M_{1,i}^{-1} \left( A_{12,i}e_2(t) + E_{1,i}d(x, u, t) \right) \quad (52)$$

Computing the  $L_2$  norm of (52) yields

$$\begin{aligned} &\|v_{1eq,i}(t) - f_a(t)\|_2 \\ &= \left\| M_{1,i}^{-1} \left( A_{12,i}e_2(t) + E_{1,i}d(x, u, t) \right) \right\|_2 \\ &\leq \|M_{1,i}^{-1} A_{12,i}\|_2 \|e_2(t)\|_2 + \|M_{1,i}^{-1} E_{1,i}\|_2 \|d(x, u, t)\|_2 \\ &\leq \|M_{1,i}^{-1} A_{12}\|_{\max} \|e(t)\|_2 + \|M_{1,i}^{-1} E_1\|_{\max} \|d(x, u, t)\|_2 \\ &\leq \left( \sqrt{\gamma_1} \|M_1^{-1} A_{12}\|_{\max} \sigma_{\max}(H^{-1}) \right. \\ &\quad \left. + \|M_1^{-1} E_1\|_{\max} \right) \|d(x, u, t)\|_2 \quad (53) \end{aligned}$$

where  $\|M_1^{-1} A_{12}\|_{\max} = \max_{i=1, \dots, k} \left( \|M_{1,i}^{-1} A_{12,i}\|_2 \right)$  and  $\|M_1^{-1} E_1\|_{\max} = \max_{i=1, \dots, k} \left( \|M_{1,i}^{-1} E_{1,i}\|_2 \right)$ . since  $\|e(t)\| \leq \sigma_{\max}(H^{-1}) \sqrt{\gamma_1} \|d(x, u, t)\|$ . It follows that:

$$\sup_{\|d\| \neq 0} \frac{\|v_{1eq,i}(t) - f_a(t)\|_2}{\|d(x, u, t)\|_2} = \sqrt{\gamma_1} \beta_1 + \beta_2 \quad (54)$$

where  $\beta_1 = \|M_1^{-1} A_{12}\|_{\max} \sigma_{\max}(H^{-1})$  and  $\beta_2 = \|M_1^{-1} E_1\|_{\max}$ . Thus for a small  $\sqrt{\gamma_1} \beta_1 + \beta_2$ ,  $f_a(t)$  can be estimated as:

$$\hat{f}_a(t) \cong \sum_{i=1}^k \mu_i \left( (\hat{\rho}_a + l_{a,i}) \frac{M_{1,i}^T P_1 \left( C_{11,i}^{-1} S_{11,i} v_1(t) - \hat{z}_1(t) \right)}{\left\| M_{1,i}^T P_1 \left( C_{11,i}^{-1} S_{11,i} v_1(t) - \hat{z}_1(t) \right) \right\| + \delta_a} \right) \quad (55)$$

Similarly, we can get

$$\sup_{\|d\| \neq 0} \frac{\|v_{2eq,i}(t) - f_s(t)\|_2}{\|d(x, u, t)\|_2} = \sqrt{\gamma_1} \|N_2^{-1} C_{22}\|_{\max} \sigma_{\max}(H^{-1}) \quad (56)$$

Therefore for small  $\sqrt{\gamma_1} \|N_2^{-1} C_{22}\|_{\max} \sigma_{\max}(H^{-1})$ ,  $f_s(t)$  can be estimated as:

$$\hat{f}_s(t) \cong \sum_{i=1}^k \mu_i \left( (\hat{\rho}_s + l_{s,i}) \frac{N_{2,i}^T P_{03} e_3(t)}{\left\| N_{2,i}^T P_{03} e_3(t) \right\|} \right) \quad (57)$$

## 5 Fault Tolerant Controller Design

Define corrected output as:

$$y_c(t) = \sum_{i=1}^k \mu_i \left( C_i x(t) + N \left( f_s(t) - \hat{f}_s(t) \right) \right) \quad (58)$$

System (2) becomes:

$$\begin{cases} \dot{x}(t) = \sum_{i=1}^k \mu_i \left( A_i x(t) + B_i u(t) + M_i f_a(t) + E_i d(t) \right) \\ y_c(t) = \sum_{i=1}^k \mu_i \left( C_i x(t) + N e_{f_s}(t) \right) \end{cases} \quad (59)$$

where  $e_{f_s(t)} = f_s(t) - \hat{f}_s(t)$ .

A SOFFTC law [35] is designed as follows:

$$u(t) = \sum_{i=1}^k \mu_i (\bar{K}_i y_c(t) - \bar{G}_i \hat{f}_a(t)) \quad (60)$$

where  $K_i$  and  $G_i$  are gains matrices to be determined.

Substituting (60) in (59), we have

$$\begin{aligned} \dot{x}(t) &= \sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j (A_i x(t) + B_i (\bar{K}_j C_j x(t) \\ &\quad + \bar{K}_j N e_{f_s(t)} - \bar{G}_j \hat{f}_a(t)) + M_i f_a(t) + E_i d(x, u, t)) \\ &= \sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j (A_i x(t) + B_i \bar{K}_j C_j x(t) + B_i \bar{K}_j N e_{f_s(t)} \\ &\quad - B_i \bar{G}_j \hat{f}_a(t) + M_i f_a(t) + E_i d(x, u, t)) \end{aligned} \quad (61)$$

The gain  $G_i$  is designed so that  $B_i G_i = M_i$  where  $B_i^+$  is the pseudo inverse of  $B_i$  [36]. It follows that

$$\begin{aligned} \dot{x}(t) &= \sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j ((A_i + B_i \bar{K}_j C_j) x(t) + B_i \bar{K}_j N e_{f_s(t)} \\ &\quad + M_i e_{f_a(t)} + E_i d(x, u, t)) \end{aligned} \quad (62)$$

where  $e_{f_a(t)} = f_a(t) - \hat{f}_a(t)$ . Then, we get

$$\begin{cases} \dot{x}(t) = \sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j [(A_i + B_i \bar{K}_j C_j) x(t) + \bar{B}_{ij} \varphi(t)] \\ y_c(t) = \sum_{i=1}^k \mu_i (C_i x(t) + N e_{f_s(t)}) \end{cases} \quad (63)$$

where  $\bar{B}_{ij} = [B_i \bar{K}_j N \quad M_i \quad E_i]$  and  $\varphi(t) = [e_{f_s}^T(t) \quad e_{f_a}^T(t) \quad d^T(x, u, t)]^T$ . The control purpose in this paper for the closed-loop fuzzy system (63) is to design a SOFFTC (60) such that

(i) The closed-loop fuzzy system (63) with  $(\varphi(t) = 0)$  is asymptotically stable .

(ii) For a given scalar  $\gamma_2 > 0$ , the following  $H_\infty$  performance is satisfied:

$$\int_0^L \|y_c(t)\|_2^2 dt < \gamma_2 \int_0^L \|\varphi(t)\|_2^2 dt \quad (64)$$

for all  $L > 0$  and  $\varphi(t) \in \mathcal{L}_2 [0, \infty)$  under zero initial conditions.

**Theorem 3.** The closed-loop system (63) is asymptotically stable and satisfy the  $H_\infty$  performance index (64), if there exist matrix  $\bar{P}_x > 0$ , matrices  $R, \bar{S}_i$  and scalar  $\epsilon > 0$ , such that:

$$\begin{cases} \Psi_{ii} < 0, & 1 \leq i \leq k \\ \frac{2}{r-1} \Psi_{ii} + \Psi_{ij} + \Psi_{ji} < 0, & 1 \leq i \neq j \leq k \\ C_i \bar{P}_x = R C_i, & 1 \leq i \leq k \end{cases} \quad (65)$$

where

$$\Psi_{ij} = \begin{bmatrix} \Phi_{ij} & 0 & M_i & N & \bar{P}_x C_i^T & B_i \bar{K}_j & 0 \\ * & -2\gamma_2 \bar{P}_x + \gamma_2 I & 0 & 0 & \bar{P}_x N^T & 0 & \epsilon \bar{P}_x C_i \\ * & * & -\gamma_2 I & 0 & 0 & 0 & 0 \\ * & * & * & -\gamma_2 I & 0 & 0 & 0 \\ * & * & * & * & -\gamma_2 I & 0 & 0 \\ * & * & * & * & * & -\epsilon I & 0 \\ * & * & * & * & * & * & -\epsilon I \end{bmatrix}$$

with

$$\Phi_{ij} = A_i \bar{P}_x + \bar{P}_x A_i + B_i S_j C_i + B_i^T S_j^T C_i^T$$

The controller gains are obtained by:

$$\bar{K}_i = R^{-1} \bar{S}_i$$

**Proof.** Choose  $V_x(t) = x^T(t) P_x x(t)$ , where  $P_x > 0$ . Its derivative is:

$$\begin{aligned} \dot{V}_x(t) &= \sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j (x^T(t) ((A_i + B_i \bar{K}_j C_j)^T P_x \\ &\quad + P_x (A_i + B_i \bar{K}_j C_j)) x(t) + 2x^T(t) P_x \bar{B}_{ij} \varphi(t)) \end{aligned} \quad (66)$$

Let

$$J_x(t) = \dot{V}_x(t) + y_c^T(t) y_c(t) - \gamma_c \varphi^T(t) \varphi(t) \quad (67)$$

where

$$\begin{aligned} y_c^T(t) y_c(t) &= \sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j ((C_i x(t) + N e_{f_s(t)})^T \\ &\quad (C_j x(t) + N e_{f_s(t)})) \\ &= \sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j (x^T(t) C_i^T C_j x(t) + x^T(t) C_i^T N e_{f_s(t)} \\ &\quad + e_{f_s}^T(t) N^T C_j x(t) + e_{f_s}^T(t) N^T N e_{f_s(t)}) \end{aligned} \quad (68)$$

Define  $Z = [N \quad 0 \quad 0]$ , then

$$\begin{aligned} y_c^T(t) y_c(t) &= \sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j (x^T(t) C_i^T C_j x(t) \\ &\quad + \varphi^T(t) Z^T Z \varphi(t) + 2x^T(t) C_i^T Z \varphi(t)) \end{aligned} \quad (69)$$

So we obtain

$$\begin{aligned} J_x(t) &= \sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j (x^T(t) ((A_i + B_i \bar{K}_j C_j)^T P_x \\ &\quad + P_x (A_i + B_i \bar{K}_j C_j)) x(t) + 2x^T(t) P_x \bar{B}_{ij} \varphi(t) \\ &\quad + x^T(t) C_i^T C_j x(t) + \varphi^T(t) Z^T Z \varphi(t) \\ &\quad + 2x^T(t) C_i^T Z \varphi(t) - \gamma_2 \varphi^T(t) \varphi(t)) \\ &= \sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j \begin{bmatrix} x(t) \\ \varphi(t) \end{bmatrix}^T \Theta_{ij} \begin{bmatrix} x(t) \\ \varphi(t) \end{bmatrix} \end{aligned}$$

where

$$\Theta_{ij} = \begin{bmatrix} \Upsilon_{ij} & P_x \bar{B}_{ij} + C_i^T Z \\ * & -\gamma_2 I + Z^T Z \end{bmatrix} \quad (70)$$

with  $\Upsilon_{ij} = (A_i + B_i \bar{K}_j C_i)^T P_x + P_x (A_i + B_i \bar{K}_j C_i) + C_i^T C_j$ .

Thus,  $J_x(t) < 0$ , if

$$\sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j \Theta_{ij} < 0 \quad (71)$$

By applying Schur complement, (70) can be written as:

$$\Theta_{ij} = \begin{bmatrix} \tilde{\Upsilon}_{ij} & P_x B_i \bar{K}_j N & P_x M_i & P_x E_i & C^T \\ * & -\gamma_2 I & 0 & 0 & N^T \\ * & * & -\gamma_2 I & 0 & 0 \\ * & * & * & -\gamma_2 I & 0 \\ * & * & * & * & -\gamma_2 I \end{bmatrix} \quad (72)$$

where  $\tilde{\Upsilon}_{ij} = (A_i + B_i \bar{K}_j C_i)^T P_x + P_x (A_i + B_i \bar{K}_j C_i)$ .

Premultiplying and postmultiplying by  $\Pi = \text{diag}\{P_x^{-1}, P_x^{-1}, I, I, I\}$  and its transpose in (72), then we obtain

$$\sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j \left( \begin{bmatrix} \Omega_{ij} & B_i \bar{K}_j N \bar{P}_x & M_i & N & \bar{P}_x C_i^T \\ * & -\gamma_2 \bar{P}_x \bar{P}_x & 0 & 0 & \bar{P}_x N^T \\ * & * & -\gamma_2 I & 0 & 0 \\ * & * & * & -\gamma_2 I & 0 \\ * & * & * & * & -\gamma_2 I \end{bmatrix} \right) < 0 \quad (73)$$

where  $\Omega_{ij} = A_i \bar{P}_x + \bar{P}_x A_i + B_i \bar{K}_j C_i \bar{P}_x + \bar{P}_x^T C_i^T \bar{K}_j^T B_i^T$  and  $\bar{P}_x = P_x^{-1}$ . Based on Lemma 1, it is easy to obtain that

$$\bar{P}_x + \bar{P}_x \leq \bar{P}_x \bar{P}_x + I \quad (74)$$

From  $\gamma_2 > 0$ , (74) can be expressed as:

$$-\gamma_2 \bar{P}_x \bar{P}_x \leq -2\gamma_2 \bar{P}_x + \gamma_2 I \quad (75)$$

Then, (73) can be rewritten as:

$$\sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j \Sigma_{ij} < 0$$

where

$$\Sigma_{ij} = \begin{bmatrix} \Omega_{ij} & B_i \bar{K}_j N \bar{P}_x & M_i & N & \bar{P}_x C_i^T \\ * & -2\gamma_2 \bar{P}_x + \gamma_2 I & 0 & 0 & \bar{P}_x N^T \\ * & * & -\gamma_2 I & 0 & 0 \\ * & * & * & -\gamma_2 I & 0 \\ * & * & * & * & -\gamma_2 I \end{bmatrix} \quad (77)$$

Notice that the matrix inequality  $\Omega_{ij} < 0$  is a Bilinear Matrix Inequalities (BMIs). Denoting  $C_i \bar{P}_x = RC_i$  and  $\bar{K}_j R = \bar{S}_j$ , so that  $K_j C_i \bar{P}_x = \bar{S}_j C_i$ . Substituting the result into (77) yields

$$\sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j \Xi_{ij} \leq 0 \quad (78)$$

where

$$\Xi_{ij} = \begin{bmatrix} \Phi_{ij} & B_i \bar{K}_j N \bar{P}_x & M_i & N & \bar{P}_x C_i^T \\ * & -2\gamma_2 \bar{P}_x + \gamma_2 I & 0 & 0 & \bar{P}_x N^T \\ * & * & -\gamma_2 I & 0 & 0 \\ * & * & * & -\gamma_2 I & 0 \\ * & * & * & * & -\gamma_2 I \end{bmatrix} \quad (79)$$

with

$$\Phi_{ij} = A_i \bar{P}_x + \bar{P}_x A_i + B_i \bar{S}_j C_i + B_i^T \bar{S}_j^T C_i^T \quad (80)$$

The gains of controller are given by

$$\bar{K}_i = R^{-1} \bar{S}_i \quad (81)$$

Furthermore,  $\Xi_{ij}$  can be further decomposed as below:

$$\Delta_{ij} + UV + (UV)^T < 0 \quad (82)$$

where

$$\Delta_{ij} = \begin{bmatrix} \Phi_{ij} & 0 & M_i & N & \bar{P}_x C_i^T \\ * & -2\gamma_2 \bar{P}_x + \gamma_2 I & 0 & 0 & \bar{P}_x N^T \\ * & * & -\gamma_2 I & 0 & 0 \\ * & * & * & -\gamma_2 I & 0 \\ * & * & * & * & -\gamma_2 I \end{bmatrix}$$

$$U = \begin{bmatrix} \bar{K}_j^T B_i^T & 0 & 0 & 0 & 0 \end{bmatrix}^T$$

$$V = \begin{bmatrix} 0 & N \bar{P}_x & 0 & 0 & 0 \end{bmatrix}$$

By using Lemma 3, it is follows that

$$\Delta_{ij} + UV + (UV)^T \leq \Delta_{ij} + \epsilon^{-1} U U^T + \epsilon V^T V \quad (83)$$

From (83), (78) is equivalent to

$$\sum_{i=1}^k \sum_{j=1}^k \mu_i \mu_j \Psi_{ij} < 0 \quad (84)$$

where

$$\Psi_{ij} = \begin{bmatrix} \Phi_{ij} & 0 & M_i & N & \bar{P}_x C_i^T & B_i \bar{K}_j & 0 \\ * & -2\gamma_2 \bar{P}_x + \gamma_2 I & 0 & 0 & \bar{P}_x N^T & 0 & \epsilon \bar{P}_x C_i \\ * & * & -\gamma_2 I & 0 & 0 & 0 & 0 \\ * & * & * & -\gamma_2 I & 0 & 0 & 0 \\ * & * & * & * & -\gamma_2 I & 0 & 0 \\ * & * & * & * & * & -\epsilon I & 0 \\ * & * & * & * & * & * & -\epsilon I \end{bmatrix} \quad (76)$$

If (84) is verified, we have  $J_x(t) < 0$ , which can guarantee the closed-loop system is asymptotically stable.  $\square$

## 6 Simulation Example

In this example, The inverted pendulum with cart system [37] is considered:

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = \frac{g \sin(x_1(t)) - m l a x_2^2(t) \frac{\sin(2x_1(t))}{2} - b a \cos(x_1(t)) x_4(t) - a \cos(x_1(t)) (F - f_c)}{\frac{4l}{3} - m l a \cos(x_1(t))^2} \\ \dot{x}_3(t) = x_4(t) \\ \dot{x}_4(t) = \frac{-m g a \frac{\sin(2x_1(t))}{2} + \frac{4m l a}{3} x_2^2(t) \sin(x_1(t)) - b a x_4(t) + \frac{4a}{3} (F - f_c)}{\frac{4}{3} - m a \cos(x_1(t))^2} \end{cases} \quad (78)$$

where  $x_1(t)$  and  $x_2(t)$  are the angular position and velocity, respectively;  $x_3(t)$  and  $x_4(t)$  are the cart position and velocity, respectively;

$m$  is the pendulum mass,  $M$  is the cart mass,  $g$  is the gravity constant, and  $a = 1/(m + M)$ . The values of the parameters used in this simulation are  $m = 0.2kg$ ,  $M = 0.8kg$ ,  $l = 0.5m$ , and  $L = 2m$ .

Then, we get the TS fuzzy system with the following fuzzy rules:

**Rule1: If  $x_1(t)$  is about 0, Then**

$$\begin{cases} \dot{x}(t) = A_1x(t) + B_1(u(t) + f_a(t)) + E_1d(t) \\ y(t) = C_1x(t) + Nf_s(t) \end{cases}$$

**Rule1: If  $x_1(t)$  is about  $\pm \frac{\pi}{4}$ , Then**

$$\begin{cases} \dot{x}(t) = A_2x(t) + B_2(u(t) + f_a(t)) + E_2d(t) \\ y(t) = C_2x(t) + Nf_s(t) \end{cases}$$

where

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{g}{3} & 0 & 0 & \frac{ba}{3} \\ 0 & 0 & 0 & 1 \\ \frac{-mga}{3} & 0 & 0 & \frac{-ba}{3} \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{g}{3} & \frac{2\sqrt{2}}{3} & 0 & \frac{ba\sqrt{2}}{3} \\ 0 & 0 & 0 & 1 \\ \frac{-mga}{3} & 0 & 0 & \frac{-ba}{3} \end{bmatrix}$$

$$B_1 = E_1 = \begin{bmatrix} 0 \\ \frac{-a}{3} \\ 0 \\ \frac{4a}{3} \end{bmatrix}, B_2 = E_2 = \begin{bmatrix} 0 \\ \frac{-a\sqrt{2}}{3} \\ 0 \\ \frac{4a}{3} \end{bmatrix}$$

$$C_1 = C_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, N = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$$

we set  $d(x, u, t) = -f_c + m l x_2^2(t) \sin(x_1(t))$  where  $f_c = \rho \text{sign}(x_4(t))$  with  $\rho = 0.05$ . The membership functions for rules 1 and 2 are chosen based on the method of sector nonlinearity [38] as follows:

$$\begin{aligned} \mu_1(x_1(t)) &= \frac{1 - \frac{1}{1 + \exp(-14(x_1(t) - \frac{\pi}{8}))}}{1 + \exp(-14(x_1(t) + \frac{\pi}{8}))} \\ \mu_2(x_1(t)) &= 1 - \mu_1(x_1(t)) \end{aligned}$$

By Choosing  $H_1 = 1$ ,  $H_2 = I_3$  and  $H_3 = I_2$ , we can solve the optimization problem of Theorem 1 using Matlab LMI Toolbox and we obtain  $\gamma_1 = 0.3614$ ,  $A_{11,1}^s = -3.2457$ ,  $A_{11,2}^s = -3.2457$ ,  $P_1 = 1.62$  and

$$P_{01} = \begin{bmatrix} 0.164 & 0.043 & -0.3 \\ 0.033 & 0.109 & -0.454 \\ -0.5 & -0.404 & 11.248 \end{bmatrix}, P_{02} = \begin{bmatrix} 0.463 & 0 \\ 0 & 0.804 \end{bmatrix}$$

The observer gains can be calculated as follows:

$$\begin{aligned} L_1 &= \begin{bmatrix} 5.0126 & 1.0616 \\ 17.4978 & 20.7737 \\ -1.3711 & -7.8575 \end{bmatrix}, L_2 = \begin{bmatrix} 5.0066 & 1.0558 \\ 14.7712 & 20.7980 \\ -1.0564 & -11.1295 \end{bmatrix} \\ K_1 &= \begin{bmatrix} -2.47 & 0.075 \\ 0.053 & -4.381 \end{bmatrix}, K_2 = \begin{bmatrix} -2.341 & 0.014 \\ 0.042 & -3.862 \end{bmatrix} \end{aligned}$$

Solving the optimization problem in Theorem 3 results the following controller gain matrices for a minimum attenuation level  $\gamma_2 = 1.6128$ :

$$\begin{aligned} \bar{K}_1 &= \begin{bmatrix} -2.7838 & -6.8067 & -11.1030 \end{bmatrix} \\ \bar{K}_2 &= \begin{bmatrix} -2.1520 & -7.5055 & -11.8265 \end{bmatrix} \end{aligned}$$

We simulate the closed-loop system by choosing  $\sigma_1 = 10$ ,  $\sigma_2 = 15$ ,  $\delta_1 = 0.01$ ,  $\delta_2 = 0.02$ ,  $l_{a,1} = l_{a,2} = 10$  and  $l_{s,1} = l_{s,2} = 12$ , initial conditions  $x_{10} = \pi/20$ ,  $x_{20} = 0$ ,  $x_{30} = 2$  and  $x_{40} = 0$ .  $f_a(t)$  and  $f_s(t)$  are assumed that, respectively

$$f_a(t) = \begin{cases} 0 & t < 6 \\ \sin(\pi(t-6)) & t \geq 6 \end{cases}, f_s(t) = \begin{cases} 0 & t < 8 \\ 0.5 & t \geq 8 \end{cases}$$

The simulation results of this example are given in Figures 1-5. Figures 1 and 2 show that the proposed adaptive SMO can estimate the actuator and sensor faults simultaneously despite the presence of uncertainties.

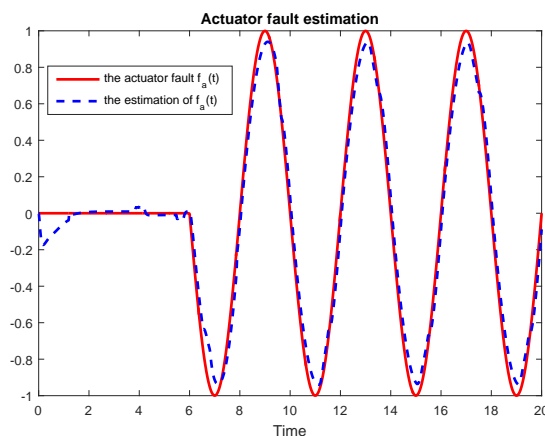


Figure 1: The actuator fault  $f_a(t)$  and its estimation.

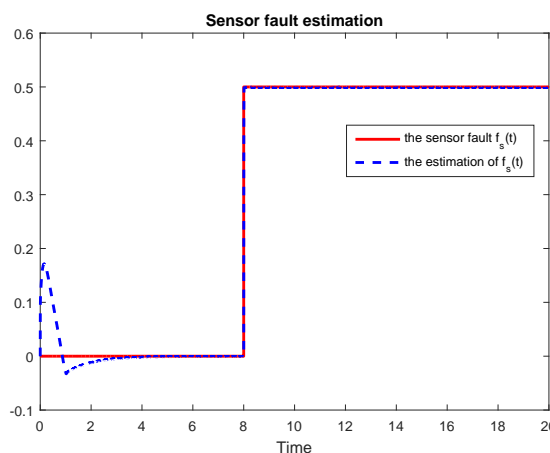


Figure 2: The sensor fault  $f_s(t)$  and its estimation.

Simulation results for the systems outputs response are illustrated in Figures 3-5. It is clear to see that the outputs without SOFFTC do not converge to the outputs of the fault-free model (i.e.



without any fault). However, the outputs's trajectories of the system with SOFFTC reach the outputs of nominal model.

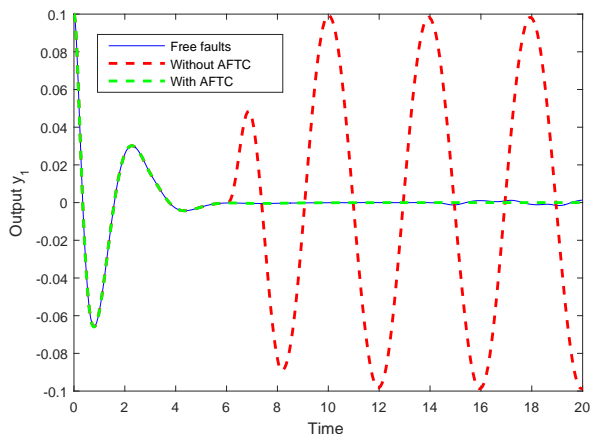


Figure 3: Output  $y_1(t)$  under the static output feedback FTC.

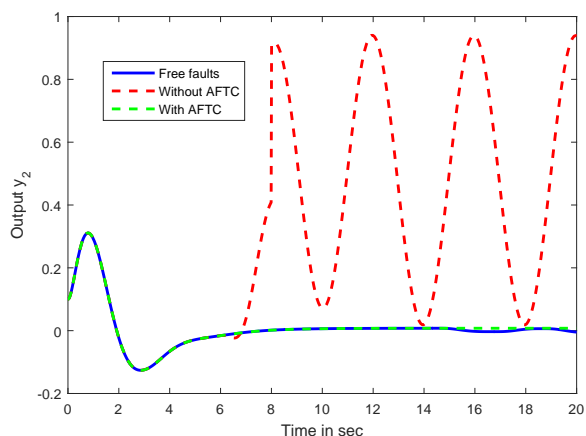


Figure 4: Output  $y_2(t)$  under the static output feedback FTC.

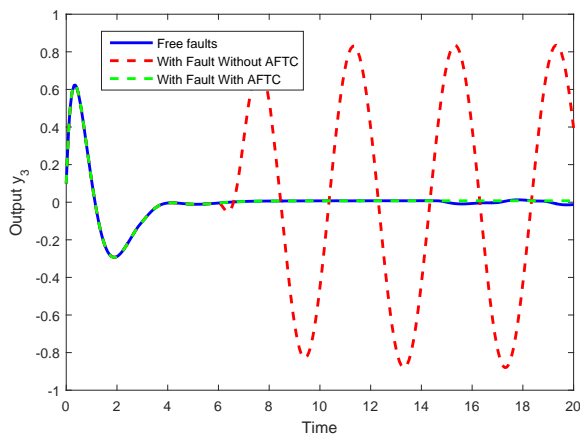


Figure 5: Output  $y_3(t)$  under the static output feedback FTC.

## 7 Conclusion

In this paper, we have considered the problems of FE and FTC for T-S fuzzy systems with uncertainties, actuator and sensor faults, simultaneously. Using the  $H_\infty$  optimization technique, an adaptive fuzzy sliding mode observer has been firstly designed to estimate the system state, actuator and sensor faults, simultaneously. Secondly, using the information of online fault estimates, a novel SOFFTC has been developed to compensate the faults and stabilize the closed-loop system. Thus, sufficient condition for the existence of the proposed ASMO and SOFFTC has been formulated in terms of LMIs. Finally, a simulation example was used to show the effectiveness of the proposed methods.

## References

- [1] S. Ding, "Integrated design of feedback controllers and fault detectors" *Annual Reviews in Control*, **33**(3), 124-135, 2009. <https://doi.org/10.1016/j.arcontrol.2009.08.003>
- [2] R. Isermann, "Fault-Diagnosis Applications : Model-Based Bondition Monitoring : Actuators, Drives, Machinery, Plants, Sensors, and Fault-Tolerant Systems", Springer Science and Business Media, 2011.
- [3] H. Alwi, C. Edwards, C. Tan, "Fault Detection and Fault-Tolerant Control Using Sliding Modes", Springer Science and Business Media, 2011.
- [4] S. Dhahri, A. Sellami, F. Ben Hmida, "Robust  $H_\infty$  sliding mode observer design for fault estimation in a class of uncertain nonlinear systems with LMI optimization approach" *International Journal of Control, Automation and Systems*, **10**(5), 1032-1041, 2012. <https://doi.org/10.1007/s12555-012-0521-3>
- [5] M. S. Shaker, R. J. Patton, "Active fault tolerant control for nonlinear systems with simultaneous actuator and sensor faults" *International Journal of Control Automation and Systems*, **11** (6), 1149-1161, 2013. <https://doi.org/10.1007/s12555-013-0227-1>
- [6] Z. Gao, C. Cecati, S. Ding, "A survey of fault diagnosis and fault tolerant techniques-part i : fault diagnosis with model-based and signal-based approaches" *IEEE Transactions on Industrial Electronics*, **62**(6), 3757-3767, 2015. <https://doi.org/10.1109/TIE.2015.2417501>
- [7] J. Wang, " $H_\infty$  fault-tolerant controller design for networked control systems with time-varying actuator faults" *International Journal of Innovative Computing, Information and control*, **11**(4), 1471-1481, 2015.
- [8] H. Azmi, M. J. Khosrowjerdi, "Robust adaptive fault tolerant control for a class of lipschitz nonlinear systems with actuator failure and disturbances" *Journal of Systems and Control Engineering*, **230**(1), 13-22, 2015. <https://doi.org/10.1177/0959651815606628>
- [9] Y. Tian, F. Zhu, "Fault estimation and observer based fault-tolerant controller in finite frequency domain" *Transactions of the Institute of Measurement and Control*, **40**(5), 1659-1668, 2017. <https://doi.org/10.1177/0959651815606628><https://doi.org/10.1177/0959651815606628>
- [10] T. Takagi, M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control" *IEEE Transactions on Systems, Man and Cybernetics*, **15**(1), 116-132, 1985. <https://doi.org/10.1109/TSMC.1985.6313399>
- [11] K. Zhang, B. Jiang, P. Shi, "Fault estimation observer design for discrete time TakagiSugeno fuzzy systems based on piecewise Lyapunov functions" *IEEE Trans. Fuzzy Syst.*, **20**(1), 192200, 2012. <https://doi.org/10.1109/TFUZZ.2011.2168961>
- [12] Q. Jia, W. Chen, Y. Jin, Y. Zhang, H. Li, "A New Strategy for Fault Estimation in TakagiSugeno Fuzzy Systems via a Fuzzy Learning Observer" in *Proceeding of the 11th World Congress on Intelligent Control and Automation (WCICA)*, Shenyang, China, 3228-3233, 2014. <https://doi.org/10.1109/WCICA.2014.7053248>
- [13] C. Sun, F. Wang, X.Q. He, "Robust fault estimation for a class of TS fuzzy singular systems with time-varying delay via improved delay partitioning approach" *J. Control Sci. Eng.*, **2016**, 2016. <https://doi.org/10.1155/2016/6305901>

- [14] D. Ding, X. Du, X. Xie, M. Li, "Fault estimation filter design for discretetime TakagiSugeno fuzzy systems" *IET Control Theory Appl.*, **10**(18), 2456-2465, 2016. <https://doi.org/10.1049/iet-cta.2016.0318>
- [15] F. You, S. Cheng, K. Tian, X. Zhang, "Robust fault estimation based on learning observer for TakagiSugeno fuzzy systems with interval timevarying delay" *Int. J. Adapt. Control Signal Process.*, **34**(1), 92-109, 2020. <https://doi.org/10.1002/acs.3070>
- [16] S. Liu, X. Li, H. Wang, J. Yan, "Adaptive fault estimation for T-S fuzzy systems with unmeasurable premise variables" *Advances in Difference Equations* **2018**, **105**(2018), 2018. <https://doi.org/10.1186/s13662-018-1571-5>
- [17] Z. Gao, X. Shi, S.X. Ding, "Fuzzy state/disturbance observer design for TS fuzzy systems with application to sensor fault estimation" *IEEE Transactions on Cybernetics*, **38**(3), 875-880, 2018. <https://doi.org/10.1109/TSMCB.2008.917185>
- [18] Q. K. Shen, B. Jiang, V. Cocquempot, "Fault tolerant control for TS fuzzy systems with application to nearspace hypersonic vehicle with actuator faults" *IEEE Trans. Fuzzy Syst.*, **20**(4), 652665, 2012. <https://doi.org/10.1109/TFUZZ.2011.2181181>
- [19] X. H. Li, F. L. Zhu, A. Chakrabarty, "Nonfragile fault tolerant fuzzy observer-based controller design for nonlinear systems" *IEEE Trans. Fuzzy Syst.*, **24**(6), 16791689, 2016. <https://doi.org/10.1109/TFUZZ.2016.2540070>
- [20] M. Liu, X. B Cao, and P. Shi, "Fault Estimation and Tolerant Control Fuzzy Stochastic Systems" *IEEE Trans. Fuzzy Syst.*, **21**(2), 221229, 2013.
- [21] A. Navarraf, M. J. Khosrowjerdi, "Fault-tolerant controller design with fault estimation capability for a class of nonlinear systems using generalized Takagi-Sugeno fuzzy model" *Transactions of the Institute of Measurement and Control*, **41**(15), 42184229, 2019. <https://doi.org/10.1177/0142331219853687>
- [22] M. Liu, X. Cao, P. Shi, "Fuzzy-model-based fault-tolerant design for nonlinear stochastic systems against simultaneous sensor and actuator faults" *IEEE Trans. Fuzzy Syst.*, **21**(5), 789799, 2013. <https://doi.org/10.1109/TFUZZ.2012.2224872>
- [23] M. Sami, R. J. Patton, "Active fault tolerant control for nonlinear systems with simultaneous actuator and sensor faults" *Int. J. Control Autom. Syst.*, **11**(6), 11491161, 2013. <https://doi.org/10.1007/s12555-013-0227-1>
- [24] S. Makni, M. Bouattour, A.E. Hajjaji, M. Chaabane, "Robust fault tolerant control based on adaptive observer for Takagi-Sugeno fuzzy systems with sensor and actuator faults: Application to single-link manipulator" *Transactions of the Institute of Measurement and Control*, **2020**, 2020. [doi.org/10.1177/0142331220909996](https://doi.org/10.1177/0142331220909996).
- [25] P. Aboutalebi, A. Abbaspour, P. Forouzannezhad, A. Sargolzaei, "A novel sensor fault detection in an unmanned quadrotor based on adaptive neural observer" *Journal of Intelligent and Robotic Systems*, **90**, 473484, 2018. <https://doi.org/10.1007/s10846-017-0690-7>
- [26] H. Li, Y. Gao, P. Shi, H. K. Lam, "Observer-based fault detection for nonlinear systems with sensor fault and limited communication capacity" *IEEE Transactions on Automatic Control*, **61**(9), 27452751, 2016. <https://doi.org/10.1109/TAC.2015.2503566>
- [27] A. Valibeygi, A. Toudeshki, K. Vijayaraghavan, "Observer based sensor fault estimation in nonlinear systems" *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, **230**(8), 759777, 2016. <https://doi.org/10.1177/0959651816654070>
- [28] M. Buciakowski, M. Witczak M, V. Puig, et al., "A bounded-error approach to simultaneous state and actuator fault estimation for a class of nonlinear systems" *Journal of Process Control*, **52**, 1425, 2017. <https://doi.org/10.1016/j.jprocont.2017.01.002>
- [29] M. Bataghva, M. Hashemi, "Adaptive sliding mode synchronisation for fractional-order non-linear systems in the presence of time-varying actuator faults" *IET Control Theory Appl.*, **12**(3), 37783, 2018. <https://doi.org/10.1049/iet-cta.2017.0458>
- [30] H. Schulte, E. Gauterin, "Fault-tolerant control of wind turbines with hydrostatic transmission using TakagiSugeno and sliding mode techniques" *Annual Reviews in Control*, **40**, 8292, 2015. <https://doi.org/10.1016/j.arcontrol.2015.08.003>
- [31] H. Moodi, M. Farrokhi, "On observer-based controller design for sugeno systems with unmeasurable premise variables" *ISA Transactions*, **53**(2), 305316, 2014. <https://doi.org/10.1016/j.isatra.2013.12.004>
- [32] T. Dang, W. Wang, L. Luoh, C. Sun, "Adaptive observer design for the uncertain takagisugeno fuzzy system with output disturbance" *IET Control Theory Applications*, **6**(10), 13511366, 2011. <https://doi.org/10.1016/j.isatra.2013.12.004>
- [33] M. Corless, J. Tu, "State and input estimation for a class of uncertain systems" *Automatica*, **34**(6), 757764, 1998. [https://doi.org/10.1016/S0005-1098\(98\)00013-2](https://doi.org/10.1016/S0005-1098(98)00013-2)
- [34] V.I. Utkin, *Sliding modes in control optimization*, Springer, Berlin, 1992.
- [35] K. Tanaka, H. Wang, *Fuzzy control systems design and analysis: a linear matrix inequality approach*, Wiley, New York, USA, 2001.
- [36] Z. Gao, P. J. Antsaklis, "Stability of the pseudo inverse method for reconfigurable control systems" *Int. J. of Control*, **53**, 717-729, 1991. <https://doi.org/10.1080/00207179108953643>
- [37] K. Ogata, *Modern control engineering*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1997.
- [38] M. C. M. Teixeira, S. H. Zak, "Stabilizing controller design for uncertain nonlinear systems using fuzzy models" *IEEE Transaction on Fuzzy Systems*, **7**(2), 133-142, 1999. <https://doi.org/10.1109/91.755395>

## Maximum Power-Point Tracking and Stall Control with Eddy Current Brake System on Small-Scaled Wind Turbines and its Application on Agricultural Harvesting

Anupa Koswatta<sup>1</sup>, Faramarz Alsharif<sup>2</sup>, Yasushi Shiroma<sup>1</sup>, Shiro Tamaki<sup>1,\*</sup>, Junji Tamura<sup>2</sup>

<sup>1</sup>Graduate school of Engineering, University of the Ryukyus, 903-0213, Japan

<sup>2</sup>School of Earth, Energy and Environmental Engineering, Kitami Institute of Technology, 090-8501, Japan

### ARTICLE INFO

Article history:

Received: 21 May, 2020

Accepted: 21 June, 2020

Online: 12 July, 2020

Keywords:

Small-scaled Wind Turbine

MPPT control

Eddy current brake

Over-rotation

Strong wind

### ABSTRACT

This research aims to enhance the generated power of the small-scaled wind turbine using the eddy current brake system and Maximum Power Point Tracking (MPPT) control method. We analyzed the behavior of the generated power and power factor, with and without the MPPT control which implemented by eddy current brake system. Also, the feasibility of the system investigated using different wind conditions such as strong and calm wind conditions. The load data has different voltage respond to the system since its conditions depend on the day/night loads pattern, weather conditions, soil moisture. Moreover, the analogical experiment for small-scaled wind turbine blade destruction is analyzed to determine the maximum penetration value of mechanical power in order to retrieve an optimal angular velocity which resulting in provides a possible maximum power to loads. At the same time, emergency break is operated when angular velocity reaches to critical speed to avoid destruction. In the simulation, we collected the real load data from a mango farm in Okinawa prefecture in Japan. The results were analyzed through simulations for the different wind conditions. In the end of simulation, we could verify that either Maximum Power Point and emergency control are activated correspondingly.

### 1. Introduction

As a renewable energy generating equipment, small-scaled wind turbines have rising demand to fulfill the modern energy requirements. The small-scaled wind turbines have several merits such as less area needed for setting up the entire system, low cost maintenance and easy to assemble which are advantages for household use [1]. Comparing to the conventional large-scale wind turbines, small-scaled wind turbines produces small amount of power while large scale wind turbines produces maximum 8MW power. Therefore small-scaled wind turbines are suitable for household and small-scaled green houses and farms.

To supply a stable electricity output, small-scaled wind turbines uses several control methods such as pitch control [2-3], yaw control, brake control [4-6]. Considering brake control, it is mainly focus on brake pad-based system for the safety of the wind turbine [7] while contactless brake systems related research is ongoing currently. In [8] we proposed to realize a contactless angular velocity in emergency case by eddy current method,

however, the optimal control was not concerned. In this research, we are focusing to enhance the power through maximum power point tracking (MPPT) using eddy current brake system and eddy current brake system for emergency rotational control.

Since the wind speed stability is always unpredictable, it is a difficult task to generate the maximum power continuously and keep the generated power in a stable state. In this regard MPPT control can be considered as an important controlling method to generate the maximum power. Many types of MPPT control methods are used for generating maximum power such as using fuzzy logic controllers [9], adaptive control [10], etc. In this research, we use the eddy current brake as the main controller tool for the MPPT control. Here, MPPT operation is to convey the maximum possible power from turbine to generator. Therefore, this operation contains power enhancement characteristics. Meanwhile, emergency break is implemented as well to avoid any destruction or malfunction of entire system since small-scaled wind turbine's angular velocity rise up drastically during the high wind penetration.

\*Corresponding Author: Shiro Tamaki, shiro@ie.u-ryukyu.ac.jp

In this study, the mango green house is a mango cultivating farm where the location is in Okinawa, Japan. The mango plants are required to be grown in a controlled environment in order to make best condition for cultivating. Also, the conditions are monitored using Internet of Things (IoT) technology. Therefore, the farmer can analyze the real time mango plants conditions such as its soil moisture, farm temperature, CO<sub>2</sub> conditions, light conditions and humidity.

In this paper, first we explain about the structure of the proposed system including proposed brake system, control diagram, simulation block diagram. Next, we explain the mathematical modeling of the system. Afterwards, we describe the simulation and results. Finally, we describe the conclusion of our research. This research is an extended work of [11].

### 2. Blade destruction analysis

This section explains the analogical experiment of the blade destruction. We performed this experiment to analyze the blade destruction for different penetration values. In other words, we did this experiment to decide the safety margins of the small-scaled wind turbine. In Figure 1, the experimented broken small-scaled wind turbine blade is displayed. In the following equation (1)  $F_{max}$  means the maximum allowable force on the wind turbine blade and  $m$  means the mass of the turbine's blades. Also,  $\omega_{max}$  represents the blade destruction angular velocity. Table 1 shows the specification of the blade. The material of the blade is made using Fiber Reinforced Plastic (FRP).



Figure 1. Broken wind turbine blade

$$F_{max} = m r \omega_{max}^2 \tag{1}$$

Table 1. Specification of the blade

Parameter	Value	Unit
$F_{max}$	4.5	kN
$m$	2.57	kg
$r$	2.5	m
$\omega_{max}$	26.45	rad/s

In Figure 2, the A point stands for the maximum force safety margin. If the penetration exceeds the point A then the turbine's blades start to brake. When the penetration value reaches nearly 28kN (point D), the turbine's blades were destroyed. Therefore, to

protect the turbine, we put c as the primary safety margin which has penetration value of 2.5kN. B means the secondary safety margin that 4kN is penetrated. For secondary safety margin, the angular velocity is 25rad/s and angular velocity for primary safety margin is 20rad/s. Hence, for the simulation purpose, we set the primary safety margin angular velocity  $\omega_{lim}$  as follows:

$$\omega_{lim} = 20\text{rad/s.}$$

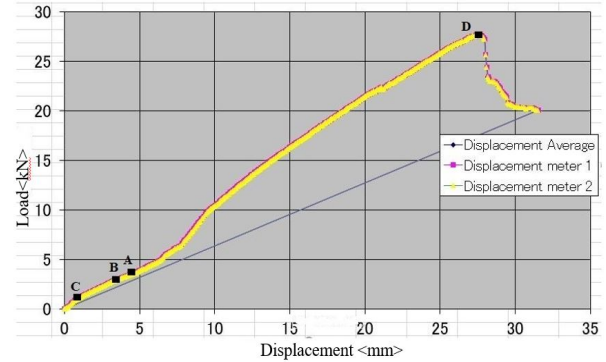


Figure 2. Blade destruction chart

### 3. Structure of the system

It is necessary to consider the electrical and mechanical aspect of the proposed system structure. For the mechanical section, we consider the brake system controlling structure. Here, the emergency rotation control eddy current brake system and MPPT control eddy current brake works independently which is illustrated in Figure 3. They always control the reference value of the angular velocity  $\omega_{ref}$  and the angular velocity output  $\omega$ . When the controller receives the angular velocity of the turbine, controller will send its signal to two servo motors which operates as emergency rotation control and MPPT control unit to place the magnets nearby the copper plates to control the rotation.

When the wind turbine rotates, the signal controller will receive the angular velocity measurement and it will send the signals to the emergency rotation control eddy current brake and MPPT control eddy current brake. According to the current rotation situation, above mentioned two brake systems will control the  $f_{emergency(\omega)}$  which corresponds to emergency rotation control eddy current brake and  $f_{MPPT(\omega)}$ . Here,  $f_{MPPT(\omega)}$  corresponds to MPPT control eddy current brake. This process is illustrated in Figure 4.

Figure 5, shows the entire block diagram of the small-scaled wind turbine system for realization in hardware. Here,  $1/s$  symbol represent the integrator symbol. For this system, the generator is considered as DC generator. Figure 6, displays the equivalent circuit of DC generator including the protection diode, load and battery.

Regarding the electrical aspect of the proposed system, important conditions are explained as follows. When the battery charge becomes low, the additional current from the generator will charge the battery. In the other hand, if the generator unable to supply demanded current, then battery will supply the required current to the load. If the battery is full and the required power is supplied to the load, then the turbine rotation is reduced using the eddy current brake system.

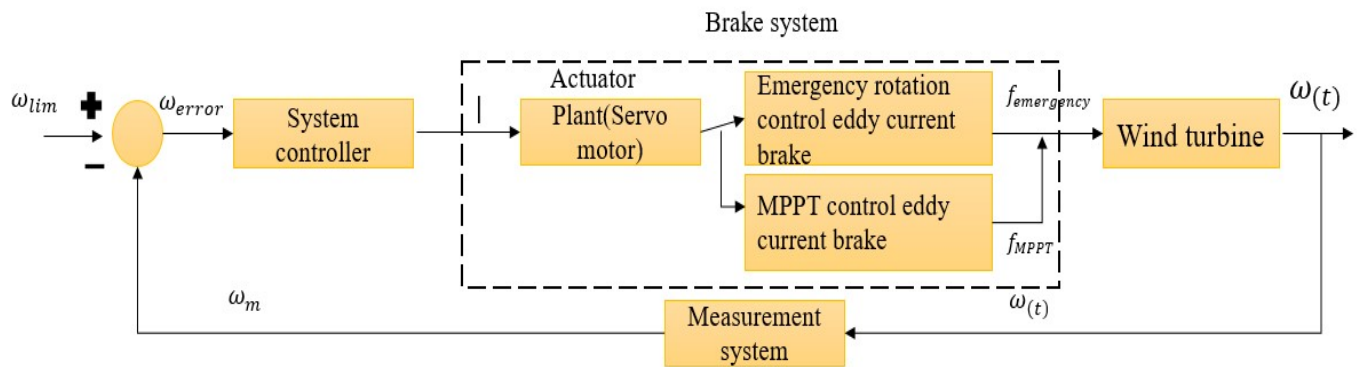
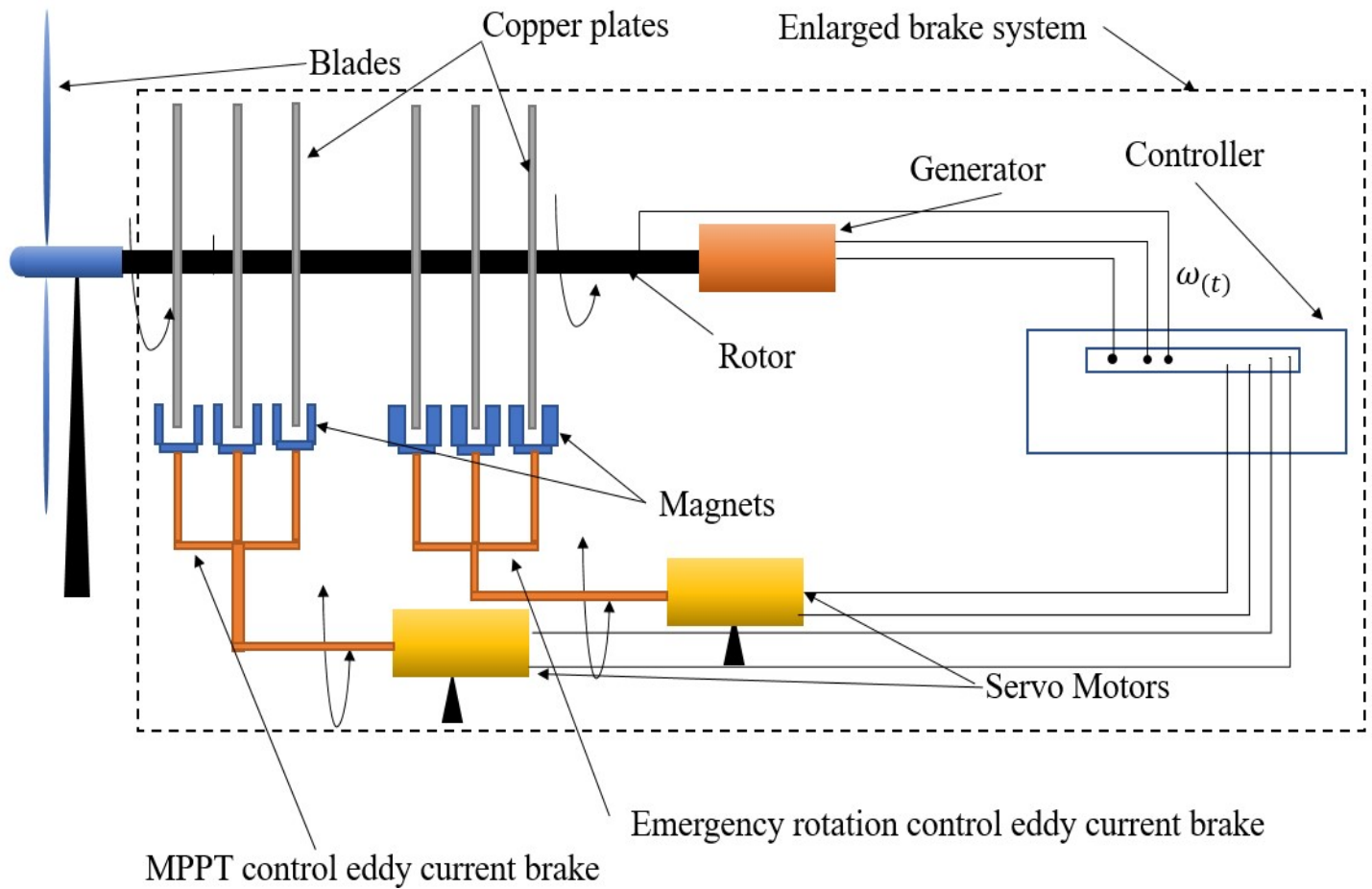


Figure 4. Control diagram



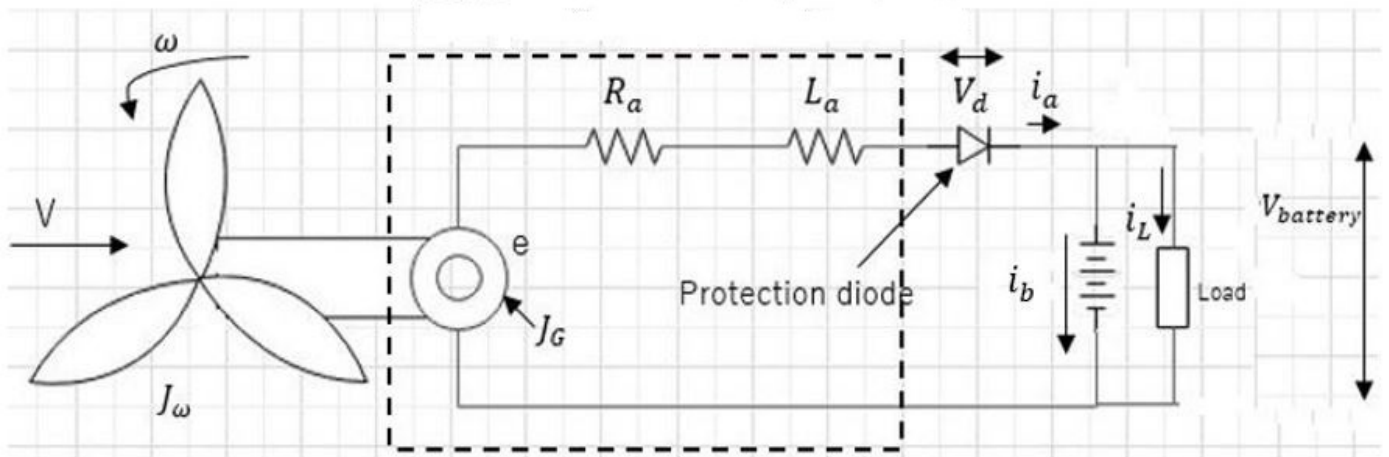
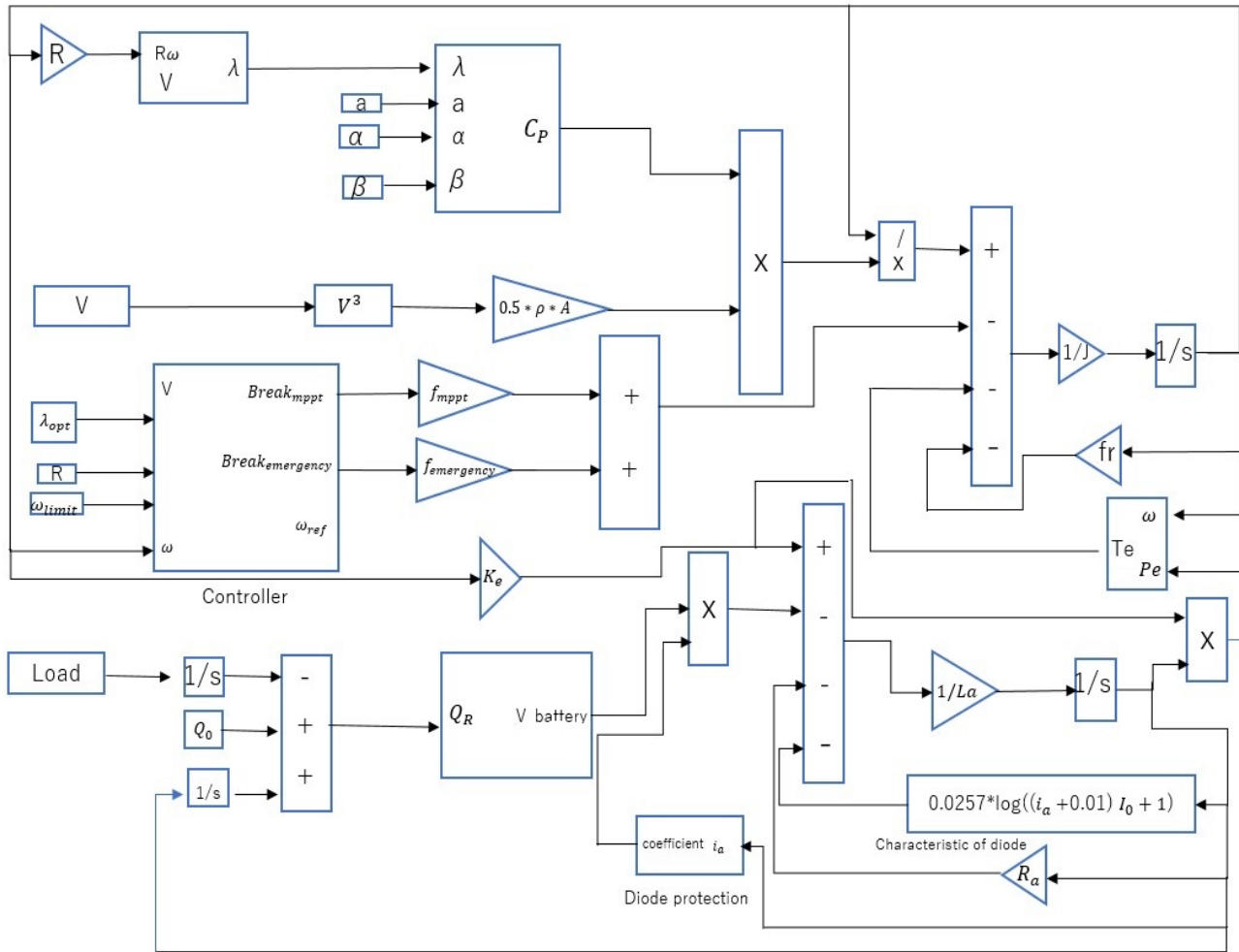


Figure 6. DC generator equivalent circuit diagram

Electromagnetic brake systems can be seen in modern world applications such as Maglev train brake system, gym instruments, and elevators, etc. Unlike the non-contactless braking operation, eddy current brake system can be considered as an efficient brake system [12].

Here, let us describe the eddy break system specifically. When the rotor shaft rotates in  $\omega$  angular velocity, the magnetic flux between the two magnets are changing due to the interaction with the copper plate. Due to this magnetic flux change, eddy current  $i_{eddy}$  will induce on the copper plate according to the faraday law of induction [13]. Due to the induced current, there will be

opposing force or in the rotational case the torque generate against the rotation. This torque will act as the eddy current brake torque. In Figure 7,  $q_m$  means the magnitude of magnetic charge on magnetic poles and  $q_c$  means magnitude of induced magnetic charge on copper plate.

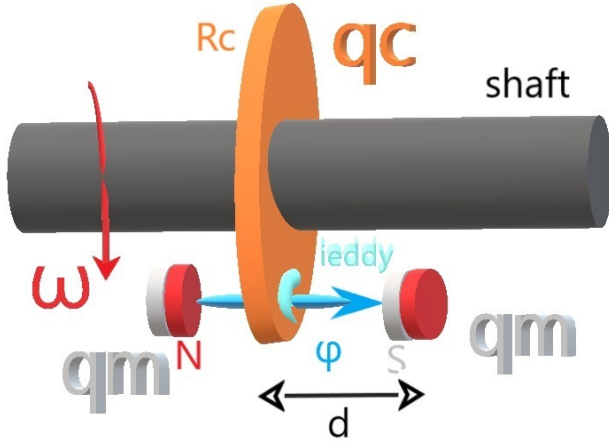


Figure 7. Eddy current brake

$$\varepsilon = - \frac{d\phi}{dt} \quad (2)$$

where,  $\varepsilon$  stands for electromotive force and  $\phi$  for magnetic flux.

Here,  $R_c$  is the resistance of the copper plate and  $I_{eddy}$  is the induced eddy current on the copper plate. The relation with the  $R_c$  and  $I_{eddy}$  is shown in (3), which is based on the ohm's law,

$$\varepsilon = R_c I_{eddy} \quad (3)$$

From above equations, we can write down the eddy current as follows.

$$I_{eddy} = - \frac{1}{R_c} \frac{d\phi}{dt} \quad (4)$$

Due to occurrence of above current, the break system is activated. Next section describes the dynamics of whole system.

#### 4. Mathematical modeling of the system

This section introduces the mathematical modeling of the entire system. Therefore, it is important to consider the dynamics representation of mechanical and electrical part of the proposed system. The parameters were used for the mechanical section is displayed in table 2 and the electrical section is displayed in table 3 along with combined electrical and mechanical equations.

Equation (5) displays the summation of the inertia moments of the wind turbine and blades

$$J = J_\omega + J_G \quad (5)$$

The conventional mechanical and electrical dynamics of the system displayed in (6)

$$J \frac{d\omega(t)}{dt} = \frac{1}{2} C_T \rho A R V_{(t)}^2 - f_r \omega_{(t)} - T_e \quad (6)$$

Table 2: Mechanical Parameters

Symbol	Name
$J_\omega$	Inertia moment of the blade
$J_G$	Inertia moment of the generator
$\omega$	Angular Velocity
$\rho$	Density of atmosphere
$A$	Area swept by blade
$R$	Length of blade
$V_{(t)}$	Speed of wind
$f_r$	Axis friction factor
$\mu$	Permeability
$q_m$	Magnitude of the magnetic field of the magnet
$q_c$	Magnitude of the magnetic field of the copper plate
$d$	Distance between copper plate and magnet
$a_{adj}$	Decision functional adjustment
$\omega_{ref}$	Reference value of angular velocity
$\omega_{lim}$	Allowed maximum angular velocity
$\omega_{stall}$	Stall controlled angular velocity
$I_{eddy}$	Induced eddy current
$C_p$	Power coefficient
$C_t$	Torque coefficient
$k$	Boltzmann's constant
$f_{MPPT}$	Induced eddy current brake torque for MPPT control
$f_{em}$	Induced eddy current brake torque for emergency brake

This formula stands for controlling the brake plate rotation in a contact manner. However, in [8] we have proposed for contactless break system that is shown as follows.

$$J \frac{d\omega(t)}{dt} = \frac{1}{2} C_T \rho A R V_{(t)}^2 - f_r \omega_{(t)} - f_m f(\omega_{(t)}) - T_e \quad (6-1)$$

where,  $f_m$  is the eddy current brake torque in following equation and  $f(.)$  is the decision function.

$$f_m = \mu \frac{2q_m q_c}{4\pi d^2} \quad (7)$$

Here,  $\mu$  stands for permeability of the space.  $q_m$  and  $q_c$  stands for magnitude of the magnetic field of the magnet and copper plate respectively. This formula is based on gilbert model magnetic force [14].

Next, we are going to model the electrical part of the system. In the following table the parameters are described.

Generated power is displayed in (8).

$$P_e = \varepsilon i_a \quad (8)$$

Electrical torque of the generator displayed in (9)

$$T_e = \frac{P_e}{\omega} \quad (9)$$

where,  $e_a$  stands for induced voltage and  $i_a$  for armature current. Equation (10) displays the battery electrical charge dynamics

$$Q_R = Q_0 + \int i_a dt - \int i_L dt \quad (10)$$

As it is shown in above equation the remaining amount of battery capacity depends on the accumulation of consuming current and armature current that provided from turbine and battery

initial capacity. Adversely, the current exchange among battery, turbine and loads can be retrieved by derivative of equation (10) which is

$$i_R = i_a - i_L \quad (11)$$

Table 3: Electrical Parameters

Symbol	Name
$i_a$	Output current of generator (armature current)
$i_b$	Charging current of the battery
$i_L$	Loads current
$\varepsilon$	Induced emf
$V_L$	Loads voltage
$V_{Battery}$	Battery Voltage
$Q_0$	Initial condition of the battery capacity
$I_0$	Saturated current
$I_d$	Flowing current
$V_d$	Voltage drop of diode
$I_s$	Saturated current
$q$	Electron charge
$T$	Temperature
$K$	Boltzmann constant

Therefore, whenever equation (11) becomes zero, the consuming and generating currents are balanced or if  $i_R$  is greater than zero it is both charging mode to battery and providing currents to loads, simultaneously. Adversely, if  $i_R$  is less than zero then the battery only provides the power to loads that is discharging mode.

From equivalent circuit of DC generator in Figure 6, the electrical dynamics of the system is obtained as follows.

$$\varepsilon - R_a i_a - L \frac{di_a}{dt} - V_d = V_L \quad (12)$$

The purpose of setting diode in the output terminal of dc generator is to avoid being motor mode for dc generator when there is no wind penetration since battery is connected to generator's terminal. Thus, diode characteristic should be considered as well. The drop voltage of diode is shown in following equation.

$$V_d = \frac{KT}{q} \ln\left(\frac{I_d}{I_s} + 1\right) \quad (13)$$

Where,  $I_d$  corresponds to armature currents as following expression (14).

$$I_d = I_s \left( e^{\frac{qV_d}{KT}} - 1 \right) \quad (14)$$

Next, the induced electromotive force (emf) of the generator is displayed in (15).

$$\varepsilon = K_e \omega \quad (15)$$

Here,  $K_e$  stands for Induced emf constant.

Up to here, the electrical model were described. From next equation the mechanical parts including the break terms are expressed.

The following equation shows the mechanical torque.

$$T_m = \frac{P_m}{\omega} \quad (16)$$

Below equation is based on the conventional break method, however,  $C_p$  is variable with respect to tip speed ratio.

$$J \frac{d\omega(t)}{dt} = \frac{C_p(\lambda) \rho A V_w^2}{2\omega} - f_r \omega(t) - f_m f(\omega(t)) - T_e \quad (17)$$

Then we can reform the above equation for proposed method by adding the eddy current brake for emergency brake and MPPT control that is shown as (18).

$$J \frac{d\omega(t)}{dt} = \frac{C_p(\lambda) \rho A V_w^2}{2\omega} - f_r \omega(t) - T_e - (f_{mp} f_{MPPT(\omega)} + f_{em} f_{emergency(\omega)}) \quad (18)$$

In above equation,  $f_{MPPT}$  and  $f_{emergency}$  has different values and  $f_{em}$  value is larger than  $f_{MPPT}$  since  $f_{emergency}$  requires more braking force to control the over-rotation of the turbine than the force use to realize MPPT.

Equation 19 task is to work as a switch to keep the output power in a maximum state which means keep the angular velocity  $\omega$  in the following condition: ( $\omega < \omega_{ref}$ ) that is in the MPPT operation status. However, if  $\omega$  exceed  $\omega_{ref}$  value ( $\omega > \omega_{ref}$ ), then the output power is not in a Maximum power state. Nevertheless, still the system is generating power since  $\omega$  value isn't reached to  $\omega_{lim}$  yet which is the safety margin of the wind turbine. If this situation going to happen then the emergency brake will start to reduce the  $\omega$  value to keep the state as  $\omega < \omega_{lim}$  condition. Therefore, equation 20 [15] task is to work as a switch to determine ( $\omega < \omega_{lim}$ ) condition.

$$f_{MPPT(\omega)} = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(\alpha_{adj}(\omega - \omega_{ref})) \quad (19)$$

$$f_{emergency(\omega)} = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(\alpha_{adj}(\omega - \omega_{lim})) \quad (20)$$

The reference value of the angular velocity is shown in (21). Here the  $\lambda_{opt}$  means the optimum value of tip speed ratio.

$$\omega_{ref} = V \frac{\lambda_{opt}}{R} \quad (21)$$

$\omega_{ref}$ , value is a function of  $\lambda_{opt}$  and  $V$ . Since this is a small wind turbine,  $\lambda_{opt}$  value require small value to obtain  $C_p^{opt}$  value. Due to the compatibility of small-scaled turbine, we designed the mathematical expression of the  $C_p$  as mentioned below in (22) and  $C_p$ - $\lambda$  relationship shown in Figure 8.

$$C_p = -a(\lambda - \alpha)(\lambda - \beta) \quad (22)$$

Where, the condition of coefficients is  $a > 0$  and  $\beta > \alpha$ .

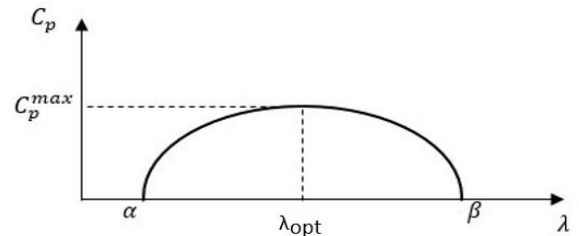


Figure 8.  $C_p$ - $\lambda$  curve

Optimum value of wind turbine power factor is shown in (23).

$$C_p^{max} = \frac{\alpha + \beta}{2} \quad (23)$$

As it is shown in Figure 7 the output terminal of generator is connected to a battery and battery is connected to the loads in parallel. Therefore, in order to verify the battery voltage, we have designed a battery capacity charging in which the battery voltage can be obtained by the Equation (24) represents the characteristic of the accumulated consumed current.

$$V_{Battery} = \frac{V_{max} - V_{min}}{0.5 - \frac{1}{\pi} \tan^{-1}(20 \times (-5))} \left( \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(I(Q_R - 5)) \right) \quad (24)$$

Notice that the load voltage and battery voltage is identical since as it has been mentioned they are connected in parallel.

### 5. Simulation and results

In this section, we conduct the simulation in two steps. First, we simulated the entire system for different average wind velocities from 7m/s, 14m/s, 21m/s and 28m/s, respectively. Then plotted the stall control charts for above velocities. We did this step to analyze the wind turbines stall control behavior for different wind velocities. Afterwards, we plotted the Mechanical power (Pm) and power coefficient of the system for different conditions. For the second step, we simulated the entire system for calm and storm days wind conditions for one day and analyzed the behavior of outputs of the system.

#### 5.1. Simulation conditions

This section explains the conditions of the simulation and system we used in order to conduct the simulation. The load data for the system is applied using real mango greenhouse system in Okinawa [16]. Load pattern is displayed in Figure 9. This load consists of Light emitting diode (LED) lights, Compact Fluorescent Light (CFL) lights, Electrical Fans. According to the Figure 9 the highest load value is 12.8A. Here, the sample time for the simulation is 0.01s. MATLAB/Simulink software platform is used for the simulation. For the simplicity, simulation time is set to be 100s from Figure 10 to Figure 21. From Figure 22 to Figure 33, simulation time is one day long. Values of the parameters for simulation are displayed specifically in Table 4.

Table 4: Parameters of simulation

Symbol	Value	Unit
$\omega$	Variable	$rad\ s^{-1}$
$\rho$	1.225	$kg\ m^{-3}$
$A$	19.625	$m^2$
$R$	2.5	m
$V(t)$	Variable	$m\ s^{-1}$
$f_r$	0.1	$kg\ m^2\ s^{-1}$
$f_{mppt}$	2200	$kg\ m\ s^{-2}$
$f_{emergency}$	5800	$kg\ m\ s^{-2}$
$\mu$	$4\pi \times 10^{-7}$	$T\ m\ A^{-1}$
$q_m$	$62.5 \times 10^9$	A.m
$q_c$	$0.1 \times 10^{-4}$	A.m
$d$	0.05	m

$a_{adj}$	10	
$\omega_{lim}$	20	$rad\ s^{-1}$
$\omega_{ref}$	Variable	$rad\ s^{-1}$
$J_\omega$	1.1	$kg\ m^2$
$J_g$	0.1	$kg\ m^2$
$L_a$	2.87	mH
$R_a$	4.2	$\Omega$
$q$	$1.602 \times 10^{-19}$	C
$k$	$1.380 \times 10^{-23}$	J/K
$I_0$	$1 \times 10^{-22}$	A
$K_e$	2.45	V.S
$Q_0$	60	A.h
$\alpha$	1	
$b$	3.5	

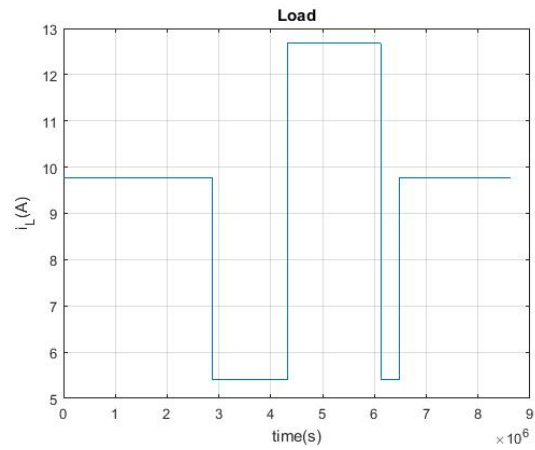


Figure 9. Load

#### 5.2. Results

From figure 10 to figure 13, they display the stall control chart for 7m/s, 14m/s, 21m/s, 28m/s, respectively.

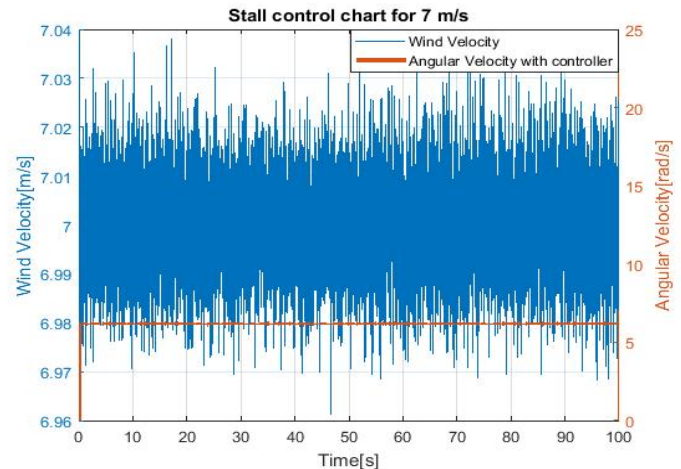


Figure 10. Stall control chart for 7 m/s



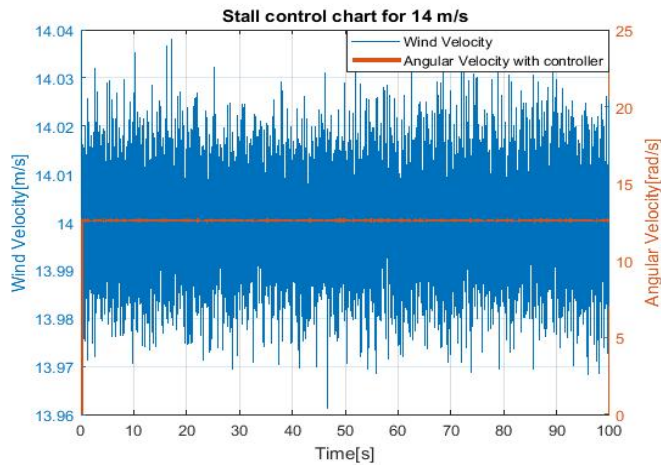


Figure 11. Stall control chart for 14 m/s

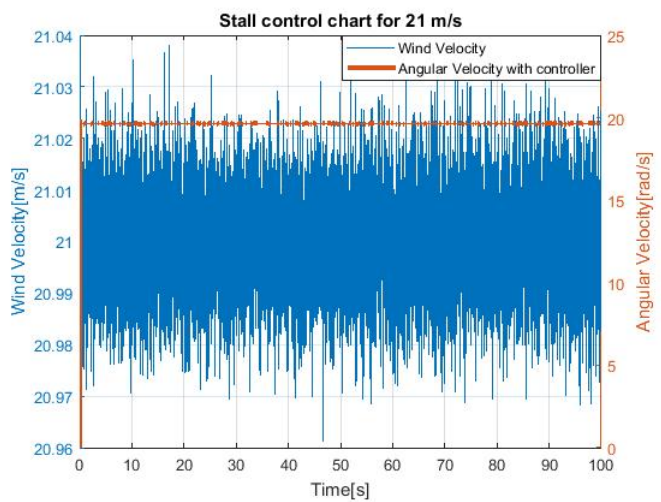


Figure 12. Stall control chart for 21 m/s

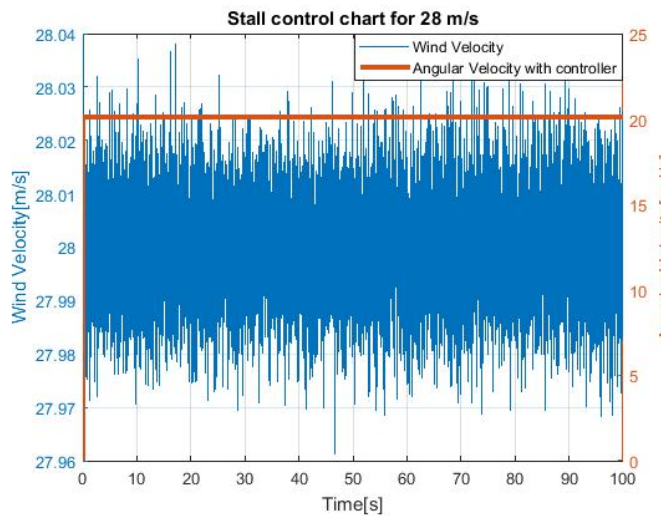


Figure 13. Stall control chart for 28 m/s

According to above stall control figures, the wind turbine angular velocity emergency control is valid up to 28m/s velocity. Therefore, the eddy current brake system can control the over-rotation of the wind turbine up to 28m/s.

Table 5 shows the  $\omega_{ref}$  which is optimal angular velocity to realize the maximum power operation and  $\omega_{stall}$  is the actual output of angular velocity.

Table 5:  $\omega_{ref}$  and  $\omega_{stall}$  comparison

Input velocity	$\omega_{ref}$	$\omega_{stall}$
7	6.299	6.1849
14	12.5999	12.5994
21	18.8999	19.6377
28	25.199	20.094

As it is obvious in the above table, the output angular velocity for input wind is lower than the optimal angular velocity since in high wind penetration emergency break is activated. Specifically, comparing to the  $\omega_{stall}$ ,  $\omega_{ref}$  value is becoming lower when wind velocity increase. That means when the input wind velocity is 28m/s the  $\omega_{ref}$  value is 25.199 the MPPT control is stopped because  $\omega > \omega_{ref}$ . However wind turbine is still operating using the emergency eddy current brake since  $\omega_{stall}$  is nearly the value of  $\omega_{lim}$ . Next, from Figure 14 to Figure 17, The instantaneous angular velocity deviation  $\Delta\omega = \omega_{stall} - \omega_{ref}$  show their behavior for 7m/s, 14m/s, 21m/s, 28m/s average wind velocities, respectively.

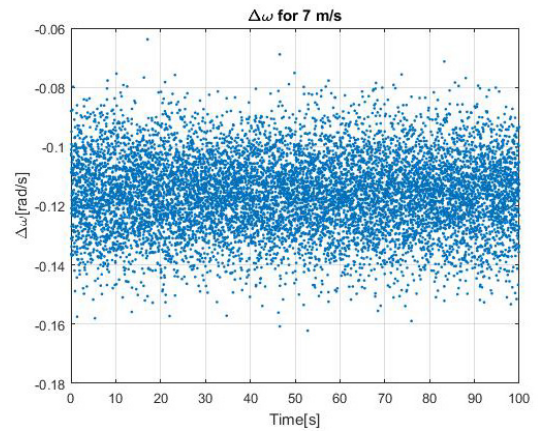


Figure 14.  $\Delta\omega$  for 7 m/s

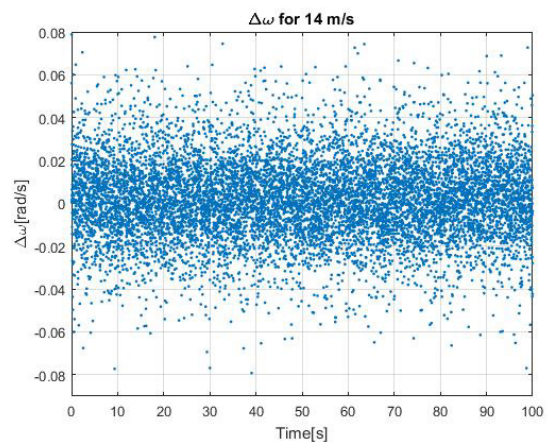


Figure 15.  $\Delta\omega$  for 14 m/s



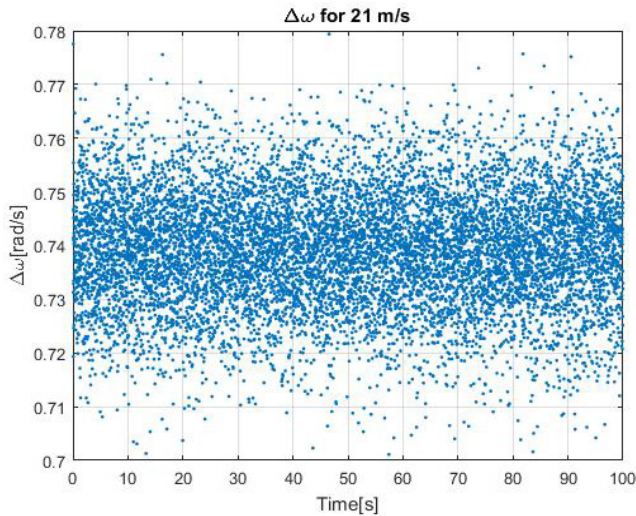


Figure 16.  $\Delta\omega$  for 21 m/s

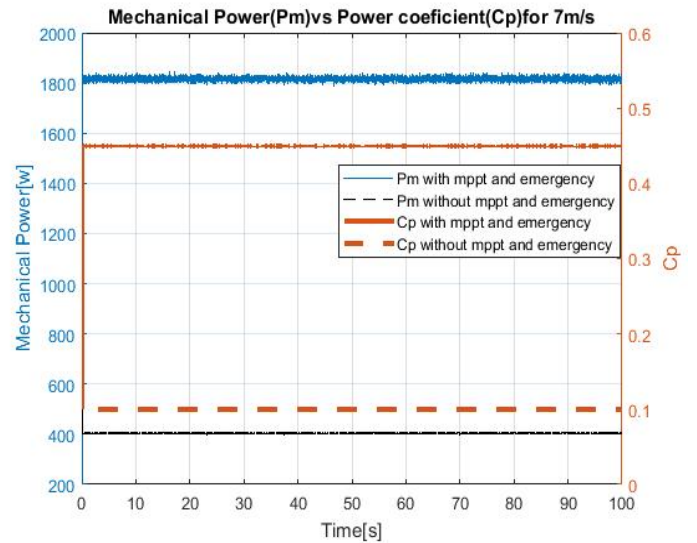


Figure 18. Pm vs Cp for 7m/s

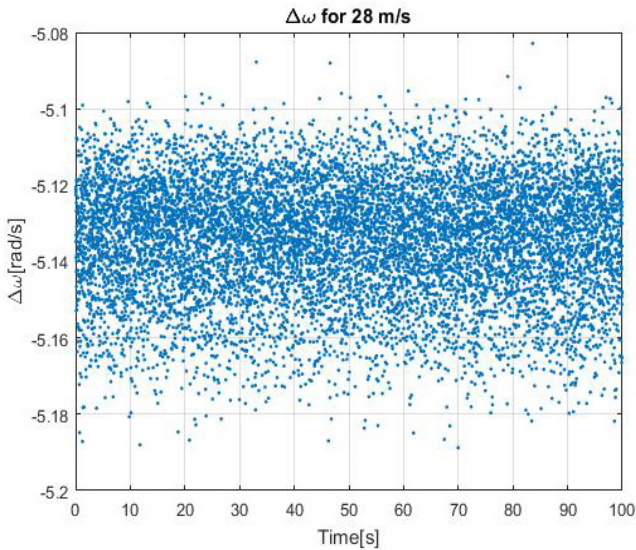


Figure 17.  $\Delta\omega$  for 28 m/s

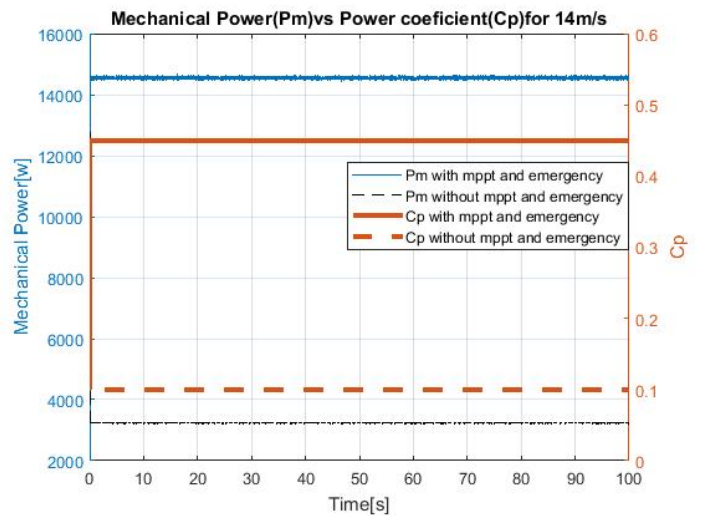


Figure 19. Pm vs Cp for 14m/s

The mechanical power comparison with and without MPPT and emergency brake and Cp behavior comparison with and without the MPPT and emergency brake are shown from Figure 18 to Figure 21.

According to the Figure 18, when the average wind velocity is 7m/s the mechanical power with the MPPT and emergency brake has higher value comparing to without the MPPT and emergency brake. As well, for Cp, when the MPPT and emergency brake is triggered it contains higher value. But, when Cp is without the control of MPPT and emergency brake, the value is low.

In Figure 19, Figure 20 and Figure 21 which are correspond to average wind velocity of 14m/s, 21m/s and 28m/s, they have the same comparison results as mentioned in above for 7m/s. Therefore, it is obvious that when the system is under the controlled state for MPPT and over-rotation, the mechanical power of the wind turbine and the power factor (Cp) have their maximum values.

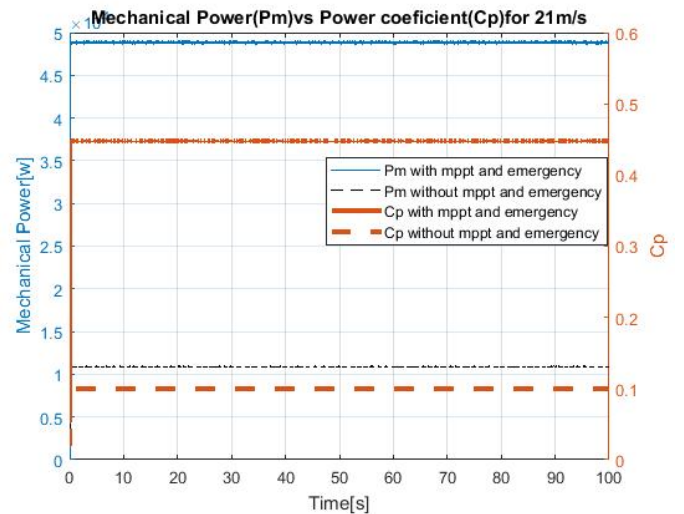


Figure 20. Pm vs Cp for 21m/s

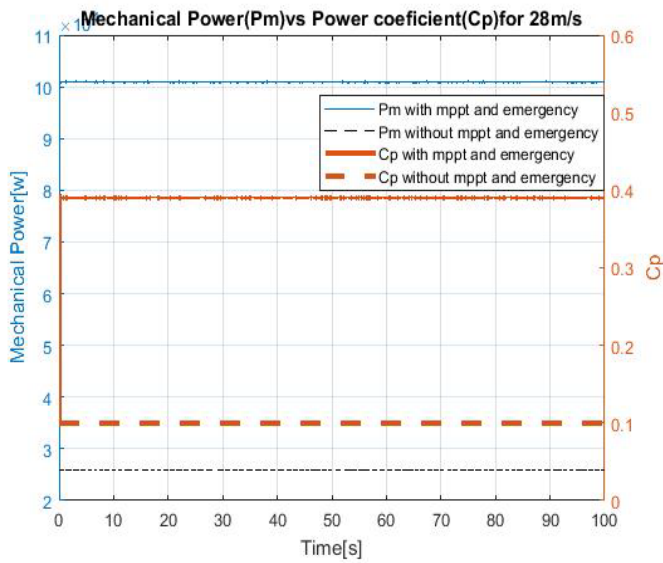


Figure 21. Pm vs Cp for 28m/s

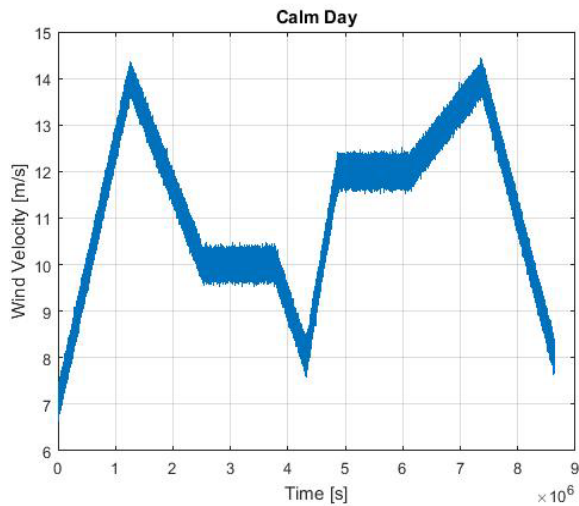


Figure 22. Calm day wind pattern

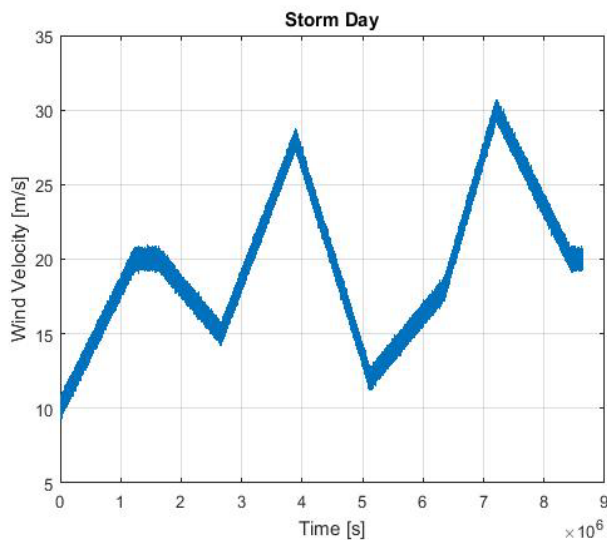


Figure 23. Storm day wind pattern

We have artificially created two wind patterns for the simulation process. First, we have done the simulation for calm day and then for the storm day condition. For calm day and storm day wind patterns are shown in Figure 22 and Figure 23. For the first phase of calm day the velocity gradually increases from 7m/s to 14m/s. Afterward, it decreases gradually from 14m/s to 10m/s. Then the 10m/s wind velocity maintain stable for few hours. Afterwards the velocity will decrease until 8m/s. Then, it will increase from 8m/s to 12m/s. Again, it will increase from 12m/s to 14m/s gradually. Finally, the velocity will decrease from 14m/s to 8m/s. For calm day the highest wind speed is approximately 14.3m/s and storm day the highest wind speed is approximately 30m/s. Lowest wind speed for calm day and storm day are 7m/s and 10m/s.

Next, Figure 24 and Figure 25 shows the angular velocity deviation  $\Delta\omega$  values which changes with respect to the wind condition. There are some peaks occurred according to the Figure 22. These peaks happen due to the emergency braking for sudden wind condition changes. Here, the  $\Delta\omega$  is nearly zero means the wind turbine produces its maximum power.

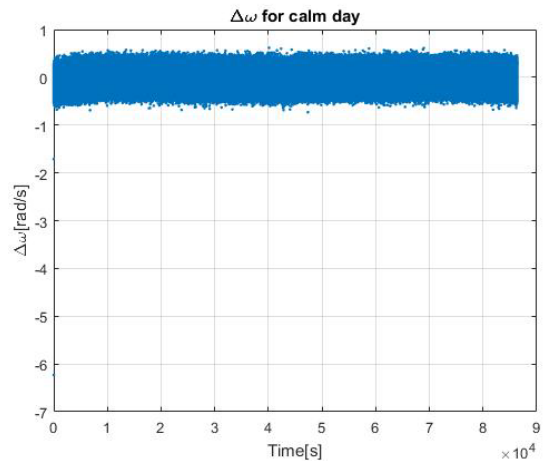


Figure 24.  $\Delta\omega$  for calm day

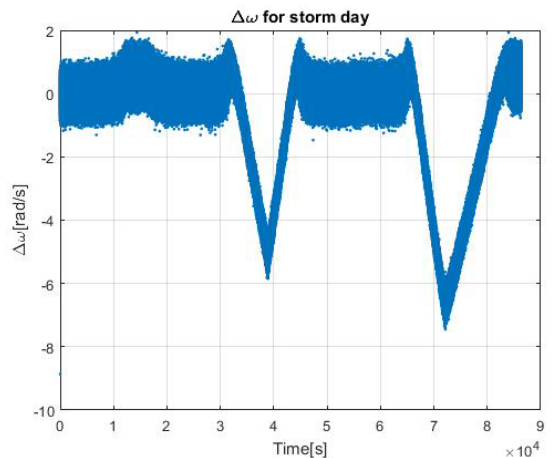


Figure 25.  $\Delta\omega$  for storm day

In calm day, the angular velocity does not achieve to critical value which is 20 rad/s. Nevertheless, according to angular velocity deviation in Figure 24, it is attempting to reach the angular velocity that operates at maximum power point. Same to Storm day, angular velocity tries to operate at its maximum power point.

However, as it can be seen Figure 25, the deviation of angular velocity has negative two peaks for some determined period. Correspondingly, it is obvious in Figure 27 that during the peaks of angular velocity deviations the output of angular velocity is suppress by 20 rad/s. Thus, the emergency brake is also activated in storm situation.

The angular velocity always kept under  $\omega_{lim} = 20$  for both calm day and storm day stall control patterns which are shown in Figure 26 and Figure 27.

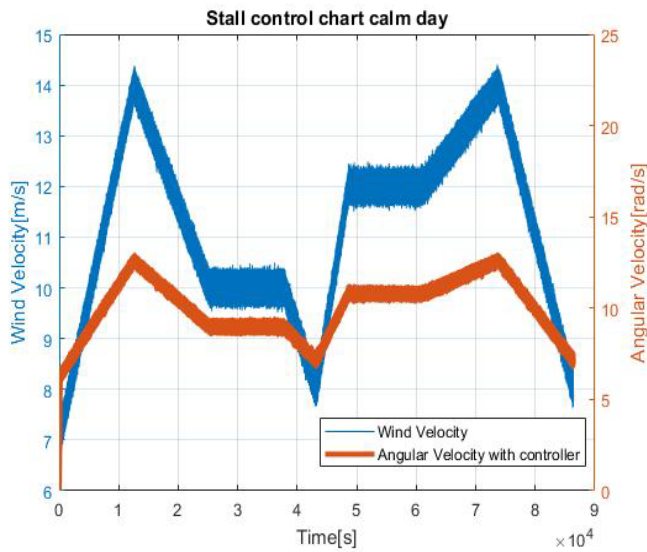


Figure 26. Stall control chart for calm day

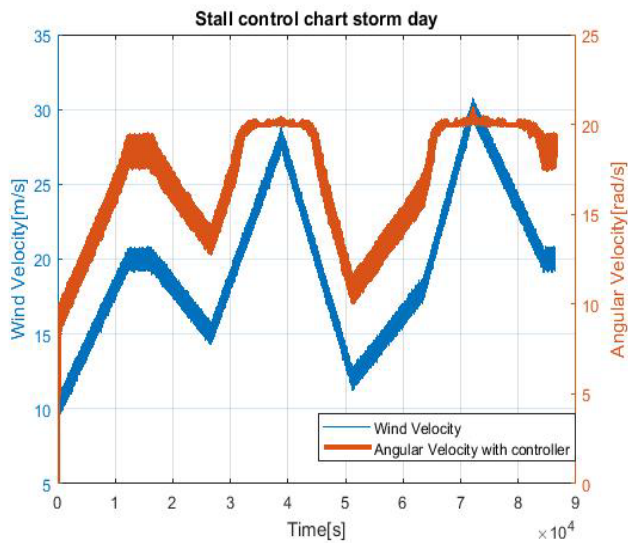


Figure 27. Stall control chart for storm day

Figure 28 and Figure 29 displays the Mechanical power (Pm), Electrical power (Pe), Electrical torque (Te) for calm day and Pe, Pm, Te, Tm for storm day. According to those figures the Pe, Pm, Te, Tm have higher values comparing to the calm day for emergency brake and MPPT brake activated state. Therefore, the small wind turbine is capable of generating high power even in a storm day.

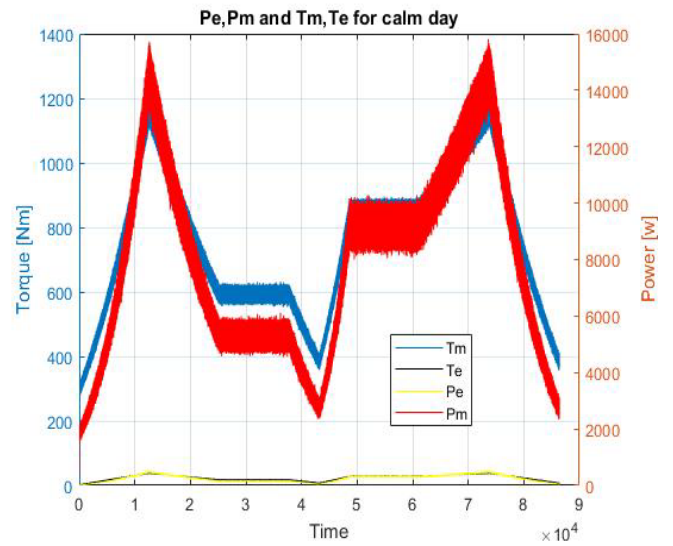


Figure 28. Pe, Pm, Tm and Te for calm day

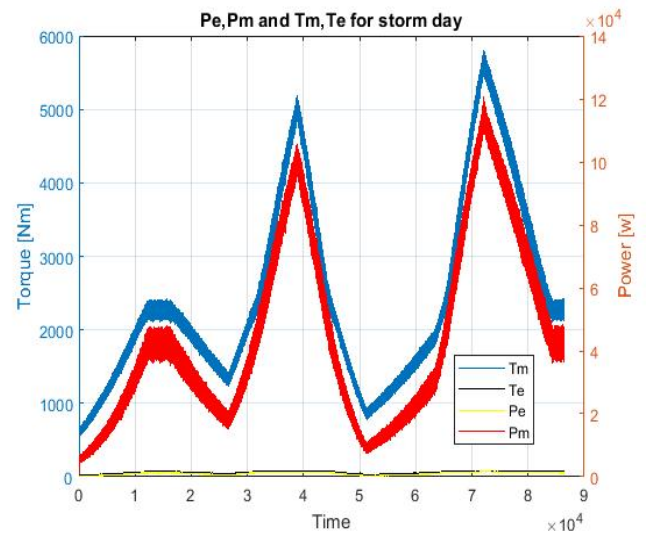


Figure 29. Pe, Pm, Tm and Te for storm day

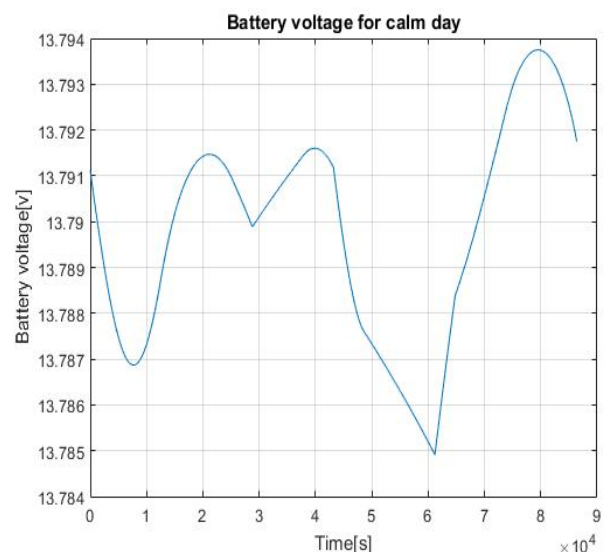


Figure 30. Battery voltage for calm day



According to Figure 30 and Figure 31, the battery voltage for storm day has good performance comparing to the battery voltage for calm day. Thus, for both cases we can say it has sufficient amount of voltage.

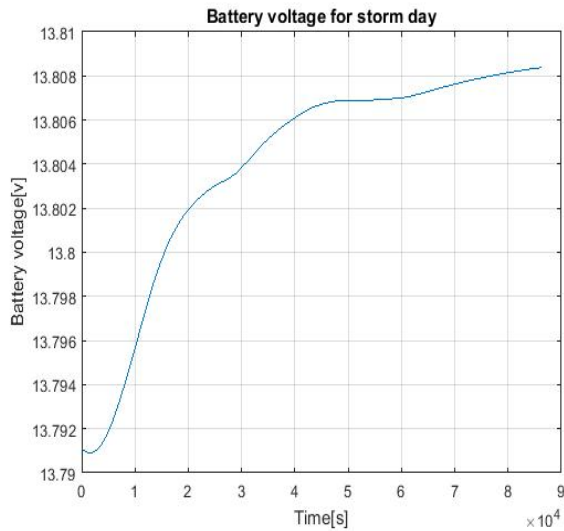


Figure 31. Battery voltage for storm day

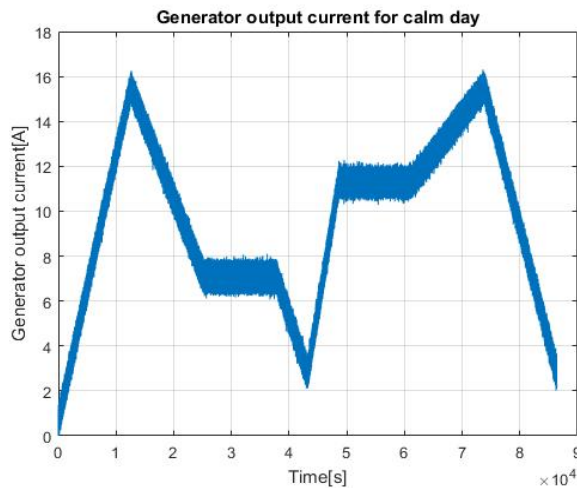


Figure 32. Generator output current for calm day

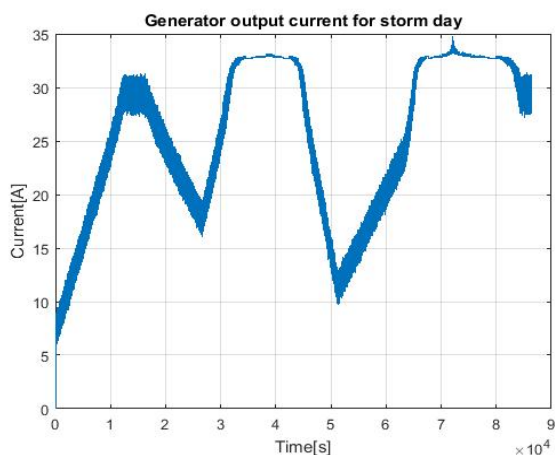


Figure 33. Generator output current for storm day

Figure 32 and Figure 33 shows the generator output current for calm day and storm day. For storm day, the generator output current has significant increase comparing to the calm day current in the MPPT and emergency brake-controlled condition. Nevertheless, even in calm day sufficient capacity of battery is maintained which is sustainable for providing the power to greenhouse environmental loads. Thus, we can conclude that efficiency of power extraction is improved with cooperative of the emergency operation.

## 6. Conclusion

We have conducted the simulation for different wind condition and analyzed the behavior of the system when the system is in control of MPPT and Eddy current brake conditions. Moreover, the analogical experiment for the small-scaled wind turbines to decide the maximum penetration value and maximum angular velocity the wind turbine blades can withhold was performed. Therefore, we can conclude that the behavior of the system is in controlled state by using above mentioned control methods. We believe the eddy current brake system can be implemented in a small-scaled wind turbine system for the purpose of MPPT and control the over-rotation by strong wind. As future works, we will add pitch control for establish the system more robust. Also, the real system will be implemented.

## Conflict of Interest

The authors declare no conflict of interest.

## Reference

- [1] J.F. Manwell, J.G. Mcgowan, A.L. Rogers, Wind Energy Explained, John Wiley and Sons, 2009.
- [2] F. Asharif, S. Tamaki, T. Nagado, T. Nagata M.R. Asharif, "Analysis of non-linear adaptive friction and pitch angle control of small-scaled wind turbine system," Future Generation Information Technology FGIT 2011, Control and Automation CA, Springer-Verlag, Lecture Note: Communication in Computer and Information Science, 26-35, December 2011. [https://doi.org/10.1007/978-3-642-26010-0\\_4](https://doi.org/10.1007/978-3-642-26010-0_4)
- [3] F. Alsharif, J. Tamura, S. Tamaki, K. Anupa, M. Futami, "Dynamics modification of passive pitch control for small-scaled wind turbine" Grand Renewable Energy 2018, Yokohama, Japan, June 2018. [https://doi.org/10.24752/gre.1.0\\_147](https://doi.org/10.24752/gre.1.0_147)
- [4] F. Asharif, S. Tamaki, H. Teppei, T. Nagado, T. Nagata, "Feasibility confirmation of angular velocity stall control for small-scaled wind turbine system by phase plane method" IEEK Transactions on Smart Processing and Computing, 2(4), 240-247, 2013.
- [5] F. Asharif, S. Tamaki, T. Nagado, M. R. Alsharif, "Comparison and evaluation of restrain control in wind turbine with various shock absorber considering the time-delay" International Journal of Control and Automation, Vol. 5(3), 111-131, September 2012.
- [6] F. Asharif, M. Futami, S. Tamaki, T. Nagado, K. Asato, "Stability and performance analysis of passive and active stall control of small-scaled wind turbine system by phase plane method" Proceedings of IEEE, International Conference on Intelligent Information and BioMedical Sciences (ICIIBMS 2015), 27-33, OIST, Okinawa, Japan, November 2015. <https://doi.org/10.1109/ICIIBMS.2015.7439483>
- [7] M. Entezami, S. Hillmansen, P. Weston, M. Ph. Papealius "Fault detection and diagnosis within a wind turbine mechanical braking system using condition monitoring" Renewable energy, 47, 175-182, November 2012. <https://doi.org/10.1016/j.renene.2012.04.031>
- [8] A. Koswatta, F. Alsharif, Y. Shiroma, M. Khosravy, S. Tamaki, J. Tamura, "Alternative braking method for small-scaled wind turbines connected DC green house with analogical experiment on blade destruction" Advances in Science, Technology and Engineering Systems Journal, 5(2), 500-511 (2020). <https://doi.org/10.25046/aj050264>
- [9] A. Z. Mohamed, M. N. Eskander, and F. A. Ghali, "Fuzzy logic control based maximum power tracking of a wind energy system," Renew. Energy, 23(2), pp. 235-245, Jun. 2001. [https://doi.org/10.1016/S0960-1481\(00\)00099-9](https://doi.org/10.1016/S0960-1481(00)00099-9)

- [10] H. Zhao, Q. Wu, C. N. Rasmussen, M. Blanke, "Adaptive speed control of a small wind energy conversion system for maximum power point tracking ", *IEEE Trans. Energy Convers.*, 29(3), 576-584, Sep. 2014. <https://doi.org/10.1109/TEC.2014.2312978>
- [11] K. Anupa, F. Asharif, S. Yasushi, K. Ken, S. Tamaki, J. Tamura, "Power analysis of eddy current brake system applied on small-scaled wind turbine for dc greenhouse" International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), Shanghai, China, 2019. <https://doi.org/10.1109/ICIIBMS46890.2019.8991519>
- [12] K.J. Lee, K. Park, "Optimal robust control of a contactless brake system using an eddy current" *Mechatronics*, 9(6), 615-631, 1999. [https://doi.org/10.1016/S0957-4158\(99\)00008-2](https://doi.org/10.1016/S0957-4158(99)00008-2)
- [13] J. D. Kraus, K. R. Carver, *Electromagnetics*, 2nd ed. McGraw Hill, 1973.
- [14] D. Gubbins, E. Herrerobervera, *ENCYCLOPEDIA of GEOMAGNETISM AND PALEOMAGNETISM*, Springer, 2007.
- [15] F. Asharif, S. Tamaki, T. Hirata, T. Nagado, T. Nagata, "Design of electromagnetic and centrifugal force pitch angle stall controller on application to small-scaled wind turbine system" International Symposium on Stochastic Systems Theory and Its Applications, 2013. <https://doi.org/10.5687/sss.2014.231>
- [16] Y. Shiroma, H. Afuso, R. Suwa, A. Kinjo, Y. Tonooka, T. Kaga, Itaru. Nagayama, S. Tamaki, G. Maharjan, "Development of higher yield and high-quality mango production system based on Internet of Things" *IEEJ Transactions on Industry Applications*, 139(2), 166-173, Feb. 2019. <https://doi.org/10.1541/ieejias.139.166>



## Multi-Agent Data Recognition System Based on Received Signal in Antenna on Board Telecom System

Chafaa Hamrouni<sup>1,2,\*</sup>

<sup>1</sup>Department of Computer Sciences, Khurma University College, Taif University, Khurma, 2935, Kingdom of Saudi Arabia

<sup>2</sup>Research Groups on Intelligent Machines Laboratory, National School of Engineering of Sfax (ENIS), Sfax University, Sfax, 3038, Tunisia

### ARTICLE INFO

Article history:

Received: 26 April, 2020

Accepted: 21 June, 2020

Online: 12 July, 2020

Keywords:

Antenna

Recognition

System

Signal

identification system

Switching System

Multi-Micro -strip Antennae

Network

Fuzzy Control System

### ABSTRACT

Information data recognition during traditional operating step in telecommunication system, as their interpretation, presents an important active research field. In this context, we propose as a solution multi-agent data identification system starting with received signal parameters in antenna network connected on board telecom system. Due to the Information Identification Data (IID) response variability that differ from one presented signal being to another, the IID remains difficult to detect and to recognize. In this paper, we presented a various problem related to IID recognition. We successfully developed a multimodal IID recognition based on two different modalities. We identify each hot moment relying on successful IID detected. Proposed solution is based on IID value caused by both information type and the power intensity value.

## 1. Introduction

Different works presented to maximize telecommunication system efficiency [1], such mathematical studies of radio quality monitoring and harmonics for antenna conception. As published references which are important to the present work, the "Petri-net Modeling and Prototype of a Multi-microstrip Antennae Network [2] Fuzzy Controlled System [3] for Pico satellites". Since telecommunication subsystems operate in High band frequency [4], we consider ferrite, as it is suitable magnetic material, to be employed in microwave applications antenna [5]. Second work which describes circular ring-shaped microstrip antenna to be used in ultra large band, another work entitled multi microstrip antenna [6] fuzzy controlled system [7] for ERPSat1. The works of S. Pulinets entitled "Electromagnetic effects in atmosphere, ionosphere and magnetosphere Pulinets", S. Pulinets, second the research work entitled "Energy Minimization for Wireless Video Transmissions With Deadline and Reliability Constraints", Seong-Ping, and third the work of Seong-Ping Chuah; Zhenzhong Chen; Yap-Peng Tan entitled "Energy Minimization for Wireless Video

Transmissions With Deadline and Reliability Constraints". Our object of this research work is, to give the telecommunication system the ability for several information types' identification (data, image, and video).

## 2. Theory and Simulation Work

This work focusses on the recognition aspect of each received signal in antenna. first step, we try to allow the operating system to be adapted to the digitalized signal state of the receiver. The adaptation can be translated into a change in the functioning of the interactive system. Frustration detection system is operated by a serial of alternatives less complicated. The process can be achieved at the interactions offered by the system. In the second step, we would propose a system for a data summarization. Information Identification Systems (IIS) are replicated, improved techniques, with more robust results and complex implementations. Faced with this development of the proposed IIS solutions, our research objective is to suggest a framework with the ability to recognize captured information type through a unifying study of the existing recognition systems, in order to capitalize and compare experiences. Our work therefore falls clearly in the engineering of interactive systems able to recognize the IID. We developed the

\*Chafaa Hamrouni, Taif University-Khurma University College-Department of Sciences Computer, Contact:00966546492338, Email:chafa.hamrouni.tn@ieec.org

conceptual and practical aspects of software design. To understand, compare and unify the experience of IIS operation and grasp the complexity and scalability of systems for received information identification, we propose an architectural conceptual framework based on the principle of functional separation.

In our work, we propose the artificial Intelligence for Multi-Agent Systems [8], it represents a new technology that offer intelligent solutions that meet human needs. Several works in which research consider, for swarms of identical mobile agents with limited communication, sensing and computation capabilities, the distributed simultaneous estimation and formation control problem. In this work, the idea of Multi-Agent System [9] is for to facilitate the Information Identification Data process, detection and recognize operation. In other researches, IID and Multi-Agent System [10] are used to develop a novel scalable algorithm that encodes the formation specifications of the swarm via geometric moment statistics, which are estimated by a distributed scheme with prescribed performance guarantees.

Based on the locally available information, we notice that in Multi-Agent System [11], each agent calculates an estimate of the global formation statistics, which is then employed by its local motion controller, thus creating a feedback interconnection between the estimator and the controller. Moreover, a minimum allowable inter-Multi Agent [12] collision avoidance is achieved.

Among the most important developments in this area, there is the adaptation of different systems to user needs which produce a personalized interaction based on information parameters and IID dimension. IID computing has managed to provide positive results in detecting information type. This dimension is not the only reliable way that can estimate the IID dimension, other indices may also provide this information, such as coded information length, or more precisely the delivered symbols that our decoder detects during system operating step. Based against this index, studies are rare and less conclusive despite the emergence of a variety of application areas among which those including remote educational platforms: educational platforms remotely, videophones. Furthermore tools that can facilitate tasks such as the Kinect sensors introduced by Microsoft are easier to implement without the feeling of being controlled. In this work is presented the architecture of the proposed system to give more precision to estimate the IID type. Below we present in figure .1 the main system operating techniques:

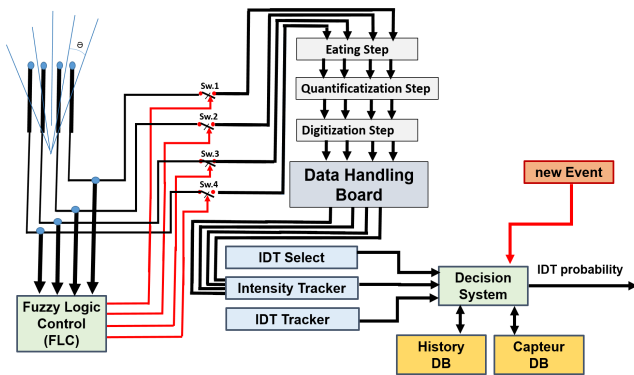


Figure 1: Main System Architecture Proposal

Unlike the classical artificial intelligence, which models the intelligent behaviors of a single entity, distributed artificial intelligence studies intelligent behavior produced by the cooperative activity of several entities generally presented by agents [13]. The evolution of the application areas of artificial intelligence allows increasing the limitations of the classical artificial intelligence approach which relies on the centralization of expertise within a single system. Such limitations appeared to the distributed used method. The target of such method to operate with is to address the short comings of the traditional approach by providing a group of agents which can work and act in a common situation and solve potential conflicts. Three fundamental axes have emerged in the field of artificial intelligence: multi-agent systems [14] distributed resolution of problems, and parallel artificial intelligence.

#### 4. Results and discussion

We introduced fuzzy control, to optimize system function, then, as a second step, we simulate the fuzzy controller and we test it. We select different powers values such power inputs such( $x_0, y_0, z_0$  and  $t_0$ ) where  $x_0=1.045w, y_0=0.967w, z_0=1.016w$  and  $t_0=0.887w$ . and for the Fuzzification step, we have:

Table 1: Selected Fuzzy Intervals

	Very Low	Low	Medium	High	Very High
$X_0=1.045w;$ $\mu(X_0)$	null	null	0.926	0.096	null
$Y_0=0.967w;$ $\mu(Y_0)$	null	null	0.926	0.096	null
$Z_0=1.016w;$ $\mu(Z_0)$	null	null	0.901	0.121	null
$T_0=0.887w;$ $\mu(T_0)$	null	null	0.9985	0.0136	null

Inputs(watts): $x_0=1.045, y_0=0.967, z_0=1.016$  and  $t_0=0.887$ ,

Outputs(watts)  $S_X, S_Y, S_Z$  and  $S_T$ .

In the inference step, we have:

$\mu_A$  (variable input): with  $A \in \{very\ low, low, medium, high, very\ high\}$

$\mu_B$  (variable output): With  $B \in \{S_{X0}, S_{X1}, S_{Y0}, S_{Y1}, S_{Z0}, S_{Z1}, S_{T0}, S_{T1}\}$ .  $\mu_{Br}$  is

we determinate, the parametres

$\mu_{Br}$  (variable output)= membership degree of variable output in each Rule (where  $r=1,.. 625$ ).

$\mu_{Br}(\text{output})=\min A \in \{very\ low, low, medium, high, very\ high\}$

$\mu_A(x_0), \mu_A(y_0), \mu_A(z_0), \mu_A(t_0)$ .

$\mu_B$  (variable output) =  $\max (\mu_{Br}$  (variable output))

Table 2: Selected Fuzzy Intervals

Rule	$\mu_a(x_0)$	$\mu_a(y_0)$	$\mu_a(z_0)$	$\mu_a(t_0)$	$\mu_a(s_x)$	$\mu_a(s_y)$	$\mu_a(s_z)$	$\mu_a(s_t)$
R <sub>188</sub>	0.926	0.901	0.9985	0.32	0.32	0.32	0.32	0.32
R <sub>189</sub>	0.926	0.901	0.0235	0.32	0.0136	0.0136	0.0136	0.0136
R <sub>193</sub>	0.926	0.121	0.9985	0.32	0.12	0.12	0.12	0.12
R <sub>194</sub>	0.926	0.121	0.0235	0.32	0.0235	0.0235	0.0235	0.0235
R <sub>213</sub>	0.096	0.901	0.9985	0.32	0.096	0.096	0.096	0.096
R <sub>214</sub>	0.096	0.901	0.0235	0.32	0.0235	0.0235	0.0235	0.0235
R <sub>218</sub>	0.096	0.121	0.9985	0.32	0.096	0.096	0.096	0.096
R <sub>219</sub>	0.926	0.121	0.0235	0.32	0.0235	0.0235	0.0235	0.0235
R <sub>313</sub>	0.926	0.901	0.9985	0.71	0.701	0.701	0.701	0.701
R <sub>314</sub>	0.926	0.901	0.0235	0.71	0.0235	0.0235	0.0235	0.0235
R <sub>318</sub>	0.926	0.121	0.9985	0.71	0.113	0.113	0.113	0.113
R <sub>319</sub>	0.926	0.121	0.0235	0.71	0.0235	0.0235	0.0235	0.0235
R <sub>338</sub>	0.096	0.901	0.9985	0.71	0.096	0.096	0.096	0.096
R <sub>339</sub>	0.096	0.901	0.0235	0.71	0.0235	0.0235	0.0235	0.0235
R <sub>343</sub>	0.096	0.121	0.9985	0.71	0.096	0.096	0.096	0.096
R <sub>344</sub>	0.096	0.121	0.0235	0.71	0.0235	0.0235	0.0235	0.0235

In the defuzzification step, we have:

$$\mu_{res}(\text{output variable}) = \max(\mu_B(\text{output variable}))$$

The value of output is the projection of  $\mu_{res}$  (variable output) of B where  $\mu_B$  (variable output) is maximum.

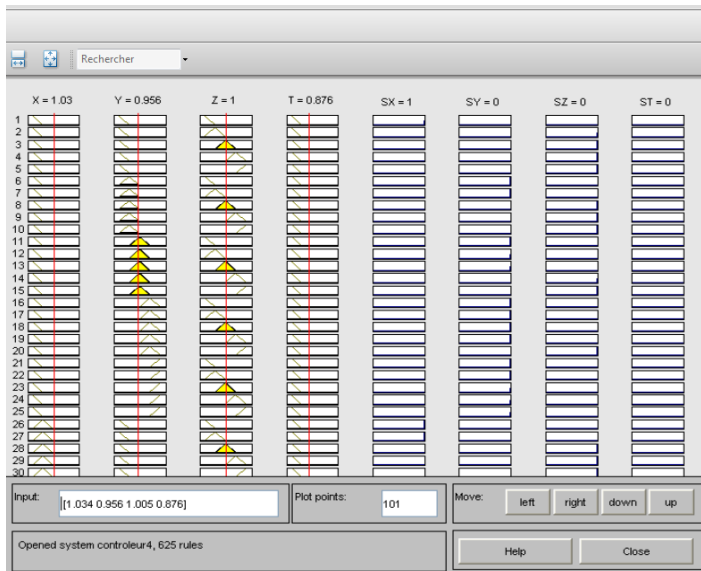


Figure 2: Simulated results for inputs  $x_0$ ,  $y_0$ ,  $z_0$  and  $t_0$

Results:  $s_x=1$ ,  $s_y=0$ ,  $s_z=0$ ,  $s_t=0$

For more efficiency and precision, we studied the queue issued in the switching subsystem for our telecom system as well as function optimization. We obtain a queue in the input which represents received signal power energy related to M.Ant.1, M.Ant.2, M.Ant.3 and M.Ant.4. In the input of the fuzzy logic system which is the switching module system, is presented a multiple queues, indeed. When arrived, the information presented in the received signal has to wait (dt) until it will be selected and judged by the Fuzzy logic controller for telecommunication

establishment circuit making. In that situation all operation is validate but depending in different parameters: Needed time by a queue to select captured signal related to the suitable antenna presented, in addition is added: the time needed to serve available received signal presented in queue 1 during dt1; the time needed to serve available received signal presented in queue 2 during dt2; the time needed to serve available received signal presented in queue 3 during dt3; and the time needed to serve available received signal presented in queue 4 during dt4. Different measures should be taken in consideration in order to evaluate the performance MAB system, such as: time taken by the processor to response, average number operation microstrip antenna.

The agent is a virtual entity or a program which operates resulting autonomously as well as autonomously and continuously on signal received by any element in the antenna network [15]. Various information represented by voice, data, image and video, are represented by agent. For example, an agent may archive various computer files or retrieve electronic messages on a regular schedule. Such simple tasks barely begin to tap the potential uses of agents. So, an intelligent agent can observe different behavior and learn to anticipate their needs, or at least their repetitive actions. It has only a partial environment representation with its expertise and its service. Having the ability to reproduce, whose behavior tends to meet its objectives, taking into account the resources and available skills, its representations and its received communication information signal [16]. A reactive agent has a 'Stimulus / Response' behavior type but not its historical memory, or an explicit representation of the environment and other agents. In a purely reactive system, many of these developed agents coexist [17] [18]. A cognitive agent has an intelligent behavior. It is capable of executing a plurality of tasks, in a plane, in order to achieve a particular objective. This agent has knowledge of other agents[19], its self and the environment as well has the memory of the past. The so-called cognitive systems contain some agents. It is because of the complexity of implementation that each cognitive agent is an expert system. However, they cooperate to find a consistent solution to a general problem. It is usually a multi-expert system. Multi-agent

systems [20] is a system that may include an environment  $E$  that is to say a space with a general metric, a set of objects  $O$  which are located to associate in a position  $E$  at a given time, several agents representing specific objects is considered system active elements, a set of relations that bind objects together.

All operation made by agents are based on a set of operations  $Op$ . For that's why agent is able to consume, produce, perceive, manipulate objects, transform and then represent the application.

In practice, three forms of interaction are distinguished: the cooperation which means that a group of agents [21] work together to reach a common goal. In general, agents cooperate because they need to share resources or skills to accomplish their goals, the coordination that refers to the organization of the work of agents in order to avoid negative interactions and enhance beneficial interactions. Some situations require cooperation agents [22] coordinate to achieve their overall common goal, and the communication which means the exchange of information between agents and makes possible cooperation and coordination. Our approach defined in [23] is to build on the concepts of multi modality in human-computer interaction and to create an architectural model suitable for signal representing information identification in telecommunication system. So, a multi-agent architecture model for multi-modal signals identification. Our approach is based on the observation of existing systems and the identification of recurring design methods and constraints that such systems have to solve. We look at the overall design of the identification of the different agents required, and how to arrange them. This paradigm is based on three-step process Capture, Analysis, and Interpretation. The capture level is the lowest level of the system. As its name suggests, it is at this level that data are captured from data base in real time.

The analysis level includes the extraction of the different carrier characteristics of signal representing information. Finally, the interpretation level involves the various interpretations of characteristics to estimate Information Data Type. As an architectural model for IDT identification, our architecture model must allow the integration of three elements such as capture devices, characteristics of any channel of affective communication, and methods of interpretation. In case of capture devices, an exhaustive list of existing devices is impossible. In addition, our architecture model should enable the creation of architectures for the integration of devices not yet conceived. Where for the characteristics of any channel of affective communication do not including identified or validated characteristics. The methods of interpretation and especially consider the possibility of the implementation of the various theories and models of IDT representation. The architecture model should also allow the merger of Capture, Analysis, and Interpretation. Every dimension consists of an attributes set. These attributes can be simple or complex. In a related work [24], the authors have identified keys dimensions which can be considered for an IDT.

The more keys we present the precise model we get. The IDT is identified by mixing the parameters indicated above achieved by attributing a weight to all parameters. The obtained value will be itemized according to the significance of the parameter. Reading that in some instance, it can be restricted to a few parameters if the found value does not correspond to any known state by state basis.

The IDT agent parameters ( $A^*$ ) provides the current IDT founded on the parameters values (such as code, power level, etc.), frequency, duration and intensity of this detected IDT. Decision maker agent ( $A^{**}$ ) which returns the final IDT is based on analyzing the results which obtained with its frequency and intensity. Supervisor agent ( $S$ ): It is a cognitive agent able to make specific tasks in order to reach a particular objective. The cited agent has more information compared with other agents, it has a memory that recites the related actions. In our context, the supervisor agent can communicate with other agents, which manage information collected to determine the user IDT at the time of consulting income statement. The agent named supervisor boost the values analysis returned by the IDT agents and marks to the manager the final result (IDT agent found with the detection intensity and frequency or unknown).

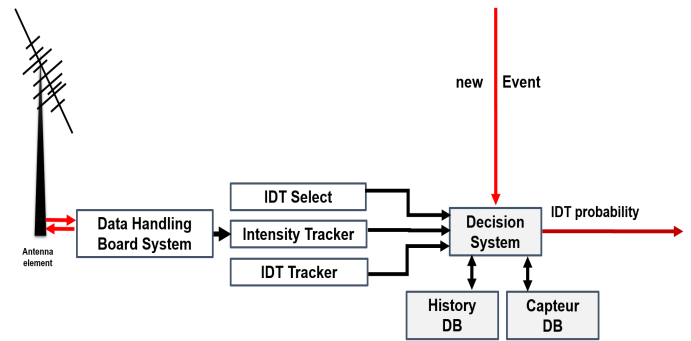


Figure 3: Developed IDT Architecture Model

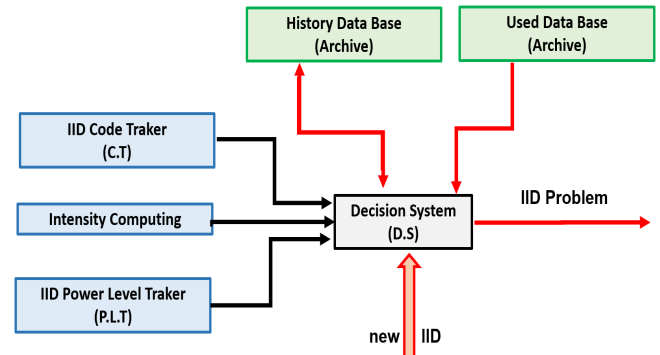


Figure 4: Real-Time Multi-Agent Architecture

We propose a three-functional level architecture of our IDT identification system. The system architecture is made up of three successive blocks, which are the capture level, the analysis level and finally the interpretation level. The scope of our system is introduced by the field of IDT identification that can detect their parameters values without being control. Our framework imposes a constraint of real-time IDT identification and the ability to communicate the detected IDT. The developed system as in figure.4 justifies our choice of multi-agent architecture that can provide real-time and easy communication between its own systems. Figure.4 shows the proposed architecture of our system which is composed of several processing blocks, following the chronological order of the different interpretations that has to the detected signal undergo to give as a result a probability of IDT type. Our work falls within the context of designing intelligent IID recognition system. Our



approach involves the introduction of techniques from different areas but with the desire to keep the logic of near to human reasoning. We have presented our developed results using the Kinect sensor.

## 5. Conclusion

In this paper, a new approach based on Multi-Agent Data Identification System for telecommunication was presented. We demonstrate that the Information data recognition during traditional operating step in telecommunication system, as their interpretation, can ameliorate system efficiency during transmission step. The proposed solution is defined by starting with received signal parameters in antenna network, which is connected on board telecom system. We demonstrate that IID response differs from one presented signal being to another. Since the IID remains difficult to detect and to recognize, we successfully developed a multimodal IID recognition based on two different modalities that solves a various problem related to IID recognition. Successful simulation result and triggered IID are considered a change and used on the IID state which can be distinguished by its real impact on the IID state. Telecommunication system may reveal a steady IID state along operating steps. The IID intensity value decreases to zero. In IID models, a decay function is generally defined to determine the IID intensity and its value at time  $t$  to time  $t+1$ .

## References

- [1] Q. Yu, L.S. Yan, S. Lee, Y. Xie, A. Willner, "Loop synchronous polarization scrambling technique for simulating polarization effects using recirculating fiber loops", *Journal of Lightwave Technology*, **21**, 1593-1600, Jul, 2003. DOI:10.1109/JLT.810103
- [2] N.S. Murthy and P. Srikanth Gautham, "Adaptive Equalizers in Smart Antenna Systems", *IETE Journal of Education*, **50**, 111-117, 2009. DOI:10.1080/09747338.2009.10876058
- [3] A. H. Levy and A. S. Willsky, "An efficient maximum entropy technique for 2-D isotropic random fields" *IEEE Journals & Magazines*, **2**, 741-744, 1988. DOI:10.1109/29.1589
- [4] P. Seong Chuah, C. Zhenzhong and Y. C. Peng Tan, "Energy Minimization for Wireless Video Transmissions With Deadline and Reliability Constraints", *IEEE Transactions on Circuits and Systems for Video Technology*, **23**, 467-481, 2013. DOI:10.1109/TCSVT.2210655
- [5] M. Vesperini, F. Breon and D. Tanre, "Atmospheric water vapor content from space borne POLDER measurements", *IEEE Transactions on Geoscience and Remote Sensing*, **37**, 1613-1619, 1999.
- [6] A. Verma, N. Srivastava and B. K. Kanaujia, "Analysis of Gunn Loaded Rectangular Microstrip Antenna", *IETE Journal of Education*, **53**, 80-87, 2012.
- [7] K.L. Rao Member, "A Simplified Approach to the Antenna Problem", *IETE Journal of Education*, **23**, 7-8, 1982.
- [8] C. Natalia, A. Estefania, B. Vicente, "THOMAS: An agent platform for supporting normative multi agent systems", *Journal of Logic and Computation*, **23**(2), 309-333, 2013. DOI: 10.1093/logcom/exr025
- [9] L. Frank, Lewis, B. Cui, Tiedong Ma, Yongduan Song, Chunhui Zhao, "Heterogeneous Multi-Agent Systems: Reduced-Order Synchronization and Geometry", *IEEE Transactions on Automatic Control*, **61**(5), 1391-1396, 2015. DOI: 10.1109/TAC.2015.2471716.
- [10] Z. Guessoum, "Adaptive agents and multiagent systems", *IEEE Distributed Systems*, **5**, 7, 2004. DOI: 10.1109/MDSO.2004.10
- [11] H. Zhang, G. Feng, H. Yan, Q. Chen, "Distributed self-triggered control for consensus of multi-agent systems", *IEEE/CAA Journal of Automatica Sinica*, **1**(1), 40-45, 2014. DOI: 10.1109/JAS.2014.7004618.
- [12] M. Guo, M.M. Zavlanos, D.V. Dimarogonas, "Controlling the Relative Agent Motion in Multi-Agent Formation Stabilization", *IEEE Transactions on Automatic Control*, **59**(3), 820-826, 2014. DOI: 10.1109/TAC.2013.2281480.
- [13] P.K. Singhal, P.S. Tomar and N. Verma, "Topologies and Applications of Meander Line Antenna", *IETE Journal of Education*, **48**, 121-131, 2007.
- [14] P. Gebhard, "A Layered Model of Affect", In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS '05*, 29-36, New York, NY, USA, 2005.
- [15] M. Dammak, M. Ben Ammar and A.M. Alimi. "Real-time analysis of non-verbal upper-body expressive gestures", In *Multimedia Computing and Systems (ICMCS)*, 2012 International Conference, 334-339, 2012.
- [16] Q. Yu, L.S. Yan, S. Lee, Y. Xie and A. Khilla and I. Wolff, "Field Theory Treatment of H-Plane Waveguide Junction with Triangular Ferrite Post", *IEEE Transactions on Microwave Theory and Techniques*, **41**, 274-281, 1993.
- [17] K. Ghamen and A. Caplier, "Estimation of facial expression intensity based on the belief theory", In *Proceedings Funchal, Portugal*, 59-60, 2008.
- [18] M. Dammak, A. Wali and A.M. Alimi, "Video summarization using viewer affective feedback", In *Hybrid Intelligent Systems (HIS)*, 13th International Conference, 279-284, 2013.
- [19] M. Neji, M. Ben Ammar and A.M. Alimi. Towards an intelligent information research system based on the human behavior: recognition of user emotional state. *12 th IEEE/ACIS International Conference on Computer and Information Science*, 56-58, 2013.
- [20] T. Xiong, Z. Pu, J. Yi, X. Tao, "Fixed-time observer based adaptive neural network time-varying formation tracking control for multi-agent systems via minimal learning parameter approach", *IET Control Theory & Applications journal*, **14**(9), 1147-1157, 2020. DOI:10.1049/iet-cta.2019.0309
- [21] M. Li, S. Chen, "Adaptive Learning: A New Decentralized Reinforcement Learning Approach for Cooperative Multiagent Systems", *Jing CIEEE Control Systems Letters*, **4**(4), 898-903, 2020. DOI: 10.1109/LCSYS.2020.2995756hen
- [22] X. Luo, X. Li, X. Li, J. Yan, X. Guan, "Globally Stable Formation Control of Nonholonomic Multiagent Systems With Bearing-Only Measurement", *IEEE Systems Journal*, **14**(2), 2901-2912, 2019. DOI: 10.1109/JSYST.2019.2935162.
- [23] S. Zheng, H. Liu, "Improved Multi-Agent Deep Deterministic Policy Gradient for Path Planning-Based Crowd Simulation", *IEEE Access journal*, **7**, 147755-147770, 2019, Electronic, 2019. DOI: 10.1109/ACCESS.2019.2946659
- [24] D. Xu, D. Chen; R. Zhang, "Performance and optimisation of agent compression and forwarding system for correlated sources » *IET Communications, IET Control Theory & Applications journal*, **13**(16), 2523-2532, Date: October 2019. DOI: 10.1049/iet-com.2018.5704



## Smart Grid Users Benefits Based on DSM Algorithm Mathematical Optimization Problems Studied

Chafaa Hamrouni<sup>1,2,\*</sup>

<sup>1</sup>Department of Computer Sciences, Khurma University College, Taif University, Khurma, 2935, Kingdom of Saudi Arabia

<sup>2</sup>Research Groups on Intelligent Machines Laboratory, National School of Engineering of Sfax (ENIS), Sfax University, Sfax, 3038, Tunisia

### ARTICLE INFO

Article history:

Received: 26 April, 2020

Accepted: 21 June, 2020

Online: 12 July, 2020

Keywords:

Distributed Energy Storage

Convex Programming

Electric Vehicle

Demand Response

Distributed Energy Generation

Emend Side Management

DMES Daily Maximum Energy

Scheduling

Distributed Sources Management

### ABSTRACT

The purpose of this research paper is to demonstrate that optimization of energy consumption, distributed generation and storage contribute towards mutually beneficial and satisfactory Demand Side Management algorithms that can be installed into consumer smart meters or in Home Energy Storage. A new solution based on an Energy Scheduling and Distributed Storage (ESDS) and Microgrid Energy Management Distributed Optimization Algorithm Demand Side Management (MEM-DOA DSM) algorithms Microgrid Energy Management Distributed Optimization Algorithm Demand Side Management that offers benefits to consumers, utility providers, policy makers and the environment Smart grid, Demand Side Management and mathematical optimization techniques which were studied. A successful development operation of a Demand Side Management Algorithms is made by using appropriate mathematical programming methods depending on the nature of their objective functions, tests results are accomplished.

## 1. Introduction

The Distributed Energy Storage (DES) refers to the location of energy storage devices in consumer premises (commercial, industrial), storage stations or sub-stations for purposes such reducing consumer energy consumption from the grid, providing emergency energy supplies, and spinning reserves. In our study the owners of storage devices is a consumers. DES device operates by taking energy from grid during off-peak periods and stores the energy to be discharged to consumers at a later scheduled time, especially during peak periods and/or to sell to the grid during peak periods or as requested. The aim of such scheduling includes energy and financial savings, a reduction in peak demand from the grid. DES device is installed in one of two modes – standalone or grid-connected mode. The standalone DES receives energy from the grid or a DEG and discharges the same locally to meet consumer demand. On the other hand, a grid-connected DES performs all the functions of the standalone DES, with inclusion of the capability to propid energy. DES has the potential to be located throughout the entire chain of the energy network from generation

right into consumer premises. The location of a DES device in the smart grid can be at consumer premises (be it residential, commercial or industrial), community or microgrid, substation or utility generation site. The DSM problems can have deterministic or stochastic techniques depending several elements and variables. Stochastic optimization is used for many years that depends on historical data. Hence, many DSM methods are designed with robust optimization.

We demonstrate that the good method that can be used in order to reduce the crosstalk effect in Digital Subscriber Line, is Dynamic Spectrum Management method. This work describes a series of Dynamic Spectrum Management algorithms including Optimal/ Iterative/ Autonomous Spectrum Balancing, Iterative or Selective Iterative Water-filling, Successive Convex Approximation for low complexity. The important point, is the existence of compared in terms of performance and computational complexity.

The purpose of the DES device determines where it will be located in the smart grid. In our research, we find out multi-benefits of Distributed Energy Storage in the smart grid in a various locations: for consumer premises such as residential/

\*Chafaa Hamrouni, Taif University-Khurma University College-Department of Siences Computer, Contact:00966546492338,Email:chafa.hamrouni.tn@ieec.org

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050413>

commercial and industrial, their benefits are leverage on time-varying energy pricing (TOU, CPP and RTP), reduction in demand charge by utility, DR and DSM capabilities, enhances installation of DEGs, and offers support during critical load or critical peak period. In case of community or microgrid, we have a reduces congestion in distribution network, less need and frequency of electricity cable replacements, upgrades, offer support to distributed generation through connection with DEGs, mitigate load shedding, electricity outages and blackouts. Reduces congestion in distribution network offers support to distributed generation through DERs and power quality to consumers.

The utility generation site can serve as spinning reserve to the utility grid, offers ramping advantage to the grid and support to distributed generation through connection with DEGs, applicable for line frequency control and used for black start generation. Electric Vehicles (EVs), also known as Electric Drive Vehicles (EDVs), are automobiles that use at least one of electric or traction motors for propulsion. EVs are basically of three types namely: Battery Electric Vehicle (BEV), Hybrid Electric Vehicle (HEV) and Plug-in Hybrid Electric Vehicle (PHEV) or Plug-in Electric Vehicle (PEV). The BEVs have only battery as the source of power. They can be charged during low price periods and discharged at high price periods or when the energy is needed. HEV is made up of an internal combustion engine and a small battery. It uses the battery to optimize the operation of the internal combustion engine and recharges the battery from the gasoline engine and regenerative braking. Studies demonstrates that EVs could lead to reduced dependence on fossil fuels and CO<sub>2</sub> emissions in the transport sector by offering an environmental-friendly and sustainable alternative to traditional internal combustion engine vehicles. EVs can be used to supply peak demand, balance power in feed prediction error in renewable integration [1][2], offer Demand Side Management functionalities [3], grid and tariff stability [4].

The possibilities that EVs can provide electricity to the grid when parked, as dispatchable energy sources, have led to vehicle-to-grid (V2G) technologies [5] [6]. Such parking locations could be in consumer premises, parking lots, shopping malls or other public and easily accessible locations. The technical and financial motivations behind V2G in the smart grid offer benefits to the utility provider, consumer, policy makers and the environment. However, increasing penetration of EVs can increase the grid peak demand due to their charging patterns and load and consumer driving behavior hence, the need to schedule EV loads. We propose for the three major component systems in DSM to be modelled by applying mathematical optimization techniques and can be used to obtain the optimum demand, supply in addition storage profiles in the smart grid for the benefits of all stakeholders in the energy industry. This research paper proves the importance of mathematical optimisation role in DSM algorithms.

## **2. Mathematical Optimization**

A mathematical optimisation provides mathematical basics needed for solving several optimisation problems in economics, science and engineering because it can be applied to study system performance, efficiency and cost effectiveness [7][8]. In our study, we classify nine several classes of mathematical optimisation problems [9]: the first is based on the existence of constraints: This

includes constrained optimisation and unconstrained optimisation problems. When there are no constraints attached to the objective function of an optimisation problem, it's called an unconstrained optimisation problem and vice versa, the second based on the nature of the design variables as parameter (or static) and trajectory (or dynamic) optimisation problems, the third based on the physical structure of the problem which includes optimal control and non-optimal control optimisation problems, the fourth based on the nature of the equations involved, it is a popular way to classify optimisation problems and includes nonlinear programming (NLP), geometric programming (GMP), quadratic programming (QP), and linear programming (LP) problems, where is the classification based on the permissible values of the design variables, it includes integer programming (IP) and real valued programming problems. We have also a classification based on the deterministic nature of the variables including stochastic programming and deterministic programming problems. Stochastic programming involves problems with uncertainty variables while deterministic programming involves problems with known parameters or variables, for the classification based on the separability of the functions: This includes separable programming and non-separable programming problems, and the classification based on the number of objective functions: it includes single objective programming and multi-objective programming problems. Most optimisation problems can belong to more than one category of these classifications. A careful study of an optimisation problem is therefore needed to be able to choose the appropriate solution method and tool.

## **3. Mathematical Optimization**

A mathematical optimization provides mathematical basics needed for solving several optimization problems in economics, science and engineering because it can be applied to study system performance, efficiency and cost effectiveness [7][8]. In our study, we classify nine several classes of mathematical optimization problems [9]: the first is based on problems including constrained and or unconstrained optimization problems.

When there are no constraints attached to the objective function of an optimization problem, it's called an unconstrained optimization problem and vice versa, the second based on the design nature variables as parameter (or static) and trajectory (or dynamic) optimizations problems, the third based on the physical problem structure. It includes optimal control and non-optimal control optimizations problems, the fourth based on the nature of the equations involved, it is a popular way to classify optimization problems and includes nonlinear programming (NLP), geometric programming (GMP), quadratic programming (QP), and linear programming (LP) problems, where is the classification based on the permissible values of the design variables, it includes integer programming (IP) and real valued programming problems. We have also a classification based on the deterministic nature of the variables including stochastic programming and deterministic programming problems. Stochastic programming involves problems with uncertainty variables while deterministic programming involves problems with known parameters or variables, for the classification based on the separability of the functions: This includes separable programming and non-separable programming problems, and the classification based on the number of objective functions: it includes single objective

programming and multi-objective programming problems. Most optimization problems can belong to more than one category of these classifications. A careful study of an optimization problem is therefore needed to be able to choose the appropriate solution method and tool.

#### 4. Mathematical Formulation of Optimization Problem

Mathematical optimization involves the search for the best element or optimal solution from a set of available elements and solutions respectively subject to certain conditions. It involves maximizing, minimizing or maximizing and minimizing a real function by selecting different input values from a set of values and obtains the optimal value that best satisfies all the constraints. Basically, an optimization problem comprises optimization variable, objective function and constraint function(s) and is of the general form:

$$\begin{aligned} \min(x) \quad & fo(x) & (1) \\ \text{Subject to} \quad & fi(x) \leq bi, \quad i=1,2,3,\dots,m \end{aligned}$$

where  $x = [x_1, x_2, x_3, \dots, x_n]^T$  is the optimization variable vector,  $fo(x): \mathbb{R}^n \rightarrow \mathbb{R}$  is the objective function/cost function and is subject to constraint function  $fi(x)$ , where

$$fi(x): \mathbb{R}^n \rightarrow \mathbb{R} \quad (2)$$

$i=1,2,3,\dots,m$  and  $b_1, b_2, b_3, \dots, b_m$  are constants that sets limits or bounds to the constraint [10]. The problem in (1) describes a minimisation problem that chooses the optimal vector  $x$  subject to the given constraints. The domain  $D$  of the optimisation problem is the set of points where the objective function and constraints are defined by:

$$D = \bigcup_{i=0}^m \text{dom } fi \quad (3)$$

The problem in (1) can be said to be feasible if there exists a subset of points  $x \in D$  which satisfies all the constraints. Therefore, a vector  $x_n$  is called the optimal solution of (1) if it has the smallest objective value among all other possible vectors that satisfy the constraints:

For any

$$z \text{ with } f_1(z) \leq b_1, b_2, \dots, f_m(z) \leq b_m \quad (4)$$

So that

$$f_0(z) \geq f_0(x^*), \quad \forall x \in D \quad (5)$$

Examples of minimization problems include optimization problems that attempt to minimize the expenditure of a consumer on certain commodities, the running capital of a company, the runtime of delivery of certain services to customers, energy consumption and expenditure.

On the contrary, maximization problems look for optimal largest objective value in their solutions (maximization of profit by a utility or manufacturing company), satisfaction or comfort for consumers etc. A general format for expressing a maximization problem is given by :

$$\text{Max}(x) \quad fo(x) \quad (6)$$

$$\text{Subject to} \quad fi(x) \leq bi, \quad i=1,2,3,\dots,m$$

where  $x = [x_1, x_2, x_3, \dots, x_n]^T$  is the optimisation variable vector,  $f_0(x): \mathbb{R}^n \rightarrow \mathbb{R}$  is the objective function/cost function and is subject to constraint function:

$$fi(x): \mathbb{R}^n \rightarrow \mathbb{R} \quad (7)$$

$i=1,2,\dots,m$  and  $b_1, b_2, \dots, b_m$  are constants that sets limits or bounds to the constraint [11]. In addition, there could also be combination of minimisation and maximisation in some problems. Such problems could either be minimax or maximin problem. A minimax problem is formulated to minimise the maximum value of a number of decision variables, while a maximin problem maximises the minimum value of a number of decision variables. For example, solving for  $x^*$  in:

$$\max_{y \in G} F(X^*, Y) = \min_{x \in H} \max_{y \in G} F(X, Y) \quad (8)$$

where  $H$  is a convex closed subset of  $E_n, G$  is a bounded closed subset of  $E_m, X$  is a known parameter and  $F(X, Y)$  is the cost function. If  $F(X, Y)$  is linear, then (2) is a linear minimax problem. However, if  $H \neq E_n$  then (2) is a constrained minimax problem [12]. Similarly, the maximin of (2) can be expressed as:

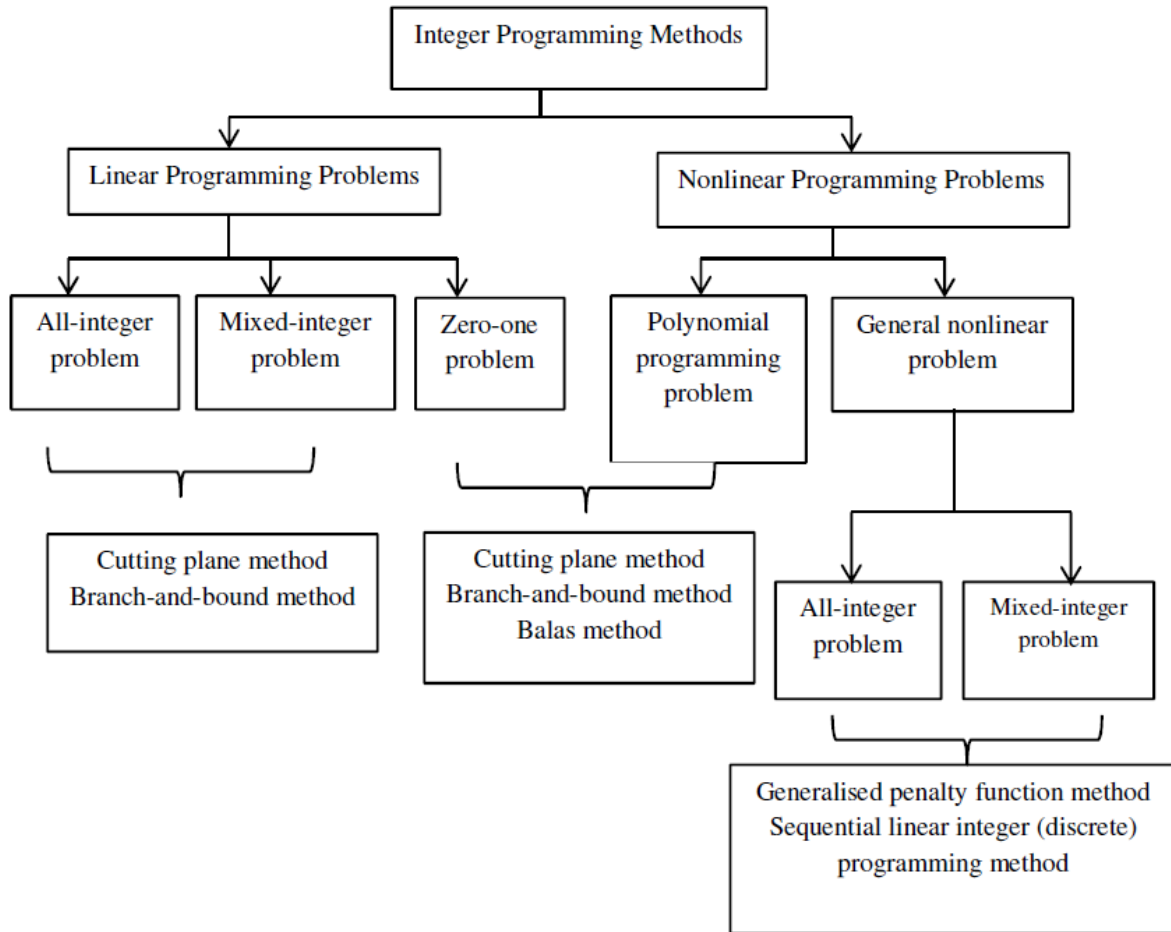
$$\min_{x \in H} F(X, Y) = \max_{y \in G} \min_{x \in H} F(X, Y) \quad (9)$$

#### 5. Solving Mathematical Optimization Problems

The size and nature of the objective function, variables and constraints determines what type of mathematical programming method to be applied and solver to be used for its solution. For instance, linear programming problems can be effectively solved using the popular Dantzig Simplex method [13], [14] or interior-point method [15] or through the use of an embedded computer application solver like the CVX toolbox for Matlab [16]. LP can also be solved using Karmarkar's method, which was presented to be fifty times faster than the Dantzig Simplex method [17]. For example, the problem in (1) can be said to be a LP problem provided the objective and constraint functions are all linear functions.

Integer and mixed integer programming problems can generally be solved by choosing the best solution from all possible solutions, but when the number of variables is large, it becomes very difficult to solve. Although there are several techniques for solving integer and mixed integer programming problems, the cutting plane algorithm of Gomory [18] and the branch-and-bound algorithm of Land and Doig [19] are the most popular. Balas in [20] went on to develop the Balas algorithm for solving zero-one LPs even though they could also be solved using the cutting plane and branch-and-bound methods. The various methods for solving different integer programming problems are presented in Figure.1

MILP problems are optimisation problems with linear objective function  $C^T X$  constraints where  $C$  is a column vector of constants and  $T$  is a column vector of unknowns with restrictions on some components of  $T$  to take on integer values. MILP problems can be conveniently solved using branch-and-bound method [21] and can be generally represented as:



$$\text{Min } C^T X \quad (10)$$

$$\text{s.t. } Ax = b \quad (11)$$

$$x \geq 0, x_i \in \mathbb{Z}, \forall i \in I. \quad (12)$$

CP has attractive theoretical properties and has found applications in the development of efficient and reliable numerical algorithms. It unifies and generalises some common optimisation problems such as least square, quadratic programming and linear programming problems. A CP problem is of the general form:

$$\min f(x) \quad (13)$$

$$\text{s.t. } g_i(x) \leq 0, i=1, \dots, m$$

Where functions  $f(x), g_1(x), \dots, g_m(x): \mathbb{R} \rightarrow \mathbb{R}$  are convex functions satisfying the condition:

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) \quad (14)$$

$$\alpha \in [0,1], x,y \in \mathbb{R}.$$

If  $f(x)$  is a convex function then  $-f(x)$  is a concave function. CP problems are commonly solved using the interior-point method. Convex and concave functions are illustrated in Figure.2:

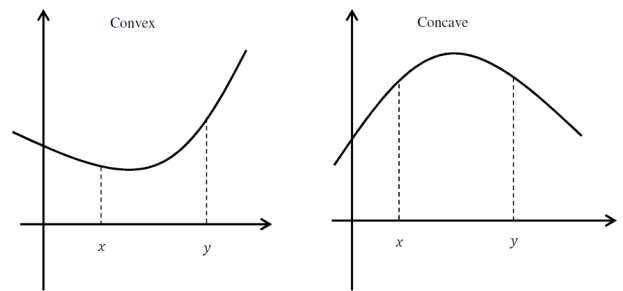


Figure 2: Illustration of convex and concave

$$\begin{aligned} &\text{minimize } \frac{1}{2} x^T Qx + c^T x \quad (15) \\ &\text{subject to } Ax \leq b. \end{aligned}$$



Suppose an  $n$ -dimensional vector is determined by  $w$ , and  $n \times n$ -dimensional real symmetric matrix is presented as  $Q$ ,  $b$  is an  $m$ -dimensional real vector and  $n$ -dimensional real matrix.

Stochastic programming involves optimisation problems that have uncertain objective or constraint functions, which are characterised by a probability distribution on the variables.

Examples of mathematical programming methods found in DSM literature include linear/ integer/ binary/ mixed integer linear programming [22] [23], quadratic programming (QP), convex programming (CP), nonlinear programming (NLP) [24], stochastic programming (SP), heuristics and metaheuristic programming, and mixed integer nonlinear programming (MINLP).

Additional sensitivity analyses were carried out to study the robustness of the ESDS algorithm [25] [26], on PAR demand, dissatisfaction cost. First, the effect of battery capacity on consumer dissatisfaction was investigated. It was discovered that battery capacity has an indirect relationship to dissatisfaction of energy consumption. That is, the higher battery capacity value which can be acquired by consumers as iHES device, the less the consumer will be dissatisfied by appliance scheduling to an extent.

However, capacity of battery can never indefinitely increase; so, the diminishing law returns and set on savings point and then battery pay-back period. Simulation result of battery capacity assumed to be 7 kWh [27]. Therefore, 4 kWh and 10 kWh battery capacities were simulated and the results obtained are presented in Figure.3.

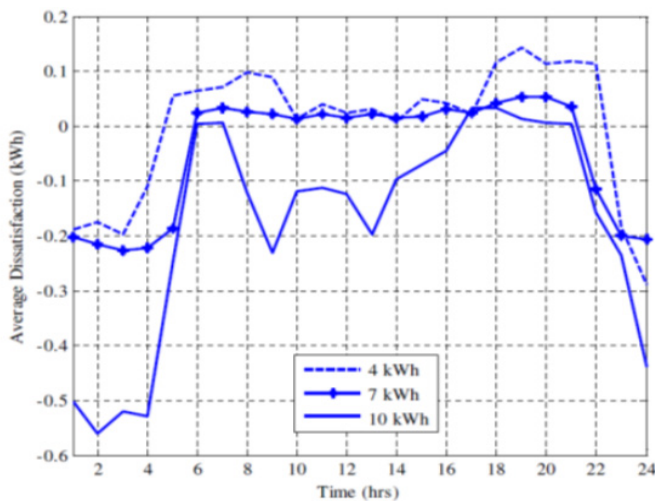


Figure 3: Relation between battery capacity and dissatisfaction coast

[28], [29], is effective in terms of robustness and convergence of algorithm. The performance of optimization techniques depends on user interactions and time scale.

The algorithm determines the energy value stored in the battery, it purchases from the grid and schedules demand to be met within the capacity of stored energy, at mentioned price. The results showed that peak period dissatisfaction, which is a common feature in DSM programs, and based on incorporation

distributed energy storage system, it can be reduced. Consumer premises applying above constraints. The effect of DSM scheduling on dissatisfaction cost is shown in Figure.3.

## 6. Conclusion

Smart grid, Demand Side Management and mathematical optimisation techniques were reviewed in this work. Optimisation of energy consumption, distributed generation and storage can contribute towards mutually beneficial and satisfactory Demand Side Management Algorithms, which can be installed into consumer smart meters. These Demand Side Management Algorithms can be designed using any appropriate mathematical programming methods depending on the nature of their objective functions, variables and constraints. The DSM algorithms proposed in this paper are DMES, DOS-EUP.

## Conflict of Interest

The author declares no conflict of interest.

## Acknowledgment

The Authors would like to acknowledge the financial support of this work by Taif University- Khurma Univerity Collegue- Kidom of Saoudia Aribya.

## References

- [1] W. Kempton and A. Dhanju, "Electric vehicles with V2G: storage for large-scale wind power", Windtech International, 2006.
- [2] M. D. Galus, and G. Andersson, "Balancing Renewable Energy Sources using Vehicle to Grid Services controlled by MPC in a Metropolitan Area Distribution Network", in Proc.CIGRE ELECTRA, 2012.
- [3] P. Mesaric and S. Krajar, "Home demand side management integrated with electric vehicles and renewable energy sources," Energy and Buildings, **108**, 1-9, 2015. Doi:10.1016/j.enbuild.2015.09.001
- [4] M. D. Galus, M. G. Vayá, T. Krause, and G. Andersson, "The role of electric vehicles in the smart grid," Wiley Interscience Interdisciplinary Reviews of WIRES, Energy and Environment (WENE), **2**, 4, 384-400, 2013. doi:10.1002/wene.56
- [5] W. Kempton and J. Tomic, "Vehicle-to-grid power fundamentals: calculating capacity and net revenue," Journal of Power Sources, **144**, 1, 268-279, 2005. doi.org/10.1016/j.jpowsour.2004.12.025
- [6] J. Suk-ye, Domestic electric vehicle market showing rapid growth, Business Korea/10126-governmental-support-domestic-electric-vehicle-market-showing-rapid-growth, 2016.
- [7] W. H. Fleming and R. W. Rishel, Deterministic and Stochastic Optimal Control, Springer, 1975.
- [8] A. K. Dixit, Optimisation in economic theory, Oxford, UK: Oxford University Press, 1990.
- [9] K. Schmedders, Numerical optimisation methods in economics, The New Palgrave Dictionary of Economics, 2008.
- [10] S. S. Rao, Engineering Optimisation Theory and Practice, John Wiley and Sons Inc, Hoboken, New Jersey, 4th Edition, 2009.
- [11] V. F. Dem'yanov, V. N. Malozemov and D. Louvish, Introduction to Minimax, Keter House Jerusalem Ltd., 2014.
- [12] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, CVX Research Inc., 2015.
- [13] S. Boyd and L. Vandenberghe, Convex Optimisation, Cambridge University Press, New York, United States of America, 7th Edition, 2009.
- [14] G. B. Dantzig and M. N. Thapa, Linear programming 1: Introduction, Springer, 1997
- [15] Y. Nesterov and A. Nemirovski, Interior-point polynomial algorithms convex programming, Studies in Applied Mathematics, SIAM, Philadelphia, 1994.
- [16] R. E. Gomory, "An Algorithm for the Mixed Integer Problem," Rand Report R.M. 25797, July 1960.
- [17] J. Nocedal, S. Wright and J. Stephen, Numerical Optimisation, 2nd Edition, Springer- Verlag, Berlin, New York, 2006.
- [18] A. H. Land and A. Doig, An Automatic Method of Solving Discrete Programming problems, Econometrica, 1960.
- [19] A. Shapiro, D. Dentcheva and A. Ruszczyński, Lecture notes on stochastic programming – modeling and theory, Siam, 2009.



- [20] B. Daryanian, R. Bohn, and R. Tabors, "Optimal demand-side response to electricity spot prices for storage-type customers," *IEEE Transactions on Power Systems*, **4**, 3, 897-903, 1989. Doi: 10.1109/59.32577
- [21] D.S. Chen, R. G. Batson and Y. Dang, "Applied Integer Programming: Modeling and Solution," John Wiley and Sons, 2010.  
S.P.Bradley, A.C.Hax and T.L.Magnanti, "Applied Mathematical Programming," Addison-Wesley, 1977.
- [22] Z. Mao and X. Wang, "Efficient optimal and suboptimal radio resource allocation in ofdma system," *IEEE Transactions on Wireless Communications*, **7**, 2, 440-445, 2008. Doi: 10.1109/TWC.060546
- [23] R. E. Moore, *Global optimisation to prescribed accuracy*, Computers and Mathematics with Applications, Eldon Hensen, G.Willamwalster, 1991.
- [24] D. Chen, X. Ziqi, L. Ximeng, Y. Yin, Y. Yang, G. Wenzhong, "Dual al-Search Artificial Bee Colony Algorithm for Engineering Optimization", *IEEE Access*, **7**, 24571 –24584, 2019, Electronic ISSN: 2169-3536, INSPEC Accession Number: 18501113, Doi: 10.1109/ ACCESS.2019.2899743
- [25] G. Cyrielle, B. Renaud, P. Marc, "Automatic Detection of Elevation Changes by Differential DSM Analysis: Application to Urban Areas", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **7**, 10, 2014, : 4020 – 4037, Doi: 10.1109/JSTARS.2014.2300509
- [26] C. Hamrouni, A. Bsissa, R. Hamza, N. Abdelkarim, "A new MEM-DOA proposal for DSM in a grid connected smart microgrid", *Progress In Electromagnetics Research Symposium-Spring (PIERS)*, 22-25 May 2017. Doi: 10.1109/PIERS.2017.8262039
- [27] W. Han ; X. Sean Lu ; M. Zhou ; X. Shen ; J. Wang ; J. Xu," Evaluation and Optimization Methodology for Efficient Power Plant Programs", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *IEEE Transactions on Systems, Man, and Cybernetics Systems*, **50**, 2, 707 – 716, 2020. Doi: 10.1109/TSMC.2017.2714198
- [28] W. Lanneer, P. Tsiaflakis, J. Maes, M. Moonen," $\alpha$ -Fair Dynamic Spectrum Management for QRD-Based Precoding With User Encoding Ordering in Downstream G.Fast Transmission", *IEEE Transactions on Communications*, **67**, 4, 2939 – 2950, 2019. Doi: 10.1109/TCOMM.2018.2890237.
- [29] D. Li. C. Wei, S. Hongjian, H.P. Vincent,"Multi-objective Optimization for Demand Side Management Program in Smart Grid", *IEEE Transactions on Industrial Informatics*, **14**,4, 1482 – 1490, 2018. Doi: 10.1109/TII.2017.2776104.

## Contribution in Private Cloud Computing Development based on Study and KPI Analysis

Chafaa Hamrouni<sup>\*1,2</sup>, Slim Chaoui<sup>3,4</sup>

<sup>1</sup>Department of Computer Sciences, Khurma University College, Taif University, Khurma, 2935, Kingdom of Saudi Arabia

<sup>2</sup>Research Groups on Intelligent Machines Laboratory, National School of Engineering of Sfax (ENIS), Sfax University, Sfax, 3038, Tunisia

<sup>3</sup>College of Computer and Information Sciences, Department of Computer Engineering and Networks, Jouf University, Skaka, 72311, Kingdom of Saudi Arabia

<sup>4</sup>Unit- Laboratory of Sciences of Electronics, Technologies of Information and Telecommunications, Sfax University, Sfax, 3038, Tunisia

### ARTICLE INFO

Article history:

Received: 26 April, 2020

Accepted: 26 June, 2020

Online: 12 July, 2020

Keywords:

QoS

Virtual Infrastructure Manager

Virtual Machine

Virtual Private Network

Quality of Service

### ABSTRACT

The preset paper describes a contributes in the implementation of private Cloud Computing. Exhaustive study made validates Cloud environment creation and successful deployed Solution Configuration choice. Related solution ensures a good investment that reduces energy costs of the data center, minimizes network traffic cost proven as an NP-hard problem. In addition, each virtual machine (VsM) depends on energy within the data center. Used genetic algorithm allows the emergence of new generations, the use of Tabu search algorithm demonstrates a divergence state, however an identification technique, based on a hybrid algorithm, ensures which optimized virtual machine.

## 1. Introduction

In our research work, we illustrate the simulation tools, main functions in addition the private cloud domain [1]. A developed a solution to solve virtual machine placement problem and memetic algorithm with a costly function. An hybrid algorithm as a combination of genetic algorithm and the Tabou algorithm, and then an experimental development based on metaheuristic algorithms.

## 2. Design and implementation Step

The CloudSim simulation is divided into two main categories of classes: the classes that model the entities shown in blue as the Data Center, the Broker, the Cloudlet. Each class has its own specific functionality [2]. The first class is the Data Center that simulates the service kernel infrastructure that is offered by resource providers in a cloud computing environment. It encapsulates a set of computing machines that must be

homogeneous or heterogeneous with respect to their resource configurations [3]. While Data Center Broker represents the second class. It models the broker who is responsible for modeling between users and service providers according to customers' QoS conditions and deploys service tasks across the clouds. The third SAN Storage class provides a simple interface for accessing or storing data at any time.

Access to files in a runtime SAN causes additional delays in the required data files across the internal data center network. The fourth class is Virtual Machine [4]. It consists of modeling a virtual machine instance, whose management during its life cycle is a responsibility of the machine (Host). The fifth class Cloudlet models the core application services of Nails deployed in data centers. CloudSim consults the complexity of an application in terms of its IT needs [5]. Each application component has a predefined instruction size and a quantity of data transfer flows. The sixth class is BWProvisioner [6]. This is an abstract class that models the policy of supplying bandwidth to virtual machines that are deployed on a Host component. While the seventh class Memory Provisioner represents the provisioning policy for

\*Chafaa Hamrouni, Taif University-Khurma University College-Department of Sciences Computer, Contact:00966546492338,Email:chafa.hamrouni.tn@ieee.org

assigning memory to virtual machines. If the Memory Provisioner component determines that the Host has the amount of free memory then the virtual machine can run and deploy to a Host. For the eighth VM Provisioner class, this abstract layer expresses the provisioning policy that a virtual machine engine uses to allocate VMs to hosts. For the last class VMM Allocation Policy [7] is an abstract implemented by a host component that models policies (shared space and time share) motivated to allocate processing power to virtual machines. Among the principles of Cloud Sim, it minimizes the number of entities in the system by using only the Data Center and the user as an inherited actuator of the Sim Java entities. This design method assists Cloud Sim in modeling a large-scale simulation environment on a machine with balanced processing capability. Virtual machines, provisioning policies, and Hosts are the key components of Cloud Sim that are instantiated as stand-alone, lightweight objects and do not compete with computing power. To take advantage of the benefits of virtualization [8], a private cloud operates on standardized architectures, so you must ensure that hosted applications are developed for these environments. Deploying a private cloud requires several steps:

-Step 1:

It consists in checking security processes proposed by the service provider (private cloud) or the IT department (internal cloud) in the event of a security breach.

-Step 2:

It consists in setting up an encrypted protection of the data.

-Step 3: It allows the measurement of the trust granted to the service provider (private cloud).

### 3. Design and implementation Step

The CloudSim simulation is divided into two main categories of classes: the classes that model the entities shown in blue as the Data Center, the Broker, the Cloudlet. Each class has its own specific functionality [2]. The first class is the Data Center that simulates the service kernel infrastructure that is offered by resource providers in a cloud computing environment. It encapsulates a set of computing machines that must be homogeneous or heterogeneous with respect to their resource configurations [3]. While DataCenterBroker represents the second class. It models the broker who is responsible for modeling between users and service providers according to customers' QoS conditions and deploys service tasks across the clouds. The third SAN Storage class provides a simple interface for accessing or storing data at any time.

Access to files in a runtime SAN causes additional delays in the required data files across the internal data center network. The fourth class is VirtualMachine [4]. It consists of modeling a virtual machine instance, whose management during its life cycle is a responsibility of the machine (Host). The fifth class Cloudlet models the core application services of Nails deployed in data centers. CloudSim consults the complexity of an application in terms of its IT needs [5]. Each application component has a predefined instruction size and a quantity of data transfer flows. The sixth class is BWProvisioner [6]. This is an abstract class that models the policy of supplying bandwidth to virtual machines that are deployed on a Host component. While the seventh class Memory Provisioner represents the provisioning policy for

assigning memory to virtual machines. If the Memory Provisioner component determines that the Host has the amount of free memory then the virtual machine can run and deploy to a Host. For the eighth VM Provisioner class, this abstract layer expresses the provisioning policy that a virtual machine engine uses to allocate VMs to hosts. For the last class VMM Allocation Policy [7] is an abstract implemented by a host component that models policies (shared space and time share) motivated to allocate processing power to virtual machines. Among the principles of Cloud Sim, it minimizes the number of entities in the system by using only the Data Center and the user as an inherited actuator of the Sim Java entities. This design method assists Cloud Sim in modeling a large-scale simulation environment on a machine with balanced processing capability. Virtual machines, provisioning policies, and Hosts are the key components of Cloud Sim that are instantiated as stand-alone, lightweight objects and do not compete with computing power. To take advantage of the benefits of virtualization [8], a private cloud operates on standardized architectures, so you must ensure that hosted applications are developed for these environments. Deploying a private cloud requires several steps:

-Step 1:

It consists in checking security processes proposed by the service provider (private cloud) or the IT department (internal cloud) in the event of a security breach.

-Step 2:

It consists in setting up an encrypted protection of the data.

-Step 3: It allows the measurement of the trust granted to the service provider (private cloud).

### 4. Implementation Step of the Meta- heuristic Algorithms

We introduce our method by identification of our data sets, then the parameters used for the different algorithms used as well as the results obtained. Finally, we will evaluate the algorithms used. Initial number of virtual machines is the most important parameter for creating an environment that can simulate a real bus. In that case, the genetic algorithm is considered an algorithm that treats a population and contains a set of generations. It begins by setting some parameters namely the size of the initial population P with which the genetic algorithm starts as well as the number of iterations. The number of iterations that correspond to the number of generations created is chosen arbitrarily. But, in our case, we have to apply this algorithm on a single generation that contains a number of individuals chosen systematically. Using the parameter of the genetic algorithm as previously stated for the virtual machine placement problem, the chromosome is arbitrarily entered by the system. Then, each one is converted into a binary number. The calculation of the optimization cost of each individual is necessary to achieve the best optimization. A list is then drawn up to follow the path of each individual on a chromosome. We proceed by a proportional selection that allows us to distinguish individuals with a large fitness. The result obtained by this procedure is subjected to a second selection by tournaments at one principle to make a draw with delivery of two individuals of population, and one makes them. The one with the highest fitness wins with a probability p between 0.5 and 1. This process is repeated n times in order to obtain the n individual population p that will serve as parents. The variance of this method is high and

increasing or decreasing the population value decreases or increases the selection effects [9]. We choose the first individual from the populations on the list to mark the one with the highest value. The result obtained is converted into binary in order to apply the two main methods of the genetic algorithm. Adopted method consists on:

- Choose individuals of the same size.
- compare the first gene of the first individual with the first gene of his binomial. This treatment is applied until the end of the individual list.
- Looking for crossing points when the value is different from one gene to another.
- Reverting the crossing requires a threshold that oscillates between 0.8 and 0.5. The list of the final cross depends on the method followed in paragraph.
- Its filling is carried out with the birth of two individuals C and D. the individual C contains the first part of the individual A and the last part of the individual B. As for the individual D, it contains the first part of the individual B and the second part of the individual A, based on the random crossing. One second method is applied after the crossing process consists of:

-Part of the crossing list obtained and choose two successive individuals.

-Compare the binary numbers of C and the numbers of D. If there is a change between the genes, the number of mutations increases but this number should not exceed 50%. The mutation method is applied once between two unique individuals. This result is stored in a list called mutation list.

During implementation of the memetic algorithm steps, we considered the memetic algorithm is a combination of the genetic algorithm and the taboo search algorithm, the algorithm uses the solution found in the genetic algorithm. We introduce the solutions in a list open only to each descending elements. We release each individual from minimum and place it in a new list. It contains the global and local minimums. We consider the minimum value of this list is exclusively global. Then the same number of individuals used in genetic algorithm for initial population and the choice of the size for the taboo list depends on each configuration given.

In addition, the search space size directly influences this parameter. If in a given configuration, multiple private Data Centers have free space for hosting one or more virtual machines, then the search space increases. In that case, more the movement possibilities numbers increase as the taboo list size, much more the taboo movements limited we have. It contains information about the latest taboo movements and this information includes the moved virtual machine, the old server that hosted it, and the new server. The performance of memetic algorithm is calculated in terms of network-traffic cost, it takes longer to be executed. Indeed, and each iteration, the taboo search algorithm is operated on each chromosome. Both genetic and taboo search algorithms represent the layout of virtual machines in private data centers, and

therefore, its size is relative to the number of VMs. Increasing the number of virtual machines therefore reduces the memetic algorithm operation time.

### 5. Experimental results

During genetic algorithm implementation, we obtained accepted result. We operate a local research algorithm based on Taboo research algorithm. Time execution increased slightly increases. Optimized virtual machines organization in distributed Data center is determined with accepted QoS. We studied the network traffic possibility, based on memetic algorithm. During execution, we present the Virtual machines number: 18, and to each virtual machine (annoyance) is affected an individual size randomly, that produces genetic algorithm action step. (Table.1 and Table.2).

Table 1: Virtual Machine conversion size list

Virtual Machine Conversion Size (Binary)	1	1	0	0	0	1	1	1	1	1	0	0	1	1
		0	0	0	0	0	1	0	1	1	0	1		1
		1	1	0	0		0	1	0	1	1			1
		1	0				0	1	0	1	0			0
		0	1				1	1	1	0	0			0
		1	0				1	1		1	0			1
			0				0	0			0			
							1	1						
Virtual Machine Conversion Size (Decimal)	1	4	2	0	0	3	2	4	5	6	3	1	1	1
		5	0				8	4	7	1	2			1

Table 2: Number of selected Virtual Machine with its size list

Virtual Machine Number	18									
Virtual Machine Size	1	6	7	3	3	2	8	8	7	
Virtual Machine Number										
Virtual Machine Size	2	2	6	6	6	6	8	2	7	

During application step operating, the system created from the virtual machines size list of the individuals according to its determined size (randomly), in a binary elements; then process to convert them subsequently into decimal. we apply the proportional selection method to fix the first integer of the list, and compare it with the other individuals, in order to select those which have the highest value in decimal, as a practice example, the maximums in decimal are: 45, 20, 2, 205, 189, 46, 3, 28, 44, 57, 61, 32, 114. Subsequently, individuals of the same size are selected to apply the main methods of the genetic algorithm. The crossing method is considered the basis of algorithm operation. We test the percentage of growth that does not exceed 80% compared to the percentage found between two individuals. If founded percentage is less than 80%, the crossover position between the two individuals will be randomly selected. Those beyond to the percentage is excluded. Result after this process will be getting two new one. In this case, the percentage is equal to 50% so we can apply this method to create a new element. We evaluate the used algorithms. So, for the memetic algorithm, it presents the best results for both small and large instances. The execution time of the memetic algorithm is high, explained by the intervention of the Taboo search algorithm on each individual in each generation. A taboo search algorithm greatly contributes to improvement the solution except that execution several times in the same iteration, slightly increases the processing time of the memetic algorithm. The compromise

between the execution time of the algorithm as well as the improvement of the solution obtained is.

## 6. Conclusion

Working on Data Center virtual machines mission are successfully accomplished, so infrastructure access facilitates virtual machines distribution and design operation. Identification and optimization of virtual machine number is based on genetic and taboo algorithms used for simulate by cloudsim. Mimetic Algorithm adjustment operations, that result minimization of virtual machines number and energy consumption, are validated. During the present research work, we successfully contribute in the implementation of private Cloud Computing. The study is approved per experimental tests in terms of Cloud environment creation as well as deployed Solution Configuration choice. Solution. We use genetic algorithm that allows the emergence of new generations, and Tabu search algorithm demonstrates an convergence state. Proposed heuristic approach contributes in virtual machines location and minimizes energy consumption.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

The Authors would like to acknowledge the financial support of this work by Taif University- Khurma Univerity Collegue- Kidom of Saudi Arabia.

## References

- [1] A. Khilla and I. Wolff, Field "Theory Treatment of H-Plane Waveguide Junction with Triangular Ferrite Post", *IEEE Transactions on Microwave Theory and Techniques*, **41**, 274-281, 1993. Doi:10.1109/TMTT.1978.1129365
- [2] N.S. Murthy and P. Srikanth Gautham, "Adaptive Equalizers in Smart Antenna Systems", *IETE Journal of Education*, **50**, 111-117, 2009. Doi: doi.org/10.1080/09747338.2009.10876058
- [3] A. H. Levy and A. S. Willsky, "An efficient maximum entropy technique for 2-D isotropic random fields", *IEEE Journals & Magazines*, **2**, 741\_744, 1988. Doi: 10.1109/29.1589
- [4] P. Seong Chuah, C. Zhenzhong and Y. C. Peng Tan, "Energy Minimization for Wireless Video Transmissions With Deadline and Reliability Constraints", *IEEE Transactions on Circuits and Systems for Video Technology*, **23**, 467-481, 2013. Doi:10.1109/TCSVT.2012.2210655
- [5] M.Vesperini, F. Breon and D.Tanre, "Atmospheric water vapor content from space borne POLDER measurements", *IEEE Transactions on Geoscience and Remote Sensing*, **37**, 1613\_1619, Doi: 10.1109/36.763275, Feb. 1999.
- [6] A.Verma, N.Srivastava and B. K. Kanaujia, "Analysis of Gunn Loaded Rectangular Microstrip Antenna", *IETE Journal of Education*, **53**, 80-87, 2012. Doi:10.1080/09747338.2012.10876096
- [7] K.L.Rao Member, "A Simplified Approach to the Antenna Problem", *IETE Journal of Education*, **23**, 7-8, 1982. Doi:1080/ 09747338.11450499
- [8] P.K. Singhal, P.j Singh Tomar and N.Verma, "Topologies and Applications of Meander Line Antenna", *IETE Journal of Education*, **48**, 121-131, 2007. Doi:10.1080/09747338.2007.11657877
- [9] Q. Yu, L.S. Yan, S. Lee, Y. Xie and A. E. Willner, "Loop synchro. polarization scrambling technique for the simulating polarization effects using recirculating fiber loops", in the *Journal of Lightwave Technology*, **21**, 1593-1600, 2003. Doi: 10.1109/JLT.2003.810103



## Optimization of Dual Motion Mechanism with Double Grooved Cams for High-voltage Gas Circuit Breaker

Masanao Terada<sup>1,2,\*</sup>, Yuki Nakai<sup>3</sup>, Hiroaki Hashimoto<sup>1</sup>, Daisuke Ebisawa<sup>3</sup>, Hajime Urai<sup>1</sup>, Yasunobu Yokomizu<sup>2</sup>

<sup>1</sup>Research & Development Group, Hitachi, Ltd., 319-1292, Japan

<sup>2</sup>Department of Electrical Engineering, Nagoya University, 464-8603, Japan

<sup>3</sup>Energy Business Unit, Hitachi, Ltd., 316-8501, Japan

### ARTICLE INFO

Article history:

Received: 02 April, 2020

Accepted: 19 June, 2020

Online: 12 July, 2020

Keywords:

Circuit Breaker

Capacitive Current Switching

Dual Motion

Grooved Cam

### ABSTRACT

*A novel design of a dual motion mechanism for a high-voltage gas circuit breaker is presented. The contact motion of the circuit breaker due to the operating mechanism increases capacitive current switching (CCS) performance. CCS is one of the interrupting duties of the circuit breaker, where high voltage is applied during the half cycle from contact separation. The dual motion mechanism drives two contacts in opposite directions from each other. Operating energy is reduced because the maximum displacement of the moving parts linked to the operating mechanism is shortened. To increase CCS performance at lower operating energies, the contact on the opposite side of the contact linked to the operating mechanism requires quick motion in the CCS period with a short displacement. The dual motion mechanism reported here is composed of two grooved cams that cross each other (double grooved cams). A pin positioned at the intersection point of the grooved cams rotates a lever linked to both contacts while changing the lever ratio that shortens the path length of the pin movement. To implement an optimized displacement curve with low operating energy and low mechanical stress while keeping the CCS performance high, a shape optimization method was developed that uses a direct search to minimize the local contact forces acting on the contact positions between the grooved cams and the pin. In order to maintain the stability of the pin in motion, a position holding part was designed by considering size of the gaps between the grooved cams and the pin. The measured displacement curve was in good agreement with the ideal one. In addition, a full-scale prototype was fabricated that successfully passed a 10,000-motion test.*

## 1. Introduction

This paper is an extension of the work originally presented at the 2019 5th International Conference on Electric Power Equipment - Switching Technology (ICEPE-ST) [1]. The changes from the conference paper are as follows. First, in section 2, the role of a high-voltage gas circuit breaker is described in Figure 1 to clarify the relationship with the target motion mechanism. Second, in section 3, the design method of the displacement curve is described in Figure 4. The method aims at achieving reductions in the operating energy and mechanical load while maintaining high capacitive current switching (CCS) performance as is required for the high-voltage gas circuit breaker. An optimal

solution is determined from the trade-off relationship between the operating energy and the mechanical load, as shown in Figure 8. Third, section 4 describes the method for optimizing the shape of the double grooved cams in detail by referring to the flowchart shown in Figure 9 and the position-holding parts for stabilizing the movements of the cams shown in Figure 14. In addition, the critical intersection angle of the double grooved cams is described in Figure 13. Furthermore, a comparison of optimization results for two different initial configurations and an analysis of the convergence process of the optimal calculation are given in Figure 15. Finally, section 5 describes the method used to verify the operation of the double grooved cams in Figure 17.

This paper contributes to design studies of general motion mechanisms by proposing the motion mechanism suitable for the

\*Corresponding Author: Masanao Terada, Email: [masanao.terada.br@hitachi.com](mailto:masanao.terada.br@hitachi.com)

gas circuit breaker, optimizing the displacement curve that contributes to low operating energy, and providing the shape optimization method of the motion mechanism.

High-voltage circuit breakers require low operating energy for the sake of high reliability and low cost. Several manufacturers have proposed motion mechanisms to increase separation speed of contact with low operating energy [2-4]. Reference [2] describes a double-speed mechanism which moves a movable-side contact linked to the operating mechanism quickly with an auxiliary cam. Moreover, reference [3,4] describes dual motion systems that drive two contacts in opposite directions from each other. The dual motion mechanisms can shorten the moving distance of the movable-side (drive-side) contact, which speeds up the heavy drive-side moving parts. By driving a light contact on the opposite side (driven-side) parts, which is fixed in the conventional single motion system, the relative speed between the driven-side contact and drive-side contact becomes sufficiently high with lower operating energy.

CCS is one of the interrupting duties, which requires withstanding of a recovery voltage to be applied during a half-cycle from contact separation. In particular, CCS requires a high gas density inside the insulation nozzle that covers the driven-side contact and a low electric field between the contacts. To increase the gas density between contacts, the density drop inside the nozzle due to flow acceleration caused by the supersonic flow should be suppressed [5-11]. To decrease the electric field between the contacts, their separation in the period of the half-cycle from contact separation has to be increased; that is, the contacts must move at a fast speed. Reference [12] shows the optimal opening speeds for CCS depending on rated voltage. Reference [13] indicates that faster motions resulting in a reduced electric field and a longer nozzle throat length that avoids a density drop are effective ways to increase CCS performance. The first goal of this paper is to present the optimal motion properties for increasing CCS performance and the second one is to describe a motion mechanism that realizes these properties, presupposing a nozzle with a sufficiently narrow cross-sectional area between the inner nozzle wall and the driven-side contact so as not to affect the density drop.

To optimize the displacement curve for higher CCS performance, it is important to control the speed at every moment and the timing when the driven-side contact starts to move. Reference [14] shows that CCS performance is improved when the driven-side contact starts to move 3 ms before contact separation. Moreover, if the CCS performance with the optimum displacement curve is the same as that of the initial displacement curve, the operating energy with the optimum displacement curve is smaller than that with the initial displacement curve. To ensure reliability, it is also important to reduce mechanical stress. The higher the speeds of the contacts are, the higher the load acting on the dual motion mechanism becomes. This requires that the acceleration be limited at the beginning and the end of the motion. Therefore, it is desirable for a motion mechanism to be able to express any displacement curve. Accordingly, a new type of dual motion mechanism is proposed, which is composed of two arbitrarily shaped grooved cams crossing each other. A cam design method using an arbitrarily shaped spline function has been proposed [15]. Reference [16] applied this cam design method to a rotating cam

of an operating mechanism. The idea for the motion mechanism involving crossing several grooved cams comes from other fields and has been embodied in refuelling devices [17].

This paper describes the design of a displacement curve on the driven-side of the high-voltage circuit breaker in a way that minimizes the operating energy and mechanical load acting on the mechanism while keeping CCS performance high. Besides, it describes an optimal design and a mounting structure for a dual motion mechanism with double grooved cams. Furthermore, the operation of the mechanism is verified with a small-sized test setup and a real-scaled prototype of the high-voltage circuit breaker.

## **2. Proposed Dual Motion Mechanism**

Figure 1 shows the high-voltage gas circuit breaker with the dual motion mechanism. During normal operation in electric power transmission, the main drive-side and driven-side contacts are in contact with each other to conduct the normal load current. Upon interrupting a short circuit current, the drive-side main contact linked to the operating mechanism is pulled in the opening direction, and the contacts separate. Briefly, after the main contact separation, the arcing contacts separate. At this time, an arc discharge is generated between the arcing contacts. The arc is extinguished by gas blowing due to the pressure rise resulting from mechanical compression and heating by an arc in the pressure build-up chamber. The dual motion mechanism drives the driven-side arcing contact in the opposite direction of the driving direction of the drive-side arcing contact. If the relative speeds of the drive-side and driven-side contacts are similar to the speed of the drive-side arcing contact only, the operating energy can be lowered and the cost of the operating mechanism can be reduced because of the reduction in the drive-side speed.

Figure 2 shows schematic drawings of the proposed dual motion mechanism and its motion. The displacement curves of the motion in Figures 2 (a)-(d) are shown in Figure 3. The operating mechanism pulls a drive rod with a grooved cam (the movable grooved cam). Until the drive pin reaches the curve of the movable grooved cam, in the period between the beginning of the drive-side motion and timing (a), the lever does not rotate because the drive pin does not apply any driving force to the lever. Therefore, the driven rod is motionless. When the curved surface of the movable grooved cam comes into contact with the drive pin, the lever rotates and the driven rod starts to move. While the drive pin moves along the stationary grooved cam, which is cut into a guide plate, the lever continues to rotate, changing the lever ratio of the driven-side arm length to drive-side arm length. When the drive pin reaches the end of the stationary grooved cam, the motion of the driven rod is finished because the lever stops rotating. After timing (d), only the drive rod moves because the drive pin does not apply any driving force to the lever.

This dual motion mechanism with double grooved cams has three advantages. The first one is that an arbitrarily shaped displacement curve can be expressed. The second one is that the driving force of the drive pin can be converted efficiently into rotational force of the lever because the drive pin's moving direction matches the vertical direction of the lever for adjusting the shape of the cam curves. The third one is that the overall size of the dual motion mechanism can be reduced because the path

length of the pin movement is shortened due to the lever ratio changing during the lever rotation.

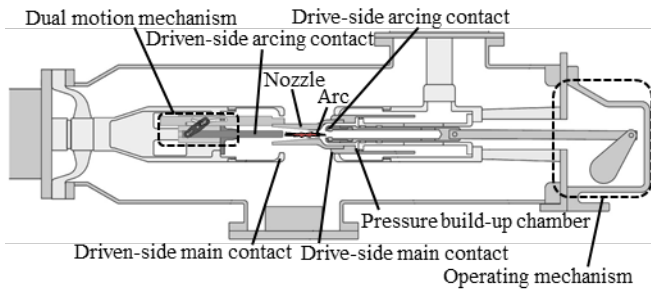


Figure 1: Overview of high-voltage gas circuit breaker with dual motion

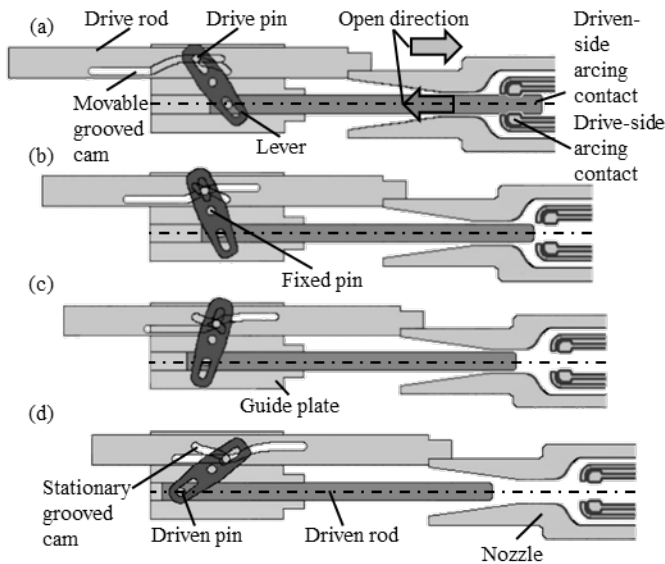


Figure 2: Schematics of the proposed dual motion mechanism

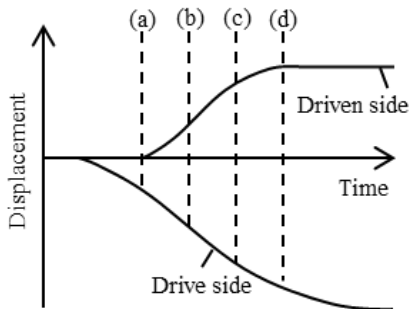


Figure 3: Drive-side and driven-side Displacement curves

### 3. Optimization of Driven-side Displacement Curve

To minimize the operating energy and reduce the mechanical load while keeping CCS performance high, it is desirable for optimizing the driven-side displacement curve.

#### 3.1. Method of Designing Driven-side Displacement Curve

The steps in the design method for the driven-side displacement curve is shown in the flowchart of Figure 4 and described in detail here.

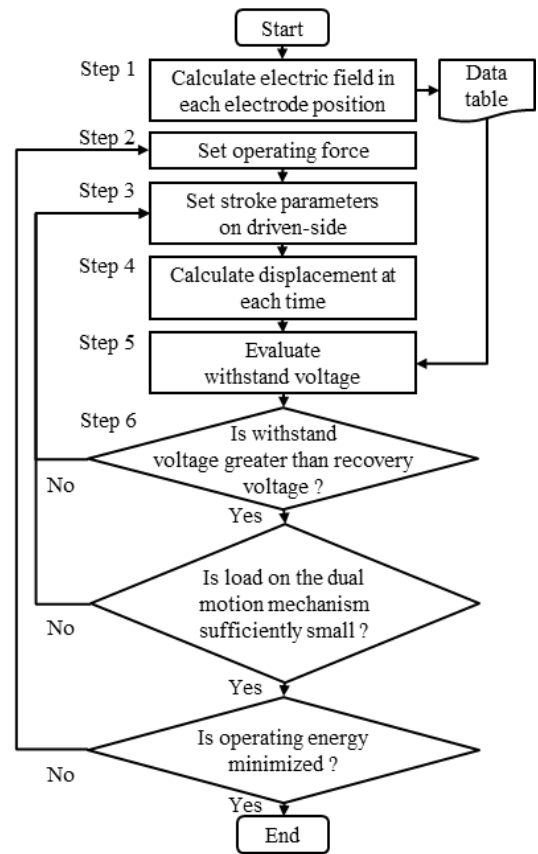


Figure 4: Design flow of displacement curve

Step 1: Calculate the electric field with the specified pitch from the position of the contact separation to the position of the half-cycle of applied voltage with power frequency, based on the electrode shape needed to withstand the lightning impulse voltage in the open position. The calculation results are saved as a data table.

Step 2: Set the operating force of the operating mechanism, which is defined by a spring force at the start position of the drive-side motion. The operating energy is calculated from the spring forces at the beginning and the end of the drive-side motion and the drive-side maximum displacement. The operating energy is proportional to the operating force which is calculated with the spring constant.

Step 3: Set the parameters of the displacement curve on the driven side in Figure 5 (a). Here, an intermittent motion with straight sliding parts of the movable grooved cam enables the start timing of the driven-side motion to be adjusted. The start timing and the end timing of the driven-side motion are the first and  $n$ -th motion points, respectively. The start timing and the end timing of the drive-side motion are the 0-th and  $(n + 1)$ -th motion points, respectively. To set an arbitrarily time-varying displacement curve, the ratio of the speed of the driven side to that of the drive side is expressed with a Catmull-Rom spline of degree 3 [18] in Figure 5 (b). The control points of the spline curve are the design parameters. The speed ratio  $l(t)_j$  between the  $j$ -th and  $(n + 1)$ -th control points with the Catmull-Rom spline curve is expressed as follows,

$$l(t)_j = \left(-\frac{1}{2}l_{j-1} + \frac{3}{2}l_j - \frac{3}{2}l_{j+1} + \frac{1}{2}l_{j+2}\right)t^3 + \left(l_{j-1} - \frac{5}{2}l_j + \frac{5}{2}l_{j+1} + \frac{1}{2}l_{j+2}\right)t^2 + \left(-\frac{1}{2}l_{j-1} + \frac{1}{2}l_{j+1}\right)t + l_j, \quad (1)$$

where  $l_j$  is the  $j$ -th control point of the speed ratio and  $t$  is a real number in the range from 0 to 1. When the width of the drive-side displacement is  $\Delta X_d$ , the  $i$ -th speed ratio  $L_i$  between  $l_4$  and  $l_5$  in Figure 5 (c) can be expressed as follows,

$$L_i = \alpha \frac{(x(t_{k+1})_4 - X_i)l(t_k)_4 + (X_i - x(t_k)_4)l(t_{k+1})_4}{x(t_{k+1})_4 - x(t_k)_4}, \quad (2)$$

where  $X_i$  is the  $i$ -th drive-side displacement,  $x(t_k)_4$  is the  $k$ -th drive-side displacement between  $x_4$  and  $x_5$ ,  $l(t_k)_4$  is the  $k$ -th speed ratio between  $l_4$  and  $l_5$ .  $\alpha$  is a coefficient to adjust the total amount of driven-side displacement, which is divided by the sum of  $L$  from the starting to the end of the displacement. The expression for  $\alpha$  is

$$\alpha = \frac{X_{dn,n}}{\sum_{i=1}^n L_i \Delta X_d}. \quad (3)$$

Step 4: Evaluate the time-variant displacement on the drive side and driven side given by the operating force  $F$  from the law of conservation of energy,

$$\frac{1}{2}M_d V_d^2 + \frac{1}{2}M_{dn} V_{dn}^2 = \int F \cdot V_d dt, \quad (4)$$

where  $V_d$  and  $V_{dn}$  are the drive-side speed and driven-side speed, respectively.  $M_d$  and  $M_{dn}$  are the drive-side mass and driven-side mass, respectively.  $V_{dn}$  is the product of  $V_d$  and  $L$ . The load acting on the dual motion mechanism is calculated from the driven-side mass and the acceleration obtained from Equation (4).

Step 5: Evaluate the withstand voltage  $V_{BD}$  at each contact position with the displacements of both sides obtained from Step 3 and the electric field at each position with the data table made in Step 1.  $V_{BD}$  is calculated using Equation (5) [19],

$$V_{BD} = 100 \cdot \frac{AP^n}{E_0}, \quad (5)$$

where  $E$  (%/mm) is the calculated value,  $P$  (MPa) is absolute pressure, and  $A$  and  $n$  are coefficients that vary in accordance with the gas condition and the shape of the gas circuit breaker.

Step 6: Optimize the parameters of the displacement curve on the driven side until the withstand voltage is greater than the recovery voltage and the load acting on the dual motion mechanism is sufficiently small.

Step 7: Minimize the operating force.

### 3.2. Design Example of Displacement Curve

The design parameters are the controlled points of the speed curve on the driven side. The calculation conditions are as follows:

- Frequency: 60 Hz
- Gas pressure: 0.6 MPa
- Maximum displacement ratio of driven- to drive-side: 0.28

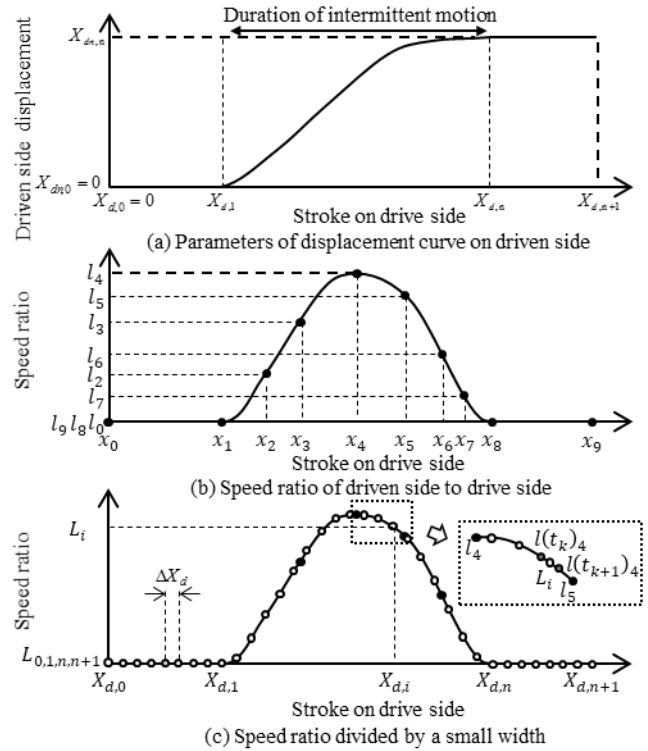


Figure 5: Expression of driven-side displacement curve

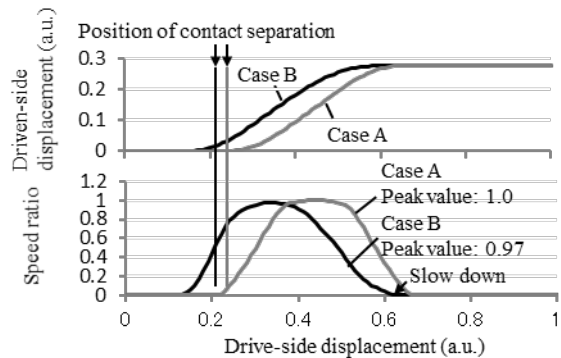


Figure 6: Input conditions for case A and case B: displacement curve (upper graph) and speed ratio (lower graph).

Figure 6 shows the input conditions for case A and case B. In case A, the movement starts at the same time as contact separation. In case B, the movement starts before contact separation, and the maximum speed ratio and acceleration in the latter half of the driven-side displacement decrease in comparison with case A. Figure 7 shows the results of the calculation. The operating energy in case B is lower than that in case A. The withstand voltages against recovery voltage for CCS in case A and case B are similar. This is because the driven-side speed at contact separation in case B is higher than in case A with the same operating energy. The mechanical load acting on the driven side in case B is lower than in case A because of the lower operating energy. To show that case B is the optimum solution, Figure 8 illustrates the relationship between the operating energy and the mechanical load. The driven-side speeds of the plotted squares are 1.2 times and 1.5 times faster than in case A, with the condition of the same timing of the motion. As the driven-side speed increases, the operating energy decreases



to have the same CCS performance. However, the mechanical load increases. In contrast, both operating energy and mechanical load are compatibly low in case B.

An optimized solution to improve CCS performance can be obtained by adjusting the driven-side displacement. The driven-side displacement of case B is assumed in section 4 to describe the design of the dual motion mechanism with the double grooved cams.

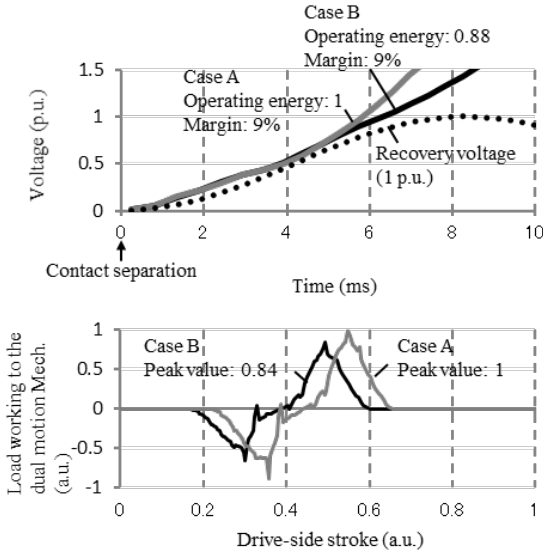


Figure 7: Results of calculation for case A and case B. The upper graph shows withstand voltages for CCS. The lower graph shows mechanical loads acting on the dual motion mechanism.

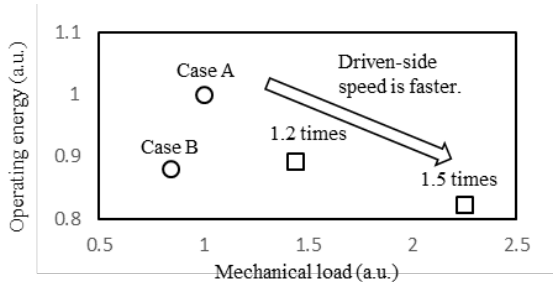


Figure 8: Relationship between operating energy and mechanical load acting on the dual motion mechanism

#### 4. Shape Optimization of Double Grooved Cams

##### 4.1. Method of designing the shape of double grooved cams

The dual motion mechanism with double grooved cams can be easily implemented in any shape because they have a high degree of design freedom. On the other hand, it is important to optimize the shape of the grooves because the motion mechanism must be compact and it must have mechanical strength and stability. In the following, the design method of the double grooved cams is described by referring to the flowchart shown in Figure 9.

Step 1: Input the drive-side and driven-side displacements shown in Figure 5.

Step 2: Set the dimensions for the dual motion mechanism with double grooved cams, which are the drive-side initial arm length of the link lever  $L_{d,0}$ , the initial lever angle  $\theta_0$ , the driven-side arm

length of the link lever  $L_{dn}$ , and the grooved cam width  $W$  shown in Figure 10. The end points  $P_0$  and  $P_m$  of the stationary grooved cam are determined from these structural parameters and the driven-side maximum displacement  $X_{dn,n}$ .

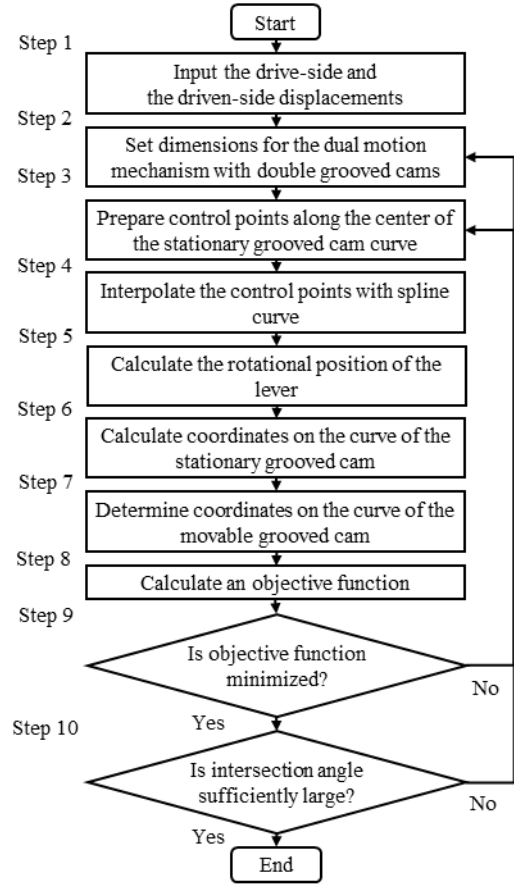


Figure 9: Design flow for shape of the dual motion mechanism with double grooved cams

Step 3: Prepare control points along the center line of the stationary grooved cam curve.  $P_i$  is the  $i$ -th control point of the stationary grooved cam ( $0 \leq i \leq m$ ). To smooth the shape of the stationary grooved cam, the shape of the curve from  $P_0$  to  $P_m$  must be a monotonic function. Therefore, the constraints between two neighbouring points are  $P_{i,x} > P_{i+1,x}$  and  $P_{i,y} > P_{i+1,y}$ , where  $P_{i,x}$  is the  $x$ -component of  $P_i$  and  $P_{i,y}$  is the  $y$ -component of  $P_i$ .

Step 4: Interpolate the control points with a Catmull-Rom spline of degree 3. The point  $P$  between  $P_i$  and  $P_{i+1}$  is described in accordance with Equation (1), using  $P_{i-1}, P_i, P_{i+1}, P_{i+2}$  and  $t$  which is a real number in the range of 0 to 1.

Step 5: Calculate the rotational positions of the link lever on the driven-side  $Q_{dn,j}$ , the rotational angle  $\theta_j$ , and the driven-side displacement  $X_{dn,j}$ , shown on the left side of Figure 11 by using Equations (6) and (7),

$$Q_{dn,j,x} = S_{dn} \sin \theta_j - X_{dn,j}, \quad (6)$$

$$Q_{dn,j,y} = -S_{dn} \cos \theta_j, \quad (7)$$

where  $\theta_j$  is expressed as



$$\theta_j = \arcsin\left(\frac{S_{dn} \cos \theta_0 - X_{dn,j}}{S_{dn}}\right). \quad (8)$$

Step 6: Calculate the intersection point of the link lever and the stationary grooved cam  $Q_{f,j}$  in accordance with the rotational points  $Q_{dn,j}$ , as shown on the right side of Figure 11.  $Q_{f,j}$  is described using the k-th point  $P(t_k)_i$  and k+1-th point  $P(t_{k+1})_i$  between the control points  $P_i$  and  $P_{i+1}$  and a parameter  $\alpha$  as follows,

$$Q_{f,j} = \alpha\{P(t_{k+1})_i - P(t_k)_i\} + P(t_k)_i, \quad (9)$$

$$\alpha = \frac{P(t_{k+1})_{i,x} - P(t_k)_{i,x}}{-P(t_k)_{i,x} + \tan \theta_j P(t_k)_{i,y}} - \tan \theta_j \frac{P(t_{k+1})_{i,y} - P(t_k)_{i,y}}{-P(t_k)_{i,x} + \tan \theta_j P(t_k)_{i,y}}. \quad (10)$$

Step 7: Determine the intersection point of the stationary grooved cam and the intersection point of the movable grooved cam  $Q_{d,j}$ , which is expressed as

$$Q_{d,j} = Q_{f,j} - X_{d,j}. \quad (11)$$

Step 8: Calculate an objective function to minimize the contact forces. In this dual motion mechanism, the drive pin is inserted at the intersection point of the stationary grooved cam and the movable grooved cam shifts while the operating mechanism pulls the movable grooved cam. Therefore, contact forces act at the points of contact between the drive pin and the surface of the grooved cams. For mechanical reliability, it is desirable for the contact force to be low. This mechanism has two possible ways to increase the contact force. The first is that the intersection angle between the double grooved cams becomes too small. The smaller the intersection angle is, the larger the contact force becomes, resulting in the possibility of a pin lock occurring during operation. This is called the wedge effect, as explained in section 4.2. The second possibility is that the contact stress becomes too large. The smaller the local curvature radius is, the larger the contact stress becomes because of the decreasing contact area. The intersection angle  $\varphi$  is determined to be the acute angle between the tangential line of the stationary grooved cam and the tangential line of the movable grooved cam. The local curvature radius  $R_s$  of the stationary grooved cam is determined as the radius of a circle calculated from three adjacent control points. The local curvature radius  $R_m$  is determined in the same way.  $\varphi$ ,  $R_s$ , and  $R_m$  are shown in Figure 12. The dimensionless objective function  $O$  is expressed as

$$O = G_1 \left(\frac{\varphi_t}{\min \varphi}\right) + G_2 \left(\frac{R_t}{\min R_f}\right) + G_2 \left(\frac{R_t}{\min R_m}\right), \quad (12)$$

where  $\varphi_t$  and  $R_t$  are criteria and  $G_1$  and  $G_2$  are weight coefficients. It is desirable for  $G_1 \gg G_2$  because a small intersection angle has the severe effect of increasing the contact force. The design described in section 4.3 sets  $G_1$  to 100 and  $G_2$  to 1. In addition,  $G_1$  is set to zero if the intersection angle is larger than the criterion. The criterion of the intersection angle is also evaluated in section 4.2.

Step 9: Iterate the procedures of Step 3 to Step 7 to minimize the objective function. The design variables are the  $x$  and  $y$

coordinates of the control points  $P$  on the stationary grooved cam. The Rosenbrock method [20] is used for optimization, which is a type of direct search. This method leads quickly to a convergent and stable solution, but there is a possibility that a local minimum solution is reached. To avoid a local minimum solution, the best one is selected from several solutions in accordance with several initial conditions.

Step 10: Perform the calculation by changing the initial dimensions if the criterion of the intersection angle is not exceeded.

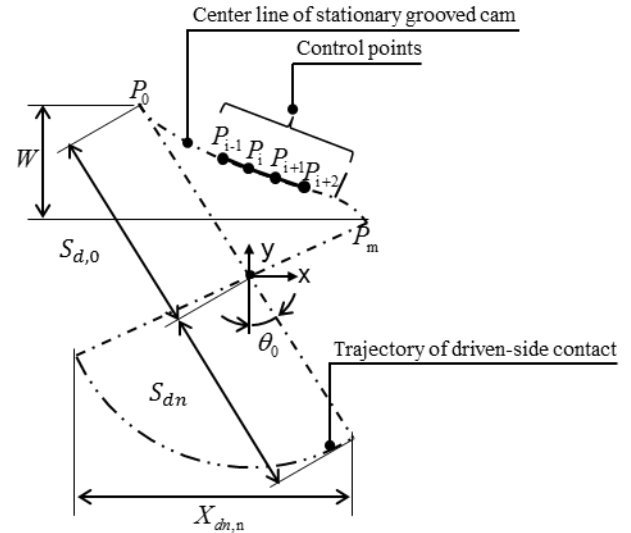


Figure 10: Structural parameters and control points for stationary grooved cam

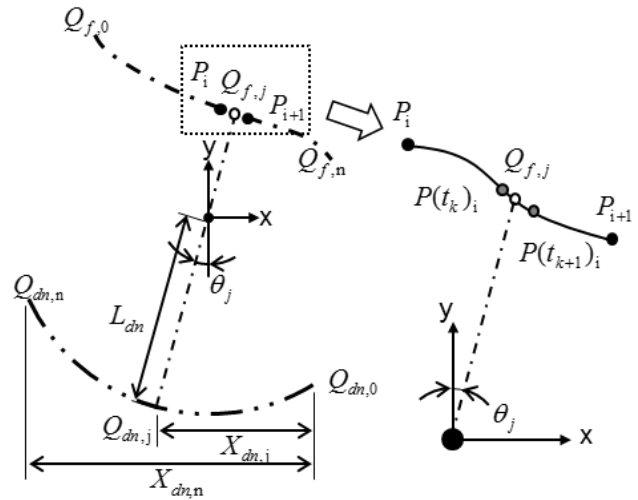


Figure 11: Moving positions on driven side and angles

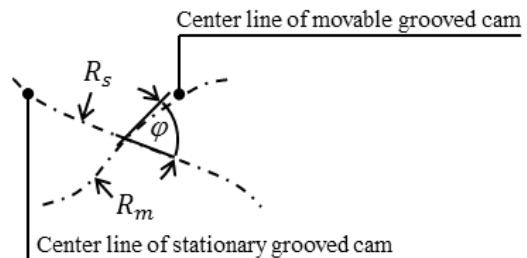


Figure 12: Intersection angle and curvature radii

4.2. Criterion of intersection angle between grooved cams

To design reliable double grooved cams, a criterion for the intersection angle of the stationary grooved cam and the movable grooved cam has to be established. The contact positions of the grooved cams and the drive pin generate nominal forces. The smaller the intersection angle, the larger the friction force induced by the nominal force. A large friction results in a pin lock due to the wedge effect during operation. This requires that the intersection angle be larger than a certain value.

Figure 13 defines the nominal forces and friction forces around the drive pin. When the lever (not illustrated) pushes the drive pin, the friction force acts on the contact position between the drive pin and the movable grooved cam. When the lever pushes the drive pin with a force  $F_0$  (smaller than the friction force), which is the non-slip condition, the drive pin presses the upper surface of the stationary grooved cam with a force  $F_0'$  ( $= F_0$ ). Therefore, a nominal force  $F_1'$  ( $=F_1$ ) acts on the contact position between the drive pin and the stationary grooved cam, as described by the following equation,

$$F_1 = \frac{F_0}{\sin \varphi} \tag{13}$$

Furthermore, a nominal force  $F_2'$  ( $=F_2$ ) acts on the lower surface of the movable grooved cam,

$$F_2 = \frac{F_0}{\sin \varphi} \cos \varphi = \frac{F_0}{\tan \varphi} \tag{14}$$

The friction force  $F_3$  is

$$F_3 = \mu F_2' = \mu \frac{F_0}{\tan \varphi} \tag{15}$$

where  $\mu$  is the friction coefficient. If  $F_3 > F_0$ , pin lock occurs during operation. Therefore, the locking condition is:

$$\tan \varphi < \mu. \tag{16}$$

If  $\mu$  is 0.3,  $\varphi$  must be more than 17 degrees. The criterion  $\varphi_t$  is decided with some margin for mechanical reliability.

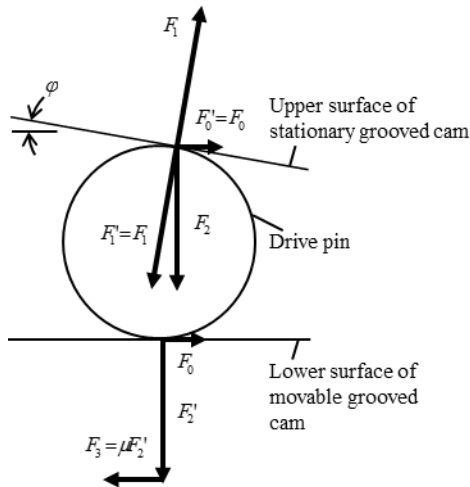


Figure 13: Contact forces and nominal forces on each contact surface

4.3. Method of structural design of double grooved cams

A method for designing the pin position holding part is presented for stabilizing the dual motion of the cam mechanism described in section 4.1. Figure 14 is a schematic drawing for distortion of the lever position caused by inclination of the drive pin. Pin rotation around the vertical axis in relation to the front view of Figure 14 caused by backlash in the grooved cams is considered. The inclination of the drive pin is defined by  $\delta$ . A small inclination of the drive pin occurs due to backlash between the other pins and holes. However, when  $\delta$  is greater than the backlash described above, a force resulting from the inclination of the drive pin is transmitted to the lever. Accordingly, the torque between the drive pin and the lever distorts the fixed pin and the driven pin and increases their stress. Finally, when the stress greatly exceeds the yield point, the pins deform plastically, resulting in destruction of the parts of the mechanism. Therefore, the inclination of the drive pin must be reduced. To reduce the inclination, pin holding parts are assembled at both ends of the pin. Even if the drive pin is inclined, the mechanism will not be destroyed because the pin holding parts touch the surface of the lever and force it to an ordinary position.  $\delta$  and the distortion between the pin holding part and the surface of the lever  $\Delta L_p$  are described as

$$\delta = L_p \sin \psi, \tag{17}$$

$$\Delta L_p = \frac{L_p}{2} (1 - \cos \psi) + \frac{D_h}{2} \sin \psi, \tag{18}$$

where  $\psi$  is the drive-pin inclination angle,  $D_h$  is the outer diameter of the pin holding part, and  $L_p$  is the distance between the ends of the pin holding parts. Once  $\psi$  is derived from Equations (17) and (18),  $\Delta L_p$  is described as follows:

$$\Delta L_p = \frac{L_p}{2} - \sqrt{\left(\frac{L_p}{2}\right)^2 - \left(\frac{\delta}{2}\right)^2} + \frac{D_h \delta}{2L_p}. \tag{19}$$

The threshold of the inclination of the drive pin  $\delta_t$  is

$$\delta_t = \delta_f + \frac{L_d}{L_{dn}} (\delta_f + \delta_n), \tag{20}$$

where  $\delta_f$  and  $\delta_n$  are inclinations of the fixed pin and the driven pin, respectively.  $L_d$  is the distance between the centers of the fixed pin and the drive pin.  $L_{dn}$  is the distance between the centers of the fixed pin and the driven pin. From Equation (19), the dimension  $D_h$  is determined to be a range such that  $\delta$  is less than  $\delta_t$ .

4.4. Design Example of Double Grooved Cams

The double grooved cams were designed to mitigate the contact forces acting on the contact positions between the grooved cams and the drive pin while keeping the overall size as small as possible so that the decided driven-side displacement curve in case B of Figure 6 was realized. Table 1 shows the design examples. Case I is the minimum-sized structure under a constraint such as on the diameter of the drive pin. One end of the stationary grooved cam is determined from the initial rotational angle  $\theta_0$  of the lever and the initial drive-side lever length  $S_{d,0}$ . The other end is determined from the width of the grooved cams  $W$  and the driven-

side maximum displacement  $X_{dn,n}$ . Ten control points between the ends of the stationary grooved cam are prepared. Compared with the initial intersection angle, the optimized intersection angle is 25 degrees, and it does not satisfy the criterion. To increase the intersection angle,  $\theta_0$  should be smaller and  $W$  should be wider in order to meet the requirement of inclining the stationary grooved cam curve steeply against the x-axis (Case II). Figure 15 shows the convergence process. The intersection angle after convergence is 40 degrees, much larger than the criterion, and the curvature radii are larger than the initial ones. The contact forces acting on the contact positions between the grooved cams and the drive pin are sufficiently small.

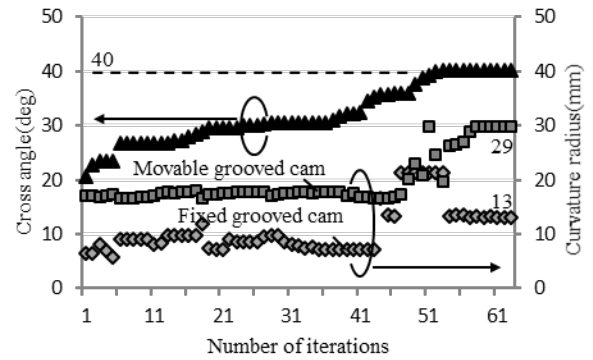


Figure 15: Convergence process

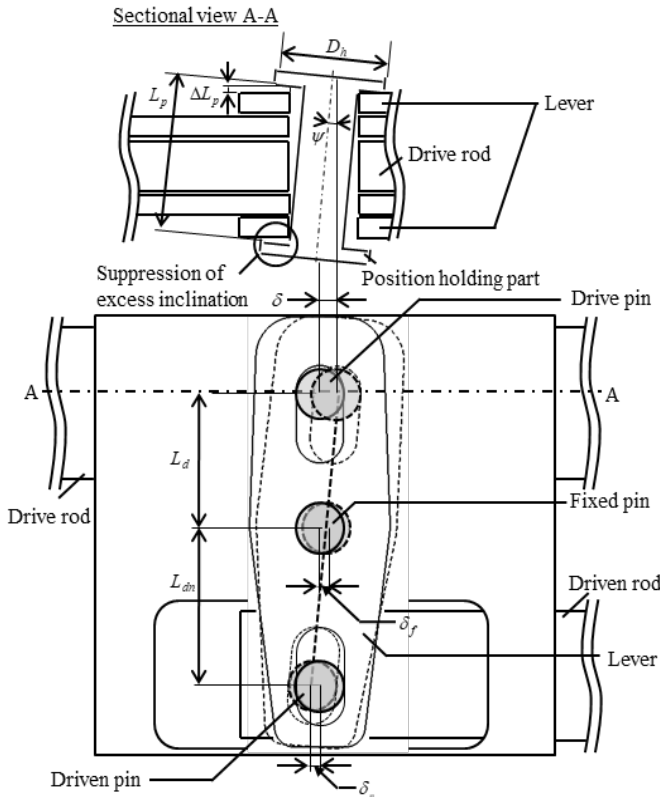


Figure 14: Schematic drawing of the double grooved cams

Table 1: Design results

Dimensions	Values			
	Case I		Case II	
	Init.	Opt.	Init.	Opt.
Lever ratio of driven-side length to drive-side length	1		1.2	
Initial angle of ling lever (degrees)	45		28.5	
Width of grooved cam (a.u.)	1		1.36	
Minimum intersection angle (degrees)	11	25	21	40
Minimum curvature radius in movable grooved cam (mm)	1	4	17	29
Minimum curvature radius in stationary grooved cam (mm)	2	2	7	13

#### 4.5. Design of pin holding parts

Figure 16 shows the relationship between the diameter of the pin position holding part and the drive-pin inclination expressed by Equation (19). As the threshold of the drive-pin inclination  $\delta_i$  is 0.0033 due to the fit tolerances of the fixed pin and the driven pin, the diameter of the position holding part is 0.157, i.e., 1.96 times larger than the drive pin.

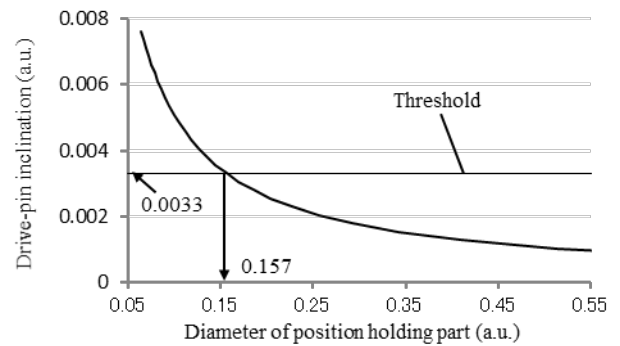


Figure 16: Drive-pin inclination versus diameter of pin position holding part

### 5. Motion verification of double grooved cams

Motion and endurance tests were performed to measure the mechanical reliability of the designed dual motion mechanism.

#### 5.1. Experimental setup

Figure 17 shows the experimental setup for the motion test. An interrupter unit is set vertically. The dual motion mechanism is assembled on the top of the interrupter unit. The operating device drives the drive-side moving parts of the interrupter unit vertically. When the drive-side moving parts move downwards, i.e., in an opening motion, the driven-side moving parts move upwards. It should be noted that the direction of the downward motion is the positive direction. The drive-side maximum displacement and driven-side maximum displacement are 150 mm and 50 mm, respectively. The operating device is a combination of a 4.4-kW servomotor and a link mechanism. The maximum speed is about half of the real-scaled high-voltage circuit breaker. The upward moving speed for returning to the initial position of the opening motion is much slower than the downward moving speed. Linear

potentiometers are set on the top of the dual motion mechanism to measure the drive-side displacement and the driven-side displacement. A strain gauge is attached on the guide plate to measure strain near the fixed pin where the maximum load acts.

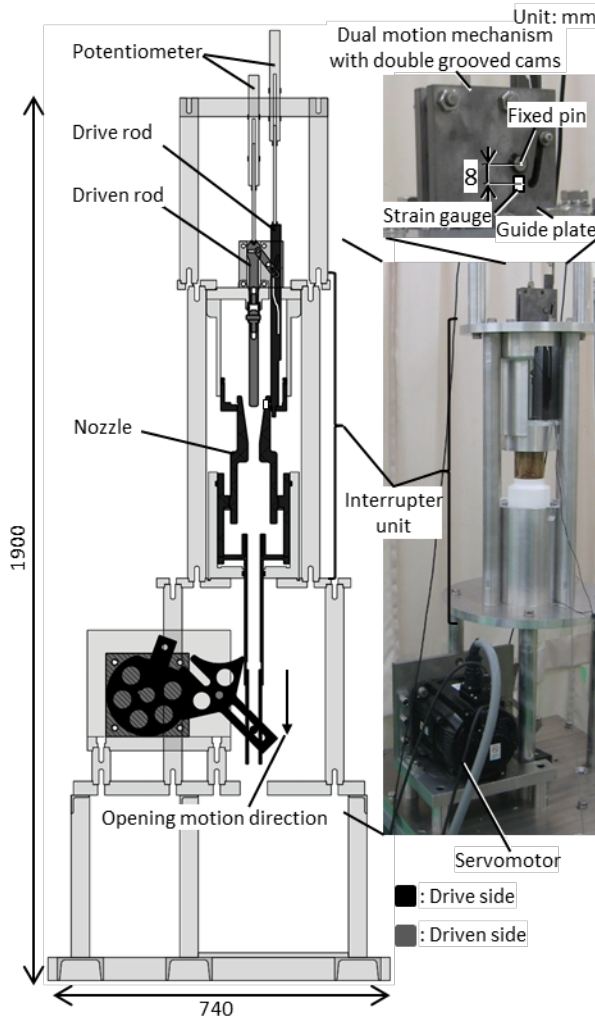


Figure 17: Experimental setup of motion test

### 5.2. Motion test results

Figure 18 compares the measured and calculated displacement curves. The results indicate that the displacement curve is in good agreement with the ideal displacement curve from 0 mm to -30 mm on the driven-side. In this period, the drive pin moves smoothly. After that, the deviation from the ideal curve is large, especially around the end of the driven-side motion. This is because the drive pin or the driven pin bends with elastic deformation due to inertia during the driven-side motion. Figure 19 shows the measured strain. Compressive strain occurred due to the reaction force at the start of operation on the driven-side. In the final stage of the operation on the driven-side, tensile strain occurs as a result of the upward force by inertia, and then compressive strain occurs as a result of downward force by gravity. The strains at the beginning and the end of the drive-side operation are almost 0%. This means that there is no residual strain, therefore the movement is not locked and it moves smoothly. Figure 20 shows deviations from the mean value of the maximum compressive strain on the guide plate in a 2,000-motion

test. Since all values fall within the range of 5%, the operation is stable and the dual motion mechanism with the double grooved cams is mechanically reliable.

In addition, an endurance test was performed on a full-scale prototype of the high-voltage circuit breaker. The prototype successfully passed a 10,000-motion test.

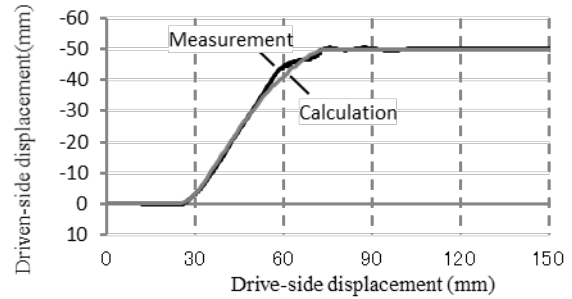


Figure 18: Comparison of measured and designed displacements

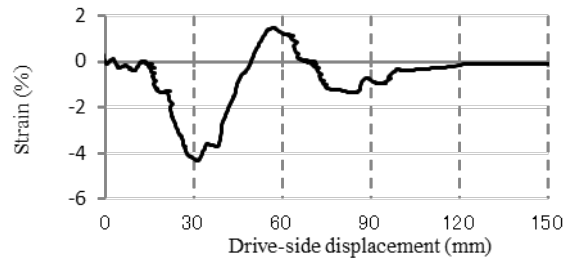


Figure 19: Tensile and compressive stress on the guide plate

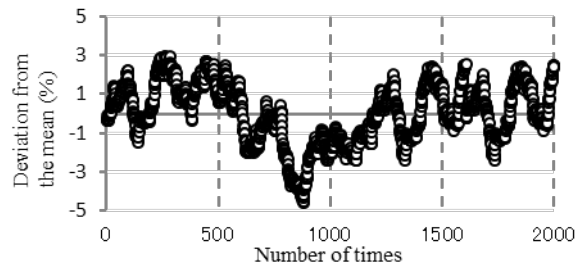


Figure 20: Deviation from the mean value of maximum compressive stress

### 6. Conclusion

The dual motion mechanism with double grooved cams that cross each other was proposed. A pin positioned at the intersection point of the grooved cams rotates a lever linked to both contacts while changing the lever ratio that shortens the path length of the pin movement. The proposed mechanism has three advantages. The first one is that an arbitrarily shaped displacement curve can be expressed. The second one is that the driving force of the pin can be converted efficiently into rotational force of the lever. The third one is that the overall size of the dual motion mechanism can be reduced. The following conclusions were obtained about: optimization of a displacement curve for the high-voltage circuit breaker, the shape of the grooved cams, and the mounting structure for the dual motion mechanism with double grooved cams.

- A displacement curve of the driven-side contact on the opposite side of the drive-side contact linked to the operating mechanism for the high-voltage circuit breaker was optimized,



as the driven-side contact operates quickly only during the capacitive current switching (CCS) period and starts to move before contact separation. The displacement curve can minimize the operating energy and mechanical load acting on the mechanism while keeping the CCS performance high.

- The optimal shape of the double grooved cams designed using the direct search algorithm can minimize the local contact forces acting on the contact positions between the grooved cams and the pin. Also, a position holding part designed in accordance with the gaps between the grooved cams and the pin can maintain the stability of the pin in motion.
- The measured displacement curve was in good agreement with the calculated displacement curve. In addition, since the variation of each measured strain in the 2,000-motion test is small, the operation is stable and the mechanism has high mechanical reliability. Furthermore, a full-scale prototype of the high-voltage circuit breaker successfully passed the 10,000-motion test.

## References

- [1] M. Terada, Y. Nakai, H. Hashimoto, D. Ebisawa, H. Urai, Y. Yokomizu, "Design of Dual Motion Mechanism Moving along Optimized Stroke Curve to Improve Capacitive Current Switching Performance for Gas Circuit Breaker" in 2019 5th International Conference on Electric Power Equipment - Switching Technology (ICEPE-ST), Kitakyushu, Japan, 2019. <https://doi.org/10.1109/ICEPE-ST.2019.8928703>
- [2] A. Ahmethodzic, R. P. P. Smeets, V. Kertesz, M. Kapetanovic, K. Sokolija, "Design Improvement of a 245-kV SF6 Circuit Breaker With Double-Speed Mechanism Through Current Zero Analysis," IEEE Transaction on Power Delivery, 25(4), October, 2010. <https://doi.org/10.1109/TPWRD.2010.2057262>
- [3] O. Hunger, "Circuit breaker with a gear having a dead point," U.S. Patent 8415578, 2013.
- [4] H. Dienemann, V. Lehmann, H. Marin, "High Voltage Circuit Breaker with Two Arcing Contacts Which Can Be Actuated in an Opposite Direction," U.S. Patent 6271494, 2001.
- [5] J. Y. Trepanier, M. Reggio, Y. Lauze, R. Jeanjean, "Analysis of the Dielectric Strength of an SF6 Circuit Breaker," IEEE Transactions on Power Delivery, 6, No. 2, April, 1991. <https://doi.org/10.1016/j.aasri.2014.05.029>
- [6] H. Joshi, A. Pandharker, G. Patil, "Optimization of High Voltage Arc Assist Interrupters," International Journal of Scientific & Engineering Research, 4(3), March, 2013. <https://doi.org/10.1109/TPWRD.2003.817536>
- [7] Q. Zhang, J. Liu, J. D. Yan, "Flow Structure Near Downstream Electrode of a Gas-Blast Circuit Breaker," IEEE Transactions on Plasma Science, 42(10), October, 2014. <https://doi.org/10.1109/TPS.2014.2309174>
- [8] J.D. Mantilla, C.M. Franck, M. Seeger, "Measurements and Simulations of Cold Gas Flows in High Voltage Gas Circuit Breakers Geometries," Electrical Insulation 2008. ISEI 2008. Conference Record of the 2008 IEEE International Symposium, 720-723, 2008. <https://doi.org/10.1109/ELINSL.2008.4570431>
- [9] H. K. Kim, K. Y. Park, C. H. Im, H. K. Jung, "Optimal Design of Gas Circuit Breaker for Increasing the Small Current Interruption Capacity," IEEE Transactions on Magnetics, 39(3), May, 2003. <https://doi.org/10.1109/TMAG.2003.810176>
- [10] H. K. Kim, J. K. Chong, K. Y. Park, "Approximation Model-Assisted Optimization Technique to Improve Capacitive Current Interrupting Performance of Gas Circuit Breaker," IEEE Transactions on Magnetics, 45, No. 3, March, 2009. <https://doi.org/10.1109/TMAG.2009.2012746>
- [11] M. T. Dhotre, X. Ye, F. Linares, P. Skarby, S. Kotilainen, "Multiobjective Optimization and CFD Simulation for a High-Voltage Circuit Breaker," IEEE Transaction on Power Delivery, 27(4), October, 2012. <https://doi.org/10.1109/TPWRD.2012.2207133>
- [12] N. Yamada, S. Taki, N. Osawa, Y. Yoshioka, "Investigation of an Automatic Design Method of the Optimum GCB Insulating Structure and Necessary Opening Speeds for Successful Capacitive Current Switching," Proceeding of the ICEE 2004, Sapporo, Japan, 7(2), 901-906, 2004. <https://www.researchgate.net/publication/267371487>
- [13] S. Delic, D. Beslija, D. Gorenc, A. Hajdarovic, M. Kapetanovic, "Capacitive current breaking capability estimation for a 145kV 40kA GIS circuit breaker," 2015 3rd International Conference on Electric Power Equipment - Switching Technology (ICEPE-ST), October 25-28, 2015. <https://doi.org/10.1109/ICEPE-ST.2015.7368412>
- [14] M. Kang, K. H. Kim, T. Ohk, Y. T. Yoon, "Analysis of Insulation Coordination for Dual Motion Mechanism Gas Circuit Breaker," 3rd International conference on Electric Power Equipment - Switching Technology (ICEPE-ST), 2015. <https://doi.org/10.1109/ICEPE-ST.2015.7368319>
- [15] D. M. Tsay, C. O. Huey, "Cam Motion Synthesis Using Spline Functions," Trans. ASME, Journal of Mechanisms, Transmissions, and Automation in Design, 110(2), 161-165, June, 1988. <https://doi.org/10.1115/1.3258921>
- [16] J. S. Jang, J. J. Lee, J. H. Sohn, H. W. Kim, B. T. Bae, W. S. Yoo, "A Cam Profile Design of a Circuit Breaker by Using Multibody Dynamics Analysis," 2015 3rd International Conference on Electric Power Equipment - Switching Technology (ICEPE-ST), October 25-28, 2015. <https://doi.org/10.1109/ICEPE-ST.2015.7368440>
- [17] A. Pourghodrat, C.A. Nelson, "A case study of designing a cam-follower mechanism with cycloidal motion," 13th World Congress in Mechanism and Machine Science, 2011. <https://www.dmg-lib.org/dmglib/handler?docum=22419009>
- [18] E. Catmull, R. Rom, A class of local interpolating splines, Computer Aided Geometric Design, Academic Press, New York, 317-326, 1974.
- [19] F. Endo, M. Sato, M. Tsukushi, Y. Yoshioka, K. Saito, K. Hirasawa, "Analytical prediction of transient breakdown characteristics of SF6 gas circuit breakers," IEEE Transaction on Power Delivery, 4(3), July, 1989. <https://doi.org/10.1109/MPER.1989.4310811>
- [20] D. M. Himmelblau, Applied Nonlinear Programming, McGraw Hill, New York, 1972.



## Design of an EEG Acquisition System for Embedded Edge Computing

Kanishk Rai<sup>1,\*</sup>, Keshav Kumar Thakur<sup>1</sup>, Preethi K Mane<sup>1</sup>, Narayan Panigrahi<sup>2</sup>

<sup>1</sup>Department of Electronics and Instrumentation, BMS College of Engineering, 560019, India

<sup>2</sup>Centre for Artificial Intelligence and Robotics, DRDO, 560093, India

### ARTICLE INFO

Article history:

Received: 17 April, 2020

Accepted: 17 June, 2020

Online: 12 July, 2020

Keywords:

Electroencephalogram (EEG)

Internet of Things (IOT)

Brain Computer Interface (BCI)

Acquisition system

Edge Computing

### ABSTRACT

The human brain is one of the most complex machines on the planet. Being the only method to get real-time data with high temporal resolution from the brain makes EEG a highly sought upon signal in the neurological and psychiatric domain. However, recent developments in this field have made EEG more than just a tool for medical professionals. The decreasing size and increasing complexity of EEG acquisition systems have brought it out of the lab and into the field where it is used for varied applications like neurofeedback, person recognition and other recreational activities. Amalgamation of the EEG signal with new developing standards of Industry 4.0 to control basic IOT devices using edge computing techniques marks the next step in the design and development our low-cost yet robust Brain Computer Interface (BCI); which is just one of the many applications that a versatile and well-built EEG acquisition system can be used for.

## 1. Introduction

What makes the human brain so complex is not just the 100 billion neurons that it is constituted of but also the 100 trillion plus synapses or the unique ways in which those neurons connect and communicate with each other. Even with more than 7.5 billion people on the planet, no two brains are alike. Therefore, designing a unified device that can capture the diverse Electroencephalogram (EEG) signals emanating from different brains is a challenging task. EEG signals can be processed to extract the characteristic patterns of the brain to study its spatio-temporal characteristics.

Though other methods of measuring brain activity like Magnetic Resonance Imaging (MRI) and Functional Magnetic Resonance Imaging (fMRI) have more spatial resolution yet, they lack temporal resolution and are often more costly to setup and operate. EEG on the other hand offers better temporal resolution and is a cost-effective solution which can directly measure brain activity. EEG can approximate spatial activity with the use of greater number of electrodes placed on the scalp, providing better resolution and thus, making it a scalable choice.

A low cost yet robust design of an EEG acquisition system was presented at the *IEEE International Symposium on Smart*

*Electronic Systems (iSES) (Formerly iNiS)*, 2019 [1]. Though the design was primarily focused on reduction of cost of manufacturing so that the system becomes affordable, yet there

were many limitations of the design. We studied the base design and propose a refined design which will overcome the followings drawbacks of the initial design:

- Overcome the limitations of low sampling rate
- To perform filtering and feature extraction from acquired EEG signal on the board itself
- To reduce and eliminate the noise during the EEG signal acquisition process by reducing wires in the system
- Better processing mechanism to extract and interpret the EEG data acquired, in real-time

The whole system has now been fabricated on a Printed Circuit Board (PCB) and the primary mode of communication between the acquisition system and the host computer is now wireless to minimize artefacts introduced due to wire movements and other Electromagnetic Interference (EMI) / Electromagnetic Compatibility (EMC) interference. The detailed design update is enumerated in the system design section.

The next section discusses current research trend in the field of BCI (Brain Computer Interface) and HMI (Human Machine Interface) using EEG, followed by the Problem Definition. Next, we briefly look at the testing results from our initial design and highlight its shortcomings. We then move on to introduce the updated block diagram and directly compare the initial and the updated components. This serves the double purpose of

\*Kanishk Rai, +91-8853389016, kr2252@gmail.com

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050416>

introducing the initial design while highlighting the updates in the improved model. We next discuss the results obtained during testing of the improved EEG acquisition system and finally, we illustrate one of the many applications of our system - which is to automatically control IoT devices such as the television, lights, wheel chair and fan depending upon the state of the subject e.g. active, semi active, asleep in real time.

## **2. Literature Review**

### *2.1. Background*

Since the first experiments of electroencephalography on humans in 1929, the electroencephalogram (EEG) of the human brain has been used mainly to evaluate neurological disorders in the clinical environment and to investigate brain functions in the laboratory. An idea that brain activity could be used as a communication channel between the subject and an IoT device has gradually emerged [2-6]. The possibility of recognizing a single message or command, considering the complexity, distortion, and variability of brain signals appeared to be extremely remote. Yet EEG demonstrates direct correlations with user intentions, thereby enabling a direct brain-computer communication channel.

Brain-Computer Interface (BCI) technology, as it is known, is a communication channel that enables users to control devices and applications without the use of muscles. The development of cognitive neuroscience field has been instigated by recent advances in brain imaging technologies such as electroencephalography, magneto encephalography and functional magnetic resonance imaging. The first BCI prototype was created by Dr. Vidal in 1973. This system was intended to be used as a promising communication channel for persons with severe disabilities, such as paralysis, Amyotrophic Lateral Sclerosis (ALS), stroke or cerebral paralysis.

Since then BCI research has been successfully used not only for helping the disabled, but also as being an additional data input channel for healthy people. It can be exploited as an extra channel in game control, augmented reality applications, household device control, fatigue and stress monitoring and other applications. BCI design represents a new frontier in science and technology that requires multidisciplinary skills from fields such as neuroscience, engineering, computer science, psychology and clinical rehabilitation.

### *2.2. Review of EEG for BCI Applications*

Controlling real-world objects with nothing more than our mind is no longer science fiction. Jaime, et al [7] have proved the same by demonstrating the fact that the steering of a tractor by use of surface EOG and EEG signals is just as accurate as manual guidance or GPS guidance with a minimal variation in response time.

Furthermore, Yuan, et al [8] have conclusively established the feasibility of using a consumer-level EEG headset to realize an online steady-state visual-evoked potential (SSVEP)-based BCI during human walking using a 14 channel commercially available EEG headset to implement a four-target online SSVEP decoding system which has in-turn helped bridge the gap between laboratory BCI demonstrations and real-life application of consumer-level EEG headsets.

Starting with some simple applications, Gerardo Rosas-Cholula, et al [9] have successfully controlled a mouse pointer

using a commercially available EEG headset. They have implemented Empirical Mode Decomposition (EMD) to reduce noise and simulated mouse clicks with the help of blinks. Next we come across articles [10-14] that focus on controlling robotic arms or wheelchairs for amputees and paraplegics using different computational techniques for identification and classification of various signals in the EEG like Event-related Potential (ERP), P300, Mu rhythm, etc. and demonstrate the accuracy and ease with which these EEG based BCIs can control simple devices.

Taking this use of EEG one step forward multiple articles [15-21] discuss integrating EEG with the Internet of Things (IOT) to control different devices in a smart home using protocols like Message Queue Telemetry Transport (MQTT) and Extensible Messaging and Presence Protocol (XMPP). These interfaces can further be used to improve healthcare and the general living standard of differently-abled people and in-turn make everyone's life easier. However, there is much work still to be done in this domain considering the new threats that continuously emerge in the realm of cyber security.

Researchers have worked upon and built multiple EEG acquisition systems [22-28] but have not used these systems to demonstrate an application. Others have demonstrated the use of EEG for various applications be it in terms of person recognition [29-32] or in the field of healthcare [33-38] or the ones mentioned above, but they have either used commercially available EEG acquisition systems or pre-existing datasets. With our research we aim to bridge this gap by not only building a robust yet low-cost EEG acquisition system but also using the same system to control IoT devices in real-time, and in doing so we not only design the hardware to acquire EEG data but also develop the software for data representation and signal processing which showcase the edge computing capabilities of our device.

## **3. Problem Definition (Limitations of Existing Designs of EEG Acquisition Systems)**

The research trends towards problems addressed in the EEG and BCI/HMI domain [1] show us that only 3 out of 30 publications focus on data acquisition and 5 out of 30 discuss EEG signal processing. However, this can paint a slightly misleading picture. To fully understand the need to design and develop a low-cost yet robust EEG acquisition system capable of edge processing we need to look at the research trends towards adaptation of technology where only 2 out of 21 publications [1] use custom-built EEG acquisition systems, while others rely on expensive alternatives available in the market or pre-existing datasets [1].

The above gaps in research are the major motivation behind this research, but the problems encountered while designing a robust yet low-cost EEG acquisition system are two-fold:

Domain specific problems are: -

- To ensure that the pickup, amplification and transmission of EEG signal is done while introducing minimum noise in the signal.
- To ensure that the sampling rate of the acquisition system is at least 2.5 times the highest frequency of input EEG signal, but at the same time not too high so as to avoid the problem of oversampling.

- To ensure that the resolution of the system is enough to register minute changes in the peak-to-peak amplitude and the frequency of the EEG signal.

Problems specific to the design of our acquisition system are:

- To build a Graphical User Interface (GUI) that accurately represents the EEG signal picked up by the acquisition system in real-time.
- To perform real-time processing and feature extraction from EEG signal in order to build a device capable of edge processing.

#### 4. System Design

After careful examination of EEG signals and other medical grade systems available in the market, we have arrived at the following design-constraints to design our EEG acquisition system.

Design constraints:

- EEG voltage range: 10  $\mu$ V to 100  $\mu$ V
- Useable EEG frequency range: 0.3 Hz to 60 Hz
- Max variation of input signal per unit time: 50 $\mu$ V
- Sampling Rate: 256 to 1k SPS

The subsequent sections talk about the step by step design of each component in the acquisition system. Beginning with an introduction to the initial system design, its results and drawbacks.

##### 4.1. Initial Block Diagram, Testing and Observations

For the initial design [1] as shown in Figure 1, gold plated, cup shaped electrodes (placed on the scalp using the international 10-20 system) were connected to the instrumentation amplifiers (INA333) which were then directly connected to the analog input pins of Arduino Mega. We only used 3 input channels; two for data and one for reference. Arduino IDE however, is neither capable of plotting data from multiple channels nor is it capable of performing complex algorithms on the acquired data.

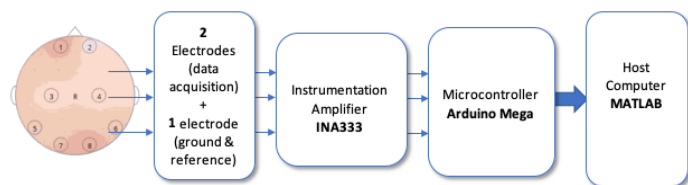


Figure 1: Initial block diagram

Therefore, to overcome the abovementioned drawbacks, we interfaced Arduino with MATLAB, which is used to plot real time frequency domain and time domain spectrum of the incoming EEG signal with further possibility of manipulating data and representing it in different formats.

Figure 2 shows the acquired EEG data from a single channel connected to analog port A0 of the Arduino at a baud rate of 9600. Subplot 1 shows the real time RAW EEG signal amplified 101 times whereas subplot 2 shows the real time Frequency plot (FFT).

A brief description of output from the initial design of our EEG acquisition system is as follows:

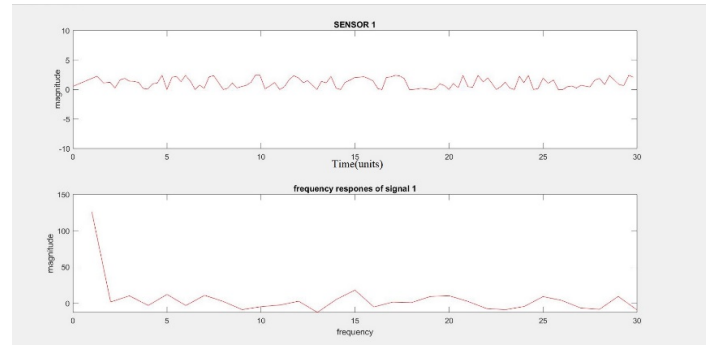


Figure 2: Data Acquisition in MATLAB interfaced with Arduino

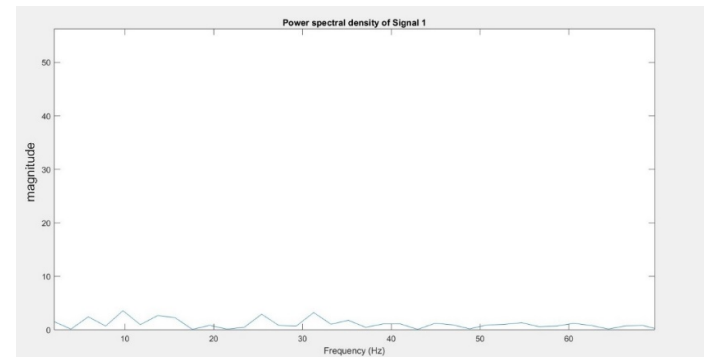


Figure 3: Power spectral density of data acquired from single sensor coming from instrumentation amplifier with a gain of 101

Once the data acquisition is complete, MATLAB then computes the power spectral density of the acquired EEG signal as observed in Figure 3. Major peaks seen around 5, 10, 25 and 32 Hz correspond to Alpha, Beta and Theta waves which contribute majorly to awake EEG patterns.

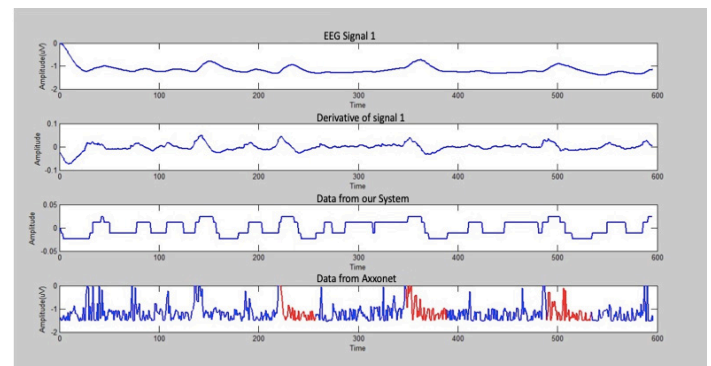


Figure 4: Comparison of data acquired from a medical grade system and our proposed system

Figure 4 shows the comparison of saccade, fix and blink data. Subplot 1 shows the data from our system, while Subplot 2 shows the data from the medical grade system. Similar for Subplot 3 and 4. It is evident from the above figure that our proposed system captures the correct data as all peaks match between both signals. The drawback here is the low sampling rate because of which the reconstructed signal does not have as much detail as is required for complex analysis – this can be seen when comparing Subplot 3 and 4. The maximum rate at which the data is captured by our initial system is about 5 to 20 Hz [1].



#### 4.2. Drawbacks of the Initial System

Due to the low sampling rates which are in-turn caused by the hardware limitations of the 10-bit ADC on Arduino Mega2560, we were only able to capture data at a maximum rate of 20 Hz. But the Practical Nyquist Sampling Theorem states that to properly capture and reconstruct incoming signal, the sampling frequency of the acquisition system must be at-least 2.5 times the maximum frequency component of the incoming signal, therefore to capture brain waves with a maximum useable frequency component of 60 Hz, the acquisition system must have a sampling rate of 150 Hz or higher. To overcome this issue of lower sampling rate, we propose the use of an ADC (ADS1299) introduced after the instrumentation amplifier and before the microcontroller as shown in the block diagram of our initial design (Figure 1).

The microcontroller Arduino Mega does not have the inherent capability to transmit data wirelessly meaning more noise is introduced in the signal due to the movement of the wires. This is overcome by replacing Arduino with NodeMCU that has an inbuilt wireless transmitter.

Moreover, the use of MATLAB for signal processing and representation poses its own set of limitations including the need for a MATLAB license to use our system, hence the same is replaced by a GUI built using open-source code in Python.

#### 4.3. Improved Block Diagram

The initial components of the system like the Electrodes (gold-plated, cup shaped, copper, non-invasive), Electrode Placement (international 10-20 system), Electrode Cap and the Instrumentation Amplifier (INA333) remain the same and are carried over from the initial design (Figure 1).

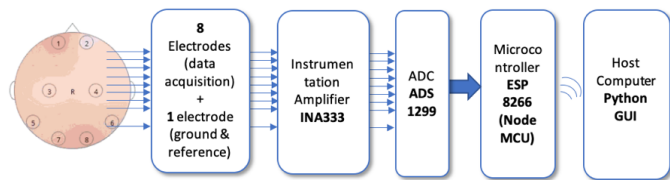


Figure 5: Updated Block Diagram

Figure 5 shows the block diagram of the updated hardware with an Analog to Digital Converter introduced in the system after the instrumentation amplifier to handle the conversion of analog data to digital domain at a much faster rate. The ADS1299 connects to the NodeMCU using SPI, which then transmits the data wirelessly to the host computer that runs a Python GUI for data processing, filtration and signal representation.

#### 4.4. Analog to Digital Converter (ADC)

The ADC chosen for our application is the TI analog front-end chip ADS1299, which is specifically designed to capture biopotential signals like EEG, ECG, EOG, EMG, among others. The chip has a 24-bit Sigma-Delta based architecture, with a maximum data rate of 16 kSPS. It boasts of a -110 dB CMRR and a Signal to Noise Ratio of 121 dB. It is a low power, low noise, high precision chip. It can further be multiplexed to use more channels and is available in 4, 6 and 8 channel alternatives.

The chip communicates with the microcontroller using SPI (Serial Programming Interface) which can be used to configure the chip to read data in Continuous mode or Single Shot mode and also to read the output data from the inbuilt registers.

#### 4.5. Microcontroller

Arduino Mega from the initial design has now been replaced with ESP 8266 NodeMCU as the choice of microcontroller for our system. The NodeMCU has a L106 32-bit RISC microprocessor running at 80 MHz. It supports 32 KB instruction RAM and 80 KB user-data RAM. It comes with integrated IEEE 802.11 b/g/n Wi-Fi connectivity and matching network WEP or WPA/WPA2 authentication. It has 16 GPIO pins and also supports SPI communication – all of which is ideal for our application. The NodeMCU does not only have more GPIO pins, better clock speed of the processor and more flash and SRAM than the Arduino, but also operates at a lower 3.3 V which translates to less power consumption for battery powered wireless use of our EEG acquisition system.

#### 4.6. Initial Observations with ADS1299

To set up and test the SPI connection between ADS1299 and NodeMCU, ADS is soldered on an adaptor (Figure 6) and the connections are made on a bread-board (Figure 7). It must be noted that for this initial testing, ADS is connected to Arduino. Once the connections are verified, the final PCB is interfaced with NodeMCU.

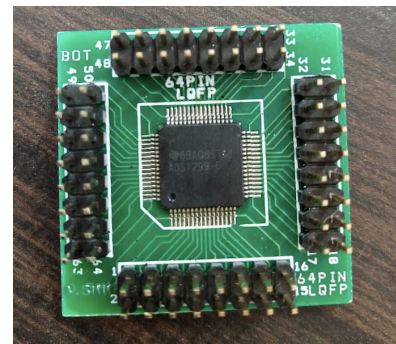


Figure 6: ADS1299 soldered on an adaptor

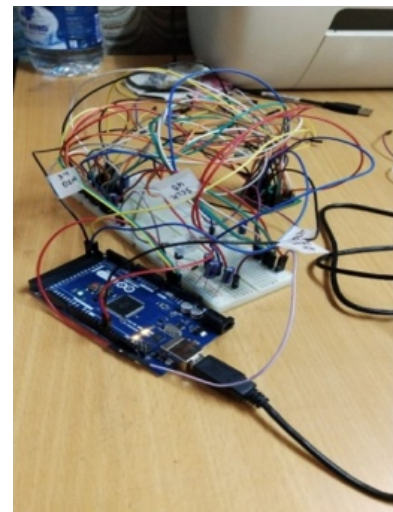


Figure 7: ADS1299 connected to Arduino Mega using SPI on a bread board

After running multiple tests Arduino Mega is able to read the correct Device ID (Figure 8). The device ID reads 1111 1100 where the last four bits correspond to the 4 channel ADS1299 chip. This means that the SPI communication was successful, but due to the inherent noise of the system which is set-up on a bread-board, the output is not consistent.

Therefore, the next logical step is to design a PCB to acquire the signals. This is carried out in the next subsection.

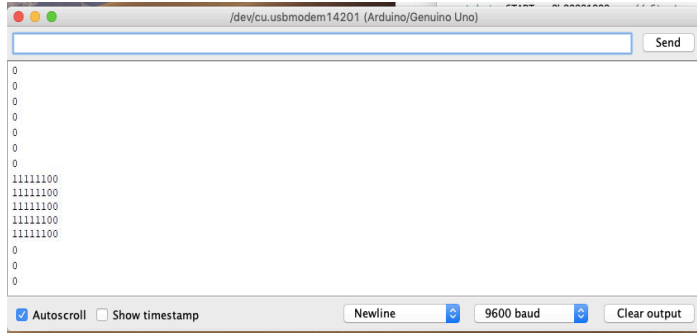


Figure 8: Correct device ID of ADS1299-4PAG displayed on serial monitor of Arduino IDE connected to Arduino

#### 4.7. Design for PCB Fabrication

The PCB is designed using an open-source software called PCB Artist. The schematic for the design is shown in Figure 10. The design comprises of a power supply unit which steps down the incoming 220V AC to 5V and 3.3V DC. Both the 5V and 3.3V lines have been taken out as well to provide power to the

microcontroller NodeMCU and for other purposes. Internally, the ADC has multiple pins that require 5V and 3.3V to power up the Analog and Digital circuits in the chip respectively.

For the input electrodes, each channel is first passed through the non-inverting pin of the instrumentation amplifier INA333 for first stage amplification. The signal is then passed through an RC high-pass filter to remove DC offsets and fed to the ADC via referential montage with one reference electrode and an internal bias. The example Schematic using the ADS1299 in an EEG Data Acquisition Application in Referential Montage from its datasheet is shown in Figure 9. The advantage of using the referential montage is that all electrodes are referenced to a single reference electrode instead of being referenced to each-other as is the case with sequential montage.

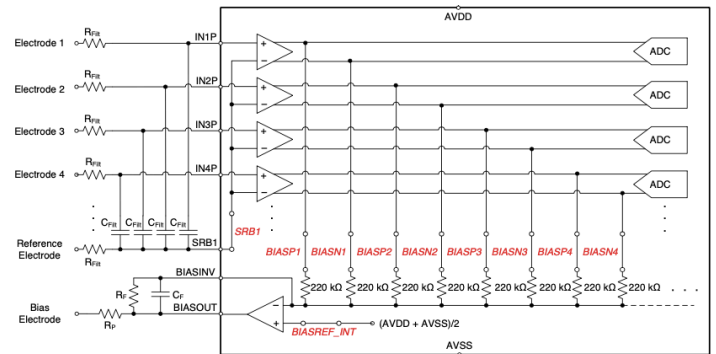


Figure 9: Use of ADS1299 in referential montage

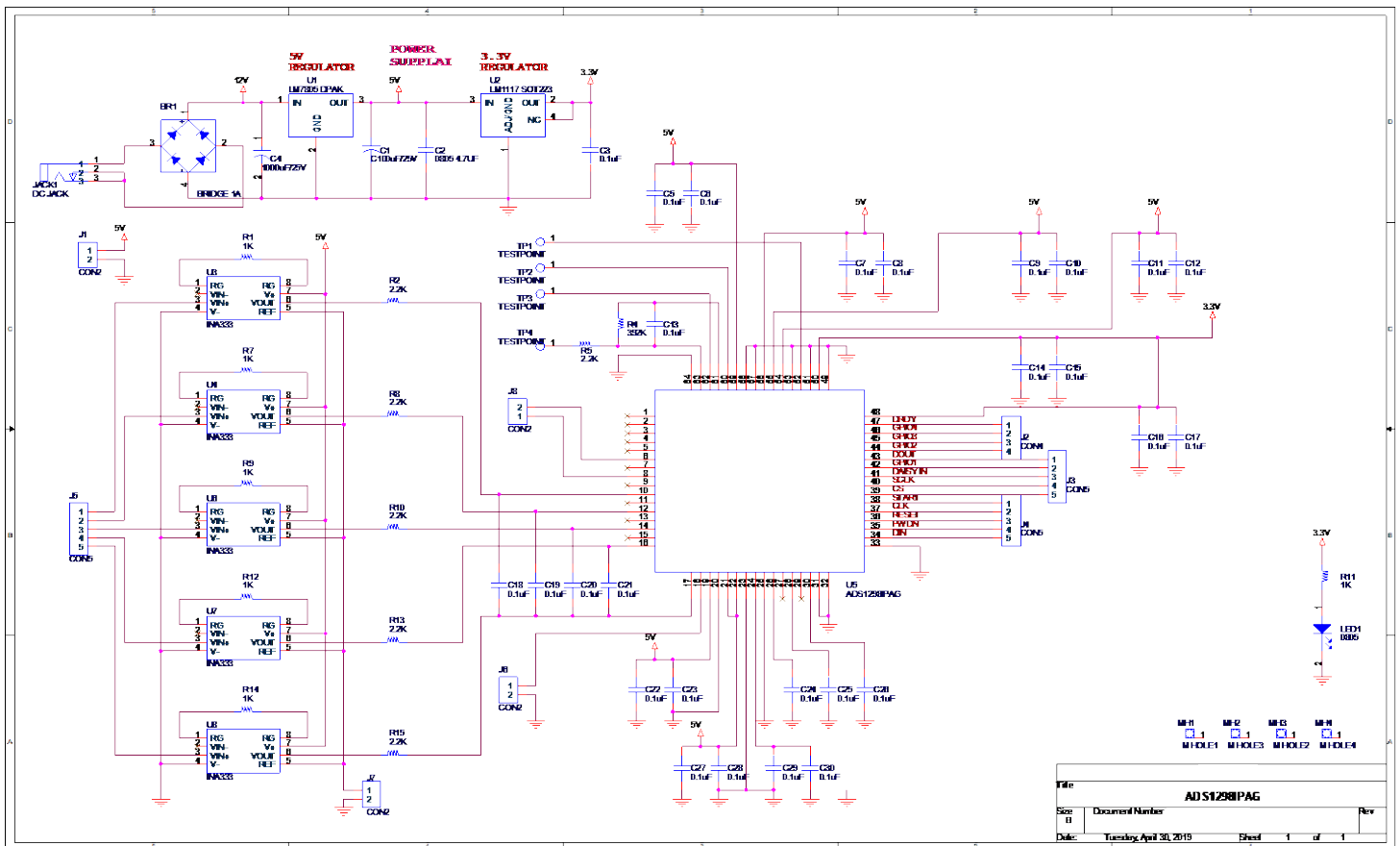


Figure 10: Schematic for PCB design



The final fabricated, two-layer PCB is shown in Figure 11 and Figure 12.

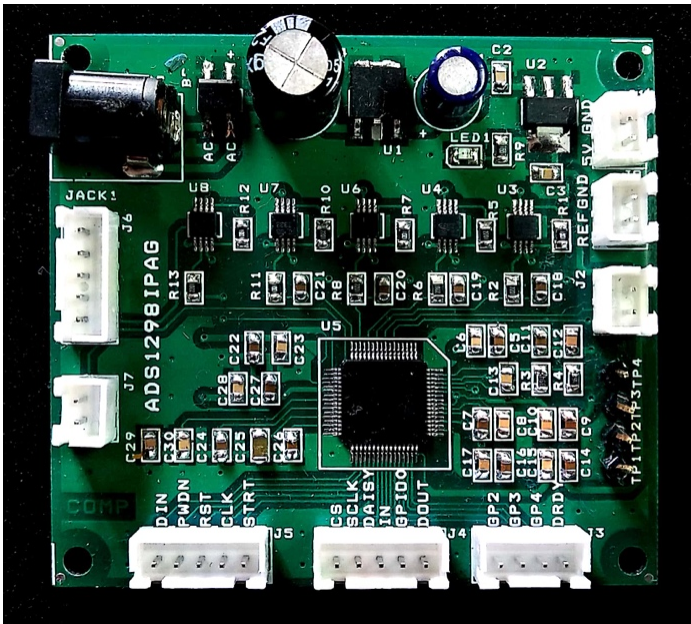


Figure 11: Fabricated PCB (top)

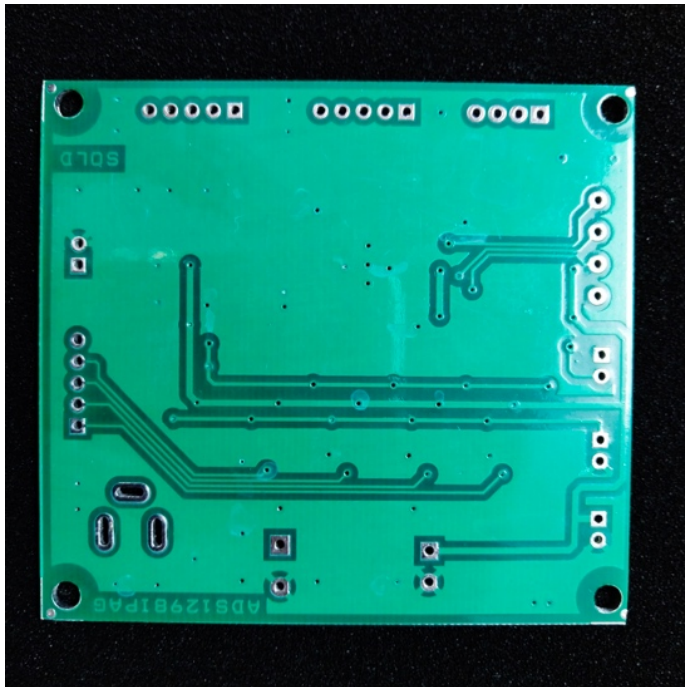


Figure 12: Fabricated PCB (bottom)

### 5. Output from PCB

Once the PCB is fabricated, it is connected to the microcontroller using SPI interface. The ADS1299 is then configured to transmit continuous data with 24-bit resolution at 250 SPS. The register values corresponding to the above configuration are shown in Figure 13 with the correct device ID. Even though the chip is capable of acquiring data at speeds up to 16 kSPS, we have not used these higher sampling rates to avoid the problem of oversampling.

```

-----
Register 0x1 modified.
-----
ID, 0x00, 0x3C, 0, 0, 1, 1, 1, 1, 0, 0
CONFIG1, 0x01, 0x0A, 0, 0, 0, 0, 1, 0, 1, 0
CONFIG2, 0x02, 0xC0, 1, 1, 0, 0, 0, 0, 0, 0
CONFIG3, 0x03, 0xC0, 1, 1, 0, 0, 0, 0, 0, 0
LOFF, 0x04, 0x00, 0, 0, 0, 0, 0, 0, 0, 0
CH1SET, 0x05, 0x07, 0, 0, 0, 0, 0, 1, 1, 1
CH2SET, 0x06, 0x07, 0, 0, 0, 0, 0, 1, 1, 1
CH3SET, 0x07, 0x8F, 1, 0, 0, 0, 1, 1, 1, 1
CH4SET, 0x08, 0xC0, 1, 1, 0, 0, 0, 0, 0, 0
CH5SET, 0x09, 0x00, 0, 0, 0, 0, 0, 0, 0, 0
CH6SET, 0x0A, 0x07, 0, 0, 0, 0, 0, 1, 1, 1
CH7SET, 0x0B, 0x08, 0, 0, 0, 0, 1, 0, 0, 0
CH8SET, 0x0C, 0x6A, 0, 1, 1, 0, 1, 0, 1, 0
BIAS_SENSP, 0x0D, 0xC0, 1, 1, 0, 0, 0, 0, 0, 0
BIAS_SENSN, 0x0E, 0xC0, 1, 1, 0, 0, 0, 0, 0, 0
LOFF_SENSP, 0x0F, 0x00, 0, 0, 0, 0, 0, 0, 0, 0
LOFF_SENSN, 0x10, 0x07, 0, 0, 0, 0, 0, 1, 1, 1
LOFF_FLIP, 0x11, 0x08, 0, 0, 0, 0, 1, 0, 0, 0
LOFF_STATP, 0x12, 0xC0, 1, 1, 0, 0, 0, 0, 0, 0
LOFF_STATN, 0x13, 0xC0, 1, 1, 0, 0, 0, 0, 0, 0
GPIO, 0x14, 0x00, 0, 0, 0, 0, 0, 0, 0, 0
MISC1, 0x15, 0x07, 0, 0, 0, 0, 0, 1, 1, 1
MISC2, 0x16, 0x08, 0, 0, 0, 0, 1, 0, 0, 0
CONFIG4, 0x17, 0xC0, 1, 1, 0, 0, 0, 0, 0, 0
-----
    
```

Figure 13: Register configuration of ADS1299

```

Millis: 224
0, C0000F, AC8A8, 1FC3E8, 3EB5AD, 50BDAD, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
1, C00007, 845F0, 1E69C4, 3CB9F6, 4F82C3, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
2, C0000F, B89FF, 1F7716, 3ECD0A, 50173E, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
3, C00007, 99422, 1F0B09, 3CC192, 4EF1B3, FFFFFFFF, FFFFFFFF, FF8000, 0
4, C0000F, B3454, 1F2A17, 3F23D8, 4F877B, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
5, C00007, 8E698, 1E9DF4, 3D8FD5, 4FB02F, FFFFFFFF, FFFFFFFF, FFFFFFFF, 800000
6, C0000F, AB00E, 1F0CC4, 3F1755, 50AA95, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
7, C00007, 8717F, 1E3768, 3CA722, 502684, FFFFFFFF, FFFFFFFF, FFFFFFFF, 800000
8, C0000F, A3709, 20393F, 3F3878, 4FFEC1, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
9, C00007, 8A828, 1EF6D1, 3C5848, 4ED02A, FFFFFFFF, FFFFFFFF, FFFFFFF0, 0
10, C0000F, BC64C, 1FB876, 3F98D5, 50B430, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
11, C00007, 8DADC, 1F3574, 3D3CA8, 4FC687, E00000, 0, 0, 0
12, C0000F, ACA48, 1FB81A, 40EC78, 4FE043, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
13, C00007, 8F07E, 1F75F8, 3D061D, 4F7899, E00000, 0, 0, 0
14, C0000F, B4FD9, 1F1D42, 3F045C, 50662E, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
15, C00007, 8CFF4, 1E9DF6, 3CB969, 4FEDED, E00000, 0, 0, 0
16, C0000F, AAF94, 1F864C, 3F1044, 50CE7D, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
17, C00007, 92530, 1E0DAC, 3C0099, 5001C8, E00000, 0, 0, 0
18, C0000F, B6B34, 203E2C, 401502, 502315, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
19, C00007, 883F2, 1EF0D1, 3CE39D, 4F2403, E00000, 0, 0, 0
20, C0000F, BCF09, 1FB840, 3F7BC7, 5018AE, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
21, C00007, 8E369, 1F1AC0, 3CD92A, 4FC144, 800000, 0, 0, 0
22, C0000F, B7E03, 1EF62A, 3FED85, 50079A, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
23, C00007, 8B601, 1EEC70, 3D2930, 4F48F9, E00000, 0, 0, 0
24, C0000F, AC008, 20DBA1, 3FC0C6, 507312, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
25, C00007, 8681F, 1E8947, 3C9FCA, 4F449B, E00000, 0, 0, 0
26, C0000F, A6233, 1F64D7, 3FA866, 50BF51, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
27, C00007, 8AD20, 1EB9A6, 3CB8EF, 4F90C1, E00000, 0, 0, 0
28, C0000F, ABC6A, 1FE306, 3DF771, 4FDBA8, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
29, C00007, 8DF90, 1F09C9, 3CE496, 4EC212, 0, 0, 0, 0
30, C0000F, B7E05, 1FE20E, 3F3FFC, 50C352, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
31, C00007, 89906, 1E0C32, 3C1DF7, 4FFE78, 0, 0, 0, 0
32, C0000F, B3618, 1F7934, 3F4059, 513716, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
33, C00007, 95F40, 1F32D4, 3D600F, 5023CA, 0, 0, 0, 0
34, C0000F, B8109, 1F8DAD, 3FB9D9, 50AEAA, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
35, C00007, 85471, 1EB142, 3CC3AD, 4EFD71, E00000, 0, 0, 0
36, C0000F, B8823, 1FCEFB, 40C8B1, 502A43, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
37, C00007, 9181B, 1E7FC9, 3D7A7A, 4FBF21, 800000, 0, 0, 0
38, C0000F, B1161, 1F81DA, 4084F3, 501EE0, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
39, C00007, 8AD14, 1F2981, 3CE11B, 4F9E57, 800000, 0, 0, 0
40, C0000F, B725D, 1EC921, 3EF51A, 503866, FFFFFFFF, FFFFFFFF, FFFFFFFF, FFFFFFFF
41, C00007, 93F25, 1F35AA, 3E734A, 4FE0C3, 0, 0, 0, 0
    
```

Figure 14: Incoming data from EEG electrodes in ADS1299

Figure 14 shows the filtered data acquired from the EEG sensors connected to the input channels of ADS1299. The RAW data is in Hexadecimal format. This is converted to Decimal format and then scaled down to convert into Volts. Since we have used a 4 channel ADS1299 chip, the data format as observed in the output is <start bit> <channel 1> <channel 2> <channel 3> <channel 4>. The last 4 bits in each row are of no use to us because those channels don't exist on an ADS1299-4PAG chip.

## 6. Data Representation using Python GUI

### 6.1. Design

The code for the GUI is burnt on the NodeMCU using the Arduino IDE after installing relevant drivers. First off, this code checks the device ID coming from the attached ADS1299 chip and if it is correct, modifies the register values to acquire data. In our case the ADS is set to acquire data at 250 SPS in continuous mode. It then creates the GUI window as shown in Figure 15.

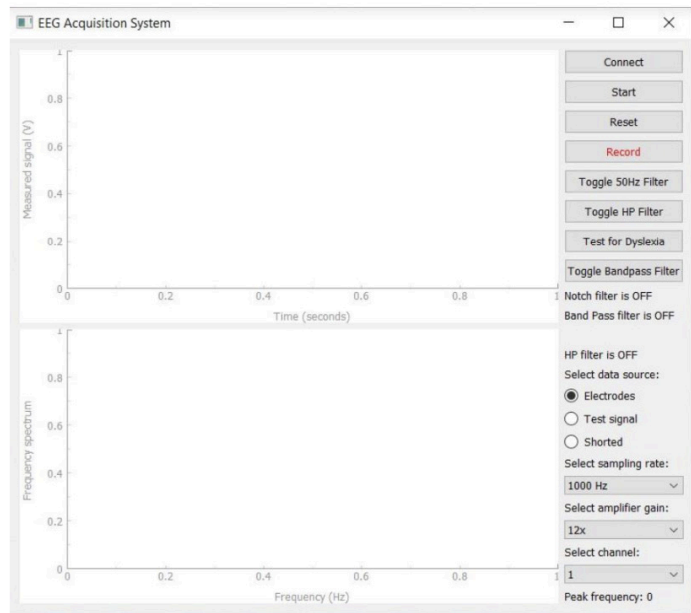


Figure 15: Python GUI window

This GUI window has options to toggle the 50 Hz filter, Low Pass filter and the band pass filter all of which are deployed in the python code using digital filtering. No hardware filters are deployed because they clip waveforms; due to the cutoff from op-amp power supply and induce more noise and drift in the system.

### 6.2. Output

A simple test to capture EEG data is run on the subject for 60 seconds. The real-time time domain plot and the frequency plot are observed in Figure 16.

The GUI has a continuous rolling time domain plot with a 4 second window (customizable) and frequency domain plot that accumulates frequency responses over the course of time that the experiment is run.

As observed in Figure 16, our acquisition system performs filtering in real-time on the PCB itself with the use of hardware high-pass filter implemented using the RC filter and low-pass & notch filter implemented in software in the code burned on to the

microcontroller. Furthermore, the splitting of acquired EEG signal into its component waves based on their frequency ranges and feature classification & identification in different frequency ranges is also implemented on the microcontroller.

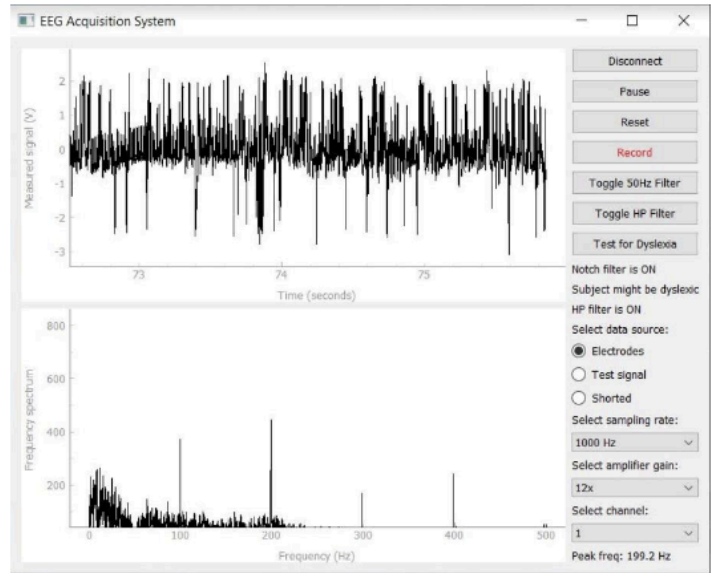


Figure 16: Real-time frequency and time domain plot of acquired EEG signal

The major peaks observed in the frequency domain plot of Figure 16 are in the range of 7 to 30 Hz which comprise mainly of Alpha (7-12 Hz) and Beta (12-30 Hz) waves that constitute a major portion of awake EEG.

## 7. Use of EEG for IOT application

Now that we have successfully designed and built a plug and play device for EEG acquisition and a GUI for data representation it is time to use it for an IOT application to highlight its edge computing capabilities.

We know in hospitals; the staff has to keep an eye on patients that cannot move to ensure that they are comfortable. The lights are off when they sleep and the fan is turned on as and when needed. Keeping this equipment running all the time does not only cause discomfort to the patient but also a higher electricity bill and waste of energy.

To tackle this issue, we use our EEG acquisition system to monitor the Alpha and Beta waves of the patient. We know that in awake EEG the Alpha waves' peak-to-peak amplitude is low when our eyes are open and high when our eyes are closed. This variation can be used to turn off lights in a room when the patient closes his/her eyes to sleep. Similar interactions are observed in the Beta waves when an individual moves or even thinks of moving their limbs to perform a certain function or action. These variations in frequency and peak-to-peak amplitude can be monitored and with the right threshold can be used to turn on or off a fan.

### 7.1. Block Diagram

In Figure 17 our acquisition system is now represented as a black-box which wirelessly transmits data to the host computer where the GUI displays EEG waves in both time and frequency domain in real-time.



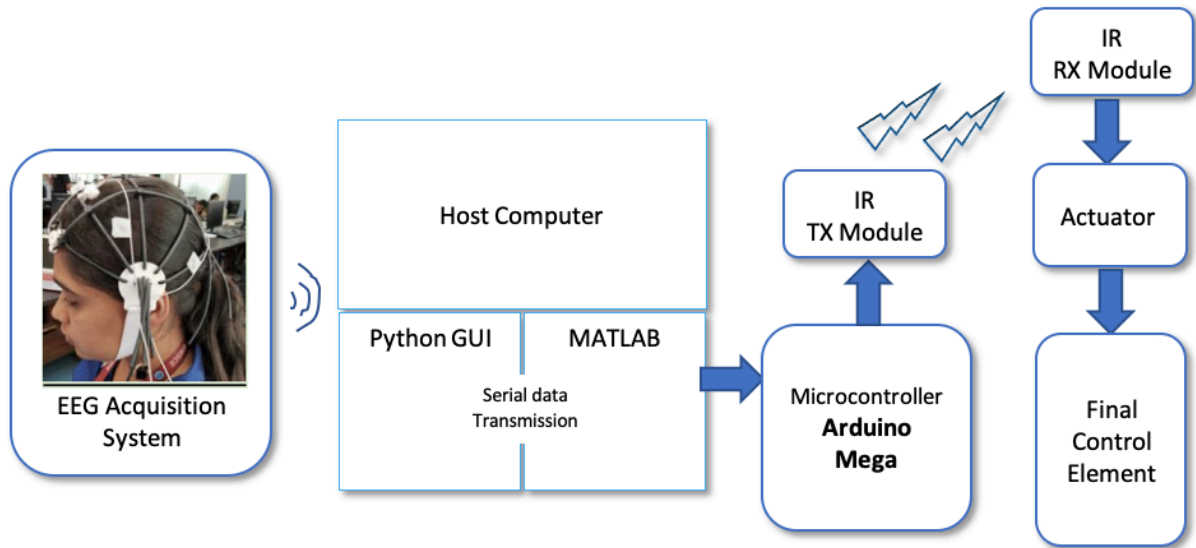


Figure 17: Block diagram for IOT application using EEG

Once the data is received by the GUI, it is transmitted to MATLAB using serial communication. MATLAB checks the peak-to-peak amplitude of the incoming signal and matches it against a set threshold in real-time. Once the threshold is met; in case of Alpha waves the peak-to-peak amplitude crosses the threshold, the program sends a command to Arduino Mega which in-turn transmits it wirelessly using an IR module to actuate or control the final control element which in this case are the lights.

### 7.2. Algorithm Work-flow

As shown in Figure 18 the incoming EEG data from the Python GUI is serially transmitted to MATLAB where an algorithm continuously scans this data and based upon the initial value of the light (on or off) compares it with the upper (light on) or lower (light off) threshold. When the EEG peak-to-peak amplitude crosses the respective threshold, the Arduino sends a command wirelessly to the actuator to change the state of the light.

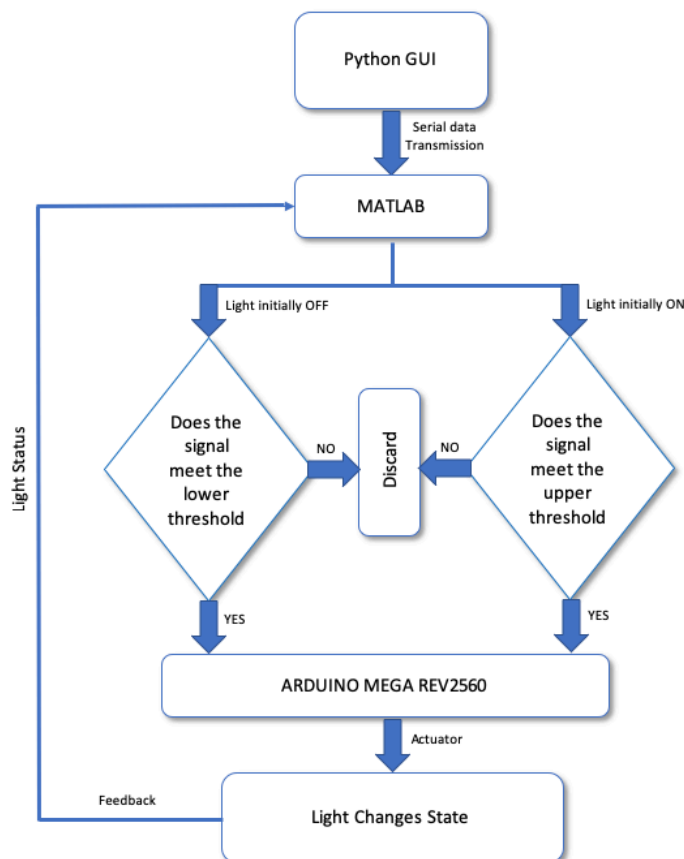
A similar algorithm is applied to control the fan, however in that case, changes in a combination of frequency and amplitude; in the Beta range of frequency (12 – 30 Hz), meeting the threshold result in the change of state of the fan. Based on further analysis of the EEG signal in various frequency ranges, multiple such IOT devices can be controlled. However, for such an application the processing module will have to apply AI, machine learning and deep learning algorithms to improve efficiency and the user will also have to be trained enough to concentrate in certain ways in order to control different objects.

## 8. Results and Discussion

Figure 19 shows the output plotted from channels F1 and F2. The x-axis is time and y-axis is amplitude in milli-volts. The experiment in this case lasted for 20 seconds, where the subject's eyes were open for the first 10 seconds and then closed for the next 10 seconds. As observed in the figure, the EEG peak-to-peak variation is less when the subject's eyes were open except for a few spikes which correspond to blinks. The variation in peak-to-peak voltage increases dramatically half way through the experiment when the subject's eyes were closed. This corresponds to the increase in Alpha Wave activity in Awake EEG when our eyes are closed.

At the 12 second mark in the experiment when the subject's eyes were closed, the lights turned off after detection of increase in Alpha wave activity at the 10 second mark. The 2 second delay is introduced to reduce the chance of false positives.

The above experiment, quite simply demonstrates the accuracy of the data coming from our EEG acquisition system and its ability



to control an IOT device (the lights) precisely in real-time. It is generally observed that the filtering of noise from the EEG signal is done post its collection. Same goes for the removal of artifacts from the EEG signal. Also, the analysis of the signal such as selection of features, classification and training of the classifier is carried out after the data is acquired and stored in a secondary device making the entire flow of denoising, analysis and classification of EEG a delayed process.

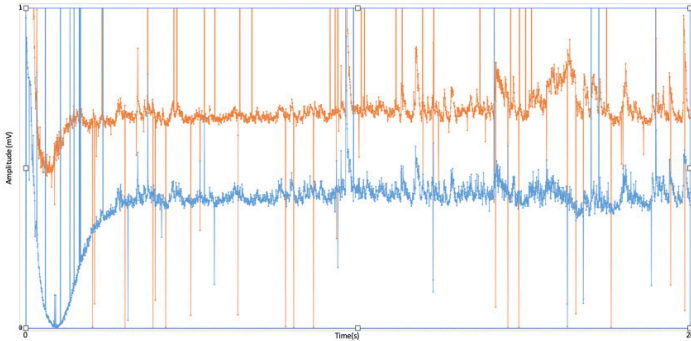


Figure 19: Plotted output from serial data communicated to MATLAB

However, our recommended design can be used to remove the noise in the EEG acquisition system and detect the features from the signal at the time of the acquisition in the device itself. Therefore, moving the denoising of the signal and classification in the EEG acquisition system to a real time system (Edge computing).

Table 1 shows a comparison of our EEG acquisition system with other commercially available systems.

Table 1: Cost comparison of EEG acquisition systems

S.no.	Product	Cost
1.	NeuroSky (1 or 3 channels)	INR 6999 (1 channels) INR 18,000 (3 channels)
2.	EMOTIV EPOC (5 or 14 channels)	USD 299 (5 channels) {INR 21,220} USD 799 (14 channels) {INR 56,707}
3.	OpenBCI (8 – 16 Channels)	USD 249.99 (4 channels) {INR 17,742} USD 499.99 (8 channels) {INR 35,485} USD 949.99 (16 channels) {INR 67,423}
4.	Neuro Insight (4,6,8 can be daisy-chained) – <b>our acquisition system</b>	INR 15,821 (4 channels) INR 16,242 (6 channels) INR 16,604 (8 channels)

The cost of our EEG acquisition system is much less compared to other commercially available EEG acquisition systems because of the fact that we have stripped down the design to a bare minimum and have only retained the most essential components that make it work efficiently and effortlessly. It must be noted that the cost mentioned for our system is applicable when single parts were ordered, the cost will further come down by about 15-20% when ordering parts and fabricating PCBs in bulk.

For our particular application we have used the ADS1299 – 4PAG chip that has 4 input EEG channels and one reference channel. For other applications that require more input channels, more ADS1299 chips can be daisy-chained to achieve a 16, 32 or even a 64 channel EEG acquisition system.

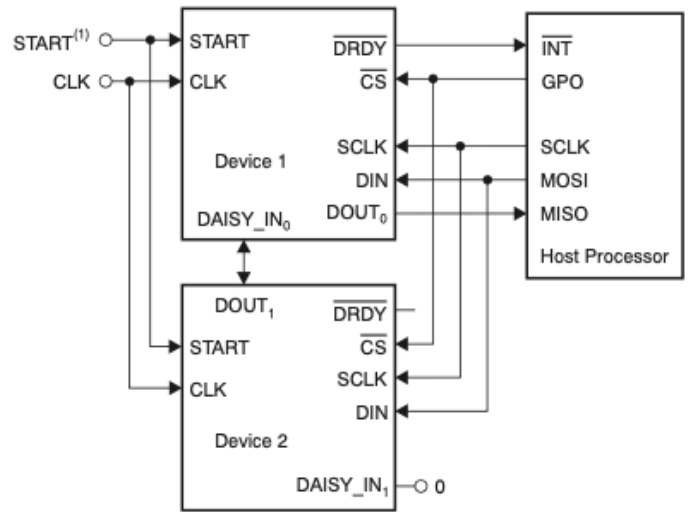


Figure 20: Daisy chain configuration of two ADS1299 chips

Figure 20 shows how two ADS1299 chips can be daisy-chained to achieve up-to 64 input channels.

Furthermore, Table 2 shows a broad specification comparison of our EEG acquisition system with others available in the market. The performance metrics considered are sampling rate, bandwidth, number of EEG acquisition channels, mode of transmission between the acquisition system and the host computer, placement of reference electrodes and the medium of conduction used between the electrode and the scalp. It is worth mentioning that all devices considered, use non-invasive electrodes for data acquisition. A direct comparison of data is also shown in Figure 4.

As evident in table 2, our EEG acquisition system is not only the most cost effective, but also happens to be superior when comparing hardware specifications with those of other commercially available products. The use of Wi-Fi for wireless transmission gives our system the maximum range of communication. Our system also gives the user the ability to choose the number of channels (by daisy-chaining ADS1299, ADS1299-4PAG and ADS1299-6PAG chips), bandwidth (by manipulating the cut-off frequency of high-pass and low-pass filters) and sampling rate (by changing the register value on ADS1299 to be as high as 16kSPS). None of these are offered by any other product in the market right now.

Table 2: Comparison of EEG acquisition systems

S. no.	EEG System	Sampling rate (Hz)	Bandwidth (Hz)	# of EEG Channels	Transmission	Reference	Conductive Mechanism
1.	NeuroSky	512	100	1 or 3	Bluetooth	Earlobe	None
2.	EMOTIV EPOC	256	43	5 or 14	Bluetooth	Left / right mastoid	Saline based
3.	OpenBCI	1000	60	8 to 16	RFD Bluetooth	Earlobe	Gel / Dry
4.	Neuro Insight (our system)	1000 (configurable)	100	4,6,8 (can be daisy-chained up to 64)	Wi-Fi	Earlobe	Gel

## 9. Conclusion

From the above design and application, the followings can be concluded:

- The proposed system captures accurate EEG signal despite the commercially available off the shelf electronic components used in designing the overall system - making it a low-cost yet robust instrument for capturing EEG data.
- In order to improve the robustness of the signal captured the ADC with the instrumentation amplifier is soldered on the same board with the microcontroller, resulting in a single smaller unit, making the overall device more portable.
- To increase the number of EEG signal channels to be captured through the proposed system more ADS1299 chips must be daisy-chained, this will further make the system modular and robust, where more input channels can be added on demand based on the needs of the user.

Further research can be carried out to compare the power consumption v/s performance of various components like the instrumentation amplifier, ADC and microcontroller to design a truly power efficient, battery powered EEG acquisition system capable of collecting data over long periods of time. The reason for moving from an AC powered device to a battery powered one is to get rid of the noise at 50 Hz and its harmonics.

On the software front, the python code can be improved to display more than one EEG channel at a time. Further computational capabilities and the ability to connect with more than one device at a time can be added to completely eliminate the use of MATLAB and finally, the software must also have the ability to export recorded data in popular EEG formats like European Data Format (EDF), CSV, etc. for further analysis and storage.

## References

- [1] K. Rai et al., "Designing Low Cost Yet Robust EEG Acquisition System" in 2019 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS), Rourkela, India, 2019 <https://doi.org/10.1109/iSES47678.2019.00096>
- [2] K.D. Nielsen, A.F. Cabrera, O.F. do Nascimento, "EEG based Brain Computer Interface - towards a better control Brain computer interface research at Aalborg university," IEEE Transactions on Neural Systems and Rehabilitation Engineering., 14(2), Article ID 1642769, 202–204, 2006 <https://doi.org/10.1109/TNSRE.2006.875529>
- [3] A. Caplier, S. Charbonnier, A. Picot, "On-Line Detection of Drowsiness using Brain and Visual Information," IEEE Trans. Syst., Man, Cybern. A, Syst., Humans, 42(3), 773-774, May 2012 <https://doi.org/10.1109/TSMCA.2011.2164242>
- [4] S. Roy, A. De, N. Panigrahi, "Saccade and Fix Detection from EOG Signal" in 2019 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS), Rourkela, India, 2019 <https://doi.org/10.1109/iSES47678.2019.00099>
- [5] N. Panigrahi, A. De, S. Roy, "A Method to Detect Blink from The EEG Signal" in Intelligent Computing and Communication, ICICC 2019, Advances in Intelligent Systems and Computing, 1034, Springer, Singapore, 2019. [https://doi.org/10.1007/978-981-15-1084-7\\_24](https://doi.org/10.1007/978-981-15-1084-7_24)
- [6] N. Panigrahi, K. Lavu, S.K. Gorijala, P. Corcoran, S.P. Mohanty, "A Method for Localizing the Eye Pupil for Point-of-Gaze Estimation" in IEEE Potentials 38(1), 37-42, 2019. <https://doi.org/10.1109/MPOT.2018.2850540>
- [7] J. Gomez-Gil, I. San-Jose-Gonzalez, L.F. Nicolas-Alonso, S. Alonso-Garcia "Steering a Tractor by Means of an EMG-Based Human-Machine Interface", MDPI Sensors, 1424-8220, 2011. <https://doi.org/10.3390/s110707110>
- [8] Y. Lin, Yijun Wang and Tzyy-Ping Jung "Assessing the feasibility of online SSVEP decoding in human walking using a consumer EEG headset", Journal of NeuroEngineering and Rehabilitation, Published online 2014 Aug 9 <https://doi.org/10.1186/1743-0003-11-119>
- [9] G. Rosas-Cholula, J.M. Ramirez-Cortes, V. Alarcon-Aquino, P. Gomez-Gil, J.J. Rangel-Magdaleno, C. Reyes-Garcia, "Gyroscope-Driven Mouse Pointer with an EMOTIV® EEG Headset and Data Analysis Based on Empirical Mode Decomposition" Sensors, 13, 10561-10583, 2013. <https://doi.org/10.3390/s130810561>
- [10] T. A. Mariya Celin & B. Preethi "Brain Controlled and Switch Controlled Robotic Leg For The Paraplegics" ITSI Transactions on Electrical and Electronics Engineering (ITSI-TEEE), 2320 – 8945, Volume -1, Issue -1, 2013.
- [11] H.A. Shedeed et al., "Brain EEG Signal Processing for Controlling a Robotic Arm", IEEE, 152-157, 2013. <https://doi.org/10.1109/ICCES.2013.6707191>
- [12] H. Kamlesh et al., "Brainwave Controlled Robot" International Research Journal of Engineering and Technology (IRJET), 02, 609-612, July 2015
- [13] Sun Chuan, Wu Chaozhong, Chu Duanfeng, Tian Fei, "A Design of Brain-computer Interface Control Platform for Intelligent Vehicle", The 3rd International Conference on Transportation Information and Safety, June 25 – June 28, 2015. <https://doi.org/10.1109/ICTIS.2015.7232147>
- [14] B.J.A. Rani, A. Umamakeswari "Electroencephalogram-based Brain Controlled Robotic Wheelchair" Indian Journal of Science and Technology, Vol 8(S9), 188–197, May 2015. <https://dx.doi.org/10.17485/ijst/2015/v8iS9/65580>
- [15] E. Mathe and E. Spyrou, "Connecting a consumer brain-computer interface to an internet-of-things ecosystem," in Proceedings of the 9th ACM International Conference on Pervasive Technologies Related to Assistive Environments, ser. PETRA '16, 90:1–90:2, 2016. <https://doi.org/10.1145/2910674.2935844>
- [16] C. P. Brennan, P. J. McCullagh, L. Galway, and G. Lightbody, "Promoting autonomy in a smart home environment with a smarter interface" in Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2015. <https://doi.org/10.1109/EMBC.2015.7319522>



- [17] B. Sujatha, G. Ambica “Eeg Based Brain Computer Interface For Controlling Home Appliances” International Research Journal of Engineering and Technology (IRJET) Volume: 02 Issue: 09, Dec-2015.
- [18] N. Abdel Ilah et al., “EEG-based Brain-computer Interface for Automating Home Appliances” JOURNAL OF COMPUTERS, 9, NO. 9, SEPTEMBER 2014. <http://dx.doi.org/10.4304/jcp.9.9.2159-2166>
- [19] K.E. Anu et al., “Non Invasive Electroencephalograph Control for Smart Home Automation” International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering 5, Special Issue 3, March 2016.
- [20] K. Alhalaseh et al., “Home Automation Application Using Eeg Sensor” Proceedings of 98th The IRES International Conference, Antalya, Turkey, 21st-22nd January, 2018.
- [21] H. Mohammad et al., “Automated Classification of L/R Hand Movement EEG Signals using Advanced Feature Extraction and Machine Learning” (IJACSA) International Journal of Advanced Computer Science and Applications, 4, No. 6, 2013. <http://dx.doi.org/10.14569/IJACSA.2013.0406>
- [22] T. Uktveris, V. Jusas “Development of a Modular Board for EEG Signal Acquisition” MDPI Sensors, Published: 3 July 2018. <https://doi.org/10.3390/s18072140>
- [23] B. Luan, M. Sun, W. Jia “Portable Amplifier Design for a Novel EEG Monitor in Point-of- Care Applications” Proc IEEE Annu Northeast Bioeng Conference, 2014. <https://doi.org/10.1109/NEBC.2012.6207127>
- [24] A.J. Bhagawati, R. Chutia “Design of Single Channel Portable Eeg Signal Acquisition System For Brain Computer Interface Application” International Journal of Biomedical Engineering and Science (IJBES), 3(1), January 2016. <http://dx.doi.org/10.5121/ijbes.2016.310>
- [25] C. Jaganathan, A. Amudhavalli, T. Janani, M. Dhanalakshmi, Nirmala Madian “Automated Algorithm for Extracting  $\alpha$ ,  $\beta$ ,  $\delta$ ,  $\theta$  Of A Human Eeg” International Journal of Science, Engineering and Technology Research (IJSETR) 4(4), April 2015.
- [26] J. Tian, W. Song “LabVIEW for EEG Signal Processing” Saudi Journal of Engineering and Technology 1(4), Oct-Dec, 2016. DOI: <https://doi.org/10.21276/sjeat.2016.1.4.10>
- [27] M. Rajya Lakshmi, Dr. T. V. Prasad, Dr. V. Chandra Prakash “Survey on EEG Signal Processing Methods” International Journal of Advanced Research in Computer Science and Software Engineering 4(1), 84-91, January 2014.
- [28] N.R. Miss et al., “Review: Wavelet transform based electroencephalogram methods” International Journal of Trend in Scientific Research and Development (IJTSRD) 2(3), Mar-Apr Page: 1777, 2018.
- [29] T. Pham et al., “A Study on the Feasibility of Using EEG Signals for Authentication Purpose” ICONIP 2013, Part II, LNCS 8227, 562–569, 2013.
- [30] T. Wilaiprasitporn, A. Dithaporn, K. Matchaparn, T. Tongbuasirilai, N. Banluesombatkul, E. Chuangsuwanich “Affective EEG-Based Person Identification Using the Deep Learning Approach” Journal of Latex Class Files, 14, August 2018. <https://doi.org/10.1109/TCDS.2019.2924648>
- [31] M. Byoung-Kyong, H. Suk, A. Min-Hee, L. Min-Ho, L. Seong-Whan, “Individual Identification Using Cognitive Electroencephalographic Neurodynamics” IEEE Transactions on Information Forensics And Security, 12(9), September 2017. <https://doi.org/10.1109/TIFS.2017.2699944>
- [32] G. Matthews, “Metrics for individual differences in EEG response to cognitive workload: Optimizing performance prediction” Personality and Individual Differences 118, 22–28, 2017 <https://doi.org/10.1016/j.paid.2017.03.002>
- [33] D. Gajic, Z. Djurovic, S.D. Gennaro, F. Gustafsson “Classification of EEG Signals for Detection of Epileptic Seizures Based on Wavelets and Statistical Pattern Recognition” Biomedical Engineering: Applications, Basis and Communications, 26(21), 2014. <https://doi.org/10.4015/S1016237214500215>
- [34] S. Mantri, V. Patil, R. Mitkar “EEG Based Emotional Distress Analysis – A Survey” International Journal of Engineering Research and Development 4(6), 24-28, 2012.
- [35] F. Al-shargie, T.B. Tang, N. Badruddin, M. Kiguchi “Mental Stress Quantification Using Eeg Signals” ResearchGate Publication 2016, from book International Conference for Innovation in Biomedical Engineering and Life Sciences: ICIBEL2015, Putrajaya, Malaysia, (15-19), 6-8 2015. [https://doi.org/10.1007/978-981-10-0266-3\\_4](https://doi.org/10.1007/978-981-10-0266-3_4)
- [36] R. Khosrowabadi “Stress and Perception of Emotional Stimuli: Long-term Stress Rewiring the Brain” Basic and Clinical Neuroscience, 9(2), 107-120, 2017. <https://dx.doi.org/10.29252/FNIRP.BCN.9.2.107>
- [37] M. Jeffrey et al., “Acute EEG Patterns Associated with Transient Ischemic Attack” Clinical EEG and Neuroscience, Article first published online; 2019. <https://doi.org/10.1177/1550059418790708>
- [38] M. Nami, S. Mehrabi, S. Derman “Employing Neural Network Methods to Label Sleep EEG Micro-Arousals in Obstructive Sleep Apnea Syndrome” Journal of Advanced Medical Sciences and Applied Technologies 3(4):221-226, 2017. <http://dx.doi.org/10.32598/jamsat.3.4.221>

# Composition of Methods to Ensure Iris Liveness and Authenticity

Ali Al-Rashid\*

Division of ICT, College of Science and Engineering, Hamad Bin Khalifa University, 00974, Doha-Qatar

## ARTICLE INFO

Article history:

Received: 08 June, 2020

Accepted: 26 June, 2020

Online: 18 July, 2020

Keywords:

Authentication

Iris Spoofing Detection

Iris Recognition

## ABSTRACT

*In a biometric system technology, a person is authenticated based on processing the unique features of the human biometric signs. One of the well known biometric systems is iris recognition, this technique being considered as one of the most secure authentication solutions in the biometric field. However, several attacks do exist that are able to spoof iris. In this paper, we propose a novel approach for securing the iris recognition system by eye liveness detection technique. The proposed system detects the eye liveness, and recognises the iris. This process includes multiple steps. As per the first step, the person opens his eye and the system reads remotely the changes in pupil size as a result of the response to the ambient illumination. Then, the system starts matching the iris with a database. The second step: the person closes his eye and the system remotely detects the heartbeats signals under the skin of the eyelid. As per the third step, the person opens his eyes again and the system reads the pupil size again and compares the results of the pupil size, and then the system matches again the iris with the above database. For the iris recognition to be validated, all the above checks have to be passed. We have conducted several experiments with our proposed system, based on a brand new dataset comprised of 40 subjects. In addition, we also used public datasets: CASIA-Interval, CASIA-Twin and Ubiris.V1. The achieved results show the quality and viability of our proposal.*

## 1 Introduction

Among all human physical biometrics, iris recognition systems are considered the most secure biometric systems that can be operated at a low false acceptance rate (FAR). The applications of iris recognition comprise personal identification cards, border controls, and other government applications [1]. Moreover, the distinct feature of the iris recognition system that makes it so popular in security is its uniqueness; even the iris patterns of twins are different.

However, iris is still subject to attacks. The most well known attacks are called spoofing—whose detection is still an open challenge. Iris recognition technology is considered as a robust authentication technology because of the difficulty of counterfeiting the human iris. But adversaries have found numerous manners by tricking the iris recognition systems through forging the iris [2]. So, researchers began to analyze the structure of the human eye to find effective countermeasures versus the adversary.

An attack technique is to utilize an artificial human iris to be placed in front of the camera sensor to grant authentication to the adversary as the legitimate person [3]. While a popular attack is by using a printed photo. The adversary captures a photo of the targeted user to impersonate him as an authorized user [4]. A countermeasure is liveness recognition, that

is employed to identify whether the person is alive or not. One of the common approaches to recognize life signs is the signal of the heartbeats. When the heartbeats occurs, then we cannot ignore the assumption that alive human is facing the sensor. The existing work in recognizing the heartbeats is remote photoplethysmography (rPPG) [5]. The rPPG detects the signals of blood flooding underneath the human skin. Furthermore, various ways to detect rPPG—the widespread technique is probably by sensing remotely the signals of the heartbeats through the face.

The Author in [6] demonstrated the possibility to remotely acquire a heart pulse through one particular camera. The experiment of capturing rPPG is achieved by detecting a certain area of the face skin and calculating a liveness score in the region of interest (ROI) within a short period [6].

In [7], the remote capturing of heartbeats is based on the detection of the color variations in the face skin of the person and recognizes the irregular bio signs. If the indication of the heartbeats under the skin does not occur, it would indicate that the analyzed face is a fake.

**Contribution** We present in our work a new contribution to detect the eye-liveness, integrated with iris recognition. The proposed methodology is articulated over a series of steps: eye liveness, iris recognition, and again eye liveness and iris recognition. The eye-liveness detection depend on

\*Ali Al-Rashid, Doha-Qatar & e-mail: alialrashid@mail.hbku.edu.qa

the reaction of pupil diameter size during the existence of the ambient light, integrated with the remote reading of the heart-beat of the eyelid; while iris recognition is based on matching the iris—with the one previously stored in a database—two times, in real-time.

posed system. Section 8 reports the limitations of our work, while conclusions are reported in Section 9.

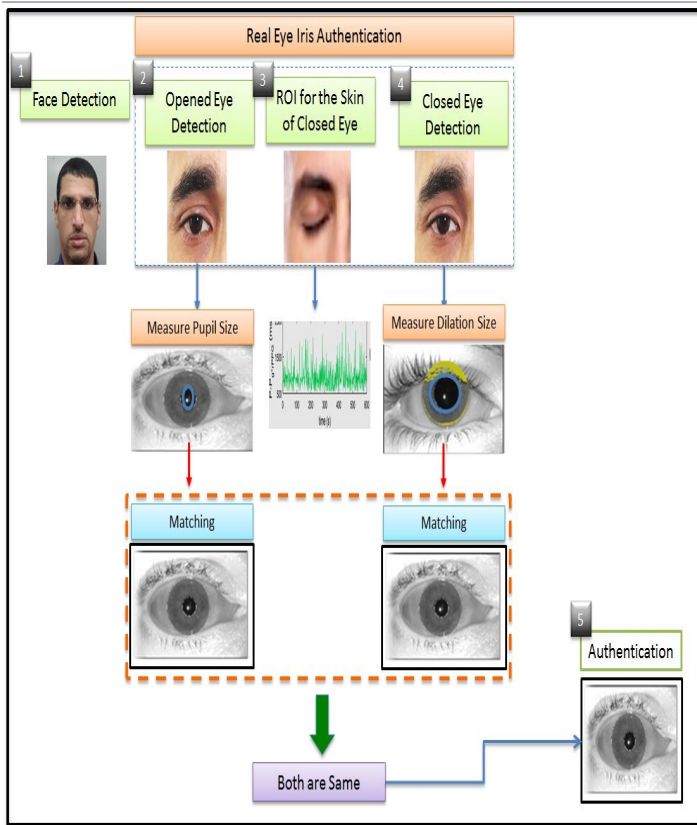


Figure 1: Proposed system

The system starts capturing the face of the supplicant, and runs a time counter. Next, the system recognizes the eye as opened during opening the eye and then reading the actual measurement of the size of pupil diameter. After that, the system initiates matching the person’s iris with the database as the first match. Then, the user closes his eyes and the system identifies the eyes as closed and then starts detecting the signals of blood flowing under the eyelid skin as heart pulses. The user subsequently opens the eyes, and the system reads the diameter size of the pupil again to recognize the dilation. The system matches again the iris with the database as a second match. Consequently, if the iris first match is equivalent to the second match, the dilation of the pupil exists, the heartbeat is recognized, and the time counter of the face did not interrupt, then eye-liveness is validated and iris is authenticated.

**Roadmap** We have structured our proposal as follows: Section 2 reviews the prior art in the field. Section 3 introduces the adversary models and illustrates the methods of spoofing that can be performed by the adversary. Section 4 presents our system and how to grant a secure authentication discriminating between an authentic and a spoofed eye. Section 5 is the iris recognition and explains the process of identifying the iris from the database. Section 6 illustrates the results of experiment. Section 7 is a discussion of the pro-

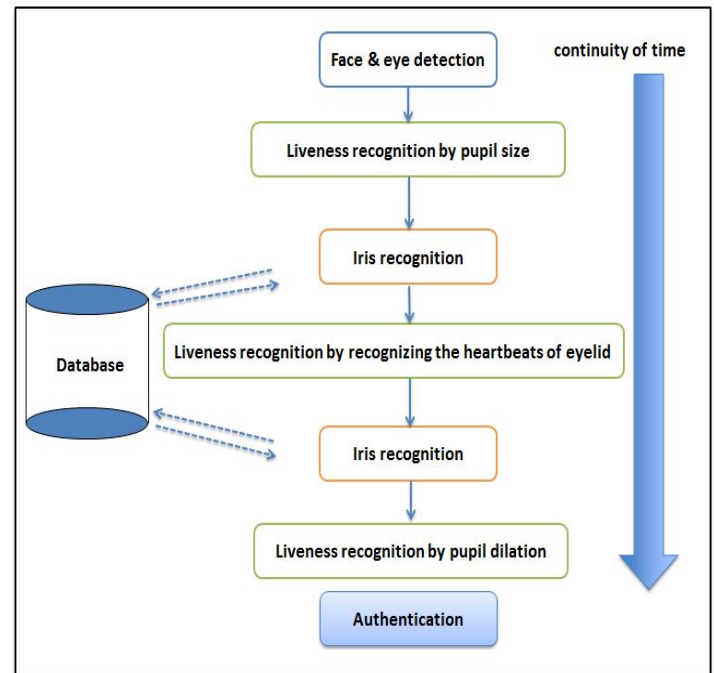


Figure 2: Flowchart for the proposed system

## 2 Related work

The concept of the iris recognition system started initially in 1987 by Flom and Safir [8]. This was demonstrated as a basic matching approach for the imaging system by using mathematical theories.

One of the famous methods for matching iris is the Daugman method. Daugman proposed an automatic segmentation approach, processes for normalization and 2D Gabor filter feature extraction, and match the data to get iris matching [9]. Where [10] made a comparison by comparing the iris recognition algorithm between Principal Component Analysis (PCA), Independent Component (ICA) and Gabor Wavelet for getting iris code.

As countermeasures for spoofing attacks, the previous proposals concentrated on liveness detection. The principle of the liveness is to guarantee that the iris of the eye has a life sign by examining the biometric characteristics of the eye [1]. Basically, illumination technique is one of the popular techniques in forged iris detection, where the eye obtains illumination by focusing light to the eye. So, the pupil reacts with the light by expanding or contraction. The change in the size of pupil confirms the eye is real.

The contributions that depend on checking the pupil radius after centering the illumination have certain restrictions. For instance, if the room is lightened by strong light, then the pupil response could be not satisfactory. Park [2] demonstrated a mechanism utilizing a texture feature to determine the counterfeited the human iris, by the pupillary light response.

Besides, the illumination method has the ability to identify the corneal whether as alive or fake. The corneal usually

reflects when directly illuminated, and creates a red circle in the center of the pupil. By examining the color that occurs in the corneal [6], an alive eye is recognized. Therefore, the attention concentrated on the reaction of the eye to the light, to discriminate between the alive eye and spoofed eye [11]. The author in [4] proposed Pupil Centre Corneal Reflection (PCCR) method relied on gaze estimation to discriminate between the signals of live and spoof eye. In addition, the vein of the sclera used to identify the liveness of the eye. The sclera is the white region inside the eye and the pattern can be extracted and compare it with the pre-stored database, such as in [8].

Regarding the rPPG, the earlier works were focused on analyzing the face skin of the human [6]. The used method is called a color-based method, it measures the heartbeat from the reflection of the colors that occur on the skin— those colors are the blood streaming under the skin [7]. This method is one of the cornerstone for detecting spoofing attacks.

The work in [12] presented a contactless heart pulse system to detect the fabricated mask of the victim. This approach has been achieved by measuring the heart pulse of the face. Therefore, when the heart pulse of the face is detected, then the face is considered alive.

Finally, the authors in [13] proposed to preventing 3D mask PAD (presentation attack detection) by using rPPG pulse detection.

### 3 Adversary models

The adversary relies on spoofing to obtain the identity of the legitimate user. In this section, we introduce the techniques of spoofing that have been used to tamper iris recognition systems.

In general, the iris system captures the iris of the person and then matches it with stored images. The major purpose of the adversary is to fool iris recognition systems by using forged iris models of the targeted user in iris recognition devices to grant an authentication. The iris spoofing attacks can be classified as follows:

- Photo Attack;
- Video Attack;
- Contact Lens Attack; and,
- Artificial Attack.

**Photo Attack:** This type of attack is one of the popular attacks in iris recognition systems due to its simplicity. The attacker uses printed photo on a piece of paper to impersonate a victim. Also, the attacker can use a stored photo in the smartphone to implement the same attack [14].

**Video Attack:** It is a recorded short video of the victim's eye that encompasses all the parts of the eye such as: iris and pupil. The attacker strives to make the video very clear to view the iris in the screen looks real by the adjusting graphics and the resolution.

**Contact lens attack:** This attack is considered a development of a photo attack. The adversary prints the iris drawing of the victim on the contact lens in a special manner. After that, he wears a fake lens to authenticate himself and bypass the system. Normally, contact lens attack is not easy to be

detected, unless the attacker shows unusual behavior in front of the iris camera.

**Artificial Eye:** The adversary uses plastic or glass to fabricate a spoofed eye. After that, the attacker place this artificial eye in iris systems to get authentication.

## 4 Proposed Method

This section details the framework of our system. The system authenticates the eye of the user and identifies spoofing attacks by recognizing a real eye from a fake eye. The proposed system in Figure 1 depicts the liveness of eye recognition. As shown in Figure 1, our system is structured over several stages, described in the following.

1. The system detects the face and starts a timer (the face must be presented facing the camera sensor and the timer will be stopped in case any interruption in the detection of the face).
2. The person opens his eyes and the system identifies the opened eye. Next, begin capturing the pupil diameter and it must be small (because of surrounding illumination) and then matches the iris with the database.
3. The person closes his eyes and the system identifies a closed eye. After that, the system extracts the vein of the eyelid skin.
4. Detecting the heart signals of the eyelid skin.
5. The person opens his eyes again and pupil size will be dilated. Then, the system immediately identifies the dilated pupil and matches the iris again with the database.

Furthermore, the process should be implemented while maintaining its continuity. If any step of the process is stopped or interrupted, then the system detects this abnormal behavior as a spoofing attack. Therefore, when the pupil size is changed before and after closing the eye and the signals are recognized as a heartbeat of the eyelid, and at the same time the iris is matched twice, the system recognizes the eye as real and then iris matching process will initiate.

### 4.1 Pupil Diameter Detection with Heartbeat Signals Recognition from the Eyelid

As an initial step, the system captures the user's face and a temporal window will set with length  $T$  [15]. The window can be configured to execute time-dependent analysis.

Subsequently, pupil diameter detection and rPPG of the eyelid will be performed into various phases, as follows: face and eye detection, pupil diameter size readings, and Region of Interest (ROI) detection with rPPG readings recognition.

#### 4.1.1 Face Detection

Capturing the face precisely is necessary to capture the eye in clear view and to determine good quality candidate regions of the eye. Consequently, the false positive rate will be reduced and also the computation time. The Haar-like feature



is typically the used approach for detecting the face due to its fast processing and accuracy [16].

#### 4.1.2 Eye Detection

The eye-detection implemented by using Haar-like features and a cascade classifiers algorithm. As explained in the previous section, if face detection is performed accurately, the efficiency of detecting the eye will be raised as well [16].

#### 4.1.3 Pupil Detection

Pupil detection is implemented by using the Hough transform algorithm [17]. There is a pre-processing step that must be achieved before implementing the Hough transform [17]. First, we have transformed the image into a grayscale. Next, we got the major data of the eye contour. To implement this stage, we used the 'Prewitt filter' to obtain the contour image [18]. It gives two types of edges  $G_x$  and  $G_y$  on the image, where  $G_x$  detects edges in horizontal direction and  $G_y$  detects edges in vertical direction as demonstrated below:

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \quad G_y = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

And the contour image as shown in Figure 3:

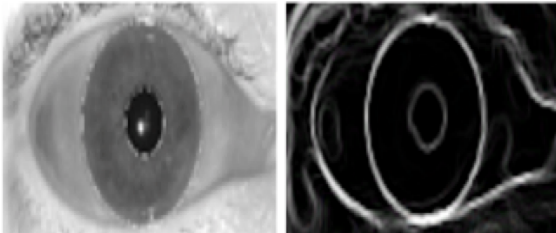


Figure 3: Prewitt filter to get the contour of the eye image

When the circle is detected on the image, the pattern of the pupil is extracted and stored. To get the circle of the pupil after this stage, we apply Hough Transform.

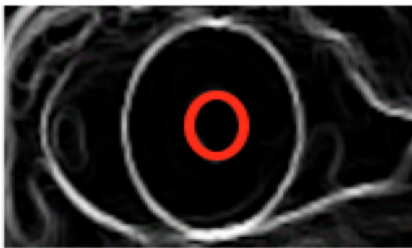


Figure 4: Hough circle

#### 4.1.4 Region of Interest (ROI) Selection

The ROI is a significant element in the captured image of the eye. It locates the skin portion of the eye [19, 20]. Furthermore, the ROI determines the external layer of the eyelid which includes the vein (inside the vein is the blood flowing, used for the detection of the heart pulses).

#### 4.1.5 rPPG Extraction

Once the ROI of the eyelid is determined in the image, we begin measuring the rPPG based on the average of the pixels which are located inside the eyelid skin [6, 7]. The measurements are accomplished independently, over three separate colored channels: Red, Green, and Blue (RGB). Accordingly, rPPG measurements appear on RGB image at every frame and the output appears on the temporal window [7].

## 4.2 Feature Extraction

#### 4.2.1 rPPG signal pre-processing

The signals generated from the rPPG contain the light variations from the heart-beats and the light surrounding the environment [12]. So, signals of the rPPG are affected by the light of the environment, and compounded with the noise from other sources. Hence, a pre-processing step is important to be performed before getting the rPPG signals. The step consists of three stages as follows [6, 7, 12, 13]:

- Detrending filter: it operates as a temporal filter that is utilized to reduce the static part of the signals that produced from the rPPG. For instance: reducing the signals that are not part of the expected heartbeats and removing the interference of light environment.
- Moving-average filter: this type of filter used to remove the random noise of rPPG signals. This random noise could occur due to incorrect parameters which are gained during the image capturing.
- Band-pass filter: commonly, the human heart rate ranges between 40-240 bpm (beats per minute), which is equivalent to a frequency between 0.6 and 4 Hz. Therefore, the frequency outside this range will be removed (because they are not heart pulses).

#### 4.2.2 Transformation of the eyelid skin to frequency

This phase has a special input signal, which is, filtered rPPG signals. The filtered rPPG is the extracted signals from the estimation of the changes in eyelid skin tone caused by the blood streaming under the eyelid skin. Now, the extracted signals will be converted to the frequency domain by using fast Fourier transform (FFT) [6, 13].

#### 4.2.3 Extracting the heart pulse

The extracted domain after the pre-processing step shows a range of frequencies; the highest frequency peaks are the heart pulses. In our state, we need to detect the heart-beat to identify the real eye or fake eye.

The method applied to the features is called transformation. The signals are converted from the domain to the frequency by applying fast Fourier transforms (FFT), and later its power spectral density (PSD) distribution is computed [12]. The patterns of PSD present the discrimination of the eye—whether it is a real or a spoofed one. An alive eye has an obvious peak in the PSD, while a spoofed eye has random noise peaks in PSD [21].

Subsequently, we extracted for every color channel two features for detecting the real eyelid.



In the feature extraction stage, PSD is visible and its maximum power response is specified as first feature  $P$ . The frequency range (expected between 0.6 - 4Hz), is the second feature  $R$  which is the ratio of  $P$  to the total power. The pair of features ( $P, R$ ) are the outcome of three channels of RGB video [21].

## 5 Iris Recognition

This section explains the stages of processing the iris image and then matching the database.

### 5.1 Segmentation

The segmentation method is used to get the contours of iris precisely, which is an inner boundary comprising the pupil and the iris and an external boundary comprises the iris with the sclera.

These boundaries are used to discriminate the pixels whether there is the iris. So, this outcome produces a binary mask that consists of on-pixels (belong to the iris) and off-pixels (do not belong to the iris) [22]. Subsequently, the Hough Transform Circle is used to determine the circular iris area. The segmentation step generates two contours (iris and pupil) that will be used in the next step which is the normalization, as depicted in Figure 5.

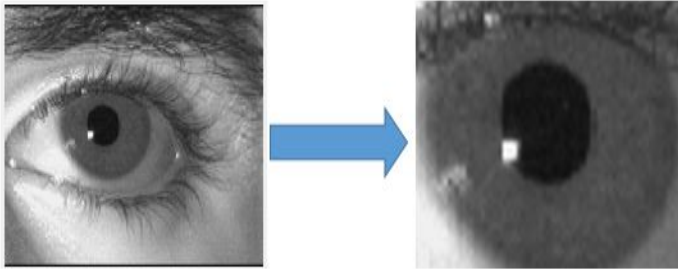


Figure 5: The left image is the original eye for a subject and the right image is the eye after segmentation

### 5.2 Normalization

The normalization converts the area of iris into a straight strip by using the Daugman's rubber-sheet technique, as depicted in Figure 7.

Daugman's rubber-sheet manner is a popular technique for iris normalization which transform the circle of iris area to a rectangular block with a fixed size [22].

$$I[x(r, \theta), y(r, \theta)] \rightarrow I \tag{1}$$

Based on equation (1), the equation converts the pixels in the circle of the iris into an equivalent location on the polar axes  $(r, \theta)$ , where  $r$  is the radial distance and  $\theta$  is the rotated angle at the corresponding radius.

### 5.3 Encoding

With encoding, the iris texture will be extracted through filtering the normalized image by using a Gabor filter method, as shown in Figure 8. The Gabor filter has the ability to generate an optimum conjoint representation of a signal in space and

spatial frequency. [23] Moreover, a Gabor filter is structured by modulating a sine/cosine wave with a Gaussian.

So, using the Daugman method we can later use the 2D Gabor filter to encode the iris data. The equation of the 2D Gabor filter over the image domain as (2):

$$G(x, y) = e^{-\pi[(x-x_0)^2/a^2+(y-y_0)^2/\beta^2]} e^{-2\pi[\mu(x-x_0)/\nu_0(y-y_0)^2]} \tag{2}$$

where,  $(x_0, y_0)$  determines the position on the image,  $(\alpha, \beta)$  determines the actual width and length, and  $(u_0, v_0)$  determines the modulation, which has spatial frequency  $t(u_0^2 + v_0^2)$ .

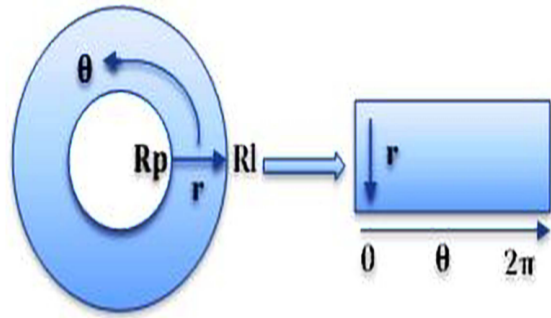


Figure 6: Daugman's model used for iris normalization.

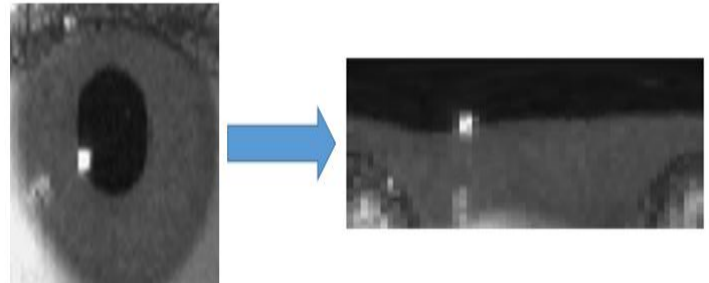


Figure 7: The subject's iris converted to strip by using normalization method.

### 5.4 Matching

The matching section is the final stage. It compares two iris codes using the Hamming distance (HD) between the binary codes corresponding to the selected points within the iris templates.

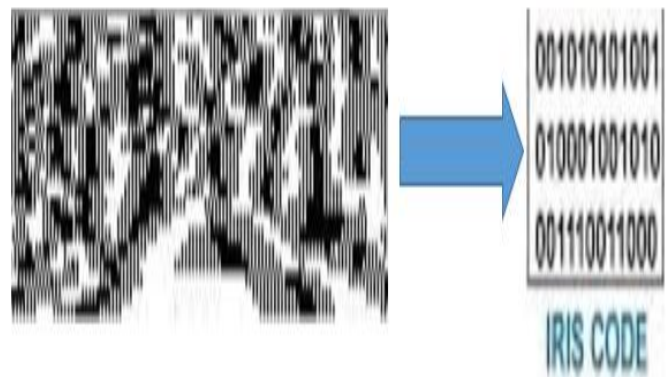


Figure 8: Encoding stage by converting the normalized iris to iris code

Indeed, if two patterns are derived from the same iris, the HD between them will be close to 0. In this stage, we match the iris two times to make sure the same user is physically present in front of the sensor. The first check, when the user opens his eye, and the second check when the user finishes from measuring the heartbeats through the eyelid while closing the eye and then open again his eye.

(infrared camera) was used for detecting the pupil diameter size and iris of the eye and the second camera (logitech C930e) for detecting the eyelid.

The application software: we designed GUI (Graphical User Interface) by using QT program, C++ for pupil diameter detection and iris matching and python for heart pulse detection through the eyelid. The first GUI is configured to camera 1 and captures the pupil from a 1-meter distance and shows on the screen the actual diameter size of the pupil depicts as in Figure 10 and matches the iris with stored database as in Figure 11. Furthermore, it determines the difference between the closed and opened eye through displaying a message.

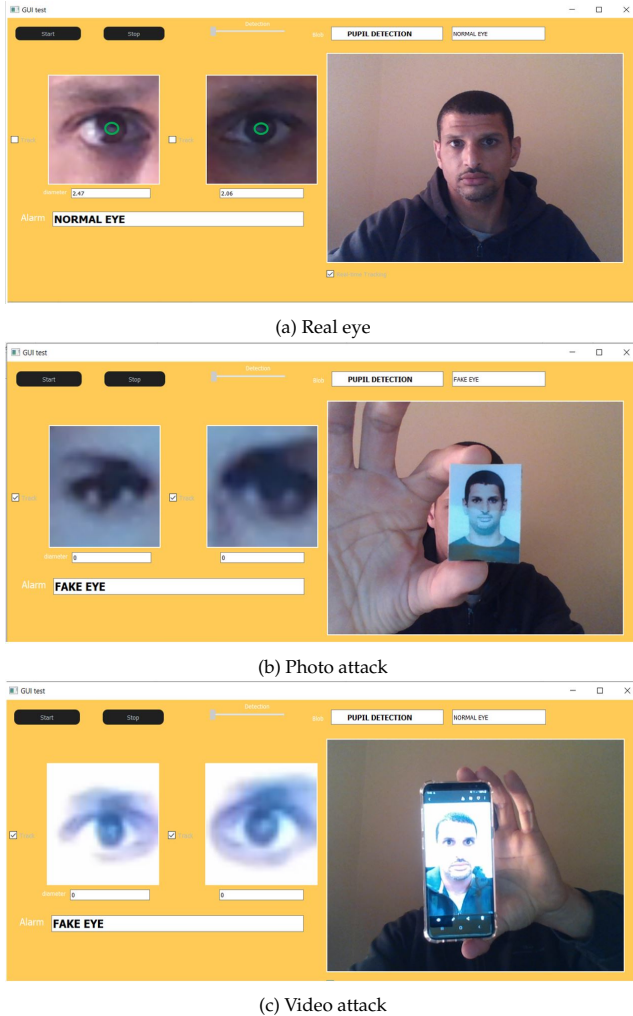


Figure 9: (a) Real eye. (b) Photo attack. (c) Video attack.

The principle of matching the iris two times with running a timer is to make sure that it is the same user present in front of the sensor, otherwise the attacker could first pass the liveness process, and then fool the system by using counterfeit iris.

## 6 Experiment and Results

### 6.1 Experiment Setup

The experiments were conducted with the resources reported in the following.  
 The hardware: 1. PC with OS Win10 ; and, 2. two cameras sensors: one infrared camera type DingDangSmart 2MP 1080P OV2710 Mini IR Webcam and one cameras type Logitech C930e. We fix the two cameras on one stand. The first camera

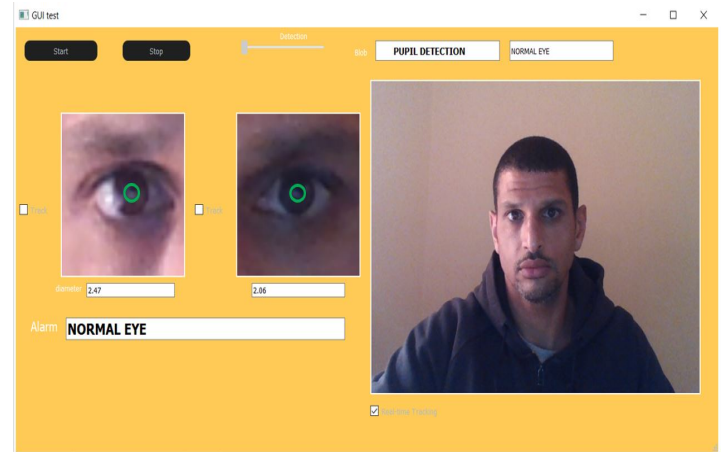


Figure 10: Pupil size recognition

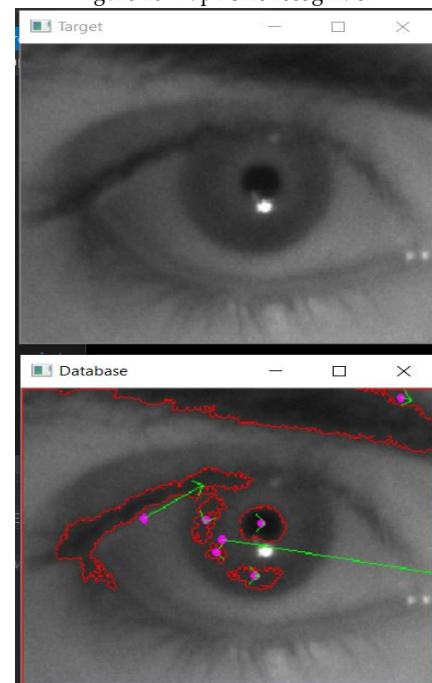


Figure 11: The first image is the scanned eye from the proposed system and the second image is recognized eye with the database.

Also, camera 1 is configured for scanning the iris and matching it with the Database. To detect the heart pulses of the eyelid, we designed another software and configured to camera 2 and measures remotely the user heart pulses of the eyelid as shown in Figure 12. The SW also displays in

Table 1: The pupil size values for 40 subjects

(a) The values of the initial state, dilation, and final state of the pupil sizes for first 20 subjects

Subjects	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Initial State of Pupil Size	2.88	2.34	2.06	2.04	2.55	2.88	2.47	2.53	2.62	2.11	2.06	2.67	2.33	2.26	2.72	2.06	2.42	2.57	2.57	2.31
Pupil Dilation	3.29	2.67	2.88	2.56	3.29	3.31	3.09	3.29	2.77	2.96	2.88	3.19	2.68	2.57	2.88	3.09	2.67	3.29	3.11	2.56
Final State of Pupil Size	2.88	2.34	2.06	2.04	2.55	2.88	2.47	2.53	2.62	2.11	2.06	2.67	2.33	2.26	2.72	2.06	2.42	2.57	2.57	2.31

(b) The values of the initial state, dilation, and final state of the pupil sizes for second 20 subjects

Subjects	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Initial State of Pupil Size	2.87	2.13	2.26	2.72	2.06	2.42	2.57	2.06	2.04	2.55	2.88	2.47	2.53	2.62	2.11	2.06	2.77	2.38	2.46	2.31
Pupil Dilation	3.47	2.67	2.67	3.09	3.91	2.67	3.11	2.67	2.47	3.11	3.21	3.09	3.09	2.88	2.22	2.67	3.67	2.79	3.07	2.56
Final State of Pupil Size	2.87	2.13	2.26	2.72	2.06	2.42	2.57	2.06	2.04	2.55	2.88	2.47	2.53	2.62	2.11	2.06	2.77	2.38	2.46	2.31

real-time the heart pulse on the screen as in Figure 13. In addition, we used certified medical oximeter device type TATRIX Medical fingertip pulse oximeter device for comparing the heartbeats signals with our system.



Figure 12: Eyelid detection

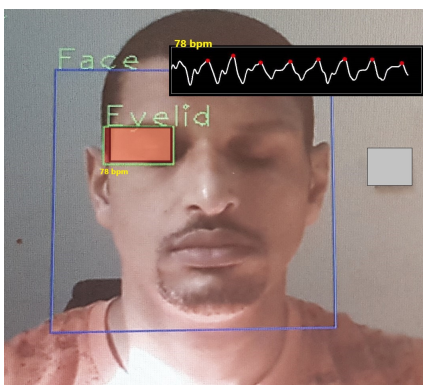
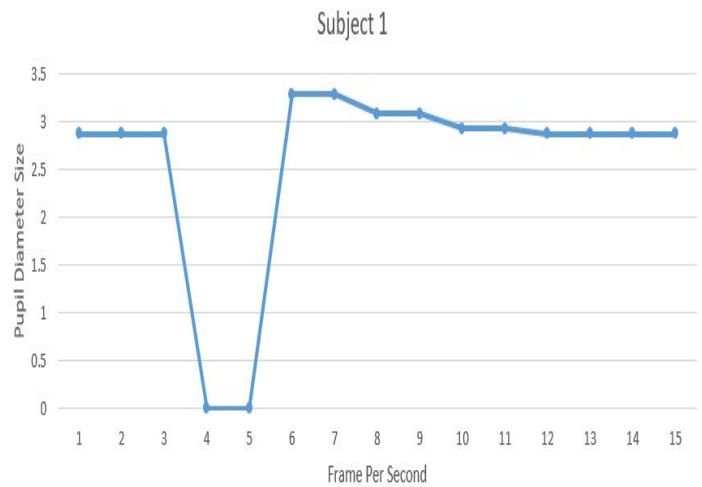


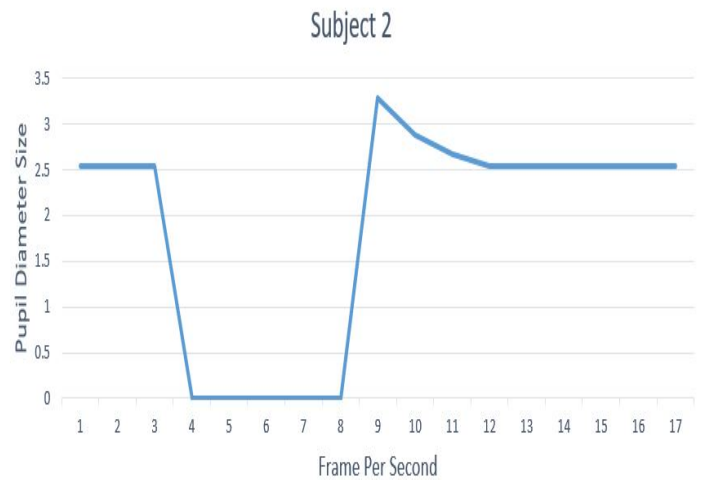
Figure 13: Heart pulse detection by eyelid

## 6.2 Data

Since, to the best of our knowledge, there is no available dataset for heartbeats signals of eyelid and pupil diameter size changes before and after closing the eye, we have created three data-sets composed of 40 subjects. The first data set is the diameter size of the eye pupil.



(a) Pupil diameter size changes for user 1



(b) Pupil diameter size changes for user 2

Figure 14: Pupil diameter size changes

The second one stores the heartbeats parameters of the eyelid. The third dataset is the iris dataset. In addition, we used external datasets from Chinese Academy of Sciences Institute



of Automation (CASIA): CASIA-IrisV3-Interval contains 2655 samples with 320\*280 image size, CASIA-IrisV3-Twins contains 3183 samples with 640\*480 image size and UBIRIS.V1 contains 1214 samples 200\*150 image size to evaluate the performance for the iris recognition [24].

---

**Algorithm 1:** The proposed system for eye liveness detection

---

**Input** : Capture the face  
**While** the same face on the screen **do**  
 Run the continuation timer  
*Begin Stage 1*  
 The system asks to open the eye for 3 seconds  
 Detect the face  
 Detect the eye  
*End of Stage 1*  
*Begin Stage 2*  
**if** The eye is opened **then**  
 | Detect the pupil and measure its diameter size;  
 | Match the iris with the database  
**else**  
 | Termination in case the eye is closed;  
**end**  
*End of stage 2*  
*Begin Stage 3*  
 The system asks to close the eye for 3 seconds...  
**if** The eye is closed **then**  
 | Detect the eyelid;  
 | Extract the vein of the eyelid  
**else**  
 | Termination in case the eye is opened;  
**end**  
*End of Stage 3*  
*Begin Stage 4*  
**if** The vein is detected and the blood streaming detected **then**  
 | Read the heart pulse signals that come through the eyelid  
**else**  
 | Termination in case not detected  
**end**  
*End of Stage 4*  
*Stage 5 (Last Stage)*  
 The systems asks to open the eye again  
**if** The eye is still closed **then**  
 | Termination  
**else**  
 | Detect the pupil and measure the changes in diameter size  
 | Pupil size must be dilated and starts to shrink gradually until it becomes smaller like the first time in step 2  
**if** The pupil size is different **then**  
 | Match the iris with the database  
 | **Output:** The eye is real and the iris authentication is triggered  
**end**  
**end**

---

### 6.3 Methods

We have divided the proposed method into 3 stages: stage 1 for opening the eye, stage 2 for closing the eye and stage 3 for opening the eye again. We created dataset for pupil diameter sizes and iris of every subject. The system captures the pupil diameter with the iris during opening the eye in stage 1 and 3. In the meanwhile the system compares the changes in pupil

diameter size and matches the iris two times .

The concept behind the two matches of the iris during the stages is to confirm the physical presence of the subject. In case the iris matched just one time, it means an error occurred in the system or the attacker tried to bypass the camera sensor by forging the iris. As per stage 2, we created a dataset for the heartbeats of the eyelid and measured it for the same 40 subjects. We let the subjects to look at camera sensor (which is dedicated for measuring the heartbeats signals) for 20 seconds. Next, the system detects and stored the heartbeats of the eyelid in database1. Furthermore, we used certified oximeter device to gather the parameters of heartbeats of the 40 subjects and stored them in database2.

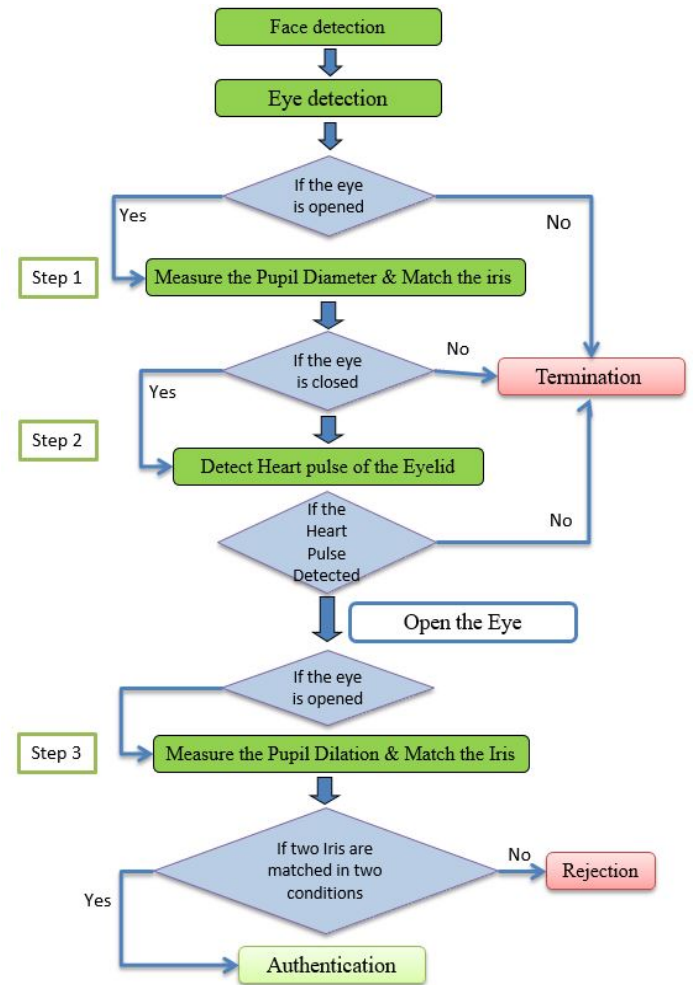


Figure 15: Proposed system

### 6.4 Results

In this section, we report the evaluation of the results for our proposal.

#### Pupil diameter detection

We calculated the pupil diameter size while the subject's eye was opened and the diameter size was smaller than while the eye was closed for a short time and then opened again. As shown in Figures 14a and 14b, we observe the curve of the pupil in a static state while the eye is opened. After that, the

curve moves toward "0" during closing the eye and then the curve increased higher than the static state. Then, the curve returns to the static state again because of the response to the surrounding light.

Moreover, we measured the average of the initial state of the pupil diameter size (when the user looks to the camera), dilation state of the pupil diameter size, and the final state of pupil diameter size (when the user closes and opens his eye again) for all the 40 subjects, as depicted in tables 1a and 1b.

**Iris recognition** In this stage we have examined 40 subjects by matching their irises two times: one before and one after closing the eye. The result of matching the same iris two times—using our algorithm—resulted in a 100% accuracy rate.

Moreover, we used outsource datasets CASIA and UBIRS to evaluate the recognition performance. In table 2 we report the parameters for evaluating the proposed system which checks the same person two times, namely: Correct Recognition Rate (CRR), false accept rate (FAR)—it calculates the probability of the user being incorrectly recognized as another user—, false reject rate (FRR)—it calculates the probability of registered users not being recognized, and Equal Error Rate (EER)—it is the value where FAR and FRR cross. Decision Threshold: when the hamming distance (HD) of two irises is lower than the threshold, then we consider the two iris the same (authorized user). However, if HD of two irises is bigger than the threshold, we consider them as different (unauthorized user). The ROC (receiver-operating characteristic) shows the relationship between the FRR and FAR.

**Eyelid signals**

We conducted a comparison between the eyelid dataset (database1) and the medical oximeter dataset (database2). Each subject was scanned for 20 seconds by our proposed system along with a medical oximeter device which was attached

physically on subject's finger.

The results that we found from database1 were very near to database2 as demonstrated in Figures 16 and 17.

According to the Figures 16 and 17, the curves depict that the parameters are close to each other. We observe that the curve of of the eyelid's heartbeat is very close to the one related to the oximeter, till some variance appears between 13 to 16 seconds because of some delay in the proposed system during the capture of the heartbeats of the eyelid. This delay considers regular because the size of the vein is small which is exists inside in the eyelid and difficult to capture this vein. Besides, the vein of the eyelid differs from subject to subject. However, the delay occurs for few seconds only (not more than 3 seconds). Overall, the parameters of the proposed system are almost the same as the certified oximeter device.

Table 2: Evaluation of our dataset, CASIA-interval, CASIA-twin and UBIRIS.v1

Database	FAR %	FRR %	CRR	ERR
Our dataset (40 irises)	1	0.06	99.25%	0.74
	0.1	3.14		
	0.01	5.67		
CASIA-Interval	1	1.27	98.85%	1.15
	0.1	4.27		
	0.01	12.34		
CASIA-twin	1	1.76	98.25	1.75
	0.1	4.32		
	0.01	8.76		
UBIRIS.V1	1	3.67	97.57%	2.43
	0.1	9.67		
	0.01	19.4		

Moreover, with reference to the heart pulse parameters, we measured the mean, the median, and the standard deviation for all the subjects, as shown in tables 3a and 3b.

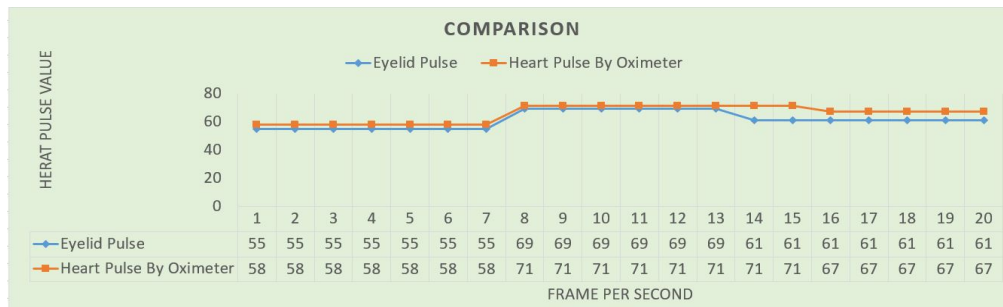


Figure 16: Comparison between heart pulse by eyelid and oximeter for user 1

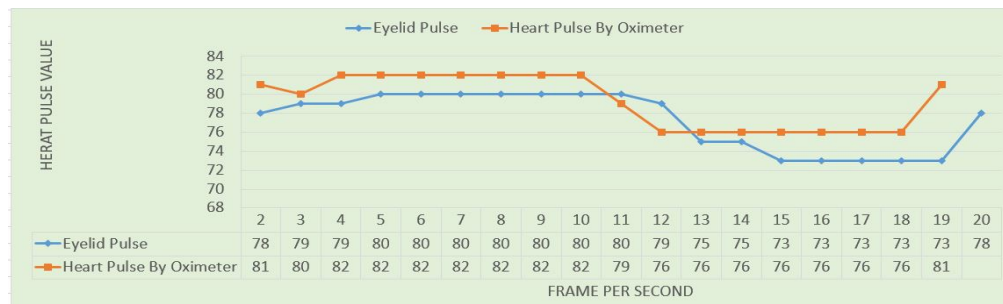


Figure 17: Comparison between heartbeats signals of eyelid and oximeter for user 2



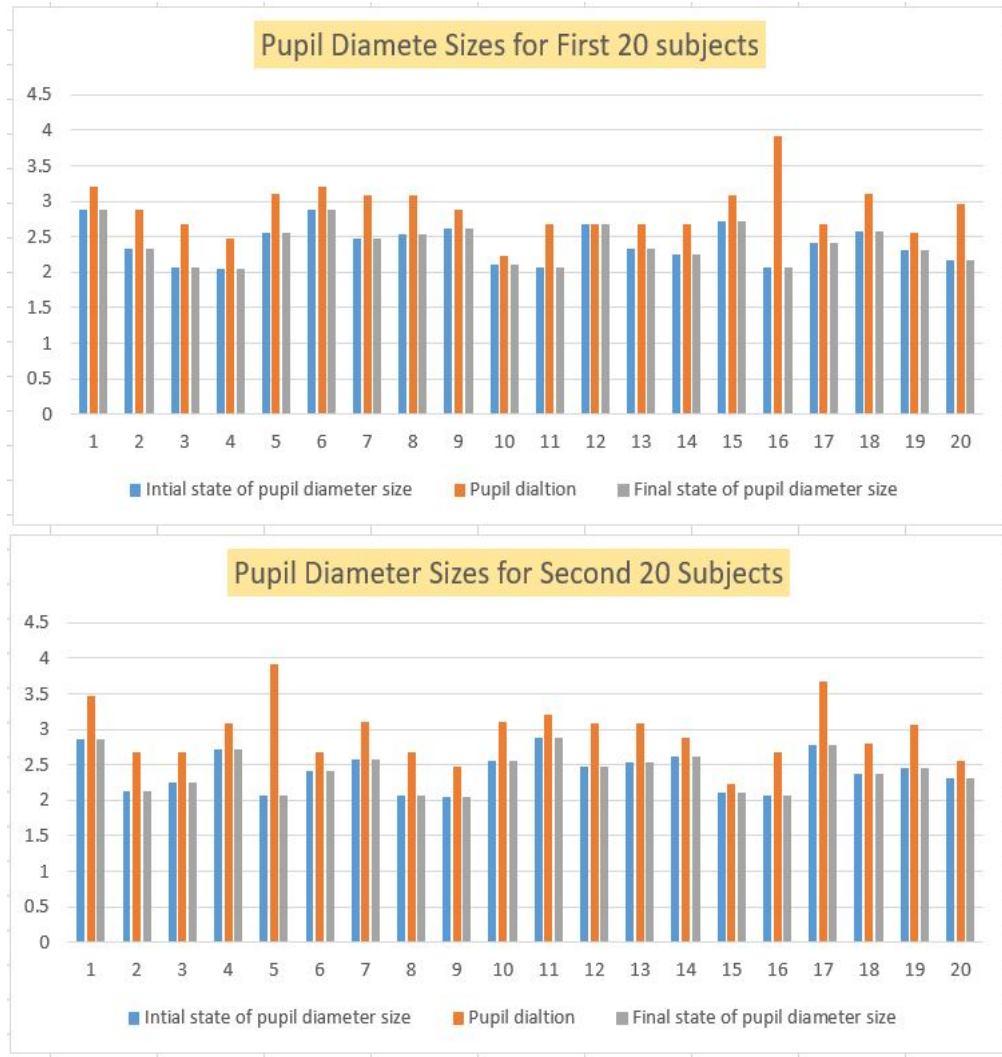


Figure 18: All subjects are grouped, and their pupil diameter sizes is reported to highlight the differentiation between the initial state and dilation state of the pupil diameter size changes.

Table 3: The heart beats parameters for 40 subjects

(a) Mean, median, and standard deviation heart pulse values for all subjects

Subjects	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Mean	55	65	55	70	71	78	55	60	78	55	53	60	59	62	57	61	59	58	72	78
Median	55	65	55	69	71	78	56	60	78	56	53	58	59	63	57	60	60	58	71	78
Standard Deviation	1	1	2	2	2	1	1	1	1	1	1	2	1	2	1	2	5	1	2	1

(b) Mean, median, and standard deviation heart pulse values for all subjects

Subjects	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Mean	55	60	78	70	55	53	60	59	53	60	59	62	59	57	61	59	58	72	65	55
Median	56	60	78	56	53	58	59	53	58	59	63	57	60	60	58	71	65	55	69	71
Standard Deviation	1	1	1	1	1	1	1	1	2	1	2	1	2	5	1	2	1	2	2	2

## 6.5 Alarms

According to our experiments, we got the average of the pupil diameter change as follows:

- The size of pupil diameter in the initial state (before closing the eye) is **2.4mm**.
- The size of pupil diameter in dilation is **2.9mm**.
- The size of pupil diameter in the final state is **2.4mm**.

Therefore, if the sizes detected at run time are not the same as the above sizes, then an exception is raised and the test would fail—that is, we detected either a spoofing attack or a system error.

The curves in Figures 19a and 19b demonstrate how the fake eyes shape differs from the expected curve. For the first fake eye in Figure 19a, we notice that the curve is in a straight line without any changes in pupil diameter.

As per the second fake eye in Figure 19b, there is no dilatation in the curve after closing the eye. While for the Figures 19c and 19d, we notice that the curve is lower than the initial state, which can be considered as an error, due to incorrect positioning of the user during the scan of the eye.

## 7 Discussion

In the following, we will discuss the significance of our proposal. Our proposed system is designed to recognize the iris and at the same time the liveness of different parts of the eye: the pupil and the eyelid.

### 7.1 Pupil analysis detection

As demonstrated in the previous section, we are concentrating on the response of the pupil to the surrounding light without using LED flashlight to decrease the pupil diameter size. The physiological characteristics of the eye make the pupil dilates and shrinks, according to on the surrounding illumination. Therefore, with just ambient illumination, we are able to capture the diameter size of the pupil during all process of our proposal.

According to the Figures 14a and 14b of the experiment, we notice that the curve for the pupil of the subjects was stable when the eye was opened and changed to 'zero' during closing the eye. Next, with opening the eye again the curve raised and then it stabilizes at around the same level of when the eye was initially opened. This experiment proves the pupil's response to surrounding light.

### 7.2 Iris recognition

We have conducted numbers of experiments to optimize the performance quality of the iris recognition system. These experiments were focused on finding the best parameters for the Gabor filter in order to raise the performance of the proposed system. As depicted in Figures 20a, 20b, 20c, and 20d, we evaluated FAR and FRR of 4 datasets: our dataset (40 irises), CASIA interval, CASIA-Twin and UBiris.V1. In our dataset, as mentioned, we matched the iris of every subject two times: before and after closing the eye. The two irises that are matched two times must be similar to each other. consequently, we enhanced the recognition accuracy in every time and obtained better results as per the FAR and the FRR,

as shown in 20a.

The external datasets (CASIA-Interval, CASIA-Twin, UBiris.V1) show different results due to their huge datasets compared with our dataset and their image quality and noise as shown in Figures 20b, 20c, and 20d.

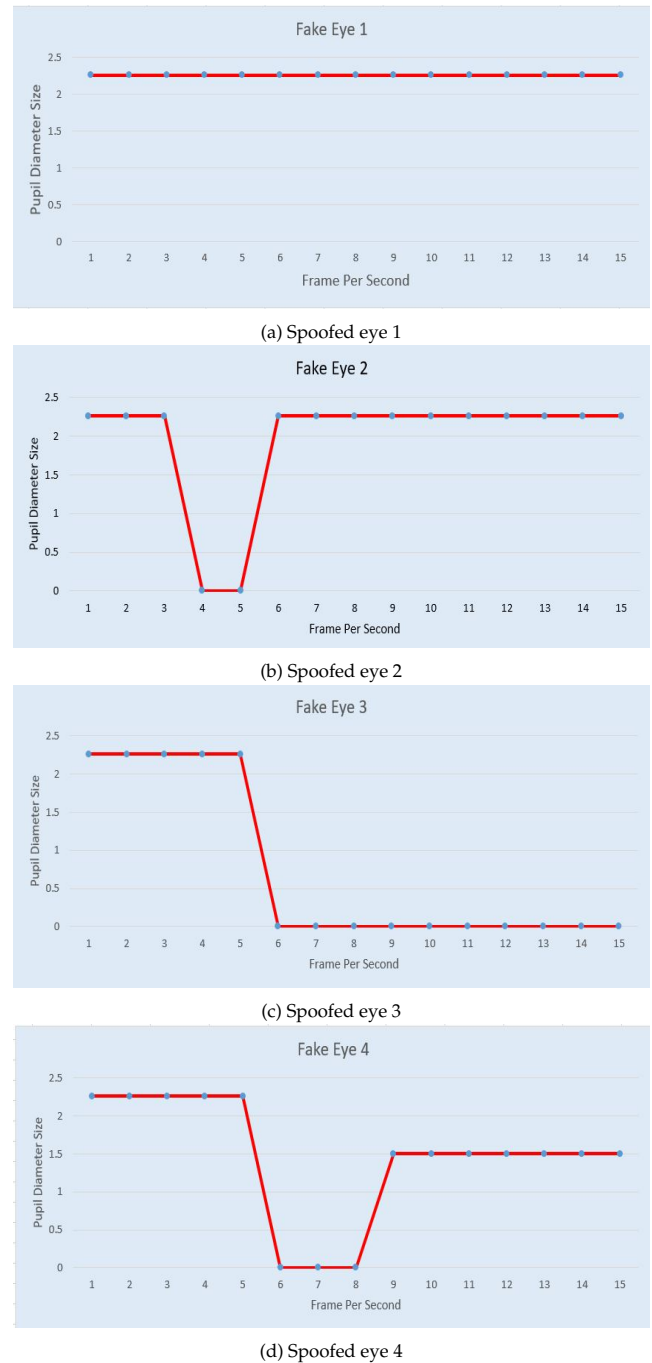


Figure 19: (a) Straight line. (b) The curve in the initial state is equal to dilation state. (c) The curve remains static line without any changes after the initial state. (d) The curve does not raise after closing.

### 7.3 Eyelid analysis detection

As per eyelid analysis, we provide a novel contribution by measuring the heartbeats of the eyelid. This technique implements liveness recognition of the human eye. Indeed, our

system detects the vein which is located inside the eyelid skin and then our system begins reading the signals of heart pulse.

In the experiment for the eyelid heartbeat, we have matched the results gathered by the proposed system with the results gathered from a certified oximeter machine. Subsequently, the results obtained shown to be very close to each other, while scanning for 20 seconds every subject by both systems, as depicted in Figures 16 and 17.

#### 7.4 Assessment of the proposed work

In this section we show our proposed solution thwarts the previously identified attacker capabilities.

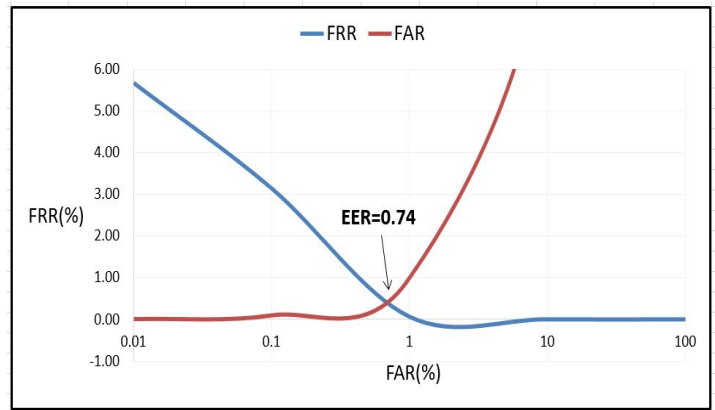
- *Pupil detection:* With the proposed method in pupil detection, the attackers will likely fail in bypassing the checks. For example: if the attacker makes a 3D model of the victim's eye, the pupil size will not change by expanding or contracting as a live eye and, subsequently, the proposed system will detect this 3D model as a fake eye. Note that the attackers could attempt to fake the pupil motion by presenting filmed video. However, these efforts will not succeed. The proposed system captures the diameter size of the pupil and its actual size changes in relation to the period of opening and closing the user's eye.
- *Eyelid detection:* By detecting the heart pulse of the eyelid, the attackers will be faced with another challenge to discover a method to create a fake heart pulse.

Thus, merging eyelid heart-beat and pupil identification into our solution will enhance the security of iris recognition. As an outcome, we have compared in table 4 our work with prior works in terms of spoofing detection capabilities.

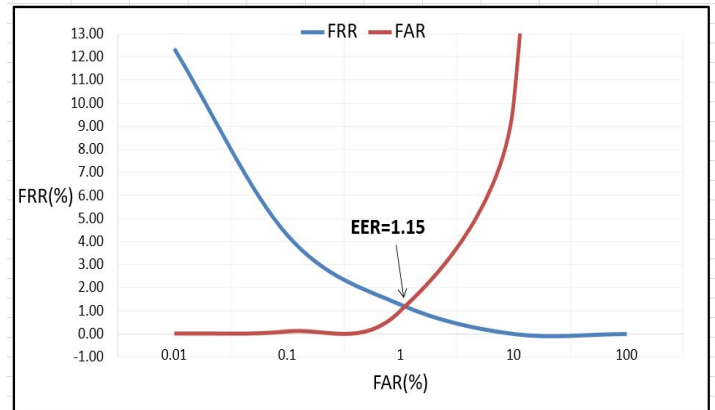
## 8 Limitations of the proposed system

One possibility for an attacker to overcome our solution would be to adopt a new approach in spoofing attacks: eye robot. The attack could be deployed as illustrated in the following.

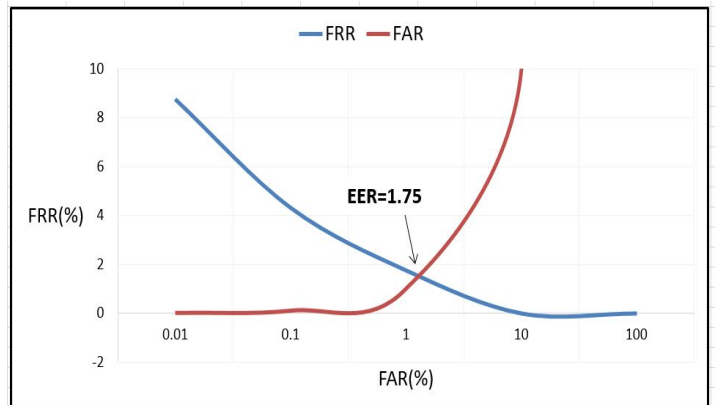
- *Pupil detection:*  
If the adversary designs an eye robot—including a synthetic pupil—the robot eye can be configured to have the pupil of normal size, similar to a real pupil, and after a period the pupil size of the robot eye may dilate and contract as a normal pupil. So, there is a probability for the adversary to bypass the proposed system.
- *Eyelid detection:*  
The eye robot could also have synthetic eyelid and the adversary can add a particular fluid inside the synthetic skin that could mimic the bloodstream—with support of a pump system. Therefore, the proposed system could be triggered into identifying the spoofed pumped-fluid as a real bloodstream.



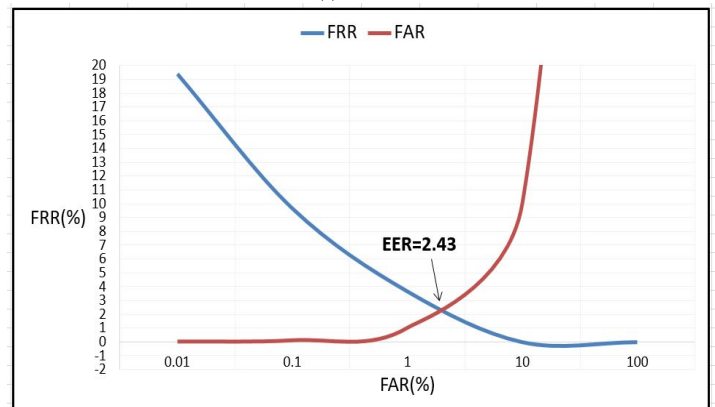
(a) Our dataset



(b) CASIA-interval



(c) UBIRIS



(d) CASIA-twin

Figure 20: The evaluation results of FAR and FRR.

Table 4: Comparison of the proposed work with previous works in terms of detecting spoofing attacks

Author	Features and Methodology	Photo attack	Video attack	Artificial Eye	Contact Lens	Fake Eyelid
Czajka [25]	Pupil dynamics	Detected	Detected	Detected	Detected	N.A.
Raghavendra [26]	Pupil dynamics	Detected	Detected	Detected	Detected	N.A.
Zhang [27]	Fake iris detection by using LBP and statistical features	Detected	Detected	Detected	Detected	N.A.
The Proposed system	Pupil size detection with heart pulse recognition by eyelid	Detected	Detected	Detected	Detected	Detected

Moreover, the ambient illumination can impact the heart pulse signals of the eyelid. For example, the adversary concentrates LED light on the synthetic eye robot to make noise in the heart beats signals that are captured by the camera sensor for the eyelid. The noise will give wrong readings of the heart pulses to the Webcam. Hence, even if unlikely, the adversary has a potential chance to bypass the check.

## 9 Conclusion

In this paper, we have provided a solution to eye biometric spoofing. Our solution relies on multiple factors. First, we examine the pupil movements during opening and closing the eye with support of just ambient illumination. Later, we recognize whether the eye is closed, and then we start measuring remotely a heartbeat signals from the eyelid to validate the liveness feature. During these checks, we match the iris against a pre-loaded database two times: the first time before closing the eye, and the second time after closing the eye. A third contribution is a database composed of 40 subjects. The same 40 subjects have contributed to test the proposed methodology: by using a camera sensor with a software program that includes all the algorithms for liveness and iris recognition. The results have shown the quality and viability of our proposal. We believe that the proposed methodology, together with the experimental results reported in this paper, other than being interesting on their own, will also encourage researchers to conduct future research in this field.

## References

- [1] A. Al-Rashid, "A Three Steps Eye-Liveness Validation System," 2019 International Conference on Cyber Security for Emerging Technologies (CSET), Doha, Qatar, 2019, pp. 1-8, doi: 10.1109/CSET.2019.8904884.
- [2] K.R. Park Robust Fake Iris Detection. In: Perales F.J., Fisher R.B. (eds) Articulated Motion and Deformable Objects. AMDO 2006. Lecture Notes in Computer Science, **4069**. Springer, Berlin, Heidelberg, 2006. doi.org/10.1007/11789239\_2
- [3] J. Connell, N. Ratha, J. Gentile and R. Bolle, "Fake iris detection using structured light," 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, 2013, pp. 8692-8696, doi: 10.1109/ICASSP.2013.6639363.
- [4] Rigas, Ioannis, and Oleg V. Komogortsev. "Eye movement-driven defense against iris print-attacks." *Pattern Recognition Letters*. Elsevier, 2015. doi.org/10.1016/j.patrec.2015.06.011
- [5] A. George and A. Routray, "Fast and accurate algorithm for eye localisation for gaze tracking in low-resolution images," in *IET Computer Vision*, **10**(7), pp. 660-669, 10, 2016. doi.org/10.1049/iet-cvi.2015.0316
- [6] M. Poh, D. J. McDuff and R. W. Picard, "Advancements in Noncontact, Multiparameter Physiological Measurements Using a Webcam," in *IEEE Transactions on Biomedical Engineering*, **58**(1), pp. 7-11, Jan. 2011, doi: 10.1109/TBME.2010.2086456.
- [7] M. P. Tarvainen, P. O. Ranta-aho and P. A. Karjalainen, "An advanced detrending method with application to HRV analysis," in *IEEE Transactions on Biomedical Engineering*, **49**(2), pp. 172-175, Feb. 2002, doi: 10.1109/10.979357.
- [8] Wayman, James, Book review: Handbook of Iris Recognition. *Biometrics*, IET. 3. 41-43, 2014. DOI: 10.1049/iet-bmt.2014.0003.
- [9] J. Daugman, "Biometric personal identification system based on iris analysis," U.S. Patent 5 291 560, Mar. 1, 1994.
- [10] J. Shi and X. Gu, "The comparison of iris recognition using principal component analysis, independent component analysis and Gabor wavelets," 2010 3rd International Conference on Computer Science and Information Technology, Chengdu, pp. 61-64, 2010. doi: 10.1109/ICC-SIT.2010.5563947.
- [11] J. Galbally and M. Gomez-Barrero, "A review of iris anti-spoofing," 2016 4th International Conference on Biometrics and Forensics (IWBF), Limassol, pp. 1-6, 2016. doi: 10.1109/IWBF.2016.7449676.
- [12] X. Li, J. Komulainen, G. Zhao, Pong-Chi Yuen and M. Pietikäinen, "Generalized face anti-spoofing by detecting pulse from face videos," 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, pp. 4244-4249, 2016. doi: 10.1109/ICPR.2016.7900300.
- [13] C. Wang, T. Pun, G.A. Chanel, "Comparative Survey of Methods for Remote Heart Rate Detection From Frontal Face Videos". *Front Bioeng Biotechnol*, **6**(33).doi: 10.3389/fbioe.2018.00033
- [14] S. Thavalengal, T. Nedelcu, P. Bigioi and P. Corcoran, "Iris liveness detection for next generation smartphones," in *IEEE Transactions on Consumer Electronics*, **62**(2), pp. 95-102, May 2016, doi: 10.1109/TCE.2016.7514667.
- [15] Królak, A., Strumillo, P. Eye-blink detection system for human-computer interaction. *Univ Access Inf Soc* **11**, 409-419, 2012. https://doi.org/10.1007/s10209-011-0256-6
- [16] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. CVPR 2001, Kauai, HI, USA, 2001. doi: 10.1109/CVPR.2001.990517.
- [17] R. G. Bozomitu, A. Păsărică, V. Cehan, C. Rotariu and C. Barabaşa, "Pupil centre coordinates detection using the circular Hough transform technique," 2015 38th International Spring Seminar on Electronics Technology (ISSE), Eger, pp. 462-465, 2015. doi: 10.1109/ISSE.2015.7248041.
- [18] A. R. Azar and F. Khalilzadeh, "Real time eye detection using edge detection and Euclidean distance," 2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI), Tehran, pp. 43-48, 2015. doi: 10.1109/KBEI.2015.7436019.
- [19] M. Heshmat, M. Girgis, W. M. Abd-Elhafiez and S. Elaw, "An efficient scheme for face detection based on contours and feature skin recognition," 2015 Tenth International Conference on Computer Engineering and Systems (ICCES), Cairo, pp. 255-260, 2015. doi: 10.1109/ICCES.2015.7393056.
- [20] W. Wang, S. Stuijk and G. de Haan, "Exploiting Spatial Redundancy of Image Sensor for Motion Robust rPPG," in *IEEE Transactions on Biomedical Engineering*, **62**(2), pp. 415-425, Feb. 2015, doi: 10.1109/TBME.2014.2356291.
- [21] J. Hernandez-Ortega, J. Fierrez, A. Morales and P. Tome, "Time Analysis of Pulse-Based Face Anti-Spoofing in Visible and NIR," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, pp. 657-6578, 2018. doi: 10.1109/CVPRW.2018.00096.

- [22] J. Daugman, "How iris recognition works. Proceedings of 2002 International Conference on Image Processing, **1**, 2002. doi: 10.1109/TCSVT.2003.818350
- [23] O. Koç, L. Tosku, J. Hoxha, A. O. Topal, M. Ali and A. Uka, "Detailed Analysis of IRIS Recognition Performance," 2019 International Conference on Computing, Electronics and Communications Engineering (iCCECE), London, United Kingdom, pp. 253-258, 2019. doi: 10.1109/iCCECE46942.2019.8941784.
- [24] CASIA-Iris, Chinese Academy of Sciences–Institute of Automation. <http://www.cbsr.ia.ac.cn/english/IrisDatabase.asp>
- [25] A. Czajka, "Pupil Dynamics for Iris Liveness Detection," in IEEE Transactions on Information Forensics and Security, **10**(4), pp. 726-735, April 2015, doi: 10.1109/TIFS.2015.2398815.
- [26] R. Raghavendra and C. Busch, "Robust Scheme for Iris Presentation Attack Detection Using Multiscale Binarized Statistical Image Features," in IEEE Transactions on Information Forensics and Security, **10**(4), pp. 703-715, April 2015, doi: 10.1109/TIFS.2015.2400393.
- [27] H. Zhang, Z. Sun, and T. Tan. Contact lens detection based on weighted lbp. In Proc. of ICPR, pages 4279–4282, 2010. doi: 10.1109/ICPR.2010.1040



# A Survey on 3D Hand Skeleton and Pose Estimation by Convolutional Neural Network

Van-Hung Le<sup>\*1</sup>, Hung-Cuong Nguyen<sup>2</sup>

<sup>1</sup>Tan Trao University, 22000, Vietnam

<sup>2</sup>Faculty of Technology and Engineering, Hung Vuong University, 35000, Vietnam

## ARTICLE INFO

Article history:

Received: 29 March, 2020

Accepted: 07 July, 2020

Online: 18 July, 2020

Keywords:

3D Hand Skeleton Estimation

3D Hand Pose Estimation

Convolutional Neural Network

## ABSTRACT

Restoring, estimating the fully 3D hand skeleton and pose from the image data of the captured sensors/cameras applied in many applications of computer vision and robotics: human-computer interaction; gesture recognition, interactive games, Computer-Aided Design (CAD), sign languages, action recognition, etc. These are applications that flourish in Virtual Reality and Augmented Reality (VR/AR) technologies. Previous survey studies focused on analyzing methods to solve the relational problems of hand estimation in the 2D and 3D space: Hand pose estimation, hand parsing, fingertip detection; List methods, data collection technologies, datasets of 3D hand pose estimation. In this paper, we surveyed studies in which Convolutional Neural Networks (CNNs) were used to estimate the 3D hand pose from data obtained from the cameras (e.g., RGB camera, depth(D) camera, RGB-D camera, stereo camera). The surveyed studies were divided based on the type of input data and publication time. The study discussed several areas of 3D hand pose estimation: (i) the number of valuable studies about 3D hand pose estimation, (ii) estimates of 3D hand pose when using 3D CNNs and 2D CNNs, (iii) challenges of the datasets collected from egocentric vision sensors, and (iv) methods used to collect and annotate datasets from egocentric vision sensors. The estimation process followed two directions: (a) using the 2D CNNs to predict 2D hand pose, and (b) using the 3D synthetic dataset (3D annotations/ground truth) to regress 3D hand pose or using the 3D CNNs to predict the immediacy of 3D hand pose. Our survey focused on the CNN model/architecture, the datasets, the evaluation measurements, the results of 3D hand pose estimation on the available. Lastly, we also analyze some of the challenges of estimating 3D hand pose on the egocentric vision datasets.

## 1 Introduction

A few recent years, Virtual Reality (VR) and Augmented Reality (AR) become promising technologies in human life. Based on computer vision techniques, they could be found in many applications, including human-computer interaction [1]-[2]; gesture recognition [3, 4]; interactive games [5]; Computer-Aided Design (CAD) [6], sign languages [7]; action recognition, etc. In those applications, real pictorial data (e.g. depth image, color image, stereo, RGB-D, and point cloud data as illustrated in Fig. 1) will be transformed into computer-based data and then could be used by the algorithms. The important work of VR/AR software is detecting the object in the environment from those data. To resolve this point, estimating the 3D skeleton of hand by the Convolutional Neural Networks (CNNs), as shown in Fig. 1 is a widely considered approach with more than 60 valuable studies over the last 4 years.

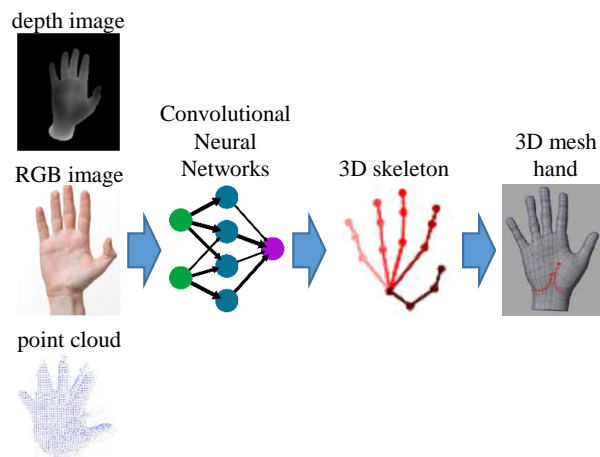
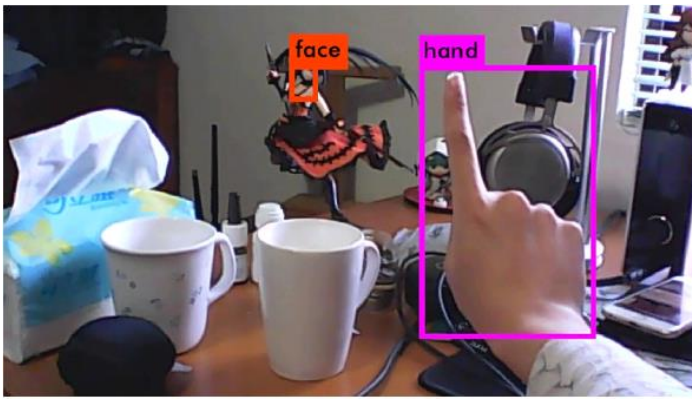
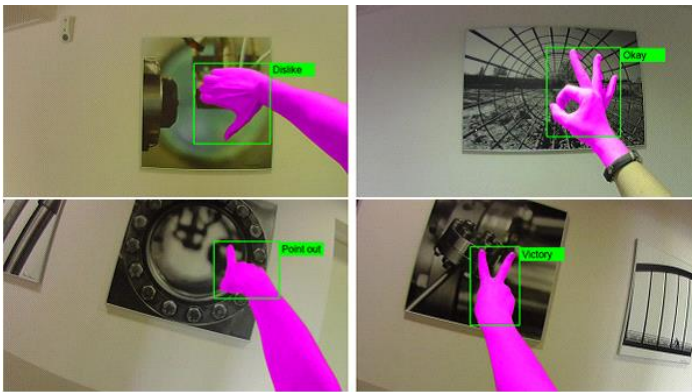


Figure 1: Illustrating the typical input data of 3D hand pose estimation and the results.

\*Van-Hung Le, Tan Trao University, & Lehung231187@gmail.com



(a)



(b)

Figure 2: (a) the result of hand detection [8], (b) the result of the hand segmentation for Gesture Recognition in egocentric vision [9].

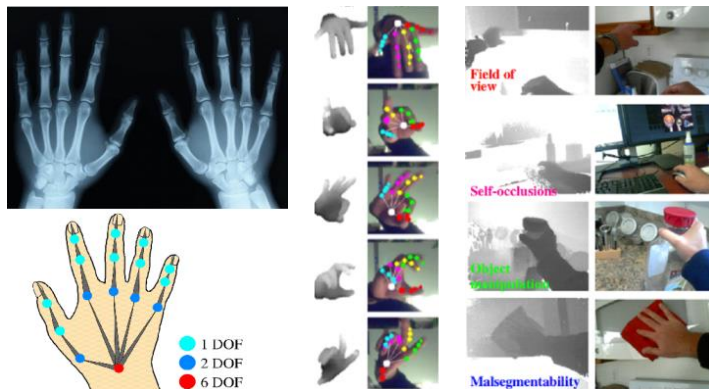


Figure 3: Top left: Hand anatomy; Bottom left: the kinematic model [10]; Right: 3D hand skeleton and pose estimation on the RGB-D image of the egocentric sensor [11].

With the strong development of the sensor/camera technology, with the appearance of depth sensors (e.g. MS Kinect v1 [12], [13], MS Kinect v2 [14], 3D Prime Sense Sensors [15], Intel Real Sense [16], Leap Motion, etc) made restoring 3D hand skeleton and pose easier and more accurate. However, the results of estimating, restoring 3D hand skeleton and pose are influenced by several factors from the captured image data such as hand motion, severe self-occlusion and self-similarity of fingers, especially the data of hand is obscured

when collected from an egocentric sensor [10], [11] as Fig. 3, before conducting further research on 3D hand detection, recognition, and full estimation of 3D hand skeleton, pose. More specific when estimating and restoring the full hand skeleton and pose in the 3D space will help recognize the grasp types, grasp attributes, object attributes [17], [18] of the object more accurately, especially can evaluate the ability to activate of the fingers [19], [20]. In this paper, we survey of the methods and results of 3D hand pose, skeleton estimation following the type input data and publication time, we only focus on the approach that applies the CNNs to estimate 3D hand pose. The input data of 3D hand pose estimation methods can be the depth image, color image, stereo, RGB-D, and point cloud data. They are illustrated in Fig. 1. In particular, we also discuss the results of methods using 3D CNNs, 2D CNNs to estimate the location of joints in the 3D space following the four issues in the 3D hand pose estimation process: The number of valuable studies about 3D hand pose estimation; The estimated results of 3D hand pose when using 3D CNNs and 2D CNNs; The challenges of the datasets which is collected from egocentric vision sensors; The methods to collect and annotate datasets from egocentric vision sensors.

The rest of the paper is organized as follows: Section 1 introduces some overview of this paper. Section 2 discusses the related work. Section 3 discusses 3D hand pose estimation by CNNs to estimate 3D hand pose from some types of data, including the depth image (Sub-section 3.1), the RGB image (Sub-section 3.2), the RGB-D image, or other camera data (Sub-section 3.3). Section 4 presents, discusses some results of 3D hand pose estimation by the CNNs. Section 5 discusses the datasets (Sub-section 5.1) and challenges (Sub-section 5.2) for 3D hand pose estimation. Section (6) concludes the paper with future work.

## 2 Related Works

Many studies of estimating and restoring the full 3D hand model, i.e. skeleton and pose, have been published in recent years. Some of them are listed comprehensively in the survey of Li et al. [21]. This paper provides the answers to many questions, including "What do we need to estimate of the hand?", "What entangles do we need to overcome?", "What is the depth sensor?", "What are the useful methods?": **The objective of 3D hand estimation** is hand detection, hand tracking, hand parsing, fingertip detection, hand contour estimation, hand segmentation, gesture recognition, etc. **The challenges of estimating hand pose** are low resolution, self-similarity, occlusion, incomplete data, annotation difficulties, hand segmentation, real-time performance. **Existing depth sensors** are also summarized by Li et al. [21], including 19 popular depth sensors produced in the last decade. They are divided into groups and illustrated in Fig. 4. The considered parameters of those sensors are depth technology, measurement range, and a maximum speed of depth data. **The methods** to solve 3D hand pose estimation are the model-based method, appearance-based method, and hybrid method. It can be considered as an extension of a review written by Erol et al. [22] that introduced two main methods (i.e. model-based method and appearance-based method) to solve this problem in the 2D space:

- The model-based method compares the hypothetical hand



Figure 4: The depth sensor groups: (1) MS Kinect group; (2) ASUS Xtion group; (3) Leap Motion; (4) Intel RealSense group; (5) SoftKinetic group; (6) Creative Interactive Gesture; (7) Structure Sensor.

pose and the actual data obtained from the cameras. The comparison is evaluated based on an objective function that measures the discrepancy between the actual observations and the estimated data that are generated from the model of the hand.

- The appearance-based method based on learning the characteristics of from the observations to a discrete set of the annotated hand poses. This method uses a discriminative classifier or regression model to describe invariant characteristics of hand pose as a map of the joints of the fingers.

However, the survey of Li et al. [21] is listed only without the presentation of methods, datasets, and evaluation methods.

Another good survey of hand pose estimation is taken by Barsoum [23]. In this study, the author also discusses three methods to perform hand pose estimation from the depth image. The appearance-based method is shown in Fig. 3, the model-based method, the hybrid method is mentioned in Fig. 2. Barsoum focuses on the hand segmentation because its outcome affects the accuracy of algorithms. Facing this problem, the discussed methods are Color or IR skin based; Temperature-based; Marker-based; Depth based and Machine learning-based. In more detail, the author presents the limitation of applying deep learning in hand pose estimation with only two publications ([24] and [25]) in two years 2014 - 2015. From point of view of the limitations as mentioned before, we summarize a survey about the state of the art of hand pose estimation that uses Deep Learning (DL) / Convolutional Neural Network (CNN) in recent years.

### 3 3D Hand Pose Estimation by CNNs

In recent years, using CNNs in detection, recognition, and estimation objects is one of the most successful approaches in computer vision. Human hands are used in the applications of VR/AR and human-computer interaction because human hands can create many different states to execute control states. As Fig. 3(bottom left) is shown 26 degrees of freedom (DOF). To build control and interaction applications using human hands, firstly, human hands need to

be fully and accurately estimated joints in the 3D space. Therefore, this issue is interested in research, especially with the success of CNN in computer vision.

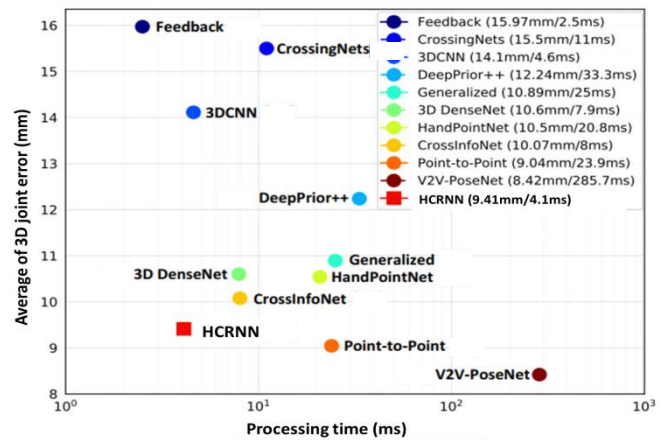


Figure 5: The results of some typical CNNs on the NYU dataset [26].

Firstly, we conducted a survey of methods, results and, discussions of estimating 3D hand pose by CNNs based on the type input data and publication time (*i*). There are 60 studies in the period 2015 - 2019 as shown in Tab. 1. Some of the prominent results are illustrated in Fig. 5. As presented in Tab. 2, the studies in Tab. 1 is published in leading conferences and journals in the field of computer vision. The input data of CNNs to estimate 3D hand pose are the color image, the depth image, and the point cloud. These are data sources that can be collected from common image sensors. Therefore, building VR/AR and human-computer interaction applications in the 3D space can use low-cost sensors and have accurate results (average of 3D joint error 8-16mm as shown in Fig. 5).

The estimated methods [23], i.e. discriminative method, generative method, and hybrid method, are shown in Fig. 6. In the next sub-sections, we give more details on approaches using the CNNs to estimate 3D hand pose from various input data.

Table 1: Statistics of the number of studies used CNNs for 3D hand pose estimation.

Author	CNN type			Data type						Approach				Publish
	2D	3D	No	Depth	RGB	RGB-D	Stereo	Point cloud	Gray image	appearance based	hybrid based	model based	data set	
<b>2015 (8 publications)</b>														
Oberweger[25]	✓			✓							✓			CVWW
Choi[27]	✓					✓				✓				ICCV
Poier[28]	✓			✓							✓			BMVC
Li[29]	✓			✓						✓				ICCV
Oberweger[30]	✓			✓						✓				ICCV
Sun[31]	✓			✓						✓				CVPR
Tang[32]	✓			✓								✓		ICCV
Oberweger[33]	✓			✓						✓				TPAMI
<b>2016 (7 publications)</b>														
Wan[34]			✓	✓								✓		ECCV
Ye[35]	✓			✓							✓			ECCV
Sinha[36]	✓					✓				✓				CVPR
Oberweger[37]	✓			✓						✓				CVPR
Xu[38]			✓	✓						✓				IJCV
Zhang[39]	✓			✓			✓			✓				Arxiv
Ge[40]	✓			✓						✓				TIP
<b>2017 (12 publications)</b>														
Deng[41]		✓		✓						✓				Arxiv
Yuan[42]	✓			✓						✓				CVPR
Choi[43]	✓			✓						✓				ICCV
Choi[44]	✓			✓						✓				ICCV
Wan[45]	✓			✓						✓				CVPR
Ge[46]		✓		✓						✓				CVPR
Neverova[47]	✓			✓						✓				Arxiv
Mueller[10]	✓					✓				✓				ICCV
Zhang[48]	✓			✓						✓				Arxiv
Malik[49]	✓			✓						✓				Arxiv
Zimmermann[50]	✓				✓					✓				ICCV
Oberweger[51]	✓			✓						✓				ICCV
<b>2018 (18 publications)</b>														
Baek[52]	✓			✓						✓				CVPR
Wu[53]		✓		✓							✓			TOC
Supancic[54]			✓	✓						✓				ICCV
Madadi[55]	✓			✓						✓				Arxiv
Rad[56]	✓			✓	✓					✓				CVPR
Garcia[57]			✓			✓							✓	CVPR
Ge[58]		✓		✓						✓				TPAMI
Chen[59]	✓				✓					✓				Arxiv
Zhang[60]	✓			✓						✓				VIPIC
Wohlke[61]	✓			✓						✓				Arxiv
Moon[62]		✓		✓						✓				CVPR
Ye[63]	✓			✓							✓			ECCV
Chen[64]		✓								✓				Access
Spurr[65]	✓				✓				✓	✓				CVPR
Huang[66]		✓		✓						✓				BMVC
Penteleris[67]	✓				✓					✓				WACV
Wan[68]	✓	✓								✓				CVPR
Ge[69]		✓		✓						✓				ECCV
<b>2019 (15 publications)</b>														
Zhang[70]	✓				✓					✓				Arxiv
Sharma[71]	✓				✓					✓				Arxiv
Yoo[72]	✓			✓						✓				Arxiv
Li[73]	✓								✓	✓				ICCV
Wan[74]	✓			✓						✓				Arxiv
Li [21]	✓						✓			✓				BMVC
Liu[75]	✓				✓					✓				TPAMI
Cejong[76]	✓			✓						✓				FG
Hampali[77]			✓			✓							✓	CVPR
Li[78]		✓								✓				CVPR
Zhang[79]	✓			✓						✓				TIP
Baek[80]	✓				✓					✓				CVPR
Lee[81]	✓			✓						✓				Arxiv
Ge[82]	✓				✓					✓				CVPR
Du[83]	✓			✓						✓				CVPR
<b>Total</b>	<b>46</b>	<b>10</b>	<b>5</b>	<b>41</b>	<b>10</b>	<b>5</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>52</b>	<b>4</b>	<b>2</b>	<b>2</b>	



Table 2: The explanation the names of conferences and journals in Tab. 1.

Acronym	Explanation
CVWW	Computer Vision Winter Workshop
ICCV	IEEE International Conference on Computer Vision
BMVC	British Machine Vision Conference
CVPR	IEEE Conference on Computer Vision and Pattern Recognition
TPAMI	IEEE Transactions on Pattern Analysis and Machine Intelligence
ECCV	European Conference on Computer Vision
IJCV	International Journal of Computer Vision
TIP	IEEE Transactions on Image Processing
Arxiv	arxiv.org
TOC	IEEE Transactions On Cybernetics
VIPC	Electronic Imaging, Visual Information Processing and Communication
Access	IEEE Access
WACV	IEEE Winter Conference on Applications of Computer Vision
FG	IEEE International Conference on Automatic Face & Gesture Recognition

### 3.1 Estimating by The Depth Image

Being the origin format of 3D data, the depth image is the most widely used when estimating a 3D hand pose. There are 2 branches as illustrated in Fig. 7, based on the form of intermediate data, i.e. hand point cloud and heat-map.

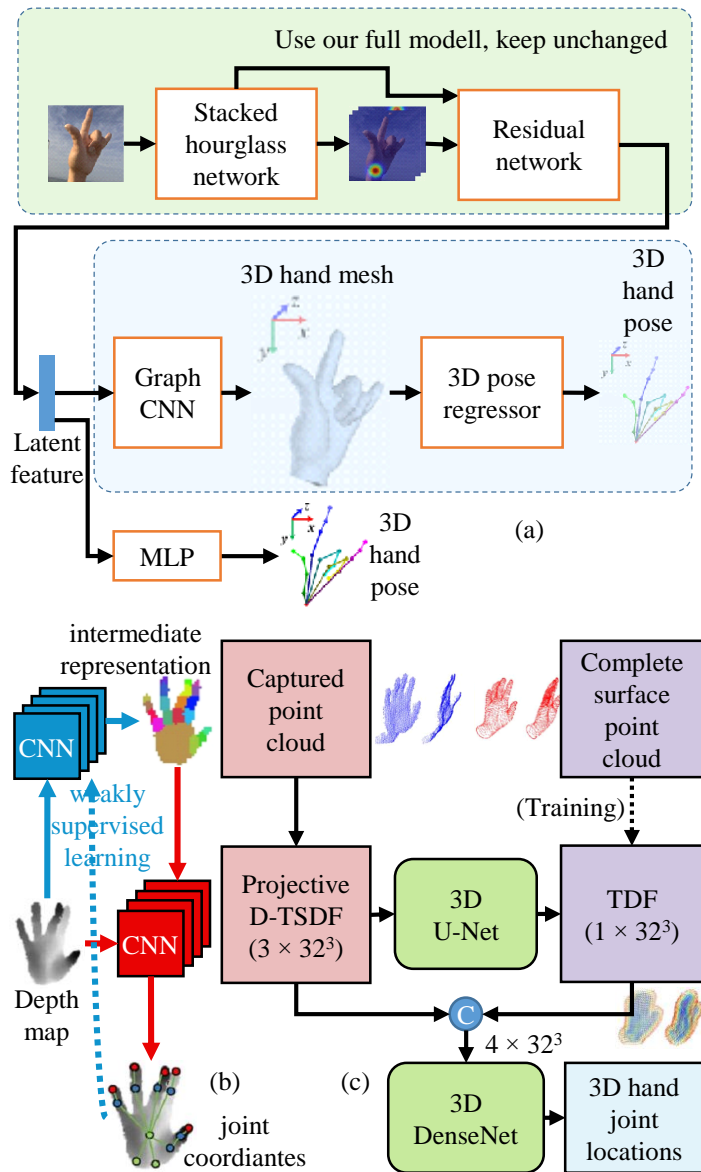


Figure 6: CNN architectures to estimate 3D hand pose as follows: (a) the RGB image [82]; (b) the depth image [47]; (c) the point cloud [58].

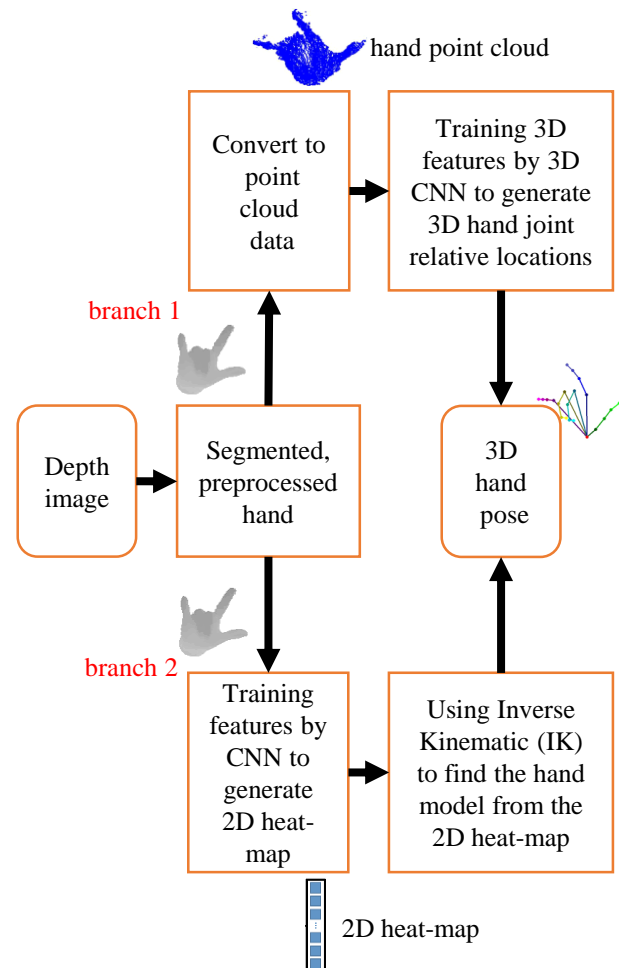


Figure 7: The CNN architecture to estimate 3D hand pose from the depth image.



### 3.1.1 Converting to Hand Point Cloud Data

The first branch of depth image approach converts collected depth data to the form of point cloud before putting it into CNN as a training data [40] [58]. Liuhaio et al. [40] proposed a multi-view regression framework for 3D hand pose estimation, as is shown in Fig. 8. This framework generates heat-maps for three views by projecting the point cloud of the hand onto three orthogonal planes, i.e.  $(x - y; y - z; z - x)$ . Then each projected image is fed into a separate CNN to generate a set of heat-maps for hand joints. This method is similar to the method of [26] to generate a set of heat-maps. After that, the combination of those three views thus contains the location distribution of the joint in the 3D space. This proposed method was evaluated by the dataset of [84]. The average estimation errors are 22.8mm and the processing time is 14.1ms when to be trained and tested on the GPUs under the system whose two Intel Xeon processors, 64GB of RAM and two Nvidia Tesla K20 GPUs. The details of time are 2.6ms for multi-view projection, 6.8ms for CNN forward propagation, and 4.7ms for multi-view fusion.

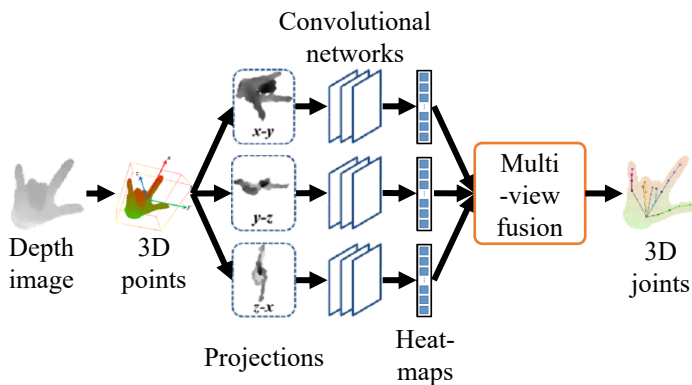


Figure 8: Multi-view regression framework for 3D hand pose estimation [40].

Wan et al. [34] proposed a Conditional Regression Forest (FCRF) which uses a set of new features. At each stage of the regression, the frame of reference is established from either the local surface normal or previously estimated hand joints of the point cloud. The normal difference feature of this method is highly robust to 3D rigid transformation because the 2.5D point cloud is projected and indexed to the image space. Therein, the hand pose estimation process is the process of estimating the joints of 5 fingers. The proposed method is evaluated on ICLV and MSRA datasets and the result of the average of joints error is about 8mm, 25mm-30mm, respectively. Ge et al. [46] proposed a simple approach for real-time 3D hand pose estimation from single depth images by using three-dimensional CNNs (3D CNNs). This 3D CNNs can effectively learn 3D features from the 3D volumetric representation. Liuhaio et al. [58] proposed Hand PointNet-based method for 3D hand pose estimation. The 3D point cloud of the hand is down-sampled and normalized in an oriented bounding box (OBB) to make the proposed method robust to various hand orientations. This method uses the estimated surface normal and normalized points of the point cloud data of the hand as the input of the hierarchical PointNet [85] and then outputs a low dimensional representation of the 3D hand joint locations. Therein, the hierarchical PointNet consists of  $L$  point set abstraction levels. The higher the level is, the smaller

the number of points. The authors evaluated the proposed method on three public hand pose datasets, including NYU [26], MSRA [84], and ICVL [86]. The experimental results when deploying in a workstation with two Intel Core i7 5930K, 64GB of RAM and an Nvidia GTX1080 GPU are:

- The per-joint mean error distances and the overall mean error distances are 10.5mm, 8.5mm, and 6.9mm, respectively.
- The average processing time of the proposed method is 20.5ms, including 8.2ms for point sampling and surface normal calculation, 9.2ms for the hand pose regression network forward propagation, 2.8ms for fingertip neighboring points search, and 0.3ms for fingertip refinement network forward propagation.

### 3.1.2 Training by CNNs to Generate 2D Heat-map

The second approach based on depth image often trains annotated joints of hand poses on a large dataset (synthesized data) by the CNNs [45]. Those datasets contain most of the actual hand poses. The estimation process evaluates the characteristics of the hand pose on the input data and finds the most fitting pose in the synthesized data as illustrated in Fig. 9. Many studies are using this method whose difference is the used of CNN to predict the position of joints. Fig. 5 illustrates the results of some prominent studies with the NYU dataset.

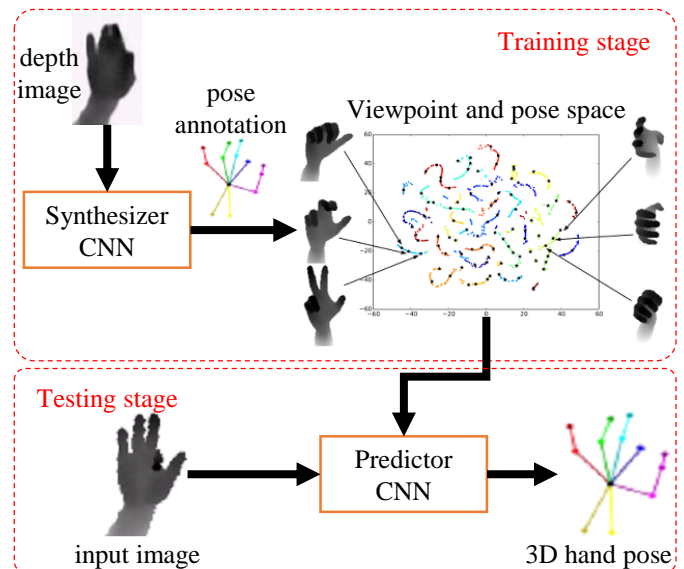


Figure 9: The estimation model of 3D hand pose from the depth image based on training and predicting the position of joints on depth images.

Oberweiger et al. [30] proposed a model called "feedback loop". This model includes Deep Networks and is optimized to use training data. This model is capable of updating the estimated hand pose and provides the experimental results with the NYU dataset as in Fig. 5. Zhang et al. [39] introduced a method for estimating the 3D pose of the human hand, mouse, and fish from the depth images. This method used CNN to predict joint locations that are represented in the manifold space by Lie group, i.e. each joint of the skeleton is represented in the manifold space by  $SE(3)$ . Five fingers are

modeled by five-subchains. And then, those chains are connected to the palm center joint and integrated into a kinematic tree as the hand skeletal model. There are two variants of this model called "deep-L2S-para" and "deep-L2S-seq", whose experimental results with the NYU dataset are 15.84mm and 14.15mm, respectively. Wan et al. [45] proposed a dual generative model that captures the latent spaces of hand poses. This model uses the variation auto-encoder (VAE) and the generative adversarial network (GAN) for estimating 3D hand pose. In more detail, this model generates the synthesized realistic depth maps of highly articulated hand poses under dramatic viewpoint changes and reduces the number of annotated training data. The model of Neverova et al. [47] allows extracting information automatically from real data by deploying a semi-supervised and weakly-supervised training algorithm. These two learning methods are trained from two different datasets, i.e. the synthetic and the real dataset. This method aims to the objective by which to perform frame-by-frame without any dynamic information. The average of 3D joints error is 14.8mm. Authors verify that this method is better than some other methods like DeepPrior [25], Hand3D [41], Crossing nets [45], etc.

The CNN model of Choi et al. [36] [44] uses paired depth images. Firstly, the position of the hand and the object in the image are determined by using CNN to predict the heat-maps. And then, those heat-maps are projected into the space of 3D hand and CAD models of the synthetic dataset. At the same time, a synthetic dataset of human grasps is also built. The next, authors then classify the hand orientations and grasp type from the multi-channel network to reduce the search space for pose estimation. The model is trained by the synthetic dataset whose number of images is 16.5K. Each grasp is captured by 500 depth maps that are rendered randomly from different objects, orientations, and backgrounds. Being evaluated additionally by a publicly available GUN-714 dataset [87], the average of 3D joints error is smaller 20mm.

In the study of Baek et al. [52], the corresponding ground-truth hand poses annotations, and the skeleton entries are the input depth maps of the training stage. The skeleton entries of each dataset are generated from separate hand pose generator (HPG) and 3D hand pose estimator (HPE). This is because the training on input depth maps and the corresponding ground-truth hand pose annotations are not enough to cover variations in poses, shapes, views, etc. CNN is trained by the skeletal hand shape model of the Big Hand 2.2M dataset [42]. The number of added skeletal poses is greater than the number of existing datasets, i.e. Big Hand 2.2M, ICVL, NYU, and MSRA. In more detail, with the Big Hand 2.2M dataset, the average of joints error reduces from 17.1mm to 12.5mm.

Madadi et al. [55] used a novel hierarchical tree-like structured CNN, whose branches are trained to become specialized in predefined subsets of hand joints, called "local poses". Being extracted from hierarchical CNN branches, local pose features are fused to learn higher-order dependencies among joints in the final pose by end-to-end training. Especially, the used loss function is also defined to incorporate appearance and physical constraints about double hand motion and deformation. This function is used to optimize network parameters during training and regression stages. The averages of joints error are 11.0mm and 9.7mm when evaluated by NYU and MSRA datasets, respectively. Rad et al. [56] use a Deep Network to predict a 3D pose from an image. This Deep Network is trained

by the features that are computed for a real image and in a synthetic image of the same pose. The average result of 3D joints error of this approach with the NYU dataset is 7.4mm. Zhang et al. [60] used the cascaded hierarchical regression in [31] to get rough locations of hand joints and proposed a refinement stage to re-estimate joint locations of stretching-out fingers. Therein, the authors used the method in [88] to predict the key joints localization. Evaluating by MSRA and ICVL datasets, the average errors for estimating all fingers are 18.02mm and 13.65mm, respectively. For estimating all fingertips, there are 20.12mm and 14.30mm, respectively. Considering the physical constraints of human hand kinematics, Wohlke et al. [61] proposed a hybrid approach that has embedded a kinematic layer into the CNN. The size of the input image is standardized over BoxNet, RotNet, and ScaleNet whose size is  $176 \times 176$  pixels. The residual network [89] is used to estimate hand parameters and a kinematic hand model layer (FKINE) forwards kinematics from hand parameters to joint locations. The hand has 61 parameters. The average of 3D joints error is 11mm with the NYU dataset.

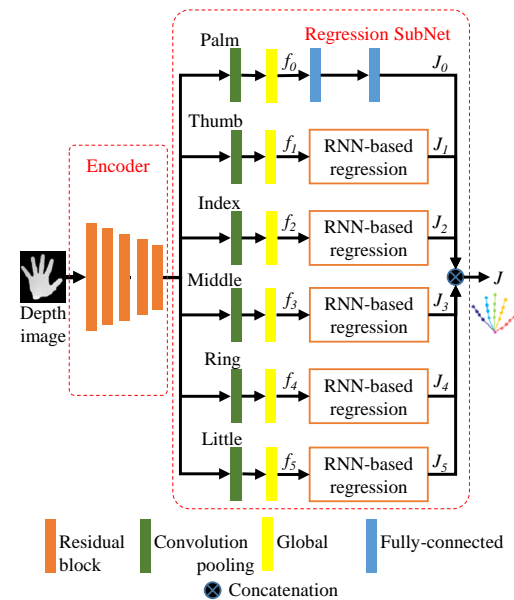


Figure 10: The HCRNN architecture for 3D hand pose estimation from a single depth image [72].

In the past year, the depth image also is used to estimate 3D hand pose in some studies. Yoo et al. [72] divide the hand into six parts, i.e. the palm and five fingers, as in Fig. 10. Then authors proposed a hierarchically-structured convolutional recurrent neural network (HCRNN) with six branches that correspond with those parts. This study exploits effectively the 2D characteristics of the depth image as input of the CNNs. Due to each branch of CNN trains and predicts a part of the hand, this approach has a very fast processing time, up to 240 fps on a single GPU. Being evaluated on the ICVL, NYU and MSRA datasets, the average of 3D joints error is 6.6mm, 9.4mm, and 7.8mm, respectively.

The presented studies are based on two methods [79] as illustrated in Fig. 7, including detection-based method (as the first branch) and regression-based method (as the second one). Facing lose spatial information of hand structure problem and lack direct supervision of joint coordinates problem, a new method of Zhang et

al. [79], called "Pixel-wise Regression", use spatial form representation (SFR) and differentiable decoder (DD). The authors explain their method as a combination of the two former above methods. Comparing with the state-of-the-art technique, this method reduces mean 3D joint error by more than 25%. Specifically, 3D joints error on the MSRA dataset is 5.186mm.

### 3.2 Estimating by RGB Image

As illustrated in Fig. 11, the 3D hand pose estimation process from RGB image usually contains five steps as follows:

- Predicting heat-maps on image space by a CNN.
- Predicting 2D hand pose.
- Training 2D hand pose and the ground truth of 3D hand pose of the synthetic dataset to generate a 3D model.
- Predicting 2D hand pose by real input data.
- Using 3D hand pose estimated model and 2D hand pose of the real data as input data to output 3D hand pose.

A few years ago, i.e. from 2015 to 2017, depth images are usually considered as input data of CNN to estimate 3D hand pose and skeleton. However, depth data is less common than color data in real-life because of the unpopularity and the expensive of depth sensors/cameras. Furthermore, CNN technologies have developed strongly. Therefore, in the last two years, researchers also use RGB images in their studies.

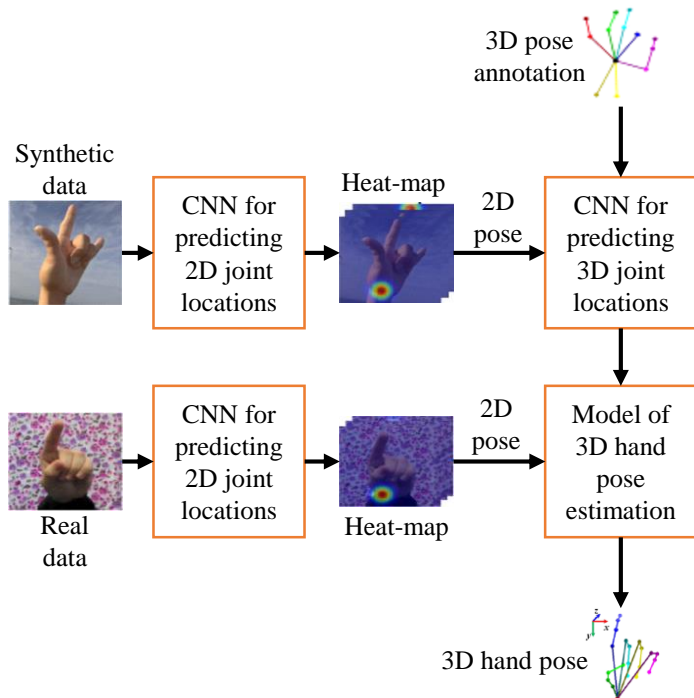


Figure 11: The model of 3D hand pose estimation from the RGB image.

In their article, Zimmermann et al. [50] estimate 3D hand poses from regular RGB images by three deep networks. The first CNN [90] provides a hand segmentation to locate the hand in the image.

The second CNN [90] models 2D hand pose in the 2D images. The last CNN [91], [92] predicts the 3D hand pose from this model. Being evaluated on the Percentage of Correct Keypoints (PCK) measurement with the RWTH German Fingerspelling dataset, the result is 32%.

Panteleris et al. [67] recover the 3D hand pose by using least-squares minimization to fits a 3D model of the hand to the estimated 2D joint positions. Those 2D data are generated by the pre-trained network of OpenPose [93] from the detected hand in the image using YOLO v2 [94]. Authors evaluate their method by three datasets, including the Stereo hand pose dataset, Synthetic dataset, and Hands in action RGB-D dataset. The result of error thresholds is less than 30mm.

A method of Spurr et al. [65] generates a 3D hand pose from an RGB image by learning a single unified latent space via an extension of the VAE (Variational AutoEncoders) framework. The data of latent space, i.e. the RGB images and 3D joint configurations, are illustrated by the blue and green colors. The Stereo Hand Pose Tracking Benchmark (STB) and the Rendered Hand Pose Dataset (RHD) datasets are used to evaluate their model.

Chen et al. [59] develop tonality-alignment generative adversarial networks (TAGANs) that have high-quality ability to generate hand pose. The working mechanism of this network is aligning the tonality and color distributions between synthetic hand poses and real backgrounds. However, hand pose datasets are not large enough to learn a stable CNN hand pose estimator. Therefore this method adopted an opensource AR simulator to produce large-scale and high-quality hand pose images with accurate 2D/3D hand-keypoint labels. The authors used the convolutional pose machine (CPM) [90] for predicting and the Hand3D [50] for estimating 3D hand pose. The experimental results are 19.9mm and 7.3mm with RHP and STB datasets, respectively.

The idea of He et al. [95] is a hand-model regularized graph CNN trained under a generative adversarial learning framework (GraphPoseGAN) that contains two modules. The first "hand model module" generates a template 3D hand pose as a prior. Its inside encoder extracts the latent code  $z$  from the input image and a parametric hand model. The second "GCN refinement module" is used to refine 3D hand pose from 3D ground truth to choose a hand pose whose parameter is the best. Being evaluated by Stereo Hand Pose Tracking Benchmark (STB) and the Rendered Hand Pose (RHD) datasets, this model gets the average error in Euclidean space between the estimated 3D joints and the ground truth joints is 12.4mm (RHD) and 4.2mm (STB).

Being introduced at CVPR in 2019, the method of Baek et al. [80] predicts 2D heatmaps from 2D feature extractor and 2D hand mask. After that, the 3D skeleton of the input data is regressed by a 2D skeleton and 3D skeleton of the supervision stage. The used datasets are the Stereo Hand pose Dataset (SHD).

### 3.3 Estimating by RGB-D Image and Other Data

Choi et al. [27] developed a real-time algorithm to use RGB-D data. This method used the local shape descriptors to retrieve nearest neighbors from the labeled dataset. And then this information is used to evaluate the unknown pose parameters by a joint matrix factorization and completion (JMFC) approach on a hand pose library.



The method of Mueller et al. [10] provides a real-time, robust, and accurate hand pose estimation that uses RGB-D data of egocentric cameras. The collected data is clutter and occlusions. Firstly, this method uses a HALNet CNN to estimate the 2D position of the hand center in the input. And then a generated normalized cropped image is fed into a JORNet CNN to regress relative 3D hand joint locations. Both of those CNNs are trained with the new SynthHands dataset. Being evaluated by the EgoDexter benchmark dataset, the lowest average error is 32.6 mm.

In their research, Iqbal et al. [96] introduce a novel 2.5D pose representation to and a solution to reconstruct the 3D pose from this 2.5D model. In this scaled and invariant representation, the 2D coordinates are the coordinates of the points on the image and the remaining 0.5D is the coordinates of the palms that are predicted from the depth. The average End-Point-Error (EPE) is 25.56mm and 31.86mm with SHP and RHP datasets, respectively. Cejneg et al. [76] used the Pose-REN method [97] to train the Hands2017 dataset.

Extending the success of detecting, identifying, and estimating objects from the color image and depth image, CNNs have been used to work with 3D data, i.e. point cloud. Li et al. [98] proposed a novel CNN for working with an un-organization point cloud data. There are 1024 points in this 3D data. This CNN computes the point-wise features from each point by the PEL (Permutation Equivariant Layer) residual network. And then those features are used by the point-to pose voting to estimate the point of hand pose. By the NYU dataset, the mean joint error is 8.99mm and 8.35mm corresponds with the single view and the three views, respectively. Besides, several CNNs use 3D points as the input data including point-wise CNN [99], Deep KD-Networks [100], Self-Organizing Net [101], and Dynamic Graph CNN [102].

### 4 Findings/Results

Based on the surveys of 3D hand pose estimation using the CNNs presented in Tab. 1. The second issue discussed in this paper (ii) is the results of it when using 2D CNN and 3D CNN. As shown above, the objective of existing methods is the location joints estimation based on the 2D, 2.5D, and 3D data. So there are two types of CNNs, i.e. 2D and 3D, as illustrated in Fig. 12. We collect the results of estimating 3D hand pose by CNNs as in the Tab. 3. The average 3D distance error when using 3D CNN is lower than using 2D CNN. This problem happens because the input data of the 3D CNNs is the 3D data. Therefore, the accuracy of 3D CNN is better than 2D CNN, as shown in Tab. 3. When training on the 3D data to generate a 3D hand pose estimation model. Therefore, the 3D data has a higher number of dimensions than 2D data, the computational time is higher, as shown in the Tab. 4.

The third issue (iii) is discussed in this paper is the results and challenges of egocentric vision datasets. Most of the proposed methods for estimating 3D hand pose are quantitatively evaluated on MSRA, NYU, ICVL, etc. In these databases, the hands are often the full joints thus the results are high accuracy. Figure 13 shows the results of 3D hand pose estimation on the egocentric vision (EgoDexter) dataset [10] and based on reading paper, the average 3D error is 32.6mm.

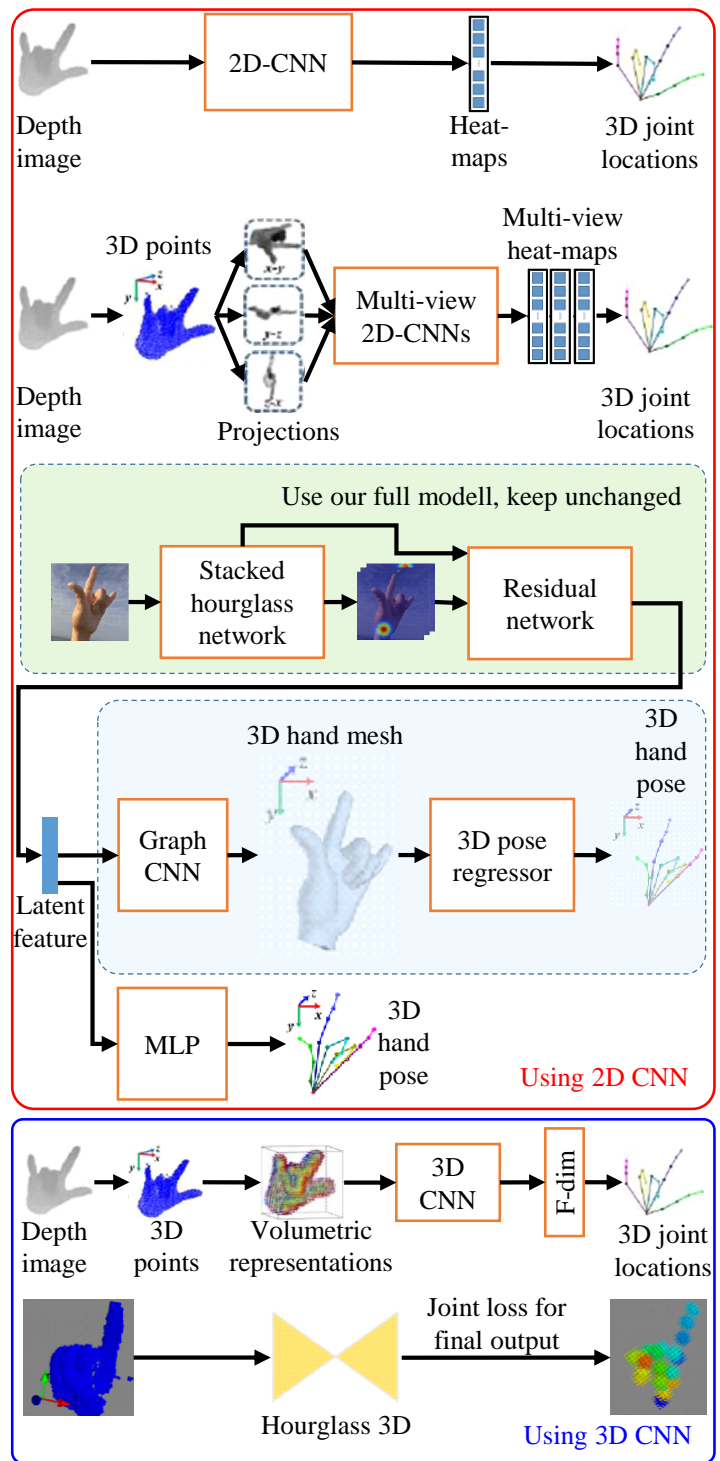


Figure 12: Illustration of two CNNs types: 2D CNN, 3D CNN.

This error is very high compared to other datasets (ICVL is 6.28 - 10.4mm, NYU is 8.42 - 20.7mm, MSRA is 7.49 - 13.1mm). Figure 14 is also shown a comparison of 3D hand pose estimation results on BigHand dataset and egocentric dataset [103]. The results on the BigHand dataset have more than 90% of 3D distance errors being less than 10mm (Fig. 14(top)). On the egocentric dataset is about 30% of 3D distance errors being less than 10mm (Fig. 14(bottom)).

Table 3: The average 3D distance error of the CNNs on the ICVL, NYU, MSRA datasets for 3D hand pose estimation [72].

Method	Mean error (mm)			Input	
	ICVL	NYU	MSRA	2D	3D
Multi-view CNNs[40]	-	-	13.1	✓	
DISCO [104]	-	20.7	-	✓	
DeepPrior [25]	10.4	19.73	-	✓	
Feedback [30]	-	15.97	-	✓	
Global2Local [55]	-	15.6	12.8	✓	
CrossingNets [45]	10.2	15.5	12.2	✓	
HBE [105]	8.62	-	-	✓	
REN (4x6x6) [106]	7.63	13.39	-	✓	
REN (9x6x6) [107]	7.31	12.69	9.79	✓	
DeepPrior++ [51]	8.1	12.24	9.5	✓	
Pose-REN [97]	6.79	11.81	8.65	✓	
Generalized [33]	-	10.89	-	✓	
CrossInfoNet [83]	6.73	10.07	7.86	✓	
HCRNN [72]	6.58	9.41	7.77	✓	
3D CNN [46]	-	14.1	9.58		✓
SHPR-Net [64]	7.22	10.78	7.96		✓
3D DenseNet [58]	6.7	10.6	7.9		✓
Hand PointNet [108]	6.94	10.5	8.5		✓
Point-to-Point [69]	6.33	9.04	7.71		✓
V2V-PoseNet [62]	6.28	8.42	7.49		✓

Table 4: The test speed of the CNNs with a single GPU [72].

Method	Test speed (fps)	Input	
		2D	3D
V2V-PoseNet [62]	3.5		✓
Point-to-Point [69]	41.8		✓
HandPointNet [108]	48		✓
3D DenseNet [58]	126		✓
3D CNN [46]	215		✓
DeepPrior++ [51]	30	✓	
Generalized [33]	40	✓	
CrossingNets [45]	90.9	✓	
CrossInfoNet [83]	124.5	✓	
Feedback [30]	400	✓	
HCRNN [72]	240	✓	

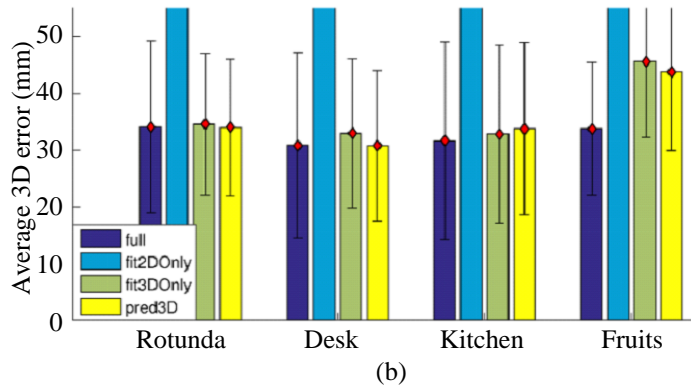
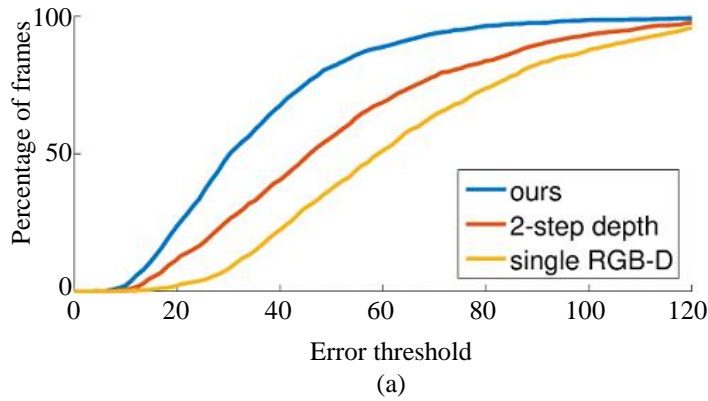


Figure 13: The distribution of 3D distance errors on the EgoDexter dataset [10]; (a) 3D distance errors of 3D hand pose estimation by the CNN that combine of HALNet and JORNet CNNs on the EgoDexter dataset; (b) the average 3D error of 3D hand pose estimation on the EgoDexter dataset.

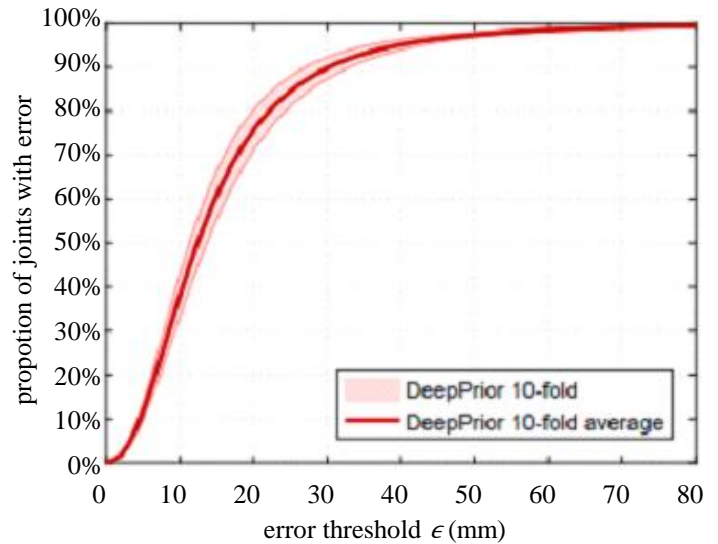
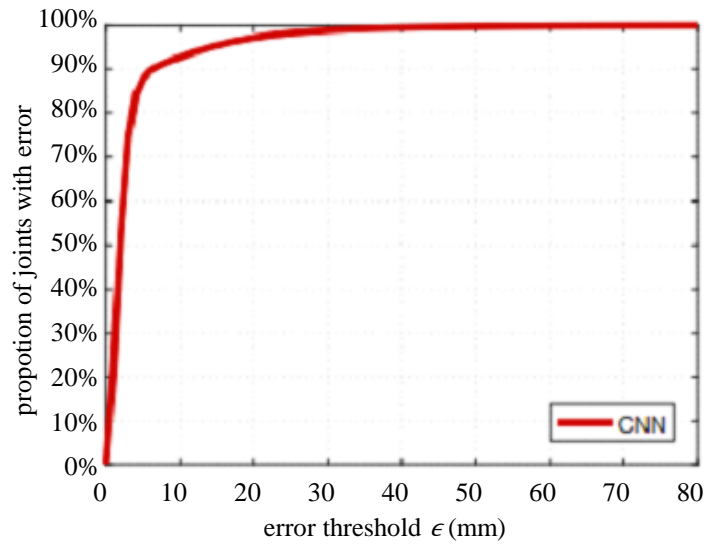


Figure 14: The distribution of 3D distance errors by baseline CNN on BigHand dataset (a) and Egocentric dataset (b) [103].

Based on the survey of 3D hand pose estimation on the egocen-



tric datasets, the estimated distance error is high because the hand is obscured by many objects or the view of the camera just looks at the palm hand, as illustrated in Fig. 15.

In reality, the real-time hand pose estimation from moving, the camera viewpoints in cluttered real-world scenes of the hand is often occluded as it naturally interacts with objects, remains an unsolved problem. In real activities as game playing, human interaction, the image data of scenes are collected from cameras mounted on the head (for VR/AR applications), shoulder, or chest. Occlusions, cluttered backgrounds, manipulated objects, and field-of-view limitations make this scenario particularly challenging. Therefore, the problem of estimating 3D hand pose on the egocentric datasets must be studied in the future.



Figure 15: The illustration data of hands are lost or obscured [109].

## 5 Discussion

To verify the proposed CNN, researchers often use some standard datasets. The greater number of those datasets lets many challenges when building CNN. The properties of datasets, i.e. the training set, testing set, validation set, and evaluation matrix, and those challenges are shown as followings:

### 5.1 Benchmark Datasets

#### 5.1.1 Obtained Datasets from A Fixed Number of Perspectives

**NYU dataset** [26] includes 72757 and 8252 images of training and testing set, respectively. Each frame consists of a pair of RGB images and depth images from three MS Kinect v1, i.e. a frontal view and two side views. Those images are annotated by the ground-truth hand-pose. The authors used the Randomized Decision Forest (RDF) to train a binary classification model by this dataset. And then this classification segments each pixel that belongs to a hand or background in the depth image. 3D ground truth includes 42 DOF of 25 joints.

76k depth images of 9 subjects' right hands are captured using Intel's Creative Interactive Gesture Camera in **MSRA dataset** [110]. Each subject has 17 gestures captured. There are about 500 frames and 21 3D ground truth hand joints per frame, including wrist, index mcp, index pip, index dip, index tip, middle mcp, middle pip, middle dip, middle tip, ring mcp, ring pip, ring dip, ring tip, little

mcp, little pip, a little dip, little tip, thumb mcp, thumb pip, thumb dip, and thumb tip. The resolution of the image is  $320 \times 240$  pixels. The camera's intrinsic parameters are also provided, i.e. principal point of the image is (160, 120) and the focal length is 241.42.

**ICVL dataset** [111] includes 22K training frames and 1.6K testing frames that captured by the Intel's Creative Interactive Gesture Camera. It also provides 3D ground truth with 16 hand joints, including palm, thumb root, thumb mid, thumb tip, index root, index mid, index tip, middle root, middle mid, middle tip, ring root, ring mid, ring tip, pinky root, pinky mid, and pinky tip.

**Stereo Hand Pose Tracking Benchmark (STB) dataset** [39] includes 18,000 stereo and depth images with the 3D ground-truth of 21 hand joints. Those truths are palm center(not wrist or hand center), little mcp, little pip, a little dip, little tip, ring mcp, ring pip, ring dip, ring tip, middle mcp, middle pip, middle dip, middle tip, index mcp, index pip, index dip, index tip, thumb mcp, thumb pip, thumb dip, and thumb tip. The stereo is captured by a Point Grey Bumblebee2 stereo camera and the depth image is captured from an Intel Real Sense F200 active depth camera. This dataset also provides the camera parameters.

**Rendered Hand Pose Dataset (RHD)** [50] provides 41258 training images and 2728 testing images whose resolution is  $320 \times 320$  pixels. The images include the RGB and depth images. This dataset also provides the 3D ground truth with 21 joint points. 214971 annotated depth images of the hands of the **Hand-Net dataset** [34] are divided into three groups. The training set includes 202198 images. The testing set contains 10000 images. The validation set has 2773 images. The used sensor is RealSense RGBD. The hand pose annotation is per pixel classes, 6D fingertip pose, and heatmap. There are 102,000 depth images of a subject in the **MSRC dataset** [112]. 100k of them belong to the training set. The resolution is  $512 \times 424$  pixels and the number of viewpoints is 3. This dataset also provides the annotation data with 22 joint points.

#### 5.1.2 Obtained Datasets from Egocentric Vision

Being captured from the Intel Creative camera mounted on the chest of humans from the right hand and left hand, the **UCI-EGO dataset** [11] provides 400 frames. 3D annotations of keypoints with 26 joint points are also provided. To annotate this dataset for evaluating 3D hand pose estimation and hand tracking the authors developed a semi-automatic labeling tool which allows to accurately annotate partially occluded hands and fingers in the 3D space by using the techniques: A few 2D joints are first manually labeled in the image and used to select the closest synthetic exemplars in the training set; A full hand pose is then created combining the manual labeling and the selected 3D exemplar; This pose is manually refined, leading to the selection of a new exemplar, and the creation of a new pose; This iterative process is followed until acceptable labeling is achieved.

**Graz16 dataset** [113] has more than 2000 depth frames of several egocentric sequences of six subjects. 3D annotations are made with 21 joint points. The size of the image is  $320 \times 240$  pixels. The authors proposed a semi-automated the application that makes it easy to annotate sequences of articulated poses in the 3D space. This application asks a human annotator to provide an estimate of the 2D re-projections of the visible joints in frames they are called

reference frames. It proposes a method to automatically select these reference frames to minimize the annotation effort, based on the appearances of the frames over the whole sequence. It then uses this information to automatically infer the 3D locations of the joints for all the frames, by exploiting appearance, temporal, and distances constraints.

**Ego Dexter dataset** is an RGB-D dataset for evaluating hand tracking and 3D hand pose estimation in the cases of occlusions and clutter. It is captured from Intel RealSense SR300. It consists of 4 sequences with 4 actors (2 female), and varying interactions with various objects and cluttered background. Fingertip positions were manually annotated 1485 frames in 3190 frames.

**Dexter+Object dataset** [114] provide 3014 frames with ground truth annotations. The frames are collected in pairs: RGB frame is captured from the Creative Senz3D color camera; The depth frame is captured from Creative Senz3D close range TOF depth camera. It consists of 6 sequences of a hand manipulating a cuboid (2 different sizes) in different hand-object configurations and grasps. The annotation of hand joints manually annotated pixels on the depth image to mark 5 fingertip positions, and 3 cuboid corners. It is illustrated in Fig. 16.

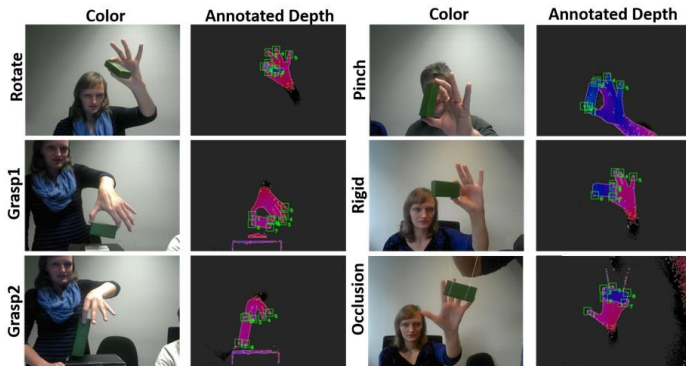


Figure 16: The illustration annotation of the Ego Dexter dataset [114].

Especially, the **BigHand2.2M dataset** [103] provides 2.2 million depth maps of ten subjects (7 males, 3 females) with accurately annotated joint locations. To determine the 3D annotations, the authors use two hardware synchronized electromagnetic tracking units, including six 6D magnetic sensors and one mid-range transmitter. The captured device is the Intel RealSense SR300 camera whose maximum speed is 60fps. The resolution is  $640 \times 480$  pixels and the number of degrees of freedom (DOF) is 31. This dataset is divided into three parts, including 1.534 million images of the prior predefined pose, 375K images of random poses, and 290K images of egocentric poses.

Hampali et al. [77] introduce a benchmark dataset with 80,000 frames of 10 different users. They manipulate one among 10 different objects from the YCB dataset. The size of both depth and RGB image is  $640 \times 480$  pixels. This dataset is synchronized from five cameras. The authors also proposed a method to automatically annotate each frame with accurate estimates of the poses, despite large mutual occlusions.

From the reality of the egocentric datasets, the data of hands

are suffering from occlusions, cluttered backgrounds, manipulated objects, and field-of-view limitations. Unlike human pose estimation, the size of a person is large so it is easier to get a standard benchmark with a hand, thus there exist no standard benchmarks for hand pose estimation, especially in egocentric datasets. As illustrated in Fig. 17, the data of the fingers is obscured, the annotated joints of these fingers are difficult. Although there are already some semi-automatic annotation methods like in [11], [113]. However, all methods have errors as shown in Table 3 [113]. Therefore, to annotate the joints for evaluating 3D hand pose estimates on the egocentric vision datasets requires further research. This is also the fourth discussion (iv) in this paper.



Figure 17: The illustration of fingers is occluded of UCI-EGO dataset [11].

### 5.1.3 Evaluation Measurements

There are three measurements to evaluate 3D hand pose estimation as follows:

- The first is **3D pose error**, which is the average error in Euclidean space between the estimated 3D joints and the ground truth joints.
- The second is **3D PCK**, as the percentage of correct key points of which the Euclidean error distance is below a threshold.
- The last is **AUC**, which is the area under the curve on PCK for different error thresholds.

## 5.2 Challenges

The 2017 Hands in the Million Challenge [115] is built on the BigHand2.2M [103] and First-Person Hand Action [109] datasets. This challenge had two tasks, i.e. 3D hand pose tracking and 3D hand pose estimation when hand interacts with different objects (e.g. juice bottle, salt bottle, knife, milk bottle, soda can, etc.). Based on this challenge, the proposed method of [116] is accepted in the IEEE Conference on Computer Vision and Pattern Recognition 2018. The tasks of the HANDS 2019 challenge [117] are Depth-Based 3D Hand Pose Estimation in the BigHand2.2M [103], Depth-Based 3D Hand Pose Estimation while Interacting with Objects in the F-PHAB [109] and RGB-Based 3D Hand Pose Estimation while Interacting with Objects in the HO-3D [77].

## 6 Conclusions

3D hand pose estimation problem is applied in many applications of computer vision and robotics: human-computer interaction; gesture recognition, interactive games, Computer-Aided Design (CAD), sign languages, action recognition, etc. The studies of 3D hand pose estimation for recognizing the grasping attributes of the objects, thereby promoting the development of robotic arms grasping objects. Before building these applications, a 3D hand pose should be fully estimated. When grasping objects, the data of the hand will be lost, missing, obscured, and are collected from cameras mounted on the head (for VR/AR applications), shoulder, or chest, thus, the process of estimating the 3D hand pose is a challenge. In this paper, we survey by the CNN methods, datasets, results of 3D hand pose estimation according to the type input data. Studies have shown that to estimate the 3D hand pose, it is necessary to use 3D hand pose libraries or 3D ground truth data to regress 3D hand pose. We also analyzed the challenges and current results of CNNs for 3D hand pose estimation on the normal benchmark datasets and egocentric datasets. In particular, we discussed internally on four issues in estimating 3D hand pose: The number of valuable studies about 3D hand pose estimation; The estimated results of 3D hand pose when using 3D CNNs and 2D CNNs; The challenges of the datasets which are collected from egocentric vision sensors; The methods to collect and annotate datasets from egocentric vision sensors. In the future, we plan to build a benchmark dataset to evaluate 3D hand pose estimation. This dataset will use the egocentric camera to collect hand data while grasping the objects. We also plan to propose a method for 3D location joints in manifold space that uses the Lie group, then extract the characteristics for training to generate an estimation model by CNNs to predict 3D location joints.

**Acknowledgment:** This research was funded by the elementary level topic of Hung Vuong University, Vietnam. The title is "Using the Lie algebra, Lie group to improve the skeleton hand presentation".

## References

- [1] P. Krejov, A. Gilbert, R. Bowden, "Guided optimisation through classification and regression for hand pose estimation," *Computer Vision and Image Understanding*, **155**, 124–138, 2016.
- [2] P. Krejov, R. Bowden, "Multi-touchless: Real-time fingertip detection and tracking using geodesic maxima," in 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, FG 2013, 2013.
- [3] Y. Zhou, G. Jiang, Y. Lin, "A novel finger and hand pose estimation technique for real-time hand gesture recognition," *Pattern Recognition*, **49**, 102–114, 2016, doi:10.1016/j.patcog.2015.07.014.
- [4] A. Spurr, "Gesture Recognition : Hand Pose Estimation," ETH Zurich, 2014.
- [5] A. T. Chan, H. V. Leong, S. H. Kong, "Real-time tracking of hand gestures for interactive game design," in IEEE International Symposium on Industrial Electronics, 98–103, 2009.
- [6] R. Y. Wang, S. Paris, J. Popovic, "6D hands: Markerless hand tracking for computer aided design," in UIST'11 - Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, 549–557, 2011.
- [7] J. Isaacs, S. Foo, "Optimized wavelet hand pose estimation for American sign language recognition," in Proceedings of the 2004 Congress on Evolutionary Computation, CEC2004, volume 1, 797–802, 2004.
- [8] YOLOv2, "Hand detection, Gesture recognition by YOLOv2," [https://www.youtube.com/watch?v=JJPshpVt\\_1A](https://www.youtube.com/watch?v=JJPshpVt_1A), 2020, [Accessed 31 May 2020].
- [9] T. Malisiewicz, "Hand Segmentation for Gesture Recognition in EGO-Vision," <https://twitter.com/quantombone/status/425560005434015744>, 2020, [Accessed 31 May 2020].
- [10] F. Mueller, D. Mehta, O. Sotnychenko, S. Sridhar, D. Casas, C. Theobalt, "Real-Time Hand Tracking under Occlusion from an Egocentric RGB-D Sensor," in Proceedings of the IEEE International Conference on Computer Vision, volume 2017-October, 1163–1172, 2017.
- [11] G. Rogez, M. Khademi, J. S. Supanovic, J. M. Montiel, D. Ramanan, "3D hand pose detection in egocentric RGB-D images," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 8925, 356–371, 2015.
- [12] Microsoft, "Microsoft Xbox 360 Kinect Launches November 4," <https://gizmodo.com/microsoft-xbox-360-kinect-launches-november-4-5563148>, 2010, [Online; accessed 7-February-2020].
- [13] J. Kramer, N. Burrus, F. Ehtler, H. C. Daniel, M. Parker, "Hacking the Kinect," Apress, 2012.
- [14] Microsoft, "Kinect v2 comes to the PC on July 15," <https://www.pcgamer.com/kinect-for-windows-v2-to-release-july-15/>, 2014, [Online; accessed 7-February-2020].
- [15] W. G. Wong, "How Microsoft's PrimeSense-based Kinect Really Works," <https://www.electronicdesign.com/technologies/embedded-revolution/article/21795925/how-microsofts-primensebased-kinect-really-works>, 2011, [Online; accessed 7-February-2020].
- [16] Intel, "Intel RealSense Technology," <https://www.intel.com/content/www/us/en/architecture-and-technology/realsense-overview.html>, 2020, [Online; accessed 7-February-2020].
- [17] M. R. Cutkosky, "On grasp choice, grasp models, and the design of hands for manufacturing tasks," *IEEE Transactions on Robotics and Automation*, **5**(3), 269–279, 1989.
- [18] M. Cai, K. M. Kitani, Y. Sato, "Understanding hand-object manipulation with grasp types and object attributes," *Robotics: Science and Systems*, **12**, 2016.
- [19] M. Teremetz, F. Colle, S. Hamdoun, M. A. Maier, P. G. Lindberg, "A novel method for the quantification of key components of manual dexterity after stroke," *Journal of NeuroEngineering and Rehabilitation*, **12**(1), 2015.
- [20] W. Yan, H. Nie, J. Chen, D. Han, "Optimal design and grasp ability evaluation of four-finger tendon-driven hand," *International Journal of Advanced Robotic Systems*, **14**(6), 1–14, 2017, doi:10.1177/1729881417748444.
- [21] Y. Li, Z. Xue, Y. Wang, L. Ge, Z. Ren, J. Rodriguez, "End-to-End 3D Hand Pose Estimation from Stereo Cameras," in BMVC, 1–13, 2019.
- [22] A. Erol, G. Bebis, M. Nicolescu, R. D. Boyle, X. Twombly, "Vision-based hand pose estimation: A review," *Computer Vision and Image Understanding*, **108**(1-2), 52–73, 2007.
- [23] E. Barsoum, "Articulated Hand Pose Estimation Review," <https://arxiv.org/pdf/1604.06195>, 1–50, 2016.
- [24] J. Tompson, M. Stein, Y. Lecun, K. Perlin, "Real-time continuous pose recovery of human hands using convolutional networks," *ACM Transactions on Graphics*, **33**(5), 2014.
- [25] M. Oberweger, P. Wohlhart, V. Lepetit, "Hands Deep in Deep Learning for Hand Pose Estimation," in Computer Vision Winter Workshop, 2015.
- [26] J. Tompson, M. Stein, Y. Lecun, K. Perlin.

- [27] C. Choi, A. Sinha, J. H. Choi, S. Jang, K. Ramani, "A collaborative filtering approach to real-time hand pose estimation," in *Proceedings of the IEEE International Conference on Computer Vision*, volume 2015 Inter, 2336–2344, 2015.
- [28] G. Poier, K. Roditakis, S. Schulter, D. Michel, H. Bischof, A. A. Argyros, "Hybrid One-Shot 3D Hand Pose Estimation by Exploiting Uncertainties," in *BMVC*, 182.1–182.14, 2015.
- [29] P. Li, H. Ling, X. Li, C. Liao, "3D hand pose estimation using randomized decision forest with segmentation index points," in *Proceedings of the IEEE International Conference on Computer Vision*, volume 2015 Inter, 819–827, 2015.
- [30] M. Oberweger, P. Wohlhart, V. Lepetit, "Training a feedback loop for hand pose estimation," in *Proceedings of the IEEE International Conference on Computer Vision*, volume 2015 Inter, 3316–3324, 2015.
- [31] X. Sun, Y. Wei, S. Liang, X. Tang, J. Sun, "Cascaded hand pose regression," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 07-12-June, 824–832, 2015.
- [32] D. Tang, Q. Ye, S. Yuan, J. Taylor, P. Kohli, C. Keskin, T.-k. Kim, J. Shotton, "Opening the Black Box : Hierarchical Sampling Optimization for Hand Pose Estimation," in *iccv*, volume 8828, 1–14, 2015.
- [33] M. Oberweger, P. Wohlhart, V. Lepetit, "Generalized Feedback Loop for Joint Hand-Object Pose Estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **14**(8), 1–1, 2015.
- [34] C. Wan, A. Yao, L. Van Gool, "Hand pose estimation from local surface normals," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 9907 LNCS, 554–569, 2016.
- [35] Q. Ye, S. Yuan, T.-k. Kim, "Spatial Attention Deep Net with Partial PSO for Hierarchical Hybrid Hand Pose Estimation," <https://arxiv.org/abs/1604.03334>, 1–16, 2016.
- [36] A. Sinha, W. Lafayette, "DeepHand : Robust Hand Pose Estimation by Completing a Matrix Imputed with Deep Features," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4150–4158, 2016.
- [37] M. Oberweger, G. Riegler, P. Wohlhart, V. Lepetit, "Efficiently creating 3D training data for fine hand pose estimation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2016-Decem, 4957–4965, 2016.
- [38] C. Xu, L. N. Govindarajan, Y. Zhang, L. Cheng, "Lie-X: Depth Image Based Articulated Object Pose Estimation, Tracking, and Action Recognition on Lie Groups," *International Journal of Computer Vision*, **123**(3), 454–478, 2016.
- [39] Y. Zhang, C. Xu, L. Cheng, "Learning to Search on Manifolds for 3D Pose Estimation of Articulated Objects," <https://arxiv.org/abs/1612.00596>, 2016.
- [40] L. Ge, H. Liang, J. Yuan, D. Thalmann, "Robust 3D Hand Pose Estimation from Single Depth Images Using Multi-View CNNs," *IEEE Transactions on Image Processing*, **27**(9), 4422–4436, 2016.
- [41] X. Deng, S. Yang, Y. Zhang, P. Tan, L. Chang, H. Wang, "Hand3D: Hand Pose Estimation using 3D Neural Network," <https://arxiv.org/abs/1502.06807>, 2017.
- [42] S. Yuan, Q. Ye, B. Stenger, S. Jain, T. K. Kim, "BigHand2.2M benchmark: Hand pose dataset and state of the art analysis," in *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, volume 2017-Janua, 2605–2613, 2017.
- [43] C. Choi, S. Kim, K. Ramani, "Learning Hand Articulations by Hallucinating Heat Distribution," in *Proceedings of the IEEE International Conference on Computer Vision*, volume 2017-October, 3123–3132, 2017.
- [44] C. Choi, S. H. Yoon, C.-N. Chen, K. Ramani, "Robust Hand Pose Estimation during the Interaction with an Unknown Object," in *IEEE International Conference on Computer Vision (ICCV)*, i, 3142–3151, 2017.
- [45] C. Wan, T. Probst, L. V. Gool, A. Yao, "Crossing Nets : Combining GANs and VAEs with a Shared Latent Space for Hand Pose Estimation," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [46] L. Ge, H. Liang, J. Yuan, D. Thalmann, "3D Convolutional Neural Networks for Efficient and Robust Hand Pose Estimation from Single Depth Images," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [47] N. Neverova, C. Wolf, F. Nebout, G. W. Taylor, "Hand pose estimation through semi-supervised and weakly-supervised learning," *Computer Vision and Image Understanding*, **164**, 56–67, 2017.
- [48] X. Zhang, C. Xu, Y. Zhang, T. Zhu, L. Cheng, "Multivariate Regression with Grossly Corrupted Observations: A Robust Approach and its Applications," <https://arxiv.org/abs/1701.02892>, 2017.
- [49] J. Malik, A. Elhayek, D. Stricker, "Simultaneous Hand Pose and Skeleton Bone-Lengths Estimation from a Single Depth Image," <https://arxiv.org/abs/1712.03121>, 2017.
- [50] C. Zimmermann, T. Brox, "Learning to Estimate 3D Hand Pose from Single RGB Images," in *IEEE International Conference on Computer Vision (ICCV)*, 2017, <https://arxiv.org/abs/1705.01389>.
- [51] M. Oberweger, V. Lepetit, "DeepPrior++: Improving Fast and Accurate 3D Hand Pose Estimation," in *Proceedings - 2017 IEEE International Conference on Computer Vision Workshops, ICCVW 2017*, volume 2018-Janua, 585–594, 2017.
- [52] S. Baek, K. I. Kim, T. K. Kim, "Augmented Skeleton Space Transfer for Depth-Based Hand Pose Estimation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 8330–8339, 2018.
- [53] Y. Wu, W. Ji, X. Li, G. Wang, J. Yin, F. Wu, "Context-Aware Deep Spatiotemporal Network for Hand Pose Estimation From Depth Images," *IEEE Transactions on Cybernetics*, 1–11, 2018.
- [54] J. S. Supancic, G. Rogez, Y. Yang, J. Shotton, D. Ramanan, "Depth-Based Hand Pose Estimation: Methods, Data, and Challenges," *International Journal of Computer Vision*, **126**(11), 1180–1198, 2018.
- [55] M. Madadi, S. Escalera, X. Baro, J. Gonzalez, "End-to-end Global to Local CNN Learning for Hand Pose Recovery in Depth Data," <https://arxiv.org/pdf/1705.09606.pdf>, 2018.
- [56] M. Rad, M. Oberweger, V. Lepetit, "Feature Mapping for Learning Fast and Accurate 3D Pose Inference from Synthetic Images," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 4663–4672, 2018.
- [57] G. Garcia-Hernando, S. Yuan, S. Baek, T. K. Kim, "First-Person Hand Action Benchmark with RGB-D Videos and 3D Hand Pose Annotations," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 409–419, 2018.
- [58] L. Ge, H. Liang, J. Yuan, S. Member, D. Thalmann, "Real-time 3D Hand Pose Estimation with 3D Convolutional Neural Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **8828**(c), 2018.
- [59] L. Chen, S.-Y. Lin, Y. Xie, H. Tang, Y. Xue, X. Xie, Y.-Y. Lin, W. Fan, "Generating Realistic Training Images Based on Tonality-Alignment Generative Adversarial Networks for Hand Pose Estimation," <https://arxiv.org/abs/1811.09916>, 2018.
- [60] C. Zhang, G. Wang, H. Guo, X. Chen, F. Qiao, H. Yang, "Interactive hand pose estimation: Boosting accuracy in localizing extended finger joints," in *International Symposium on Electronic Imaging Science and Technology*, 1, 1–6, 2018.
- [61] J. Wohlke, S. Li, D. Lee, "Model-based Hand Pose Estimation for Generalized Hand Shape with Appearance Normalization," <https://arxiv.org/abs/1807.00898>, 2018.



- [62] G. Moon, J. Y. Chang, K. M. Lee, "V2V-PoseNet: Voxel-to-Voxel Prediction Network for Accurate 3D Hand and Human Pose Estimation from a Single Depth Map," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5079–5088, 2018.
- [63] Q. Ye, T. K. Kim, "Occlusion-aware hand pose estimation using hierarchical mixture density network," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 11214 LNCS, 817–834, 2018.
- [64] X. Chen, G. Wang, S. Member, C. Zhang, K. I. M. Member, X. Ji, "SHPR-Net : Deep Semantic Hand Pose Regression From Point Clouds," *IEEE Access*, **PP(c)**, 1, 2018.
- [65] A. Spurr, J. Song, S. Park, O. Hilliges, "Cross-modal Deep Variational Hand Pose Estimation," in *CVPR*, 89–98, 2018.
- [66] F. Huang, A. Zeng, C. Science, H. Kong, H. Kong, M. Liu, J. Qin, Q. Xu, "Structure-Aware 3D Hourglass Network for Hand Pose Estimation from Single Depth Image," <https://arxiv.org/abs/1812.10320>, 1–12, 2018.
- [67] P. Panteleris, I. Oikonomidis, A. Argyros, "Using a single RGB frame for real time 3D hand pose estimation in the wild," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2018.
- [68] C. Wan, T. Probst, L. V. Gool, A. Yao, "Dense 3D Regression for Hand Pose Estimation," in *CVPR*, 2018.
- [69] L. Ge, Z. Ren, J. Yuan, "Point-to-point regression pointnet for 3D hand pose estimation," in *European Conference on Computer Vision*, volume 11217 LNCS, 489–505, 2018.
- [70] Y. Zhang, L. Chen, Y. Liu, J. Yong, W. Zheng, "Adaptive Wasserstein Hourglass for Weakly Supervised Hand Pose Estimation from Monocular RGB," <https://arxiv.org/abs/1909.05666>, 2019.
- [71] S. Sharma, S. Huang, D. Tao, "An End-to-end Framework for Unconstrained Monocular 3D Hand Pose Estimation," <https://arxiv.org/abs/1911.12501>, 1–12, 2019.
- [72] C.-h. Yoo, S.-w. Kim, S.-w. Ji, Y.-g. Shin, S.-j. Ko, "Capturing Hand Articulations using Recurrent Neural Network for 3D Hand Pose Estimation," <https://arxiv.org/abs/1911.07424>, 2019.
- [73] Y. Li, C. Twigg, Y. Ye, L. Tao, X. Wang, "Disentangling Pose from Appearance in Monochrome Hand Images," <https://arxiv.org/abs/1904.07528>, 2019.
- [74] C. Wan, T. Probst, L. Van Gool, A. Yao, "Dual Grid Net: hand mesh vertex regression from single depth maps," <https://arxiv.org/abs/1907.10695>, 2019.
- [75] J. Liu, H. Ding, A. Shahroudy, L. Y. Duan, X. Jiang, G. Wang, A. C. Kot, "Feature Boosting Network for 3D Pose Estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42(2)**, 494–501, 2019.
- [76] L. W. X. Cejnog, R. M. Cesar, T. E. De Campos, V. M. C. Elui, "Hand range of motion evaluation for Rheumatoid Arthritis patients," in *Proceedings - 14th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2019*, 2, 2019.
- [77] S. Hampali, M. Rad, M. Oberweger, V. Lepetit, "HONnotate: A method for 3D Annotation of Hand and Objects Poses," <https://arxiv.org/abs/1907.01481>, 2019.
- [78] S. Li, D. Lee, "Point-to-Pose Voting based Hand Pose Estimation using Residual Permutation Equivariant Layer," in *CVPR*, 2019.
- [79] X. Zhang, F. Zhang, "Pixel-wise Regression : 3D Hand Pose Estimation via Spatial-form Representation and Differentiable Decoder," *XX(Xx)*, 1–10, 2019.
- [80] S. Baek, T.-k. Kim, "Pushing the Envelope for RGB-based Dense 3D Hand Pose Estimation via Neural Rendering," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.
- [81] K.-w. Lee, S.-h. Liu, H.-t. Chen, K. Ito, "Silhouette-Net : 3D Hand Pose Estimation from Silhouettes," <https://arxiv.org/abs/1912.12436>, 2019.
- [82] L. Ge, Z. Ren, Y. Li, Z. Xue, Y. Wang, J. Cai, J. Yuan, "3D Hand Shape and Pose Estimation from a Single RGB Image," in *CVPR*, 2019.
- [83] K. Du, X. Lin, Y. Sun, X. Ma, "Crossinfonet: Multi-task information sharing based hand pose estimation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2019-June, 9888–9897, 2019.
- [84] X. Sun, Y. Wei, S. Liang, X. Tang, J. Sun, "Cascaded hand pose regression," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 07-12-June, 824–832, 2015.
- [85] C. Qi, L. Yi, H. Su, L. Guibas, "PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space," in *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, 5105–5114, 2017.
- [86] D. Tang, H. J. Chang, A. Tejani, T. K. Kim, "Latent regression forest: Structured estimation of 3D hand poses," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39(7)**, 1374–1387, 2017.
- [87] J. Supancic, D. Ramanan, J. S. Supan, "Understanding Everyday Hands in Action from RGB-D Images," in *2015 IEEE International Conference on Computer Vision (ICCV)*, 3889–3897, 2015.
- [88] X. Chen, C. Shi, B. Liu, "Static hand gesture recognition based on finger root-center-angle and length weighted Mahalanobis distance," *Real-Time Image and Video Processing 2016*, **9897(61271390)**, 98970U, 2016.
- [89] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [90] S.-E. Wei, V. Ramakrishna, T. Kanade, Y. Sheikh, "Convolutional pose machines," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016)* url = <http://dx.doi.org/10.1109/CVPR.2016.511>, year = 2016,.
- [91] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," <https://arxiv.org/abs/1603.04467>, 2016.
- [92] D. P. Kingma, J. L. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15, 2015.
- [93] openpose, "openpose," <https://github.com/CMU-Perceptual-Computing-Lab/openpose>, 2019, [Accessed 23 April 2019].
- [94] J. Redmon, A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, volume 2017-Janua, 6517–6525, 2017.
- [95] Y. He, W. Hu, S. Yang, X. Qu, P. Wan, Z. Guo, "GraphPoseGAN: 3D Hand Pose Estimation from a Monocular RGB Image via Adversarial Learning on Graphs," <https://arxiv.org/abs/1912.01875>, 2019.
- [96] U. Iqbal, P. Molchanov, T. Breuel, J. Gall, J. Kautz, "Hand Pose Estimation via Latent 2.5D Heatmap Regression," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 11215 LNCS, 125–143, 2018.
- [97] X. Chen, G. Wang, H. Guo, C. Zhang, "Pose guided structured region ensemble network for cascaded hand pose estimation," *Neurocomputing*, 2019.
- [98] S. Li, D. Lee, "Point-to-pose voting based hand pose estimation using residual permutation equivariant layer," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2019-June, 11919–11928, 2019.



- [99] “Pointwise Convolutional Neural Networks,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018.
- [100] R. Klokov, V. Lempitsky, “Escape from Cells: Deep Kd-Networks for the Recognition of 3D Point Cloud Models,” in *Proceedings of the IEEE International Conference on Computer Vision*, volume 2017-October, 863–872, 2017.
- [101] J. Li, B. M. Chen, “So-net: Self-organizing network for point cloud analysis,” in *Proceedings of the IEEE International Conference on Computer Vision*, 9397–9406, 2018.
- [102] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, J. M. Solomon, “Dynamic graph Cnn for learning on point clouds,” *ACM Transactions on Graphics*, **38**(5), 2019.
- [103] S. Yuan, Q. Ye, u. . h. y. Björn Stenger and Siddhant Jain and Tae-Kyun Kim, booktitle=CVPR, “BigHand2.2M Benchmark: Hand Pose Dataset and State of the Art Analysis,” .
- [104] D. Bouchacourt, P. K. Mudigonda, S. Nowozin, “DISCO Nets : DISsimilarity COefficients Networks,” in D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, R. Garnett, editors, *Advances in Neural Information Processing Systems* 29, 352–360, Curran Associates, Inc., 2016.
- [105] Y. Zhou, J. Lu, K. Du, X. Lin, Y. Sun, X. Ma, “HBE: Hand branch ensemble network for real-time 3d hand pose estimation,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, **11218 LNCS**, 521–536, 2018.
- [106] H. Guo, G. Wang, X. Chen, C. Zhang, F. Qiao, H. Yang, “Region ensemble network: Improving convolutional network for hand pose estimation,” in *Proceedings - International Conference on Image Processing, ICIP*, volume 2017-Septe, 4512–4516, 2018.
- [107] H. Guo, G. Wang, X. Chen, C. Zhang, “Towards Good Practices for Deep 3D Hand Pose Estimation,” *Journal of Visual Communication and Image Representation*, **55**, 404–414, 2017.
- [108] L. Ge, Y. Cai, J. Weng, J. Yuan, “Hand PointNet : 3D Hand Pose Estimation using Point Sets,” *Cvpr*, 3–5, 2018.
- [109] G. Garcia-Hernando, S. Yuan, S. Baek, T.-K. Kim, “First-Person Hand Action Benchmark with RGB-D Videos and 3D Hand Pose Annotations,” in *Proceedings of Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [110] H. Su, S. Maji, E. Kalogerakis, E. Learned-Miller, “Multi-view Convolutional Neural Networks for 3D Shape Recognition,” in *Proc. ICCV*, 264–272, 2015.
- [111] D. Tang, H. J. Chang, A. Tejani, T. K. Kim, “Latent regression forest: Structured estimation of 3D hand poses,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**(7), 1374–1387, 2017.
- [112] “Accurate, robust, and flexible realtime hand tracking,” in *Conference on Human Factors in Computing Systems - Proceedings*, volume 2015-April, 3633–3642, 2015.
- [113] M. Oberweger, G. Riegler, P. Wohlhart, V. Lepetit, “Efficiently creating 3D training data for fine hand pose estimation,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2016-December, 4957–4965, 2016.
- [114] S. Sridhar, F. Mueller, M. Zollhoefer, D. Casas, A. Oulasvirta, C. Theobalt, “Real-time Joint Tracking of a Hand Manipulating an Object from RGB-D Input,” in *Proceedings of European Conference on Computer Vision (ECCV)*, 2016.
- [115] ICCV, “The 2017 Hands in the Million Challenge on 3D Hand Pose Estimation,” <http://icvl.ee.ic.ac.uk/hands17/challenge/>, 2017, [Online; accessed 22-march-2020].
- [116] S. Yuan, G. Garcia-Hernando, B. Stenger, G. Moon, J. Yong Chang, K. Mu Lee, P. Molchanov, J. Kautz, S. Honari, L. Ge, J. Yuan, X. Chen, G. Wang, F. Yang, K. Akiyama, Y. Wu, Q. Wan, M. Madadi, S. Escalera, S. Li, D. Lee, I. Oikonomidis, A. Argyros, T.-K. Kim, “Depth-Based 3D Hand Pose Estimation: From Current Achievements to Future Goals,” in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [117] ICCV, “the HANDS19 Challenge,” <https://sites.google.com/view/hands2019/challenge>, 2019, [Online; accessed 22-march-2020].

## Nearest Neighbour Search in k-dSLst Tree

Meenakshi Hooda\*, Sumeet Gill

Department of Mathematics, Maharshi Dayanand University, 124001, India

---

### ARTICLE INFO

Article history:

Received: 12 May, 2020

Accepted: 26 June, 2020

Online: 18 July, 2020

---

Keywords:

Nearest Neighbour

Spatial Indexing

k-d tree

Sorted Linked List

Duplicate Keys

---

### ABSTRACT

In the last few years of research and innovations, lots of spatial data in the form of points, lines, polygons and circles have been made available. Traditional indexing methods are not perfect to store spatial data. To search for nearest neighbour is one of the challenges in different fields like spatiotemporal data mining, computer vision, traffic management and machine learning. Many novel data structures are proposed in the past, which use spatial partitioning and recursive breakdown of hyperplane to find the nearest neighbour efficiently. In this paper, we have adopted the same strategy and proposed a nearest neighbour search algorithm for k-dSLst tree. k-dSLst tree is based on k-d tree and sorted linked list to handle spatial data with duplicate keys, which is ignored by most of the spatial indexing structures based on k-d tree. The research work in this paper shows experimentally that where the time taken by brute force nearest neighbour search increases exponentially with increase in number of records with duplicate keys and size of dataset, the proposed algorithm k-dSLstNearestNeighbourSearch based on k-dSLst tree performs far better with approximately linear increase in search time.

---

## 1. Introduction

Geospatial datasets are huge and complex in structure and relationships. We need complex spatial operators to fetch the required information from spatial datasets. These complex spatial operators include intersection, overlap, adjacency etc. The traditional indexing structures can't handle queries related to spatial details efficiently. Spatial queries use spatial relationships among different geometries and make use of n-dimensional geometric data such as points, lines and polygons to retrieve required data.

A spatial query retrieves features depending on relationships of spatial data with geometry of queried data. The objective of spatial queries is to find out the spatial relationships in one or more subjects to search for spatial objects. The extracted information help in taking decisions related to various policies or for doing analysis in various fields. Spatial queries permit for the utilization of data types related to geometry like point, line etc. and take into consideration the spatial relationship in these geometries. Spatial indices are used for spatial database to optimize spatial query. The indices provide one of the optimization techniques for improving the quality of services based on location.

### 1.1 Range Query

The range queries are related with bounded areas. The output of such kind of queries contains some region that might be overlapped. The spatial objects are associated to each other within specific area or distance. These kinds of queries have associated area and require at least two parameters i.e. location and boundary limit.

Example:

Find all the hostels within 12 kilometers of a given university.

Find all towns within 90 kilometers of a given village.

### 1.2 Spatial Join Query

These queries are the combination of more than one spatial query. We need spatial join operation for retrieving required information. These queries are expensive as join condition(s) involve areas and their proximity to each other.

Example:

Find all the cities near given lake.

---

\*Meenakshi Hooda, Maharshi Dayanand University, Haryana, India & mshthebest@gmail.com

The partitioning and summarization of spatiotemporal data is significant for numerous community based applications like environmental science, public safety and health [1]. The relations among spatial objects can be categorized in different ways such as classifications based on distance, topology and direction relations. These may be further joined using logical operators to depict neighbourhood relation between spatial objects [2].

## 2. Related Work

### 2.1 Nearest Neighbour Query

Nearest Neighbour Search (NNS) is also called as Proximity, Similarity or Closest Point Search. It is an optimization problem to find nearest points in metric spaces [3]. In such kind of queries, the data closet to the queried location is retrieved. Here, we search for objects which are nearest to the particular location. We can also query for xNN i.e. x nearest neighbours, in which search is done for x objects nearest to the queried location. The result might be ordered by their respective proximity to queried spatial location. [4] introduce a method to fetch nearest neighbour data in 3-Dimensional space by making use of clustered hierarchical indexing tree structure. Studies show that this approach achieves remarkable improvement in response time analysis as compared to already existing spatial data accessing methods in databases. [5] research work introduces a progressive algorithm for a search of approximate k-nearest neighbor. The most of the KNN algorithms though utilize k-nearest neighbor libraries for many of the data analysis procedures, but the fact is that these algorithms run only after indexing of the whole dataset, which means that the datasets are not online. [6] proposed a method of parallel kd-tree construction for 3-D points on a Graphic Processing Unit. The method consists of a sorting algorithm to maintain high level of parallelism throughout the creation.

Example: Find 5 nearest hostels with respect to given location comprised of n-dimensional coordinate.

In this research paper, we are introducing the algorithm k-dSLstTreeNearest for Nearest Neighbour in k-dSLst tree structure designed in our previous work to index spatial data with duplicate keys.

### 3. Brute Force Method: Nearest Neighbour Search for Spatial Data with Duplicate Keys

Brute Force exploration process, which is also called as Exhaustive/Blind Search, is extremely generalized technique to solve any problem. It systematically enumerates all possible contenders for the problem's solution and tests to check if the contender fulfills the statement of the problem [7]. In this algorithm, we find the distance of every object in dataset from the position of the queried object. As, multiple objects can be found at the nearest neighbour location, a list is maintained to keep record of all objects with minimum distance. Algorithm 1, **Distance bruteForceNearestNeighbourSearch (SpatialDataset, Location, ResultNN\_BF \*resultNearestNeighbourBF)**, receives *spDataSet* of *SpatialDataSet* type, k dimensional *queryPosition* of *Location* type to which nearest neighbour is to be found and returns either

*INFINITY*, if no record is found in dataset or minimum distance of the nearest found object(s) from *queryPosition* in case record(s) is found. Also, a pointer *resultNearestNeighbourBF* of *ResultNN\_BF* type is passed to save the address of result node(s), if found.

If there is no record in dataset and list is empty i.e. *spDataSet* is NULL, **bruteForceNearestNeighbourSearch** will return INFINITY. On the other hand, if records exist in the list, initialize *minDistance* to INFINITY. Now, traverse every node in the *spDataSet* to find the distance between *queryPosition* and *currentRecord*. Also, update the minimum distance *minDistance*, if it is lesser than the *minDistance* calculated up to now. Now, insert object details in *resultNearestNeighbourBF* which satisfy the minimum distance condition.

---

#### Algorithm 1 bruteForceNearestNeighbourSearch

---

```

1: Begin
2: if spDataSet is NULL
3:   return INFINITY
4: end if
5: Initialize minDistance to INFINITY
6: Set distanceSqr to 0
7: Set currentRecord to spDataSet
8: loop while currentRecord is NOT NULL
9:   Find distanceSqr between currentRecord and
   queryPosition
10:  Insert calculated distaceSqr to currentRecord
   temporarily
11:  if distanceSqr < minDistance
12:    then
13:      minDistance = distanceSqr
14:    end if
15:  Move to next currentRecord
16:  Reset distanceSqr to 0
17: end loop
18: Set currentRecord to spDataSet
19: loop while currentRecord exist
20:  if distanceSqr inserted to currentRecord temporarily ==
   minDistance
21:    then
22:      add currentRecord to resultNearestNeighbourBF
23:    end if
24:  Move to next currentRecord
25: end Loop
26: return minDistance
27: End

```

---

### 4. k-dSLst Tree

In our earlier research work, we had introduced an indexing structure k-dSLst tree to index spatial data with duplicate keys. k-dSLst tree is based on k-d tree and sorted linked list. In this paper, we are introducing an algorithm to find nearest neighbour for a given n-dimensional spatial point in a spatial dataset. In k-dSLst tree, first a *kdSLstNode* is created according to n-dimensional spatial location of an object, and then we insert the data related to same object in *dataSNode* in *kdSLstNode* created above. If *kdSLstNode* related to n-dimensional point already exists, then *dataSNode* is inserted in a sorted way according to *object\_id* of the object in already existing *kdSLstNode*. When we insert the *dataSNode*, we need to find the right location for insertion by traversing the linked list at particular *kdSLstNode*. It will increase the insertion time as compared to k-dLst tree, but when comes to searching for a particular object at particular location, it will outperform k-dLst tree on an average.

Figure 1 shows the structure of 2-dSLst tree for dataset records in 2-d space stored as *kdSLstNodes* in a 2-dSLst tree, which can be generalized for k-d space to create k-dSLst tree. A rectangular space is also represented by the leaf *kdSLstNode* which is further separated into two spaces by the newly inserted point.

The root *kdSLstNode* will have discriminator 0. It will be 1 for two sons at the next level, and will be incremented for every next level until it reaches up to  $k-1$  on  $k^{th}$  level. Then it again starts with 0 for  $k+1^{th}$  level and the cycle repeats till the end of k-dSLst tree.

So,  $nextDiscriminator (level\_i) = (level\_i + 1) \bmod k$ .

**Notations:**

**K:** Keys of *kdSLstNode*

**N:** *kdSLstNode* of k-dSLst tree structure

**$K_0(N), K_1(N) \dots K_{k-1}(N)$ :** k keys of *kdSLstNode* N

***lSon(N)*:** Left branch of *kdSLstNode* N

***rSon(N)*:** Right branch of *kdSLstNode* N

***dc(N)*:** Discriminator of *kdSLstNode* N

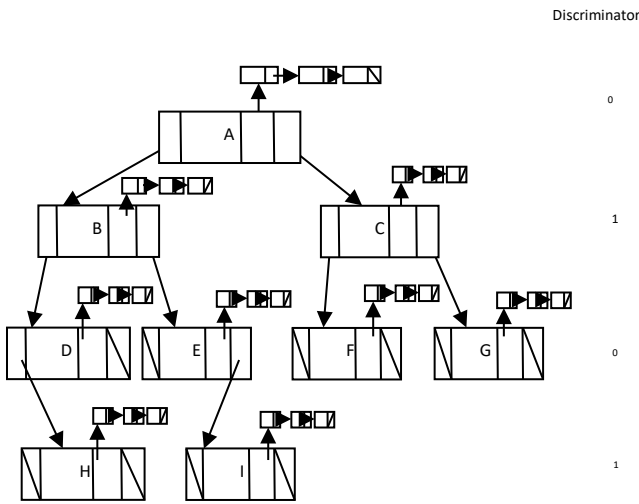


Figure 1: Structure of k-dSLst tree for 2-d keys  
(It can be generalized for k-d keys)

Now, whether to insert new *kdSLstNode* as left son or right son depends on the result of comparison of keys. Let *dc* be the discriminator for *kdSLstNode* N. If  $K_{dc}(N) \neq K_{dc}(Q)$  then the successor *kdSLstNode* Q will be inserted either on left or right side of N i.e. either *lSon(N)* or *rSon(N)*. If  $K_{dc}(N) < K_{dc}(Q)$ , then Q will be inserted on right side of N i.e. as *rSon(N)* else on left side of N i.e. as *lSon(N)*. But, if  $K_{dc}(N) = K_{dc}(Q)$ , then the keys for remaining dimensions will be compared. A superkey *SK* of *kdSLstNode* N is defined by cyclical concatenation of all keys starting with  $K_{dc}(N)$  as

$$SK_{dc}(N) = K_{dc}(N) K_{dc+1}(N) \dots K_{dc+k-1}(N) K_{dc0}(N) \dots K_{dc-1}(N)$$

Now if  $SK_{dc}(Q) < SK_{dc}(N)$  then Q will be added as *lSon(N)* else Q will be added as *rSon(N)*.

Now, in case of duplicate keys i.e. if  $SK_{dc}(Q) = SK_{dc}(N)$ , then address of new *dataNode* will be saved in the same already existing *kdSLstNode* node as a sorted linked list of *dataNodes*.

**5. Nearest Neighbour Search in k-dSLst Tree**

Nearest Neighbour Search is a challenge in various domains like computer vision, spatial data mining and machine learning. There is explosive growth of location based data on the Internet and it is becoming a challenge day by day to store and manage this available data in an efficient way. The researchers have designed many indexing data structures using spatial partitions and recursive hyperplane decomposition to index spatial data. These indexing structures also speed up the nearest neighbour search. But, when it comes to spatial data with duplicate keys, many of the indexing structures do not handle them. K-dSLst tree structure is a combination of k-d tree and sorted linked list. The k-d tree having N nodes require  $O(\log N)$  inspections to search for nearest neighbour as it requires the traversal to at least one leaf of the tree. Also, as the nearest neighbour search needs to traverse a node at most once, it will not visit more than N nodes [8]. Linked list is maintained to hold all the objects at particular spatial location rather than one only.

To find nearest neighbour for any queried n-dimensional point, the search is started at the root of indexing tree and subtrees are explored recursively using the given rule of pruning spatial subtrees: If the nearest neighbour discovered up to now is nearer than the distance between the queried point and hyperplane coordinates corresponding to the current *kdSLstNode*, we can prune the exploration of this *kdSLstNode* and its subtrees further. A *kdSLstNode* needs to be explored further only if it contains point which is nearer as compared to the best one found so far. The effectiveness of the pruning rule depends on finding a nearby point. To do this, we need to organize the recursive method in a way that if we have two probable subtrees to traverse down further, we must opt for the subtree on the same side of the splitting line as is the queried point. The nearest point searched while exploring the first subtree may enable pruning of second subtree.

In this paper, we are proposing Nearest Neighbour Search algorithm **k-dSLstTreeNearest** for *kdSLst* tree which was designed and implemented for indexing spatial data with duplicate spatial keys. Algorithms 2 and Algorithm 3 show the proposed work. The **Algorithm 2 k-dSLstTreeNearest** takes two parameters as arguments. First is the root node of the k-dSLst tree and second is the position containing N-dimensional coordinates about which the nearest neighbour is to be searched. Rather than returning a single nearest neighbour object, the algorithm returns the list of all nearest neighbours. If root node of the tree or hyper rectangle points to NULL then INFINITY is returned back to show that there is no nearest neighbour, otherwise root node is considered as nearest neighbour to start with. The variable *distanceSqr* holds the distance in between nearest point up to now and position of point about which nearest neighbour search is queried.

**Algorithm 2 k-dSLstTreeNearest**

- 1: Begin
- 2: if k-dSLstTreeRoot is NULL
- 3: return NIL;
- 4: if k-dSLstTreeHRect is NULL

```

5:   return NIL;
6:   Initialize resultOfNearest_kdSLst with k-dSLstTreeRoot.
7:   Initialize distanceSqr with 0.
8:   Find distanceSqr between resultOfNearest_kdSLst and
   queryPosition.
9:   CALL k-dSLstTreeNearestIterative(k-dSLstTreeRoot,
   queryPosition, &resultOfNearest_kdSLst, &distanceSqr, k-
   dSLstTreeHRect).
10:  if (resultOfNearest_kdSLst)
11:  then
12:    Traverse the sorted linked list of dataSNodes to show all
   objects at resultOfNearest_kdSLst
13:    Visualize resultOfNearest_kdSLst
14:  end if
15: End

```

An iterative **Algorithm 3 k-dSLstTreeNearestIterative** is called with parameters which include root node of the k-dSLst tree, position queried about for nearest neighbour, pointer to save result node, minimum distance up to now and hyper rectangle of the tree.

**Algorithm 3** k-dSLstTreeNearestIterative

```

1:  Begin
2:  Initialize currentDim to kdsLstNode's dimension
3:  Set decideLeftRight by difference in queryPosition and kdSLstNode
   coordinate values for currentDim
4:  if decideLeftRight <= 0
5:  then
6:    Set nearerSubTree to lSon(kdSLstNode)
7:    Set fartherSubTree to rSon(kdSLstNode)
8:    Update nearerHyperRectCoordinates and
   fartherHyperRectCoordinates
9:    using k-dSLstTreeHRect
10: else
11:   Set nearerSubTree to rSon(kdSLstNode)
12:   Set fartherSubTree to lSon(kdSLstNode)
13:   Update nearerHyperRectCoordinates and
   fartherHyperRectCoordinates
14:   using k-dSLstTreeHRect
15: end if
16: if nearerSubTree exists
17: then
18:   Save nearerHyperRectCoordinates to decideLeftRight
   temporarily
19:   Update nearerHyperRectCoordinates with kdSLstNode
   coordinates for currentDim
20:   CALL k-dSLstTreeNearestIterative (nearerSubTree,
   queryPositions, resultOfNearest_kdSLst,
   resultDistanceSqr, k-dSLstTreeHRect);
21:   Update nearerHyperRectCoordinates with
   decideLeftRight;
22: end if
23: Reset distanceSqr with 0
24: loop for every dimension dim
25:   Compute distanceSqr between kdSLstNode coordinates
   [dim] -queryPosition[dim]
26: end loop
27: if distanceSqr less than resultDistanceSqr
28: then
29:   Update resultOfNearest_kdSLst and resultDistSqr
   accordingly
30: end if
31: if fartherSubTree exist
32: then
33:   Save fartherHyperRectCoordinates to decideLeftRight
   temporarily
34:   Update fartherHyperRectCoordinates with kdSLstNode
   coordinates for currentDim
35: if closest point of k-dSLstTreeHRect is closer than
36: resultDistanceSqr
37: then
38:   CALL k-dSLstTreeNearestIterative (fartherSubTree,

```

```

   queryPositions, resultOfNearest_kdSLst,
   resultDistanceSqr, k-dSLstTreeHRect);
40: end if
41: Update fartherHyperRectCoordinates with decideLeftRight;
42: end if
43: end

```

Depending on the distance of query point from current point for current dimension, we decide nearer and farther sub trees and update the coordinates of nearer and farther hyper rectangles accordingly. If we find any sub tree which is nearer, we call

**Algorithm 3 k-dSLstTreeNearestIterative** iteratively with updated parameters. Now, check the distance of the queried point from the current point so far and update the value of *resultDistSqr*, if it is nearer. Now, also repeat the search for farther sub tree to test for any other nearer point on other side of the slice. The algorithm saves the nearest neighbour N-dimensional coordinates in *resultOfNearest\_kdSLst*. At last, we traverse the complete linked list of *dataSNode* at *resultOfNearest\_kdSLst* to show all of the nearest neighbours. Also, all of the nearest neighbours are displayed graphically using QGIS software.

**6. Performance Evaluation: bruteForceNearestNeighbourSearch vs k-dSLstNearestNeighbourSearch**

We have implemented the algorithms using **language C and GNU Compiler Collection (GCC) compiler - version 6.4.3 on Operating System Ubuntu-10.04.1-Desktop-amd64**. For visualization of spatial datasets and output of **algorithms bruteForceNearestNeighbourSearch** and **k-dSLstNearestNeighbourSearch**, we have used **Quantum Geographic Information System (QGIS) Desktop 2.12.1**.

In our experiments, we are using synthetic datasets which hold spatial details of vehicles. The datasets include vehical-id, spatial coordinates (latitude, longitude) and other non-spatial attributes. We have taken six datasets of different sizes listed in Table 1 given next.

Table 1: Datasets with Duplicate Keys for Performance Evaluation

S. no.	Dataset	Number of records
1	Dataset-01	215
2	Dataset-02	430
3	Dataset-03	860
4	Dataset-04	1720
5	Dataset-05	3440
6	Dataset-06	10320

The six datasets are visualized in Figure 2 through Figure 7 using QGIS Desktop 2.12.1. These figures show the spatial location of different vehicles.

While evaluating performance of searching for nearest neighbour using **Algorithm 1 bruteForceNearestNeighbourSearch** and **Algorithm 2 k-dSLstNearestNeighbourSearch**, we are using the query point Q (43, 20), with Latitude as 43 and Longitude as 20. Both algorithms



give the same results i.e. show the same Vehicles' Id as nearest neighbours for query point Q. But, the time taken by both algorithms is different. As the number of records in datasets and number of nearest neighbours increase, the algorithm `k-dSLstNearestNeighbourSearch` outperforms `bruteForceNearestNeighbourSearch`.

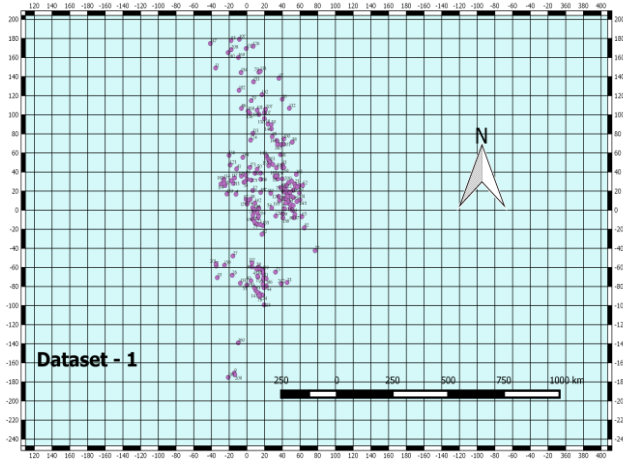


Figure 2: Vehicles' location as per Spatial Dataset – 1 (215 Spatial Records).

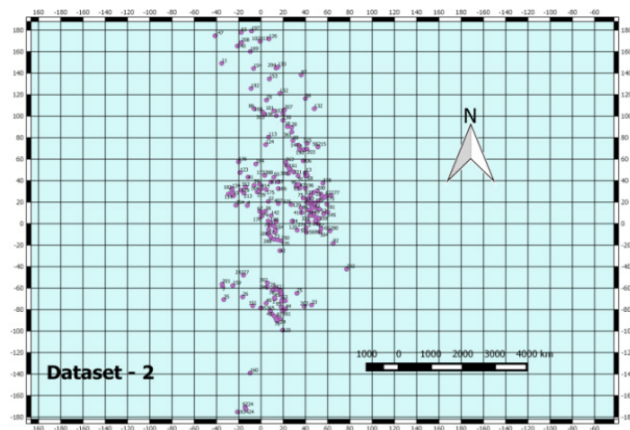


Figure 3: Vehicles' location as per Spatial Dataset – 2 (430 Spatial Records).

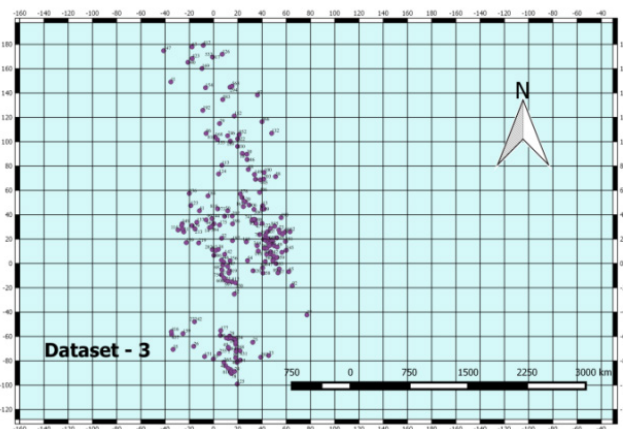


Figure 4: Vehicles' location as per Spatial Dataset – 3 (860 Spatial Records).

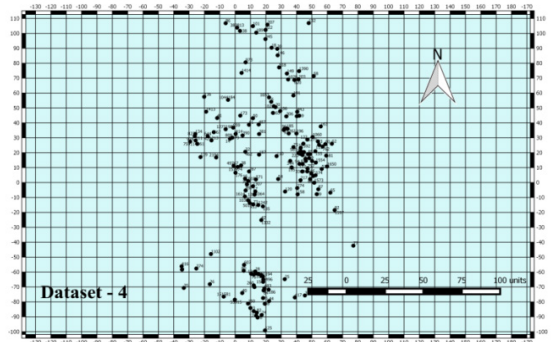


Figure 5: Vehicles' location as per Spatial Dataset – 4 (1720 Spatial Records).

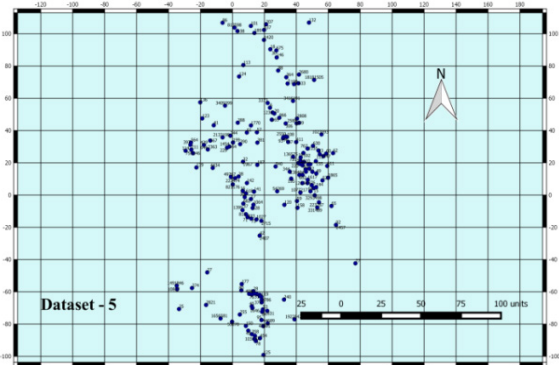


Figure 6: Vehicles' location as per Spatial Dataset – 5 (3440 Spatial Records).

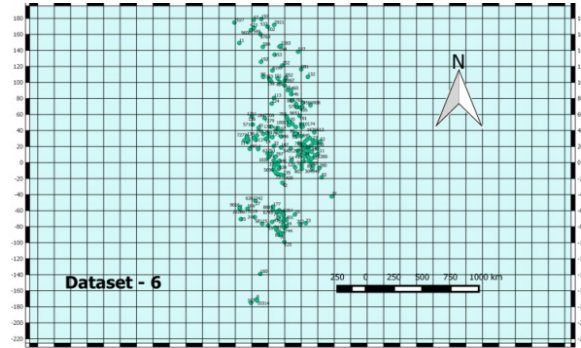


Figure 7: Vehicles' location as per Spatial Dataset – 6 (10320 Spatial Records).

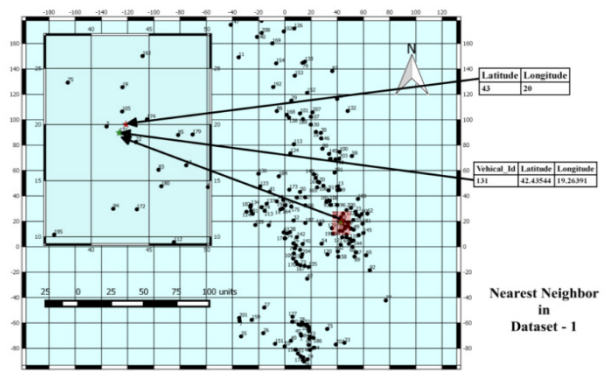


Figure 8: Nearest Neighbors of Q(43, 20) in Dataset - 1.

Figures 8 through 13 visualize the output of nearest neighbour query for query point Q (43, 20). Figures list Ids of all Vehicles found nearest to Q. Figures show that the algorithm **k-dSLstNearestNeighbourSearch** is well efficient to find multiple Vehicle Ids at the same nearest spatial location.

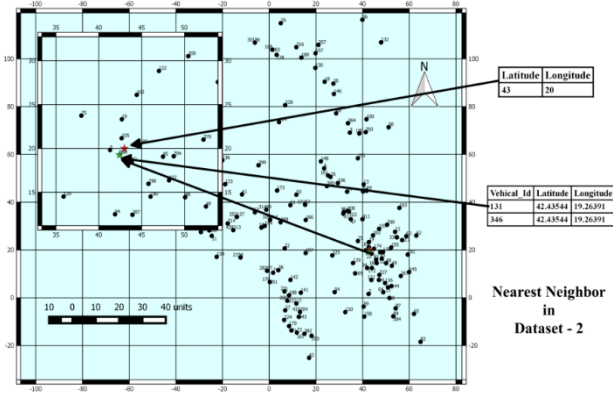


Figure 9: Nearest Neighbors of Q(43, 20) in Dataset - 2.

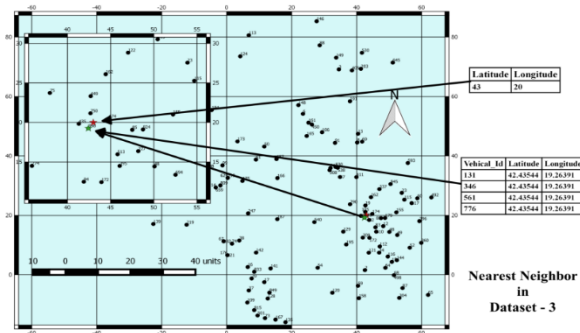


Figure 10: Nearest Neighbors of Q(43, 20) in Dataset - 3.

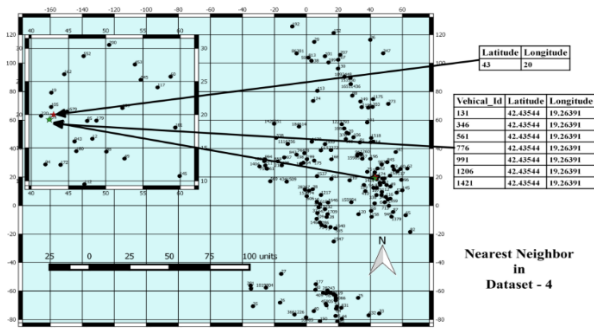


Figure 11: Nearest Neighbors of Q(43, 20) in Dataset - 4.

Table 2 shows the number of records found as nearest neighbours and the time taken by both algorithms in microseconds for searching the same. Results show that algorithm **k-dSLstNearestNeighbourSearch** is more efficient as compared to **bruteForceNearestNeighbourSearch** as the number of spatial records increases and capable of finding all nearest neighbours for

spatial datasets with duplicate spatial keys. The search time take by algorithm **bruteForceNearestNeighborSearch** is increasing continuously and there is exponential rise from Dataset-05 to Dataset-06 as number of records increase from 3440 to 10320. Search time of algorithms is also affected by the number of records found at nearest location which are 64 in case of Dataset-05 and 192 in case of Dataset-06. As compared to algorithm **bruteForceNearestNeighborSearch** which takes 407 micro secs. and 1391 micro secs. for nearest neighbor search in Dataset-05 and Dataset-06 respectively, algorithm **k-dSLstNearestNeighbourSearch** takes much less time i.e. 162 micro secs. and 223 micro secs. for corresponding datasets.

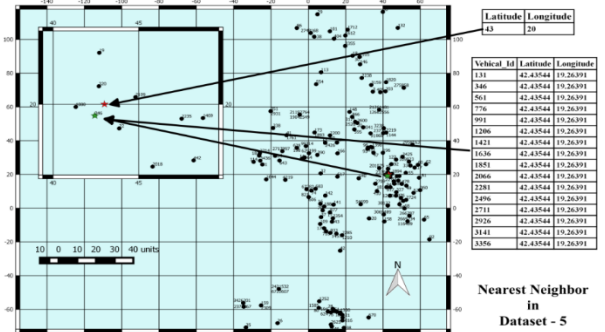


Figure 12: Nearest Neighbors of Q(43, 20) in Dataset - 5.

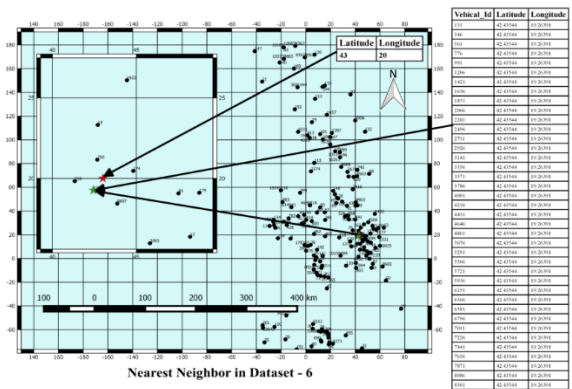


Figure 13: Nearest Neighbors of Q(43, 20) in Dataset - 6.

Table 2: bruteForceNearestNeighbourSearch vs k-dSLstNearestNeighbourSearch

S. no.	Dataset	No. of records found at Nearest location	Time taken for search (in microseconds)	
			bruteForceNearest NeighborSearch Algorithm	k-dSLstNearest NeighbourSearch Algorithm
1	Dataset-01	4	29	106
2	Dataset-02	8	46	89
3	Dataset-03	16	90	96
4	Dataset-04	32	212	110
5	Dataset-05	64	407	162
6	Dataset-06	192	1391	223

Figure 14 shows the results of **bruteForceNearestNeighbourSearch** vs **k-dSLstNearestNeighbourSearch** graphically.

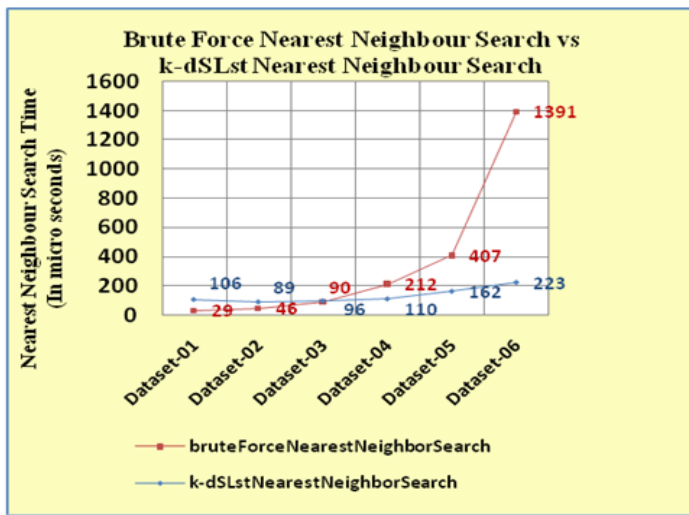


Figure 14: Search time: bruteForceNearestNeighbourSearch vs k-dSLstNearestNeighbourSearch

[8] M.W. Andrew, "Efficient Memory-based Learning for Robot Control. Technical Report No. 209, Ph. D Thesis, Computer Laboratory, University of Cambridge, 1991.

## 7. Conclusion and Future Scope

The **k-dSLstNearestNeighbourSearch** algorithm for searching nearest neighbour in spatial datasets with duplicate keys is based on **k-dSLst indexing tree structure** which is further based on k-d tree and sorted single linked list data structures. We have compared the proposed algorithm with brute-force approach. The performance evaluation shows that the search time taken by algorithm **bruteForceNearestNeighborSearch** increases exponentially as numbers on data records and nearest neighbours increases. But in case of algorithm **k-dSLstNearestNeighbourSearch** for the same datasets, the time take for search is much less and doesn't rise exponentially. The performance evaluation of both algorithms shows that the proposed work in this paper revealed better performance when compared to the conventional brute-force approach.

The work can be further expanded to search for N nearest neighbours and objects within a given range. Also, the spatial data structure k-dSLst tree can be extended to accommodate temporal data also to index spatio-temporal datasets.

## References

[1] S. Shekhar et al., Spatiotemporal Data Mining: A Computational Perspective, International Journal of Geo-Information ISSN 2220-9964, 04, 2306-2338, 2015.  
 [2] S. Geetha, S. Velavan, "Optimization of Location Based Queries using Spatial Indexing, International Journal of Soft Computing, 4, 2014.  
 [3] Verma et al., "Comparison of Brute-Force and K-D Tree Algorithm, International Journal of Advanced Research in Computer and Communication Engineering, January, 03, 2014.  
 [4] A. Suhaibaha, et al., "3D Nearest Neighbour Search Using a Clustered Hierarchical Tree, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXIII ISPRS Congress, XLI-B2, 2016.  
 [5] J. Jaemin, et al., "A progressive k-d tree for approximate k-nearest neighbors, IEEE Workshop on Data Systems for Interactive Analysis (DSIA), 2017.  
 [6] W. David, R. Rafael, "Parallel kd-Tree Construction on the GPU with an Adaptive Split and Sort Strategy, International Journal of Parallel Programming, 46, 1139-1156, 2018.  
 [7] Stoimen. Computer Algorithms: Brute Force String Matching. Stoimen's web log. [Online] March 2012.

# Modelling and Simulation of Aerodynamic Parameters of an Airship

Anoop Sasidharan Pillai, Venkata Ramana Murthy Oruganti\*

Department of Electrical and Electronics Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, 641112, India

## ARTICLE INFO

Article history:

Received: 21 May, 2020

Accepted: 17 July, 2020

Online: 22 July, 2020

Keywords:

Airship

Analytical method

Dynamic model

Aerodynamic coefficients

## ABSTRACT

The dynamic modelling of an airship is the primary requirement in designing and developing a control system for a particular application. Extracting/predicting/modelling the aerodynamic coefficients is a crucial step towards the modelling of an airship. There is a huge amount of literature on the aerodynamic modelling of airships which presents experimental as well as analytical methods. All these techniques require some experimental data such as the geometrical data, control derivatives, etc. In this work, we are investigating an analytical technique which can calculate the aerodynamic parameters for a high altitude airship. The complete airship model is implemented using the derived aerodynamic coefficients. A MATLAB® Simulink based simulation is used for the investigation.

## Nomenclature

$\alpha$  Angle of attack, deg

$\beta$  Side slip angle, deg

$\delta_a$  Aileron deflection angle, deg

$\delta_e$  Elevator deflection angle, deg

$\delta_r$  Rudder deflection angle, deg

$\eta$  Throttle ratio

$\gamma, \mu$  Angles along the wind axes, deg

$\phi, \theta, \psi$  Airship's Euler angles - roll, pitch and yaw, deg

$\rho$  Density of air, kg/m<sup>3</sup>

$b$  Wing span, m

$b_z$  Point of action of buoyant force in z direction, m

$c$  mean aerodynamic chord, m

$C_D, C_L, C_Y$  Coefficient of drag, lift and side force

$C_l, C_m, C_n$  Coefficient of rolling moment, pitching moment and yawing moment

$d_z$  Point of action of thrust in z direction, m

$I_x, I_y, I_z$  Moment of inertia components about x, y and z axes, kg/m<sup>2</sup>

$I_{xz}$  Product of inertia about y axis, kg/m<sup>2</sup>

$m$  Airship mass, kg

$m_x, m_y, m_z$  Apparent mass components in the x, y and z axes, kg

$p, q, r$  Airship angular velocities along the roll, pitch and yaw axes, deg/s

$S$  Wing planform area, m<sup>2</sup>

$t, T$  Current simulation time and gust time period, s

$T_m$  Maximum engine thrust, N

$u, v, w$  Airship linear velocities along the x, y and z body axes, m/s

$V$  Airship forward velocity, m/s

$V_m, V_g$  Mean wind and gust velocity, m/s

$x_E, y_E, z_E$  Position coordinate in the earth frame, m

**CoG** Centre of Gravity

**CoV** Centre of Volume

**EOG** Extreme Operating Gust

**LTA** Lighter-Than Air

\*Venkata Ramana Murthy Oruganti, Amrita Vishwa Vidyapeetham, ovr.murthy@cb.amrita.edu



# 1 Introduction

Man's dream of flying has been fulfilled by hot air balloons by the end of the 18<sup>th</sup> century and the realisation of a controlled flight was done by the design proposal of Jean-Baptiste Meusnier. Airship technology has a great history starting from the hot air balloons used in the medieval ages and has been widely used in the 19<sup>th</sup> century. Unfortunately, the interest in airships has been lost due to the high maintenance cost, drastic developments in air crafts and helicopters and some catastrophic tragedies such as the Hindenburg disaster [1]. By the end of the 20<sup>th</sup> century, a resurgence in the interest in airships by researchers and military agencies has occurred (as shown in Figure 1) due to the low speed, long endurance, economic, noiseless and eco-friendly operation of airships. Major applications [2] of airships include transportation of war equipment, intelligence, surveillance, reconnaissance, advertisement, telecommunication, monitoring and inspection of climate and environmental effects, inspection of man-made structures such as buildings, pipelines, railway network, and power lines, planning for urban development, research, exploration and so on.

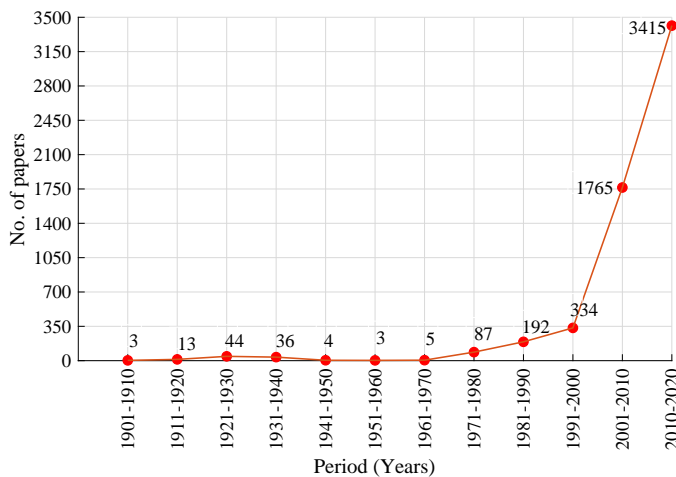


Figure 1: Number of papers published in the field of LTA vehicles (Source: Scopus database)

The airship based application has to be modelled to reflect the expected behaviour. Like any other complex system, dynamic modelling of an airship is one of the challenging areas for researchers in this field. Dynamic modelling of Lighter-Than-Air (LTA) vehicles differ from other members of aerial vehicles such as helicopters and air crafts. There is a close similarity between the dynamics of an airship and an underwater vehicle [3], which researchers exploited for modelling the airship in the early stage of research [4]. A detailed survey of the dynamic modelling of the airship is presented in [1] along with the details of wind tunnel tests; which are being utilized by researchers for the dynamic modelling and structural design of airships. A system level mathematical modelling of a high altitude airship is presented in [5]. The presented model includes the airship dynamic model, solar power system model and atmosphere models. The aerodynamic modeling of the airship was done using the geometrical data available from the literature. A linearized longitudinal and lateral dynamic models of an airship is presented

in [6]. An analysis of different stability modes for the two linear models was presented. A dynamic model is presented for an airship in [7] which utilizes the geometrical aerodynamic parameters for the model. A simulation methodology is also presented to demonstrate the effectiveness of the developed model using MATLAB®.

Most of the above cited works on the dynamic modelling of airships have utilised the available database or the method of practical extraction from the actual system for obtaining the model parameters such as the aerodynamic coefficients. There are various methods by which we can estimate the parameters included in the aerodynamic model such as CFD based methods [8], experimental methods using wind tunnel [4, 9] and analytical methods [10]. Computational Fluid Dynamics (CFD) tools can be used to extract the aerodynamic coefficients for airship geometries [11, 12]. A method of extracting the force and moment coefficients of an airship using CFD and Fourier analysis is presented in [11]. Forced sinusoidal oscillations are used in the longitudinal direction to extract the force and moment coefficients. A specialised computational tool used for the aerodynamic coefficient extraction is presented in [12]. The major drawbacks of CFD based analysis are the inability to model the geometry accurately, computational cost and time taken for simulation. The traditional way is to use wind tunnel testing for the extraction of aerodynamic coefficients [4, 9]. An analytical method for calculating the aerodynamic coefficients and other control derivatives is presented in [9]. The method is heavily dependent on the wind tunnel test results. A detailed investigation of LOTTE airship using wind tunnel is presented in [4] in which the force and moment coefficients, pressure distributions and the flow visualization were elaborately discussed. There are practical difficulties in this method such as the size limitation of the test section of the wind tunnel, difficulties in the installation and inaccurate measurements, inaccuracies due to the blockages in the tunnel, etc. Most accurate and realistic data for the calculation of those coefficients can be obtained through actual test flights [10]. But it is not always feasible to conduct flight tests because of fuel consumption, unsteady atmosphere conditions, etc. There are a lot of works carried out in this context which are trying to reduce the efforts in the calculation. A method which requires no experimental data for calculating the aerodynamic coefficients using geometrical parameters is presented in [13]. The problem with this method is the inability to generalise the method to all class of airships. The parameters specified in the study are particularly available for a certain class of airships only. The digital DATCOM software based on Fortran is reported to be useful in the calculation of the aerodynamic coefficients. Such a work is presented in [14, 15] which used the DATCOM to predict the aerodynamic coefficients for a high altitude airship to manoeuvre at an altitude of 21000 m. This method is also limited to the available datasets in the DATCOM database.

There is no doubt that the above cited works are providing enough light to the field of aerodynamic modelling of airships. As per the difficulties posed by those methods, it is evident that there is still scope for research in this field using methods which need not rely much on experimental data, complex geometry development for CFD, etc. The aim of this work is to present a study on the aerodynamic modelling of a high altitude airship. The current study is motivated by the work presented in [13] which describes a method of aerodynamic modeling for the aircraft with a little ex-



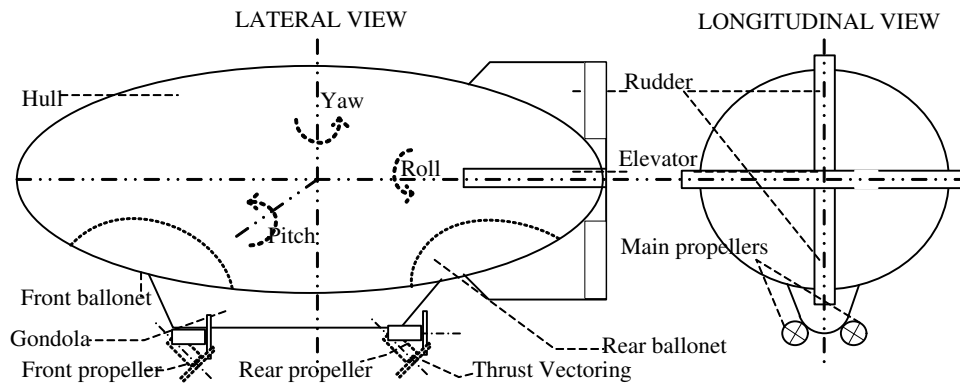


Figure 2: Basic components, actuators, and degrees of freedom of a double ellipsoid airship [16]

perimental data. We are investigating the same technique for a high altitude airship which is modelled using a modified set of aircraft equations of motion. The main advantage of this method is that the experimental data required for the modelling is considerably less. The remaining paper is presented in five sections. After the introduction and literature survey presented in Section 1, a brief description of the principle of operation of airships is given in Section 2 as a preliminary note. Different components of the airship, application of each component, principle of lift and hovering are also presented in Section 2. Section 3 discusses the complete dynamic modelling of a high altitude airship with a detailed study on aerodynamic modelling. Section 4 discusses the simulation results and their analysis. Simulation of the aerodynamic coefficients for different flight conditions and an open loop simulation of the airship are presented along with the results. Finally the conclusions and future works are presented in Section 5.

## 2 Airship Principle of Operation

The basic principle behind the operation of airships is the natural buoyancy of LTA gases such as hot air, helium and hydrogen. The concept of airship came from the basic hot air balloon, which uses the buoyancy property for lift generation. But to use it as a vehicle, there should be some modifications made on the balloon such as the streamlined body to reduce the air resistance and to enhance the aerodynamic properties. Figure 2 shows the major parts of a double ellipsoid airship [16]. The main body, which is known as the hull, occupies the large volume of LTA gas. Based on its structure, airships are classified as non-rigid, semi rigid and rigid airships. Non-rigid airships, also known as blimps, do not have any rigid structure to maintain its streamlined shape. It is maintained by the pressure inside the envelope. While rigid and semi rigid airships have a fully supported or partial mechanical structure on which the hull envelope is attached. Hull houses the pressure controlling ballonets, which can even control the altitude and pitch of airships. The pressure and thereby volume of the LTA gas may vary with altitude, which results in the expansion and contraction of the gas accordingly. As altitude increases, the helium expands which pushes the air in ballonets out. This will continue till the pressure height which is the designed maximum altitude of operation. Beyond the pressure height, the safety valve will release helium to avoid bursting.

Thrusters or propellers are the aerostatic actuators used for longitudinal and lateral movement of airships. A pair of propellers are placed on both sides of the gondola which provides the forward thrust. Thrust vectoring can be included to have a longitudinal freedom when the aerodynamic actuators are not active due to low wind. Apart from these main propellers, stern propellers may be used for providing a lateral degree of freedom when the aerodynamic surfaces cannot provide much actuation. Aerodynamic control surfaces are provided at the stern end of airships which can control the yaw, roll, and pitch degrees of freedom. These control surfaces can be either designed as a ‘+’ configuration or ‘X’ configuration. Elevators produce a pitch degree when they act together in the same direction and a roll degree can be generated by differentially actuating the elevators. Yaw is provided by rudders which are the vertical control surfaces. Another part of the airship is the ballast mechanism to control the pitch. By using solid or liquid as a movable mass, the pitch can be controlled by manipulating the location of the centre of gravity of the airship, and sometimes the gondola completely can be the ballast.

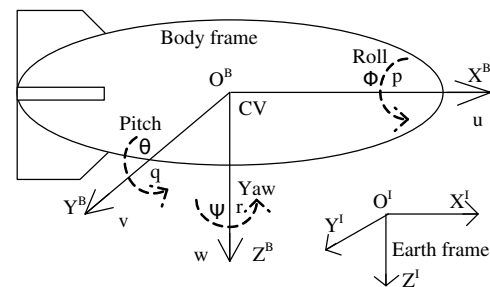


Figure 3: The reference frames used for the dynamic model along with the linear and angular position and velocities [16]

The airship considered in this study is a high altitude airship which is intended to operate at 21000m. The hull is of Gertler shape with four tail fins (as shown in Figure 2) and two thrusters. There are some assumptions based on which the following analysis is carried out. They are,

- Hull volume is constant throughout the simulation
- Structure of the airship is rigid
- The airship geometry is symmetric in the XZ plane

- Airship is at neutral buoyant condition
- A flat earth approximation is followed

### 3 Dynamic Model

As mentioned in Section 1, the dynamic modelling of an airship differs largely from that of an aircraft. The displaced Centre of Gravity (CoG) from the Centre of Volume (CoV) and the effect of the added mass of the surrounding air are the main contrasting features of LTA vehicle models. The displacement of the CoG vertically downwards of the CoV is to provide unconditional stability for the airship. The tendency of rotation will be counteracted by the downward displaced CoG. Even though the mathematical model of an airship differs from that of an aircraft model, it is possible to include the added mass, buoyancy and the aerodynamic properties of the airship to the aircraft model. The advantages of this methodology are the availability of well-established flight dynamic equations, availability of experimental data for different flight and airship configurations and more insight into the dynamic behaviour of the airship in the light of aircraft dynamics.

#### 3.1 A mathematical model for an airship

For deriving the dynamic equations of a rigid body there should be a coordinate system which can properly behold the complexity, degrees of freedom and application requirements. There are mainly three reference frames earth fixed inertial frame, wind frame and body frame. The position and orientation of an object in the atmosphere with respect to the earth can be obtained with the help of earth fixed inertial reference frame. In the same way, the forces and moments of a moving body are not only depending on the velocity with respect to earth but also to the surrounding air itself. This leads to the consideration of an atmosphere fixed reference frame i.e., wind reference frame which is a moving frame with respect to earth but not with respect to the vehicle. Another set of a reference frame is the body fixed frame which is fixed to the CoV and is moving with the vehicle. The earth fixed inertia and the body fixed reference frame are shown in Figure 3 (wind frame is not included since it is considered temporarily during the derivation). The inertia frame  $O_I X_I Y_I Z_I$  is having the orientation as  $X_I$  pointing towards the north,  $Y_I$  towards the east and  $Z_I$  towards the centre of the earth. Similarly, the body frame  $O_B X_B Y_B Z_B$  has its centre at the CoV of the airship and  $X_B$  pointing towards the nose of the vehicle,  $Y_B$  towards the right side wing and  $Z_B$  pointing downwards perpendicular to  $X_B$  and  $Y_B$ . The wind frame is used only during the derivation of equations and the final equations will be expressed in terms of the earth and body frames only. In a general way, the dynamic equations for an airship can be represented as follows,

$$\begin{aligned} \dot{V} &= \frac{1}{m_x} [T_m \eta \cos \alpha \cos \beta - \frac{1}{2} \rho V^2 S C_D - (mg - B) \sin \gamma] \\ \dot{\alpha} &= q - \frac{1}{\cos \beta} [(p \cos \alpha + r \sin \alpha) \sin \beta + \frac{1}{m_z V} (T_m \eta \sin \alpha + \frac{1}{2} \rho V^2 S C_L - (mg - B) \cos \mu \cos \gamma)] \\ \dot{\beta} &= \frac{1}{m_y V} [-T_m \eta \cos \alpha \sin \beta + \frac{1}{2} \rho V^2 S C_Y + (mg - B) \sin \mu \cos \gamma] + (p \sin \alpha - r \cos \alpha) \end{aligned} \quad (1)$$

$$\begin{aligned} \dot{p} &= \frac{1}{I_x} [(I_y - I_z)qr + I_{xz}pq + \frac{1}{2} \rho V^2 S b C_l - B b_z \sin \mu \cos \gamma] \\ \dot{q} &= \frac{1}{I_y} [(I_z - I_x)pr + I_{xz}(r^2 - p^2) + \frac{1}{2} \rho V^2 S c C_m + T_m \eta d_z \cos \alpha \cos \beta - B b_z \sin \theta] \\ \dot{r} &= \frac{1}{I_z} [(I_x - I_y)pq - I_{xz}qr + \frac{1}{2} \rho V^2 S b C_n] \\ \dot{\phi} &= p + q \sin \phi \tan \theta + r \cos \phi \tan \theta \\ \dot{\theta} &= q \cos \phi - r \sin \phi \\ \dot{\psi} &= \sec \theta (q \sin \phi + r \cos \phi) \\ \dot{x}_E &= V \cos \alpha \cos \beta (\cos \psi \cos \theta) + V \sin \beta (\cos \psi \sin \theta \sin \phi - \sin \psi \cos \phi) + V \sin \alpha \cos \beta (\cos \psi \sin \theta \cos \phi + \sin \psi \sin \phi) \\ \dot{y}_E &= V \cos \alpha \cos \beta (\sin \psi \cos \theta) + V \sin \beta (\sin \psi \sin \theta \sin \phi + \cos \psi \cos \phi) + V \sin \alpha \cos \beta (\sin \psi \sin \theta \cos \phi - \cos \psi \sin \phi) \\ \dot{z}_E &= V \cos \alpha \cos \beta (-\sin \theta) + V \sin \beta (\cos \theta \sin \phi) + V \sin \alpha \cos \beta (\cos \theta \cos \phi) \end{aligned} \quad (2)$$

where

$$\begin{aligned} \sin \gamma &= \cos \alpha \cos \beta \sin \theta - \sin \beta \sin \phi \cos \theta - \sin \alpha \cos \beta \cos \phi \cos \theta \\ \sin \mu \cos \gamma &= \cos \alpha \sin \beta \sin \theta + \cos \beta \sin \phi \cos \theta - \sin \alpha \sin \beta \cos \phi \cos \theta \\ \cos \mu \cos \gamma &= \sin \theta \sin \alpha + \cos \alpha \cos \phi \cos \theta \end{aligned}$$

where the states  $[V \alpha \beta p q r \phi \theta \psi x_E y_E z_E]^T$  are the forward velocity, angle of attack, side slip angle, body axis role, pitch and yaw rates, roll, pitch and yaw angles and position coordinate in the earth frame respectively,  $m_x, m_y$  and  $m_z$  are the components of apparent mass,  $T_m$  is the maximum engine thrust,  $\eta$  is the throttle ratio,  $\rho$  is the density,  $S$  is the wing planform area,  $c$  is the chord,  $b$  is the wingspan,  $b_z$  is the point of action of buoyant force along the  $Z_B$  direction,  $d_z$  is the point of action of thrust along the  $Z_B$  direction,  $B$  is the buoyant force,  $m$  is the airship mass,  $I_x, I_y, I_z, I_{xz}$  are the inertia components,  $C_D, C_L, C_Y$  are the drag, lift and side force coefficients,  $C_l, C_m, C_n$  are the rolling, yawing and pitching moment coefficients.

#### 3.2 Aerodynamic coefficients

The set of equations described in Section 3.1 are modified aircraft equations to demonstrate the dynamics of a high altitude airship. The aerodynamic parameters in the equation such as  $C_D, C_L, C_Y, C_l, C_m$  and  $C_n$  are the major design parameters of the airship. There is an analytical method for the calculation of the aerodynamic coefficients for aircraft using the actuator values [13]. Since the airship model used in this paper is modified from the same aircraft model, this method is investigating for the airship in this paper. The method calculates the aerodynamic coefficients as a function of  $\alpha, \beta, p, q, r, \delta_e, \delta_a$  and  $\delta_r$ . The set of equations for the calculation of aerodynamic coefficients are given by equations (3) to (8).

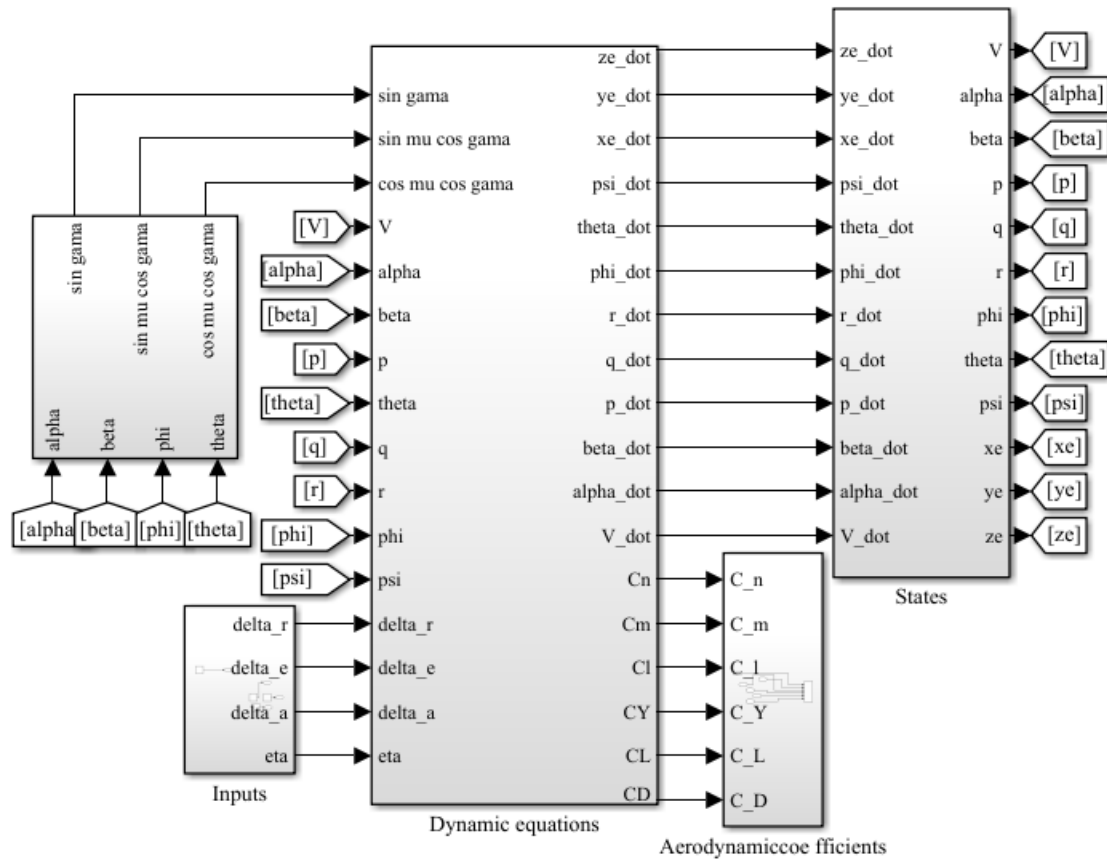


Figure 4: Simulink block diagram of the simulation of airship

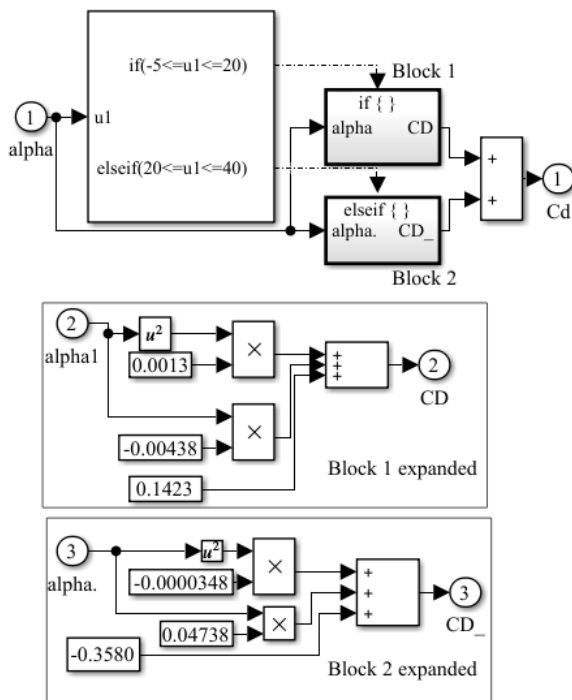


Figure 5: Simulink block diagram of the simulation of coefficient of drag

$$C_D = \begin{cases} 0.0013\alpha^2 - 0.00438\alpha + 0.1423 & \text{if } -5 \leq \alpha \leq 20 \\ -0.0000348\alpha^2 + 0.0473\alpha - 0.3580 & \text{if } 20 \leq \alpha \leq 40 \end{cases} \quad (3)$$

$$C_Y = -0.0186\beta + \frac{\delta_a}{25}(-0.00227\alpha + 0.039) + \frac{\delta_r}{30}(-0.00265\alpha + 0.141) \quad (4)$$

$$C_L = \begin{cases} 0.0751\alpha + 0.0144\delta_e + 0.732 & \text{if } -5 \leq \alpha \leq 10 \\ -0.00148\alpha^2 + 0.106\alpha + 0.0144\delta_e + 0.569 & \text{if } 10 \leq \alpha \leq 40 \end{cases} \quad (5)$$

$$C_l = C_l(\alpha, \beta) - 0.0315p + 0.0126r + \frac{\delta_a}{25}(0.00121\alpha + 0.0628) - \frac{\delta_r}{30}(0.000351\alpha - 0.0124) \quad (6)$$

where,

$$C_l(\alpha, \beta) = \begin{cases} (-0.00012\alpha - 0.00092)\beta & \text{if } -5 \leq \alpha \leq 15 \\ (0.00022\alpha - 0.006)\beta & \text{if } 15 \leq \alpha \leq 25 \end{cases}$$

$$C_m = -0.00437\alpha - 0.0196\delta_e - 0.123q - 0.1885 \quad (7)$$

$$C_n = C_n(\alpha, \beta) - 0.0142r + 0.0126r + \frac{\delta_a}{25}(0.000213\alpha + 0.00128) + \frac{\delta_r}{30}(0.000804\alpha - 0.0474) \quad (8)$$

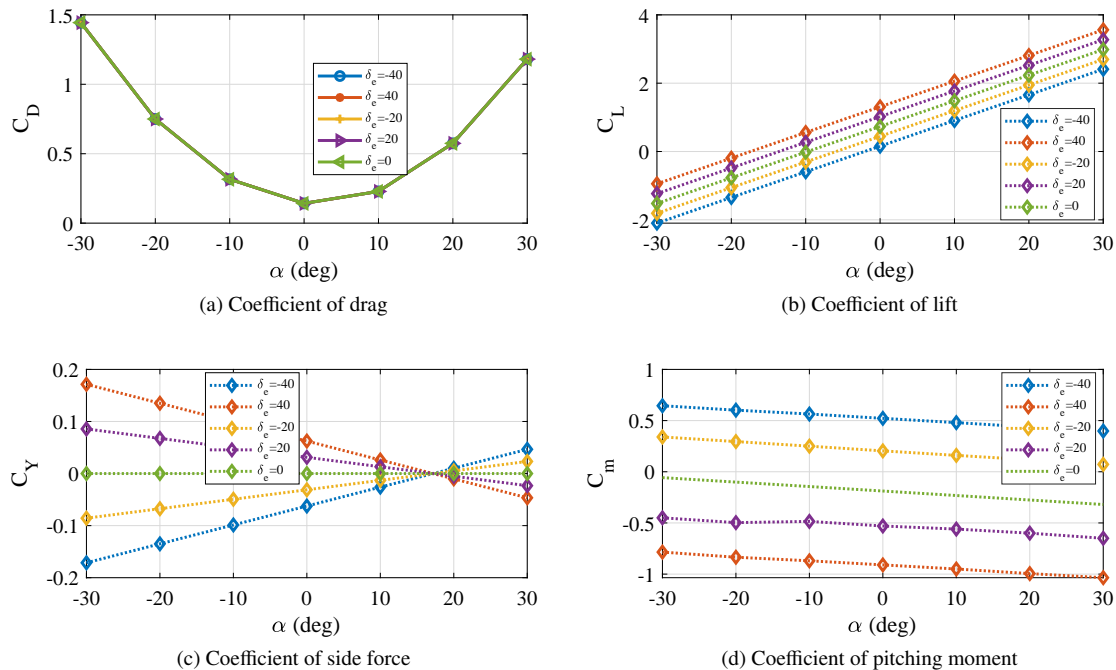


Figure 6: Simulation results of the aerodynamic coefficients given by (3)-(8) for various values of  $\delta_e$

where,

$$C_n(\alpha, \beta) = \begin{cases} 0.00125\beta & \text{if } -5 \leq \alpha \leq 10 \\ (-0.00022\alpha + 0.00342)\beta & \text{if } 10 \leq \alpha \leq 25 \\ -0.00201\beta & \text{if } 25 \leq \alpha \leq 35 \end{cases}$$

where  $\delta_e, \delta_a$  and  $\delta_r$  are the deflection of the elevator, aileron and rudder respectively.

## 4 Results and Discussion

### 4.1 Simulation

The model of the airship is implemented using MATLAB<sup>®</sup> Simulink 2018a. Simulink block diagram of the simulation is as shown in Figure 4. The ‘Dynamic equations’ block shown in Figure 4 holds the realisation of airship dynamic equations given in (1) and (2). The ‘Aerodynamic coefficients’ block in Figure 4 consists of the realisation of coefficients given in (3) - (8). As a representative figure, the simulink block diagram of the coefficient of drag expressed in (3) is shown in Figure 5. The parameters used for the simulation are given in Table 1. ODE45 solver with variable time step execution is used for the simulation. Five sets of elevator deflection angles are used for the simulation.

### 4.2 Simulation results

The method of obtaining aerodynamic coefficients explained in (3) - (8) is directly applicable to aircraft. It can be extended to airships also. Since there is no aileron for the airship, the value of  $\delta_a$  is taken as same as  $\delta_e$ . The aerodynamic coefficients represented using equations (3)-(8) are obtained for different values of  $\alpha$  and  $\delta_e$  are

shown in Figure 6.  $C_D, C_L, C_Y$  and  $C_m$  are shown in Figure 6 and  $C_l$  and  $C_n$  are not shown since their values are zero for the simulated conditions. Even though we have shown the coefficients for a wide range of angle of attack starting from  $-30$  to  $30$  degrees, the airship simulation later in this section shows that the angle of attack remains within  $10$  degrees. It is evident that the angle of attack obtained in our simulation is far below the stall angle for a stream lined airship [17]. So the problem of stalling is not in the scope of this work.

Table 1: Model parameters used for the simulation

No.	Parameter	Value	Unit
1	$m$	23145	$kg$
2	$I_x$	17400000	$kgm^2$
3	$I_y$	245264282	$kgm^2$
4	$I_z$	245264282	$kgm^2$
5	$I_{xz}$	1920000	$kgm^2$
6	$T_m$	6000	$N$
7	$\eta$	23145	$kg$
8	$m_x$	25032	$kg$
9	$m_y$	43044	$kg$
10	$m_z$	43044	$kg$
11	$T_m$	6000	$N$
12	$\rho$	0.0767	$kgm^3$
13	$S$	4748	$m^2$
14	$b_z$	16.43	$m$
15	$d_z$	14.67	$m$
16	$c$	68.9	$m$
17	$b$	68.9	$m$

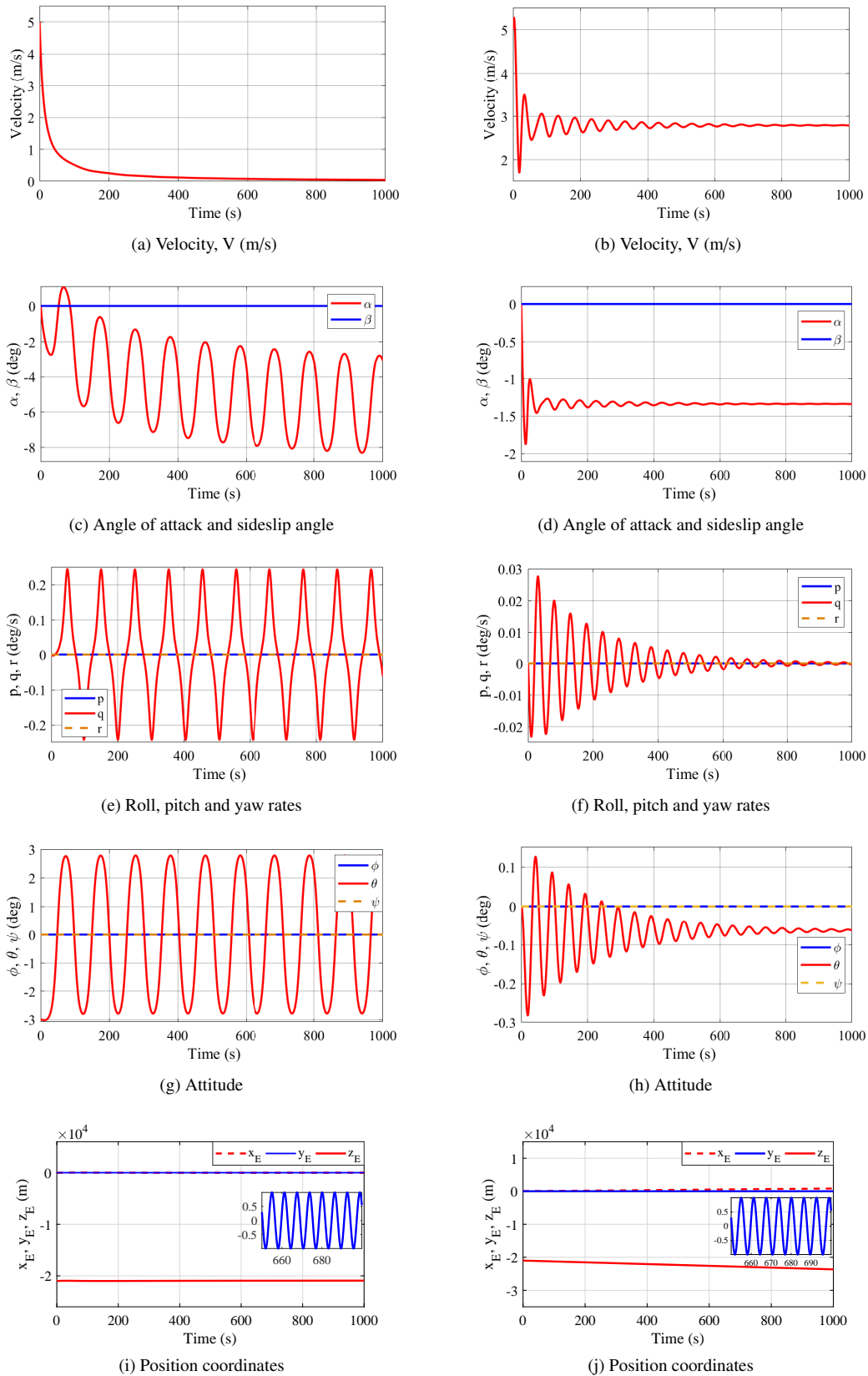


Figure 7: Simulation results for the airship model for zero thrust shown in the first column((a), (c), (e), (g), (i) and 0.1836 throttle ratio shown in the second column ((b), (d), (f), (h), (j))



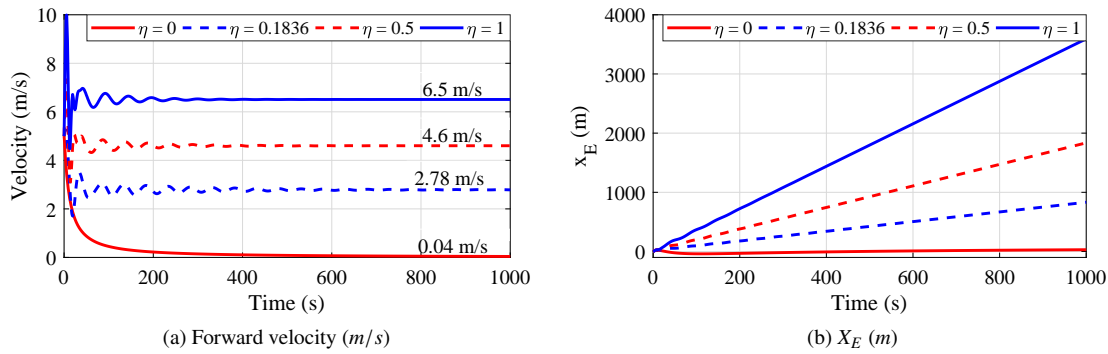


Figure 8: Simulation results of the forward velocity and position along x axis for different η values

Coefficient of drag,  $C_D$  is given in Figure 6 (a). The higher values of drag at the higher angle of attacks are as expected and it is clear from the results that there is no effect of  $\delta_e$  on the drag coefficient of the airship. Even though the coefficients are given for a wide range of angle of attacks, the normal operation (except the take off and landing) of airships will be in a small angle of attack. The drag coefficient at zero angle of attack is 0.1432.

Coefficient of lift,  $C_L$ , seems to have a wide range of deviation for the simulated angle of attacks as shown in Figure 6 (b). The model shows a positive lift at zero angle of attack and the above, which is a characteristic peculiarity of airships. Coefficient of side force,  $C_Y$ , and the coefficient of pitching moment,  $C_m$  are shown in Figure 6 (c) and (d) respectively.

These extracted aerodynamic coefficients are used for the model given in (1) and (2) and are simulated to see the response for different inputs. Some of them are given in the following paragraphs.

The initial condition for the 9 state variables used in (1) and (2) is as given as follows;

$$x_0 = \begin{bmatrix} V_0 & \alpha_0 & \beta_0 & p_0 & q_0 & r_0 & \phi_0 & \theta_0 & \psi_0 & x_{E0} \\ y_{E0} & z_{E0} & & & & & & & & \end{bmatrix}^T \quad (9)$$

$$= \begin{bmatrix} 5 & 3 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & -21000 & & & & & & & & \end{bmatrix}^T$$

A zero input simulation of the airship is shown in the first column of Figure 7. The simulation was run for 1000 seconds. The

velocity of the airship is asymptotically settling to zero as shown in Figure 7 (a). There is an oscillating negative angle of attack around  $-7$  degree which causes a slight decrease in the altitude of the airship from the initial altitude. The side slip angle remains zero throughout the simulation as shown in Figure 7 (c). Figure 7 (e) shows the roll, pitch and yaw rates of the airship. The roll and yaw rates remained at zero throughout the simulation, but the pitch rate shows an oscillation as shown in Figure 7 (e). These oscillations are reflected in the attitude of the airship also. Figure 7 (g) shows the oscillatory pitch angle and the zero roll and yaw angles. The position of the airship with respect to the inertial frame are shown in Figure 7 (i). It is clear from the figure that the airship is staying at the initial position through out the simulation. There are slight deviations in the y direction but it is negligibly small. There is a slight oscillation in the  $y_E$  of the order of 1 m/s which is shown in Figure 7 (i).

Another set of results are shown in the second column of Figure 7 with a minimum throttle ratio of 0.1836 by keeping all other inputs zero. The velocity is settling at 2.8 m/s after a prominent overshoot and oscillations as shown in Figure 7 (b). The angle of attack and side slip angle are shown in Figure 7 (d). There is a large improvement in the magnitude of the negative angle of attack. Now the angle is  $-1.3$  degree which can be expected due to the thrust generation off to the centre of gravity. The side slip angle remains zero. The pitch rate oscillation dies out asymptotically as shown in Figure 7 (f). The roll and yaw rates are remaining the same as in the first

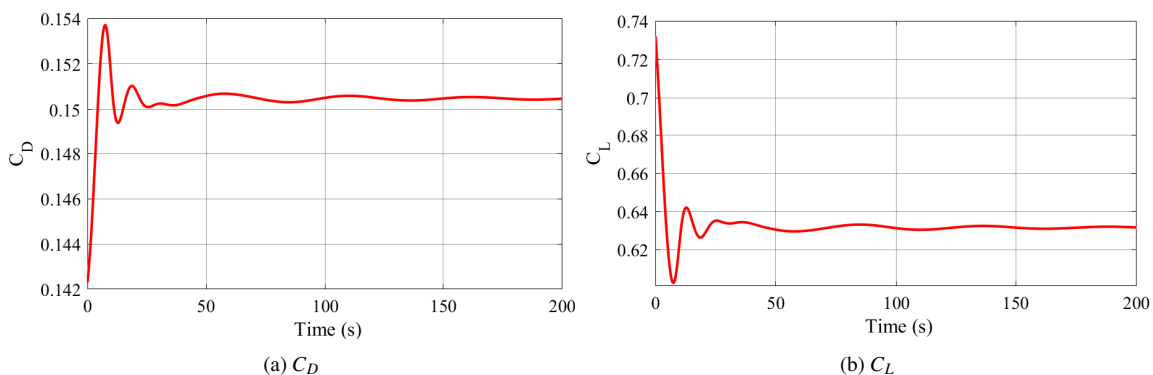


Figure 9: Transient response of the coefficient of drag and lift

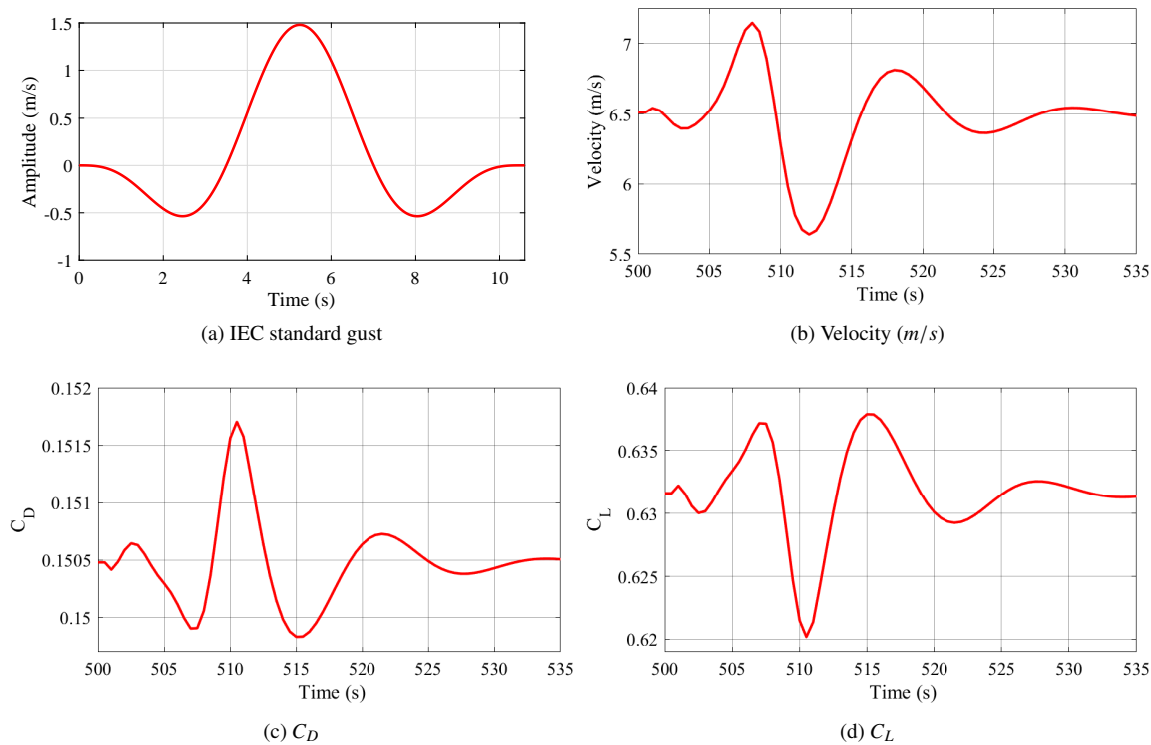


Figure 10: Gust response of the velocity, coefficient of drag and lift

case. The attitude of the airship is shown in Figure 7 (h) in which the pitch angle  $\theta$  has a negative angle of  $-0.6$  degree. This indicates the inherent dynamics of the airship due to the displaced CoG from the CoV. An active elevator control is necessary to overcome this. Otherwise, the station keeping of the airship at the desired altitude will be difficult. The roll and yaw angles remain zero. The position of the airship is shown in Figure 7 (j). The position along the  $x$  direction shows a linear motion which gives  $831m$  at the end of the simulation. There is a deviation of  $3660 m$  in the  $z$  direction also, which shows again the necessity of an active elevator control. The  $y$  direction motion which is expected to be steady at zero value shows a sustained oscillation throughout the simulation. An enlarged view of a portion of the response (650 seconds - 700 seconds) is shown in the insight of Figure 7 (j).

Figure 8 shows a comparison of the forward velocities and position along  $x$  direction for different throttle ratios. Four different cases were considered; zero throttle ratio, initial condition specified in (9), half throttle and a full throttle. It is shown that the maximum velocity this airship can attain is  $6.5 m/s$ . There are considerable overshoots and undershoots present in the velocity as the throttle ratio approaches its maximum. The position of airship also increases with the increase in input.

The transient response of the coefficient of drag and lift for a maximum thrust is shown in Figure 9. The drag shows a considerable overshoot from the initial value which causes a slight damped oscillations lasts for 50 seconds. This increase in drag is due to the motion initiated at the starting of simulation from rest. Similarly for the lift curve, there is a dip in the lift value mainly due to the induced drag.

Even though we haven't considered any closed loop control strategy for the airship, the response to a gust is analyzed. A standard wind gust model, Extreme Operating Gust (EOG) described in IEC standard [18], is used for the analysis. The velocity profile of the EOG gust is given in (10).

$$V_p = \begin{cases} V_m - 0.37V_g \sin(3\pi t/T)(1 - \cos(2\pi t/T)) & : 0 \leq t \leq T \\ V_m & : otherwise \end{cases} \quad (10)$$

A representative gust profile with 10.5 seconds period, zero mean wind velocity and  $2 m/s$  gust amplitude is shown in Figure 10 (a). Since this is an initial stage of this work, we haven't considered real wind scenario. The ability of the model to retain the steady state condition even after the occurrence of a gust is analysed. The gust is applied at 500 second which is sufficiently far from the transient region. The plot of velocity is shown in Figure 10 (b). The direction of application of gust is in the free stream direction of the airship. It is shown that after a surge the velocity falls down quickly at the maximum of the gust and retains to the steady state. Drag increases with velocity and is clearly shown in Figure 10 (c). Since a part of the lift is contributed by the aerodynamic flight, there will be a proportional increase in lift with velocity as shown in 10 (d). With this analysis the open loop stability of the airship in the presence of gust is demonstrated.

## 5 Conclusion

A study on the modelling and simulation of airships lead us to this work where an analytical methodology to model the aerodynamic

coefficients is investigated for an airship. The implementation of the aerodynamic coefficients given in (3)-(8) was successfully done on the mathematical model of the airship given in (1) and (2). All the six coefficients were simulated for five different values of the elevator deflection angles. Even though the drag coefficient at zero angle of attack is 0.1432, the large drag value ( $> 1$ ) at a higher angle of attacks demands an optimal sizing of the airship envelope. The inability to be at the initial altitude indicates the necessity of an active position control system for the airship. A gust analysis of the airship is also carried out using IEC standard gust. The open loop stability of the airship in the presence of gust is demonstrated.

The presented method of calculating aerodynamic coefficients is originally applicable to aircraft where the effect of added mass and buoyancy are negligible. Nevertheless, a general outline on the aerodynamic behaviour of the airship can be obtained from the presented analysis. The model behaves as expected using the coefficients extracted using this method.

This study can be considered as an initial investigation of the mathematical modelling of the aerodynamic parameters of airships. As a future work, an investigation on the validation of the proposed method will be done using CFD based analysis on the selected airship configuration.

**Conflict of Interest** The authors declare no conflict of interest.

**Acknowledgment** The project is funded in part by the Amrita University Fellowship.

## References

- [1] Y. Li, M. Nahon, I. Sharf, "Airship dynamics modelling: A literature review", *Progress in Aerospace Sciences*, **47**(3) pp. 217–239, 2011. doi: 10.1016/j.paerosci.2010.10.001.
- [2] F.A. d'Oliveira, F.C.L.d. Melo, T. C. Devezas. "High-altitude platforms - Present situation and technology trends", *Journal of Aerospace Technology and Management*, **8**(3), pp. 249-262, 2006. doi: 10.5028/jatm.v8i3.699.
- [3] H. Gopinath, V. Indu and M.M. Dharmana, "Development of autonomous underwater inspection robot under disturbances", in *Int. conf. on Technological Advancements in Power and Energy (TAP Energy)*, Kollam, India, pp. 1–5, 2017. doi: 10.1109/TAPENERGY.2017.8397219.
- [4] S P Jones and De Laurier J D. 'Aerodynamic estimation techniques for aerostats and airships', *Journal of Aircraft*, **20**(2), pp. 120-126, 1983. doi: 10.2514/3.44840.
- [5] R. P. Kukillaya, A. Pashilkar, "Simulink model development, validation and analysis of high altitude airship", Report. National Aerospace Laboratories (NAL), No. PDFMC/2017/1000, 2017. doi: 10.13140/RG.2.2.11844.22400.
- [6] M. V. Cook, J. M. Lipscombe and F. Goineau. "Analysis of the stability modes of the non-rigid airship", *The Aeronautical Journal*, **104**(1036), pp. 279–290, 2000. doi: 10.1017/S0001924000091612.
- [7] M Z Ashraf and M A Choudhry. "Dynamic modelling of the airship with Matlab using geometrical aerodynamic parameters", *Aerospace Science and Technology*, **25**(1), pp. 56-64, 2013. doi: 10.1016/j.ast.2011.08.014.
- [8] X. Wu, Y. Wang, C. Huang, Y. Liu, L. Lu. "Experiment and numerical simulation on the characteristics of fluidstructure interactions of non-rigid airships", *Theoretical and Applied Mechanics Letters*, **5**(6), pp. 258-261, 2015. doi: 10.1016/j.taml.2015.11.001.
- [9] P. Funk, T. Lutz, Siegfried. "Experimental investigations on hull-fin interferences of the LOTTE airship", *Aerospace Science and Technology*, **7**(8), pp. 603–610, 2003. doi: 10.1016/S1270-9638(03)00058-0.
- [10] Y. Li, M. Nahon, "Modelling and simulation of airship dynamics", *Journal of Guidance, Control and Dynamics*, **30**(6), pp. 1691-1700, 2007. doi: 10.2514/1.29061.
- [11] X. Wang. "Computational Fluid Dynamics Predictions of Stability Derivatives for Airship", *Journal of Aircraft*, **49**(3), pp. 933–940, 2012. doi: 10.2514/1.C031634.
- [12] J. L. M. Junior, J. S. Santos, M. A. V. Morales, L. C. S. Ges, S. Stevanovic, R. Santana. "Airship Aerodynamic Coefficients Estimation Based on Computational Method for Preliminary Design", *AIAA Aviation Forum*, Dallas, Texas, 2019. doi: 10.2514/6.2019-2982.
- [13] Y. Fan, F. H. Lutze, E. M. Cliff. "Time optimal lateral maneuvers of an aircraft", *Journal of Guidance, Control and Dynamics*, **18**(5), pp. 1106-1112, 1995. doi: 10.2514/3.21511.
- [14] D. Gobiha and N. K. Sinha. "Autonomous maneuvering of a stratospheric airship", in *Indian Control Conference (ICC)*, IIT Kanpur, India, pp. 318-323, January 4-6, 2018. doi: 10.1109/INDIANCC.2018.8307998.
- [15] S. Agrawal, D. Gobiha, N.K. Sinha. 'Nonlinear parameter estimation of airship using modular neural network', *The Aeronautical Journal*, pp. 1-20, 2019. doi: 10.1017/aer.2019.125.
- [16] S. Anoop, O. V. Ramana Murthy and K Rahul Sharma. 'Analysis of airship dynamics using linear quadratic regulator controller', in *15<sup>th</sup> IEEE India Council Int. conf. (INDICON)*, Coimbatore, India, 2018. doi: 10.1109/INDICON45594.2018.8987186.
- [17] Casey Marcel Lambert. 'Dynamics Modelling and Conceptual Design of a Multi-tethered Aerostat System', Master's Thesis, Dept. of Mechanical Engineering, University of Calgary, 1999. <http://www2.eng.cam.ac.uk/~hemh1/SPICE/papers>.
- [18] IEC 61400-2:2013. International Standard, Wind turbines-Part 2: Small wind turbines, 2006. <https://webstore.iec.ch/publication/5433#additionalinfo>.

## Industry 4.0 Operators: Core Knowledge and Skills

Olayan Alharbi\*

Department of Computer Science, College of Science and Humanities in Rumah, Majmaah University, Al-Majmaah, 11952, Saudi Arabia

### ARTICLE INFO

#### Article history:

Received: 23 April, 2020

Accepted: 07 July, 2020

Online: 22 July, 2020

#### Keywords:

Industry 4.0

Internet of Things (IoT)

Human Resources

### ABSTRACT

One of the most important technological changes due to the arrival of Industry 4.0, an initial, gradual, and complex process of technology transfer is taking place, which strongly relies on the integration of universities, industries, and governments. In this context, to make the Industry 4.0 approach a reality, several requirements need to be met. One of them is the need to qualify people to work in industries. This research paper aims to clarify the required knowledge and learning for a person to operate the manufacturing processes associated with some of the capabilities of Industry 4.0. Interviews were conducted with individuals who are Industry 4.0 employees, including experts of technology, education vendors, and employers who are eager to develop and improve their projects. This study provides several results, the most important of which is the focus on the rehabilitation of operators using modern technologies in alignment with the Fourth Industrial Revolution.

### 1. Introduction

The era we are living in these days is that of the digital revolution, which first took place the last quarter of the last century. This era is characterized by the fusion of all the techniques that have been developed, so that the physical, digital, and biological fields have overlapped, and the lines between them have blurred. Therefore, there was a need to move from the third revolution to a new one which is considered something unique. There are three reasons why these transformations are unique and not merely an extension of the previous revolution.

These reasons are: the speed of change, its scope, and the impact of these transformations on the prevailing systems. The speed of the current scientific breakthroughs has no parallel in human history, since there is an amazing speed in the development of digital technologies, compared to that of the previous industrial revolutions [1]. Hence, the Fourth Industrial Revolution is the impact of technology, the Internet, and computers on various sectors of development and labor, including material science, robots, nanotechnology, three-dimensional (3D) printing, unmanned aerial vehicles, digital computing, and globalization [2,3].

The Fourth Industrial Revolution is based on the digital revolution, in which technology is an integral part of the society and a link between the digital, physical, and biological worlds. It is characterized by the use of advanced technology in various fields to improve efficiency and promote development and growth [4].

The new mechanisms in the industry focus on technical education and the development of young people skills.

in different sectors of industry, which works to support the direction of political leadership in human development and devotes attention to human capital because this is the basis of any industrial revolution and development. Investing in the human workforce has become more important than ever before, as the world faces economic and social-environmental challenge [5].

The Fourth Industrial Revolution has added new pressures on the labor markets. Therefore, lifelong learning, learning reform, and new skill retraining initiatives will be the key to ensure that individuals have an effective economic opportunity to compete in the new world of work. Companies will also have the opportunity to attract the talented employees they need for future jobs. According to the World Economic Forum, at least 54% of all employees will need to be taught additional skills by 2022 [6]. Therefore, public and private sectors are required to hire new talented employees or train their employees to gain all the required knowledge to implement the latest technologies of Industry 4.0 Revolution.

The manufacturing industry is one of the sectors that are influenced by the fourth industrial revolution. Similarly, manufacturing is moved from Industry 3.0 phase into the Industry 4.0 phase. In industry 4.0 the manufacturing machines are connected to a virtual environment for each physical machine that presents the manufacturing process in real-time [7]. Therefore, factories are one of the entities that are encouraged to retrain and

\*Corresponding Author: Olayan Alharbi, Email: [O.alharbi@mu.edu.sa](mailto:O.alharbi@mu.edu.sa)

develop their operators to be able to gain all benefits of industry 4.0.

This research paper aims to clarify the required knowledge and learning for a worker to operate the manufacturing processes associated with some of the capabilities of Industry 4.0. This research focuses on operators whose factory is moving from Industry 3.0 into Industry 4.0. Interviews were conducted with individuals who work for Industry 4.0, including vendors of technology, education vendors, employers, and operators who work in factories capable of adopting Industry 4.0 technologies. This study provides several results, the most important of which is the focus on the rehabilitation of operators using modern technologies in alignment with the Fourth Industrial Revolution.

## 2. Literature Review

The influence of the educational system was obviously noticed through all the three industrial revolutions. Bryan Edward Penprase assures this point and adds that there were positive reactions in providing new educational tracks, majors, institutes for research in addition to new curriculums with the aim of preparing individuals [8]. The first industrial revolution was based on elective courses, enormous discoveries and inventions, centered on the use and adaptation of water and steam energy to the mechanization of productivity, including the gradual transformation of societies from their agricultural nature to a new identity. This new identity was characterized by industries that relied on steam engines, both for production, transport, and communication [9].

The second industrial revolution, from 1860 till 1900, was supported by educational institutes such as Pomona College (1887), University of South California (1880) and others to implement the new technologies to close the gap in operators and labor skills. The revolution was characterized by the use of electrical energy instead of steam and coal, which enabled countries with diverse natural resources to devote them to be exploited in the industry. This gave rise to industrial societies based on the principle of huge productivity. The emergence of electrical machines and their exploitation in peace and war was marked by very important inventions for the development of human civilization [10].

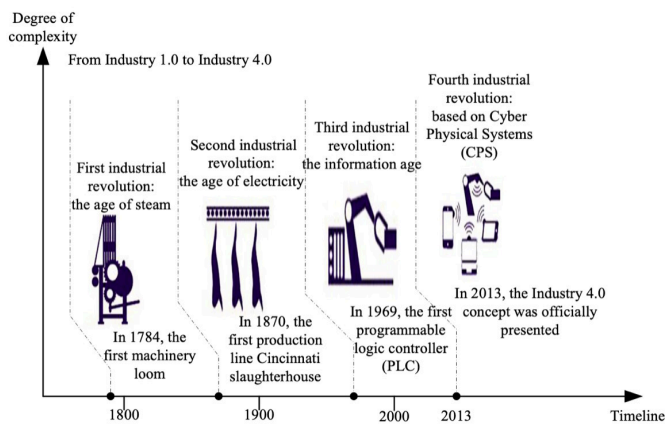


Figure 1. The four stages of the industrial revolution[11]

The third industrial revolution, during 1980s and 1990s, began with the launch of digital technology, such as computers and web-based sites on the Internet. The influence of this revolution was clear in two dimensions. These were the online based education

and new majors such as software development, engineering and automated manufacturing. Numerology was used in programming machines to function without the need to human supervision, causing machines to replace the labor force more extensively than the previous industrial revolutions, see figure 1 [11].

### 2.1. Industry 4.0 in manufacturing

The Fourth Industrial Revolution came with digital capabilities such as the digital-twin that were used in factories to read and analyze data of machines in real time. The revolution beginning was based on the remarkable achievements of the Third Industrial Revolution, particularly, the Internet and its uses, the enormous processing power, the ability to store and recall information whenever it is needed in addition to the unlimited potential for finding the required knowledge. Those achievements have paved the way to unlimited possibilities via the breakthroughs of emerging technologies in various fields. Those fields included many useful applications, inventions and theories such as: artificial intelligence (AI), robots, Internet objects, self-driving vehicles, 3D printing, nanotechnology, biotechnology, materials science, quantum computing, Blockchain, etc [12]. The Fourth Industrial Revolution, or the so-called Industry 4.0, simply means the intensive use of technology and advanced mechanization in the manufacturing processes, the activation of Internet objects, cloud computing, and robotization in the so-called Smart Factory.

The Fourth Industrial Revolution requires new skills such as data analysis and basic information technologies [13]. By these skills, the Revolution can have many impacts on business: Influencing customer expectations, product improvement, collaborative innovation, and organizational forms [11]. Regarding consumers' expectations or businesses, customers are always the centerpiece of the economy, which revolves around how to improve the process of customer service and satisfy them. Moreover, physical products and services can be improved through digital facilities that help increase the value. New technologies provide assets with better qualities and flexibility, while data and analysis change the way these assets are maintained. At the same time, the world of customers' feedback and the data-based services, asset the evaluation process through analysis, new forms of co-operation, especially the provided speed with which innovation occurs, the emergence of global platforms and other new business models which ultimately means that with talent, new culture and organizing methods will emerge [14].

The drastic shift from the simplicity of the Third Industrial Revolution to the innovation based on a combination of technologies (the Fourth Industrial Revolution) is creating a kind of pressure on companies to reconsider the way their businesses are progressing. One of the most experienced aspects of the recent Industrial Revolution (the 4<sup>th</sup>) is the concept of service-based design. This design is different from the factory settings regarding customers products and those products of companies for specific individuals, according to their specifications. On the other hand, this means that the new products will continue to be automatically and permanently produced by manufacturers [11].

The Fourth Industrial Revolution will affect people in many ways [15,16]. It will change what we do, what we are, and will also affect our identity and everything associated with it. For instance, our sense of privacy, property concepts, consumption patterns, and the time devoted to work and leisure, as well. Furthermore, it will also improve our skills and enhance our relationships. The feedback process for the manufactured products



will be self-explanatory, which will automatically help the manufacturer develop his manufacturing method.

### 2.2. *Benefits and challenges of Industry 4.0*

The mechanisms and tools of the Fourth Industrial Revolution reduce the operational costs of factories, improving the levels of corporate profits and the good employment of machines, as well as working to improve and raise the efficiency of human capital to cope with the flourishing era of advanced industry and manufacturing technology. The Fourth Industrial Revolution offers many opportunities for human societies to achieve higher rates of economic, social, and human development in general by minimizing the production costs and consequently providing services, transport, and communication, which combine high efficiency with lower prices. Other positive factors of Industry 4.0 include a contribution to better human healthcare as it shortens the time needed to the process of development, and easily disseminates its achievements throughout the world.

According to [17], the Fourth Revolution has a number of advantages. Three of these are:

1. The speed of its development at the rate of exponential growth, as modern a technology, always drives the development and the emergence of newer and stronger technology rather than the rate of a linear growth.
2. The increase in the volume of benefits per capita in this revolution is different. In the digital age, companies need a few number of employees and materials to manufacture products with great benefits and high qualities. For digital companies, storage, transport costs and product reproduction are reduced to zero level. Some technology-based companies are developing without large capital, e.g. Instagram, WhatsApp.
3. The coordination and integration of different discoveries is widely obvious in this Revolution. The profits of digital manufacturing technology and biotechnology have been shared. Recruiting designers and architects for digital design, implementing modern materials science, and industrial biology have increased the innovation and the production quality of modern products.

At the same time, Industry 4.0 imposes unprecedented challenges on human societies. Among these we find the following:

1. The dominance of large companies in industrial production and the threat of the disappearance of approximately 50% of jobs, which poses a challenge to the spread of unemployment, especially in countries that are not ready for the transformation process.
2. The possibility of employing the capabilities and advantages of the Fourth Industrial Revolution techniques in committing illegal or immoral acts that could harm society, values, or individuals, such as the growth of cybercrime, violation of privacy, spreading hatred, extremism and false news, and preventing the development of further global views.
3. Employees of factories who are moving from Industry 3.0 into industry 4.0 need to reskill themselves to utilize the benefits. This challenge relies on the effort of several stakeholders such as employers, vendors of technologies, and educational institutes.

### 2.3. *Technologies of Industry 4.0*

The Fourth Industrial Revolution works in harmony with three technologies: A) Physical technologies which are easily recognized, because they are concrete, such as self-driving cars, production line machines in factories, 3D printing machines, and sophisticated robots [14]. B) Digital technologies which create radically new devices, software and platforms that will revolutionize the way people and businesses conduct their dealings, namely, the IoT which can create the relationship between products, services, places, and people through related technologies and platforms [18]. C) Biological technologies: Industry 4.0 includes technologies to link humans to other technologies such as physical technologies. Currently, a machine can allow only operators who are eligible to work with it based on the operator's record in the factory human resources system [19].

### 2.4. *Human Capital in Industry 4.0*

Two of the most important possibilities of human capital are knowledge and education, where they support the Fourth Industrial Revolution to achieve the desired goals, which benefit institutions. The human capital theory believes that knowledge provides greater cognitive skills to individuals, thus it is necessary to impel their productivity and efficiency in a potential way to develop the related activities.

Certain education, experience, knowledge, and skillset have to be possessed by employees. As a result, these are used to create values regarding the success of an organization. Thus, we can see how experience, knowledge, skills, and education are critical for the human capital in organizations, which, in essence, underscore the importance and the role of the human capital in the Industry 4.0 revolution.

The requirements of Industry 4.0 are not only a workforce, but also human capital, nurtured in competitive education systems, which must be well prepared for creative work environments [20]. Nowadays organizations do not require physical and tangible humans as the present and future expectations seem to offer plenty of challenges to organizations and humanity. Therefore, as humans embrace the ushering in Industry 4.0, it has become imperative for nations, as well as organizations, to establish education systems that are more focused on knowledge which is beyond what the world currently addresses. This requires creative teaching particularly to children at early ages, right up to the university levels [21].

As modes of passing examinations never produce thinkers, creators, or ingenuity, there is a need to a new move away from the traditional education systems of writing, reading, and memorization. Therefore, nations need to revolutionize their education systems to produce super-humans capable of implementing the Industry 4.0 revolution. Educational revolutions require a national culture that is supportive of such initiatives of the government, where the citizens feel they have something to contribute towards achieving the goals of the Industry 4.0 revolution. Consequently, the result is the production of human capital that is capable of benefiting the Industry 4.0 revolution [22].

### 2.5. *Education in Industry 4.0*

Education based on Industry 4.0 is defined as technology literacy, information literacy, media creativity, social competence and responsibility, workplace skills, and civic engagement. This is

because the available information dramatically increases, therefore people are required to have new skills to critically access and process content to ensure the best social communication and interaction. The Industry 4.0 revolution presents an opportunity, as well as challenges, to nations' education systems, and only those nations whose education systems are anchored in inclusiveness and technological imperatives will remain competitive. To meet all of these demands for the Industry 4.0 revolution, lifelong learning is necessary to ensure that everyone can stay well-trained. Universities have to lead the research efforts not only to identify the relevant skills, but also to produce a high-caliber workforce that own the skills necessary for the Industry 4.0 revolution [23].

Barro [24] pointed out the importance of education and the accumulation of human capital, as well as their positive and moral impact on economic growth. In his study of 1991, he examined the determinants of the economic growth of a sample of 98 countries for the period 1960-1985 and concluded that the human capital has a significant direct effect on the economic growth. In his study of the impact of education on economic growth in 2013, he pointed out that applied studies were based on the neo-classical models of growth that were developed in the 1950s and 1960s. The variables of these models were often expanded to include government policies, institutions, and the accumulation of the human capital. In his 2013 study, he tried to emphasize the importance of his model of growth from to understand the possibility of maintaining long-term growth.

Besides, [25] pointed out that the 2030 Sustainable Development Plan and Goals emphasize the importance of education in promoting change in knowledge, skills, values, and behavior to achieve greater growth sustainability. The emergence of this concept was associated with the emergence of a need to a good educational system to deal with the increasing challenges.

Since the Fourth Industrial Revolution began to appear in the labor market, graduates have to prepared for it, and quality has to be achieved within the educational and research institutions in addition to applying an expansion of modern study disciplines. Curriculum development has become necessary to suit the Fourth Industrial Revolution, as well as enhancing the abilities and skills of students and faculty members in the field of scientific research and innovation.

### 3. Method

The chosen method was selected by the author based on his approach which was influenced by the phenomenological approach to sociology as it has been used in similar studies [26], [27] [28]. The method was chosen as it helps to understand all the required knowledge areas related to the 4<sup>th</sup> Revolution. To understand the required knowledge and skills for operators in the early adoption of Industry 4.0, interviews with individuals who work for Industry 4.0, including vendors of technology, education vendors, employers and employees (operators and managers) were conducted. It is believed that individuals who experienced the transition from Industry 3.0 to Industry 4.0 can identify the required skills and knowledge for the operators who are going to work in Industry 4.0.

#### 3.1. Data collection

The interviews were planned, typed up, and arranged in a way that allowed for enough freedom in discussing related issues. An introductory question was developed to warm up the discussion and to let respondents provide important information. Additional

questions were prepared for cases where the respondent was not sure what was meant by the questions or if they fell short in answering a particular question. Permission to record and take notes was always obtained before an interview.

#### 3.2. Data Analysis

The collected data from the semi-structured interviews were then tested thematically. Thematic analysis refers to the identification and analysis of patterns of meaning found in a specific set of data [29]. Specifically, it is a method used to identify, analyze and report patterns in the collected data, which helps to provide an organized, descriptive, and rich interpretation of those data. To correctly identify themes, the analyst needs to understand what counts as a theme. Themes grasp important aspects of the dataset that have a connection to the research question, representing a level of pattern or concept that is found in the analyzed data [30].

This research employed thematic analysis primarily for identifying, interpreting, and explaining the different aspects of the research topic. The adoption of this method in the research was due to its suitability for analyzing the collected data, concerning the research objectives. In addition to other advantages, thematic analysis was also an appropriate match to the research paradigm. To ensure a thorough analysis, the six steps proposed by [30] were followed. These steps are familiarization with the data, generation of initial codes, search for themes, revision of themes, definition, and naming of themes, and, finally, production of the report. The report shows the connections of the qualitative data in answering the research question.

The layered approach was used to strategically recruit participants that could provide us with insights into the required knowledge and skills for operators to work in an Industry 4.0 factory. The researcher interviewed technology vendors first because they would have had the experience of the new technology and the required skills needed for Industry 4.0.

Technology vendors provided comments about Industry 4.0 in general. While technology vendors explained their opinion about the general requirement of building Industry 4.0 solutions, the discussion was led to focus on the new skills required for operators to work with such technology. Moreover, technology vendors shed light on the essential features of Industry 4.0 and the skills required to gain the benefits of transitioning to Industry 4.0 or the quick wins of the transaction. Thus, the interview was led to a discussion on the required skills and knowledge for operators to work in factories that apply Industry 4.0. Furthermore, technology vendors provided stories about how technology has changed over time.

After clarifying the theme of the new technologies of Industry 4.0, participants who had been involved in working in factories that apply some of the Industry 4.0 technologies were recruited. In the second round, people of different positions on the employer side, starting with a chief Executive Officer (CEO) of a rigid plastic factory were interviewed. Lastly, participants from education vendors that have programs for preparing operators to find the missing knowledge in traditional operator programs and the required knowledge and skills for Industry 4.0 operators were recruited.

The first round consisted of eight interviews of technology vendors and was conducted in the summer of 2017 to find the required knowledge and skills for being an operator in Industry 4.0.

The second round consisted of 17 interviews that aimed identify the required knowledge and skills for Industry 4.0, based on the feedback of users and employers of Industry 4.0. The users and employers applied the essential basics of Industry 4.0, such as connecting the machine and starting it, to automatically retrieving production data from the machine immediately. These users are those who, for instance, use manufacturing execution systems to show dashboards for data that are retrieved from the machines directly.

The last round consisted of interviewing educational vendors that teach and prepare operators to work in manufacturing. Educational vendors, who are designing new courses to bridge the gap between Industry 3.0 and the early features of Industry 4.0 (e.g., such as data visualization and connecting production machines), were chosen. This round consisted of interviewing eight participants. In total, 27 individuals were interviewed in person and six over the phone, all of whom had been involved in Industry 4.0 in many different capacities (see Table 1 for a partial list).

Table 1. Selected interviewee roles.

Title	Number of interviews
Factory CEO	1
Factory manager	1
Production manager	1
Operator	3
IoT manager and leader	3
Industry 4.0 pre-sale architectures	3
Program coordinator	1
Instructor	2

### 3.3. Themes

The collected data were carefully reviewed, looking for the barriers, challenges, and difficulties addressed by respondents, either implicitly or explicitly. The findings were coded and clustered into main and sub-categories. The issues, concepts, and patterns were re-categorized, and refined.

Among the several steps taken to analyze the collected data, the interviews were revisited to find and list all of the mentioned challenges, which were then organized into a list. The list was then refined, excluding repeated issues and combining similar challenges. Similar basic issues were clustered to form organizing themes.

This step reduced the list into organizing themes, which were then combined under wider classifications. The themes are discussed underneath.

The first round of interviews targeted the Industry 4.0 solution vendors. Firstly, the focus of the interviewees was shifted to the new IT capabilities to be added to factories as required competencies for Industry 4.0. These capabilities include edge connectivity, cloud solutions, and microservices. Additionally, there are certain prerequisites, such as the ability to connect a production machine and the availability of process and production sensors.

However, the focus of this research was on the operators. Therefore, the interviewer led the conversation to focus on the new set of skills required for Industry 4.0 operators. After clarifying

these concerns, the technology vendors started by stating the tools and solutions, focusing on Industry 4.0 operators, as follows:

One of the essential features of Industry 4.0 is its digital twin. A digital twin is a digital replica of a living or non-living physical entity. By bridging the physical and the virtual world, data are transmitted seamlessly, allowing the virtual entity to exist simultaneously with the physical entity. Operators will be able to obtain different sources of manufacturing data, beyond what their old friend, the human-machine interface (HMI), could provide them with. The digital twin in Industry 4.0 can provide real-time data for the targeted machine and can compare it with similar machines or the process and production data of the same machine historically. Operators need to have essential knowledge of data science activities. These activities include data analysis, descriptive analysis, and problem-solving [7].

Technology providers have industrial platforms that empower users to use and rebuild micro services with minimum programming knowledge. Predix by General Electric (GE) and MindSphere by Siemens are examples of these platforms. Operators will gain access to such technology and could develop basic services for their common issues. These microservices do not require advanced programming skills to be utilized by an operator; however, basic knowledge of programming, web development, and problem-solving are required for operators who want to run the extra mile and develop their tools.

On the other hand, operators work closely with the operational technology that is provided on the machine side. Industry 4.0 will allow operators to access real and historical manufacturing data from different resources. Therefore, operators need to have the ability to work with IT solutions, such as manufacturing execution systems.

The second-round targeted employers and business owners that recently empowered some of their production lines with Industry 4.0 tools and solutions.

The first concern of an employer (a production manager) was that the operators will remain essentially the same. The operators will not be developers for Industry 4.0, but they will need to know how to work and use Industry 4.0 tools. The operators will still need to have the essential instrumentation and process knowledge. However, they will have access to new tools that will help them to improve the productivity of their machines. An example of a new tool is historian trends. This is an analytical tool that allows users to present different process data in one dashboard to analyze a certain issue, such as major downtime.

Surprisingly, a CEO and one of the production managers stated that the English language is important for empowering operators to get the most from Industry 4.0 tools and solutions. The available knowledge and most of the provided solutions are available only in English.

On the other hand, IT literacy is another important prerequisite for operators to utilize Industry 4.0 tools and solutions. IT literacy is different from digital literacy or informatics literacy/computer science literacy, as it puts more emphasis on computer programming, algorithms, and other important mathematical and computational concepts. IT literacy includes an essential area, such as information management, integrated system technology, platform technology, and system paradigms.

The last round consisted of interviews with educational institutions. All of the interviewed educational institutions had an ongoing contract involving the training of operators for local employers. One institution had contracts with food and beverage packaging manufactures and two had contracts with petrochemical, oil, and gas industries. The first theme introduced by the institution members was the IoT. New operators will work on machines that are connected to edge or cloud computing services. The existing sensors in the machine or the new ones will have the same concept of the IoT. Operators will be introduced to the concept of IoT and how platform technology can help. This new theme works smoothly with most of the industrial platform microservices, which are developed to interact with the physical assets to provide better information for operators and the rest of the factory team.

Similar to technology vendors, the interviewees from the educational institutions listed the following concepts to be added to the required skills of Industry 4.0 operators: (1) Basic programming and problem-solving, (2) web development and advanced analytical skills, and (3) basic data science activities, such as anomaly detection. Additionally, IT literacy is a topic that is required by all three different stakeholders. Table 2.0 lists all of the common required knowledge areas, as indicated by all of the stakeholders.

The educational institution interviewees agreed that operators are required to know about programmable logic controllers (PLC), instrumentation, process control, and automation. However, these are not new, and there are diplomas for operators.

In summary, the results from all of the interviews could be organized into two main categories:

1. Operation technology categories which include PLC microcontrollers, pneumatics and hydraulics, and industrial instrumentation.
2. Information technology (IT), which is the goal of this study and identifies the missing knowledge for Industry 4.0 operators as follows: digital industry platform; computer skills and information technology (IT literacy); data analytics and visualization; basic programming and problem-solving skills; introduction to web development; and industrial IoT.

#### 4. Results and Discussion

According to the conducted interviews, the study provided several results, including the following:

It is necessary to have certain knowledge for operators working in factories adopting industry 4.0. Moreover, it is not a must that the required knowledge should be of a high level. This knowledge can be provided in vocational and technical institutes that offer a two-year diploma after finishing secondary school. This was clear in the interview with one of the interviewees who mentioned that a few number of courses can be added to the current ones at vocational institutes. The new added courses, which match the conditions of the work fields, are attractive factors to students at the vocational institutes. Those students chose to study there as their GPA at secondary schools did not enable them to join universities but on the other hand, gave the chance to get a good job immediately after finishing their diplomas.

Table 2. Selected interviewee roles.

Stakeholder Title	Recommended Course/Knowledge
Technology Provider	<ol style="list-style-type: none"> <li>1. Programming and web development foundation</li> <li>2. Industrial platform technologies</li> <li>3. Analytical tools and problem-solving skills</li> </ol>
Educational Institution	<ol style="list-style-type: none"> <li>1. Information and communication technology</li> <li>2. Programming and web development foundation</li> </ol>
Employer	<ol style="list-style-type: none"> <li>1. Information and communication technology</li> <li>2. Human-machine interface (HMI)</li> <li>3. Programmable logic controllers (PLC)</li> <li>4. Instrumentation</li> <li>5. Process control</li> <li>6. Automation</li> <li>7. Familiarity with technology</li> <li>8. Programming and web development foundation</li> <li>9. Data science</li> </ol>

Furthermore, the results shed light on the role of the private sector firms in supporting the 4<sup>th</sup> revolution. One of the interviewees has mentioned that the institute was established by several companies, and added that the courses to be studied are related to the future needs of those companies.

Many modern technologies must be available to suit the Fourth Industrial Revolution, including edge connectivity, cloud solutions, and microservices. Additionally, certain prerequisites must be met, such as the ability of a production machine to be connected and the availability of process and production sensors. The operator is one of the most important factors supporting the Fourth Industrial Revolution due to the fact that he/she, in most cases, is the first user of industry 4.0 products rather than developers or designers. Consequently, he/she must be rehabilitated and trained for that purpose.

The required knowledge can be achieved through:

1. Basic knowledge of programming, web development, and problem-solving, which are required for operators who want to run the extra mile and develop their tools.
2. The ability to work with IT solutions, such as manufacturing execution systems.
3. Mastering the English language, which is important to empower operators to get the most of Industry 4.0 tools and solutions.
4. IT literacy, which is an important prerequisite for operators to utilize Industry 4.0 tools and solutions.
5. Knowledge of PLC, instrumentation, process control, and automation.



6. Knowledge of digital twins (one of the essential features of Industry 4.0). A digital twin is a digital replica of a living or non-living physical entity. By bridging the physical and the virtual world, data are transmitted seamlessly, allowing the virtual entity to exist simultaneously with the physical entity. Operators will be able to retrieve different sources of manufacturing data from those which their old friend, the human-machine interface (HMI), provided them with. The digital twin in Industry 4.0 can provide real-time data for the targeted machine and can compare it with similar machines or the process and production data of the same machine historically.
7. Essential knowledge of data science activities. These activities include data analysis, descriptive analysis, and problem-solving.
8. Knowledge of how to work and use Industry 4.0 tools, without being a developer for Industry 4.0.

## 5. Conclusion

In light of the rapid changes in financial, economic, political, and social systems, our world is on the threshold of a Fourth Industrial Revolution, which will not only change the form of industries and the methods of production but will impose new requirements on humans. This research paper provides strong evidence of the importance of the role human capital plays in the Industry 4.0 revolution in which the success or failure of most organizations largely depends on how their human capital is managed. This is because the Industry 4.0 revolution provides a space where employee-machine interactions are the order of the day. There is an interconnectedness among the various players and actors. The interfaces created to become the connecting points between workers and machines. The features of the Industry 4.0 revolution require creative and inventive workers i.e. those who are not just creative but are also knowledgeable and have the technical expertise required to work in such environments. Such workers are nurtured through an education system, where creativity, inventiveness, knowledge, and technology flourish and are entrenched in the national culture.

## References

- [1] M. Peruzzini, F. Grandi, M. Pellicciari, "Benchmarking of tools for user experience analysis in industry 4.0. *Procedia Manufacturing*", 11, 806–813, 2017. <https://doi.org/10.1016/j.promfg.2017.07.182>.
- [2] J. Bloem, M.V. Doorn, S. Duivestein, D. Excoffier, R. Maas, E.V. Ommeren, "The fourth industrial revolution", *Things Tighten* 2014, 8, 2014.
- [3] G. Li, Y. Hou, A. Wu, "Fourth Industrial Revolution: technological drivers, impacts and coping methods", *Chinese Geographical Science* 2017(27), 626–637, 2017. doi: 10.1007/s11769-017-0890-x
- [4] M. Mindas, S. Bednar, "Mass customization in the context of industry 4.0: implications of variety-induced complexity". *Advanced Industrial Engineering* 2016, 21–38, 2016.
- [5] N.W. Gleason, "Higher education in the era of the fourth industrial revolution; Springer, 2018. [https://doi.org/10.1007/978-981-13-0194-0\\_6](https://doi.org/10.1007/978-981-13-0194-0_6);
- [6] R. Samans, "Globalization 4.0 shaping a new global architecture in the age of the Fourth Industrial Revolution: A call for engagement. In *Proceedings of the World Economic Forum Report*; 2019.
- [7] T.H.J. Uhlemann, C. Lehmann, R. Steinhilper, "The digital twin: Realizing the cyber-physical production system for industry 4.0. *Procedia Cirp* 2017, 61, 335–340, 2017. doi: 10.1016/j.procir.2016.11.152.
- [8] B.E. Penprase, "The Fourth Industrial Revolution and Higher Education. In *Higher Education in the Era of the Fourth Industrial Revolution*; Gleason, N.W., Ed.; Springer: Singapore, 2018; 207–229, 2018. doi: [https://doi.org/10.1007/978-981-13-0194-0\\_9](https://doi.org/10.1007/978-981-13-0194-0_9).
- [9] D. Plinta, "New information technologies in production enterprises, *Advanced -industrial engineering, Industry 4.0, Bielsko-Bia, 7-20, 2016.*

- [10] F. Zezulka, P. Marcon, I. Vesely, O. Sajdl, "Industry 4.0—An Introduction in the phenomenon. *IFAC-PapersOnLine* 2016, 49, 8–12, 2016. doi: <https://doi.org/10.1016/j.ifacol.2016.12.002>.
- [11] K. Zhou, T. Liu, L. Zhou, "Industry 4.0: Towards future industrial opportunities and challenges. In *Proceedings of the 2015 12th International conference on fuzzy systems and knowledge discovery (FSKD)*; IEEE, 2015; 2147–2152, 2015. doi: <https://doi.org/10.1109/FSKD.2015.7382284>.
- [12] N. Benias, A.P. Markopoulos, "A review on the readiness level and cybersecurity challenges in Industry 4.0. In *Proceedings of the 2017 South Eastern European Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM)*; IEEE, 1–5, 2017. doi: 10.23919/SEEDA-CECNSM.2017.8088234 .
- [13] A. Benešová, J. Tupa, "Requirements for education and qualification of people in Industry 4.0". *Procedia Manufacturing*, 11, 2195–2202, 2017. doi: <https://doi.org/10.1016/j.promfg.2017.07.366>.
- [14] M. Hermann, T. Pentek, B. Otto, "Design principles for industrie 4.0 scenarios. In *Proceedings of the 2016 49th Hawaii international conference on system sciences (HICSS)*; IEEE, 3928–3937, 2016. doi: <https://doi.org/10.1109/HICSS.2016.488>.
- [15] M. Armstrong, S. Taylor, "Armstrong's Handbook of Human Resource Management Practice: Edition 13; Kogan page, 2014;
- [16] L. Gehrke et al., "Discussion of Qualifications and Skills in the Factory of the Future: A German and American Perspective.; 2015.
- [17] H. Kohl, "Holistic Approach for Human Resource Management in Industry 4.0, 2016. doi: <https://doi.org/10.1016/j.procir.2016.05.102>.
- [18] T. Bauernhansl et al., "Industrie 4.0 in Produktion, Automatisierung und Logistik: Anwendung-Technologien-Migration; Springer, 2014. doi: <https://doi.org/10.1007/978-3-658-04682-8>;
- [19] S. Vasin et al., "Emerging Trends and Opportunities for Industry 4.0 Development in Russia. *ERSJ* 2018, 21, 63–76, 2018. doi: 10.35808/ersj/1044.
- [20] M. Baygin, H. Yetis, M. Karakose, E. Akin, "An effect analysis of industry 4.0 to higher education", In *Proceedings of the 2016 15th international conference on information technology based higher education and training (ITHET)*; IEEE, 1–4, 2016. doi: <https://doi.org/10.1109/ITHET.2016.7760744>.
- [21] S. Weyer, M. Schmitt, M. Ohmer, D. Gorecky, "Towards Industry 4.0 - Standardization as the crucial challenge for highly modular, multi-vendor production systems", *IFAC-PapersOnLine* 2015(48), 579–584, 2015. doi: <https://doi.org/10.1016/j.ifacol.2015.06.143>.
- [22] B. Sivathanu, R. Pillai, "Smart HR 4.0—how industry 4.0 is disrupting HR. *Human Resource Management International Digest* 2018. doi: 10.1108/HRMID-04-2018-0059 .
- [23] C. Burmeister, D. Lüttgens, F.T. Piller, "Business model innovation for Industrie 4.0: Why the "Industrial Internet" mandates a new perspective on innovation. *Die Unternehmung* 2016, 70, 124–152, 2016. doi: 10.5771/0042-059X-2016-2-124.
- [24] R.J. Barro, "Education as a determinant of economic growth. 2002. doi: 10.1016/j.sbspro.2015.07.156.
- [25] A.E. Wals, Y. Mochizuki, A. Leicht, "Critical case-studies of non-formal and community learning for sustainable development; Springer, 2017. doi: <https://doi.org/10.1007/s11159-017-9691-9>;
- [26] F. Baena, A. Guarín, J. Mora, J. Sauza, S. Retat, "Learning Factory: The Path to Industry 4.0. *Procedia Manufacturing* 2017(9), 73–80, 2017. doi:10.1016/j.promfg.2017.04.022.
- [27] I. Neaga, "Applying Industry 4.0 And Education 4.0 To Engineering Education", *Proceedings of the Canadian Engineering Education Association (CEEA)* 2019. doi: 10.24908/pceea.vi0.13859.
- [28] P.A. Quezada-Sarmiento et al., "Curricular design based in bodies of knowledge: Engineering education for the innovation and the industry. In *Proceedings of the 2016 SAI Computing Conference (SAI)*; IEEE: London, United Kingdom, 843–849, 2016. doi: <https://doi.org/10.1109/SAI.2016.7556077>.
- [29] H. Joffe, "Thematic analysis. Qualitative research methods in mental health and psychotherapy, 1, 2012. doi: <https://doi.org/10.1002/9781119973249.ch15>.
- [30] V. Braun, V. Clarke, "Using thematic analysis in psychology. *Qualitative research in psychology* 2006(3), 77–101, 2006. doi: 10.1191/1478088706qp063oa.



## Contextualization of the Augmented Reality Quality Model through Social Media Analytics

Jim Scheibmeir<sup>1,\*</sup>, Yashwant Malaiya<sup>2</sup>

<sup>1</sup>*Department of Systems Engineering, Colorado State University, Fort Collins, 80523, United States*

<sup>2</sup>*Department of Computer Science, Colorado State University, Fort Collins, 80523, United States*

---

### ARTICLE INFO

*Article history:*

*Received: 15 June, 2020*

*Accepted: 12 July, 2020*

*Online: 22 July, 2020*

---

*Keywords:*

*Augmented Reality*

*Software Quality*

*Social Media Analytics*

*Quality Model*

---

---

### ABSTRACT

*Augmented Reality applications are gaining popularity while maintaining novelty. Many industries are utilizing the user interface type, and use cases are becoming repeat patterns of problem solutions. Despite this rising popularity, quality has not matured nor has the technology become mainstream. Novelty must be approached as risk, and risk must be evaluated for and tested to assure adequate levels of quality. Quality itself can also be vague and have contextual definition. For these reasons, a quality model for augmented reality was created. This work analyzes over two hundred thousand tweets, collected during 2019 and 2020, relating to augmented reality technology, and contextualizes various data points to the established AR Quality Model. The education industry had the highest mentions among the tweets within the scope of this research while the tweets labeled to the transportation industry had the highest sentiment. Furthermore, the tweets were shown to illustrate the needs of testing against the characteristics within the quality model; presence, perspective, interaction, portability and persistence.*

---

## 1. Introduction

This paper is an extension of research work producing an Augmented Reality (AR) quality model and originally presented in IEEE Ubiquitous Computing, Electronics & Mobile Communication Conference [1]. Augmented reality was first coined as a term by Caudell and Mizell in 1990 [2]. Augmented reality is accomplished through an application's user interface by merging both physical surroundings as well virtual objects, which may be models or scenes [3].

Research of AR applications covers many different use cases and industries. One study has uncovered findings such as children seem to understand how to use the augmented reality application as soon as they began to interact with the technology. Another research effort found learning English utilizing AR can help students in learning material and motivated students to learn the language. Another study proposes AR mobile applications in the household to educate parents and caregivers of potentially dangerous objects in the home that may endanger infants. Easily and quickly comprehended how innovative the space is becoming, and how useful for humans. In fact, Gartner, Inc. has predicted that "by 2021, at least one-third of enterprises will have deployed

a multiexperience development platform (MXDP) to support mobile, web, conversational and augmented reality development" [4].

### 1.1. The Augmented Reality Application Quality Model

The importance of this technology extends into many industries. AR has been utilized in applications for factory design [5]. Military equipment maintenance is another use case where AR has been applied [6]. Yet another example is in the use of AR for teaching new electronics assembly skills [7].

The effectiveness of the applications requires them to have levels of quality. One model that identifies eight characteristics of software quality is the ISO 25010 model: Functional, Performance, Compatibility, Usability, Reliability, Security, Maintainability, Portability [8]. However, AR applications have not matured, and expectations of AR application users are not currently known [9]. Because the AR interface is a novel one, the intent of traditional usability evaluation methods which have matured over time for traditional software quality are inadequate [10]. For these reasons, a quality model was established that encapsulates the characteristics for AR applications broadly as Presence, Perspective, Interaction, Immersion, and Persistence [1].

---

\*Corresponding Author: Jim Scheibmeir, Email: [jimscheibmeir@gmail.com](mailto:jimscheibmeir@gmail.com)

Beyond the characteristics of quality is the necessity to utilize context in testing. Emerging technologies and designs such as digital twins must still begin with system needs analysis through known methods such as context diagrams and objective trees [11]. And the same is true for established technologies. According to Gartner, “your strategy for API quality needs to be built with the business and environment context in mind” [12].

These contextual concerns drive much of this extended research. Utilizing over 200,000 tweets towards AR and VR, this research contextualizes quality regarding AR and various industries using text analysis with mention, time series and sentiment analysis.

## 2. Methodology

### 2.1. Collection

Beyond some popular and academic notes, this research methodology focuses on social media analytics. Data acquisition is completed through a small program written in R and executed daily beginning in August of 2019. The program continues to execute, as of the publication of this research, populating a MySQL database hosted in AWS RDS. The program uses the twitterR library as an API to Twitter social media platform. Figure 1 omits various keys and passwords yet implements the basics of importing the twitterR library and establishing connections to both Twitter and the MySQL database prior to searching on ‘#AR’ and storing the search results.

```

1 library(twitterR)
2 library(RCurl)
3 library(ggplot2)
4 library(reshape)
5 library(sentimentr)
6 library(RMySQL)
7
8 consumer_key <- ""
9 consumer_secret <- ""
10 access_token <- ""
11 access_secret <- ""
12
13
14 register_mysql_backend("", "", "", "")
15 setup_twitter_oauth(consumer_key, consumer_secret, access_token, access_secret)
16
17 dbartweets <- search_twitter_and_store("#AR", retryonRateLimit = 20, lang = "en")
    
```

Figure 1 Screenshot of example usage of twitterR library within R Studio

Many hashtags and twitter profiles names were utilized in the searches. The next table identifies some of the additional search items, extending the code found on line 17 of Figure 1.

Table 1 Search Terms used in the twitterR API Calls

'#AR'	'@AR_Maxst'
'#VR'	'@wikitude'
'#augmentedreality'	'@Vuforia'
'# omnichannel'	'#sparkar'
'#multiexperience'	'@fbplatform'
'#virtualreality'	'#arkit'
'@GoogleARCore'	'#arkitnews'
'@ArcoreGoogle'	'@AR_Maxst'

The tweets are stored with other certain meta data, made available through the twitterR library and API. Fields beyond the text collected and utilized in this research include discrete data fields towards the number of times a tweet has been marked as a favorite by a Twitter user, as well as the number of times a tweet has been re-tweeted. A date field within the table indicates the date the tweet was created in the Twitter platform.

### 2.2. Data Preparation and Processing

Social media analytics is not without data quality and research concerns. Tweets lack standards, such as in hashtag usage. A tweet that embeds #AR may be about Arkansas, a state within the United States, and not about augmented reality technology. Prior to formulating findings and information from the collected data, much data preparation and cleaning are required. Data preparation activities may account for 80 percent of time invested in data science effort [13]. The amount of data dropped during preparation from the data set can be large. This is to be expected, as with the example of #AR as Arkansas, and not augmented reality. The search algorithm sought out technology vendors, whose tweets may have been towards other products, process, marketing or communication needs. Tweets are also fixed in length, 280 characters [14]. A short message may come from a targeted technology provider, but also fail to cite key search criteria words. Within this research set, counting only the tweets that mention specifically the words “augmented” or “virtual” (not case sensitive), the original 637,384 tweets were reduced by two-thirds to 211,269 tweets. The breakdown of augmented and virtual tweet counts is illustrated in Figure 2.

The intention of data preparation and cleaning is to move the data from chaotic towards precision [13]. In this research, we create useful subsets without disrupting the raw database contents, as new hypothesis and meaning may yet to be executed and found across the raw data.

To create useful datasets, we must also create new fields, such as categorical values. Tweets have some indication to business and industry, so categorizing the tweets by industry lends toward understanding a contextual need for quality, such as HIPAA requirements and testing concerns towards health or pharmaceutical implementations of AR. A small program in R reads the tweet texts from the accumulated corpus and maps to common industries, counting the industry mentions, including if multiple industries are mentioned in a single tweet.

## 3. Results

The first finding is that among all tweets that mention either “augmented” or “virtual”, augmented reality mentions (135,319) are more frequent than virtual reality mentions (80,540), illustrated below.

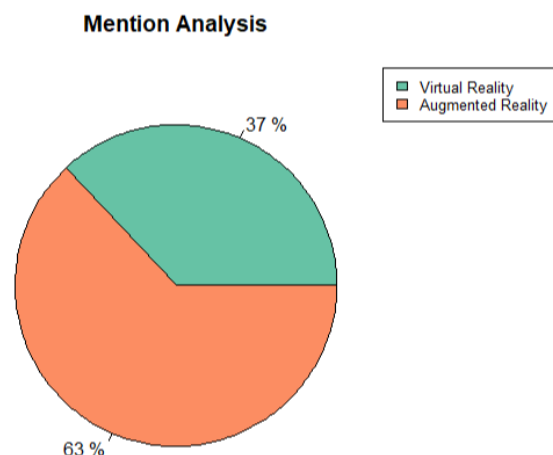


Figure 2 Percentage of "Augmented" versus "Virtual" mention in tweets

This first finding seems to contradict a 2018 survey by Gartner, where survey recipients responded towards virtual-reality apps (20%) more than augmented-reality apps (14%), in terms of apps that will have most impact on business success by year-end 2020 (shown in Figure 3). Although, the Gartner report analysts stated, “It is, however, surprising to see VR apps identified as the second most impactful type of multiexperience app (20%), as AR has more potential use cases and device support. But only 14% of the respondents thought that AR apps would be the most impactful, despite AR app development tools being more widely available. “[15]. The following chart shows where virtual-reality apps were more frequent in survey recipient response compared to augmented-reality apps for business success.

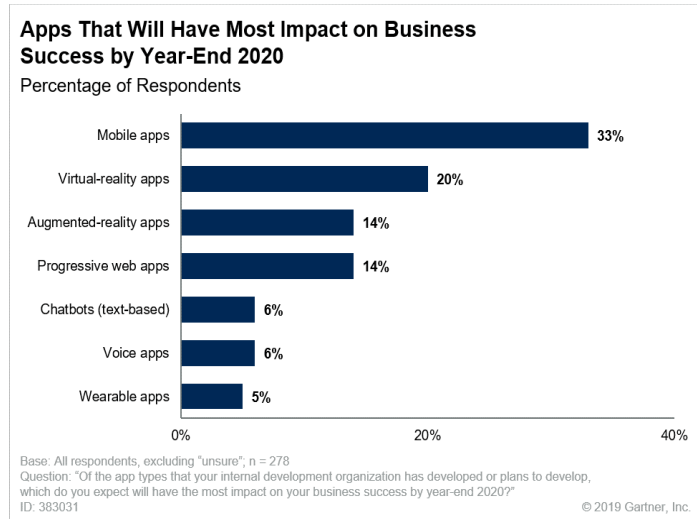


Figure 3 Gartner: Apps expected to have most impact on business success by year-end 2020

### 3.1. Most Favored and Most Retweeted Tweets about Augmented Reality

Tweets can be liked, or favorited, as well as re-tweeted. These mechanisms, favorite and retweet, are utilized to propagate, share, show appreciation of other user posting, and the number of favorites indicate the count of unique user accounts that like or agree with the content of the tweet [16]. Tweets frequently have meaning, and in this research, the gathered tweets are towards specific technologies.

Within scope of this research, the most favorited tweet is towards AR hardware, wearable glasses, and contains embedded video with a link to a Kickstarter site to raise investment through crowdfunding (tweet image found in Figure 4). The glasses, by TiltFive, are marketed to be the future of tabletop games. A new take on what appears similar to traditional board games. The digital approach gives the characters dimension, allows sharing the digital and augmented experience with other players, as well as the ability to save the game. These AR application quality characteristics are mentioned and defined within the quality model as presence, perspective and persistence. Further evaluation in section four applies this and subsequent text analysis to the previously established quality model.

Tilt Five @tiltfive · Oct 27, 2019  
 Tilt Five Augmented Reality glasses free your games from the flat screen.  
 Play holographic games solo or with your friends!  
 Only a few hours left to get discounts and free games - [kickstarter.com/projects/tiltfive](https://kickstarter.com/projects/tiltfive)  
 👍👍👍👍👍👍👍👍👍👍👍👍  
 #AR #VR #crowdfunding



Figure 4 Most favored tweet within research scope

The following tweet, posted in 2017, had at the time of this research been viewed over 212,100 times and was re-tweeted 4,247 times. The tweet embeds demonstration video to augment the physical environment and makes this statement, “This is ARKit’s A-ha moment”. ARKit is a development technology for building AR applications. Apple released ARKit with iOS 11 in 2017 [17]. More development tool analysis is found in section 4.5. Within the scope of this research, Figure 5 illustrates the most re-tweeted tweet.



Figure 5 Most re-tweeted tweet within the scope of this research



### 3.2. Mention, Sentiment and Time Series Analysis

Mention analysis identifies the frequency text occurrence. Research shows augmented reality applications being utilized across many industries; mention analysis from social media provides an additional perspective. Figure 6 indicates that Education, Entertainment and Commerce are the top three mentioned industries. This is determined by first creating a list of known industries from a source [18]. Then the algorithm reads the tweet texts and will count industry mention, including if multiple industries are mentioned in a single tweet.

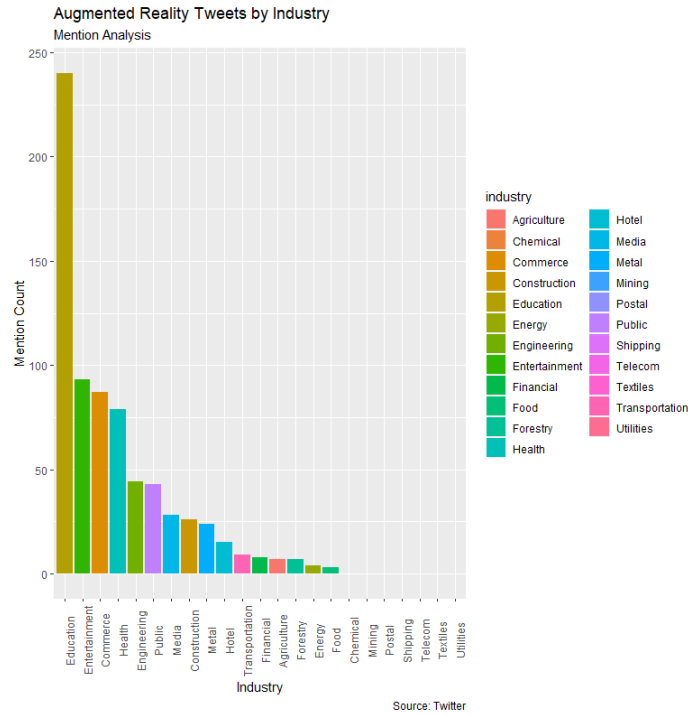


Figure 6 Industry Mention Analysis: Education, Entertainment and Commerce are top three mentioned industries

Further evaluation of the top mentioned industry, education, utilizing a time series chart seen in Figure 7. This chart indicates a slight positive trend in the smoothed line and a peak of tweets in November of 2019.

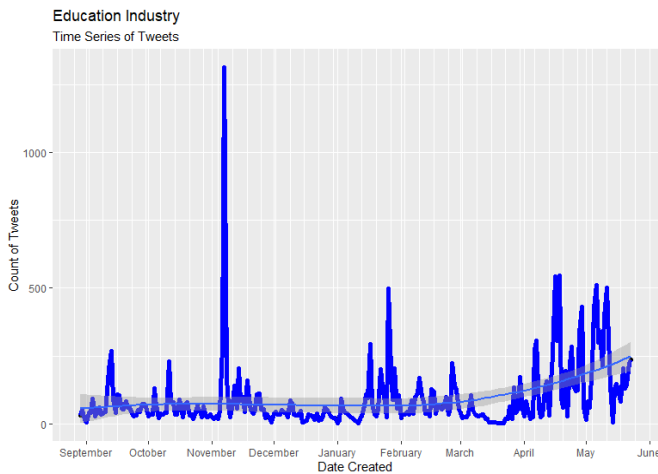


Figure 7 Time series chart of tweets towards Education industry

Figure 8 is a tweet containing embedded AR demo video of a skeletal dinosaur walking on a physical plane as an educational and awareness instrument. This tweet was retweeted over 1,000 times and was the largest contributing factor to the November peak of tweets having mention of education industry.

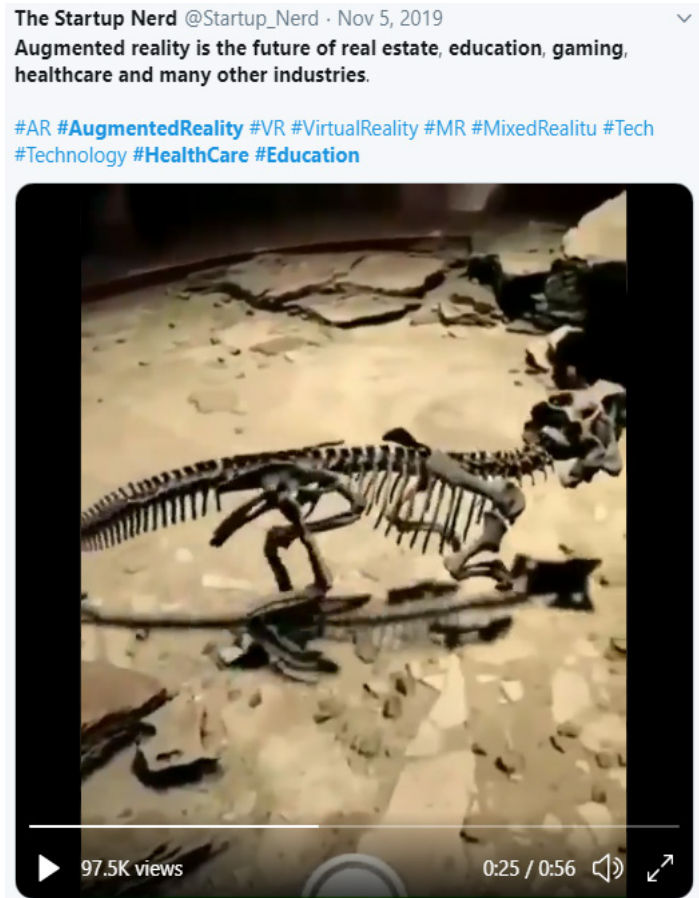


Figure 8 Highly favored tweet was large contribution to the November 2019 peak of Education industry tweets

Tweets are short in character length but offer volume of opinion which can be mined as user-driven data and classified in terms of sentiment. Opinion mining and sentiment analysis techniques can be utilized to support or drive communication plans [19]. Sentiment analysis has also been combined with other analytic methods to produce stock market prediction models with accuracy greater than 60% [20]. Tweets are opinionated text. Opinionated text is the primary resource for shoppers when making purchases [21]. We can utilize user-driven data for measuring system quality. Indeed, the quality of the system is also important in procurement, investment and communication planning.

The highest average sentiment score within the scope of this research is Transportation. Figure 9 illustrates average sentiment scores by the aggregated tweets within industry categories. A tweet which references multiple industries would have a sentiment score towards each referenced industry.

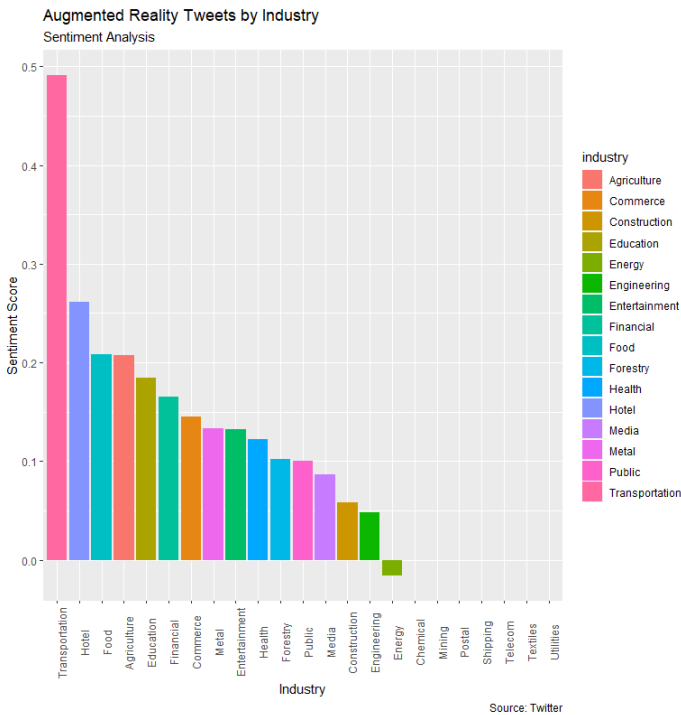


Figure 9 Sentiment Scores by Aggregated Tweet Sentiment Grouped by Industry

Further analysis of Transportation industry tweets indicates a spike in March of 2020. Figure 10 is a chart indicating the slope of the time series data as well as the aforementioned spike.

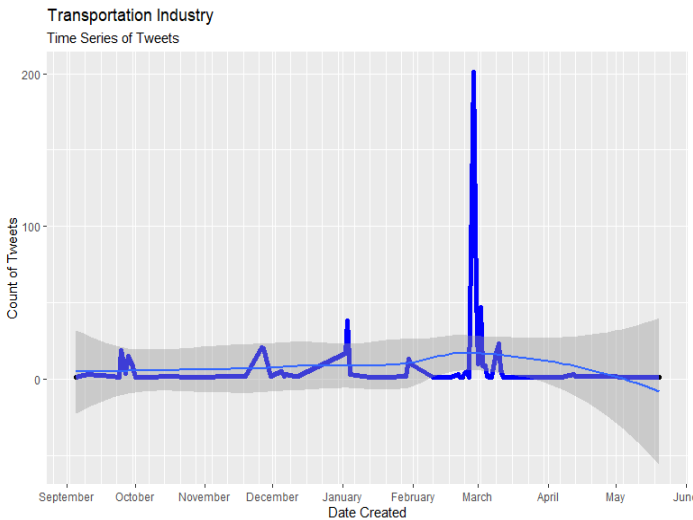


Figure 10 Time series chart of tweets towards the Transportation industry

The tweet found in Figure 11 was created near the end February of 2020 and had hundreds of retweets. This tweet is the main contributor towards the March spike of tweets categorized into the Transportation industry. The tweet references cost optimization of warehouse processing time by helping warehouse employees more quickly identify product location. Beyond cost optimization, the tweet states that AR technology can also help reduce possibility of human error within logistics.



Figure 11 Highly Leveraged Tweet and Main Contributor to the March Spike of Transportation Industry Tweets

The Sentimentr library is utilized within the R code of this research to provide sentiment scores of the collected tweets. The sentimentr library utilizes valence shifters to apply more meaning and accuracy to sentiment, such as negator words (e.g. I do not like it) [22]. The highest sentiment scored tweet, score of 1.540922, is found in Figure 12 (sentimentr library scores range between -2, 2). The tweet references an easier and more fun life due to augmented reality technology and utilizes the hashtag #CX (customer experience). While application users may be customers, they may also be internal employees. This research refers towards user experience as a more generalized approach.



Figure 12 Tweet within research scope having highest sentiment score

#### 4. Application of Analysis to the Established AR Quality Model

Augmented reality applications are novel and require levels of quality to be established to be effective for the users experience. Establishing the quality of an AR application goes beyond traditional testing practices and standards. A new model has been



established for test case focus and categorization, to build system confidence and inform system readiness and release decisions. These are critical decisions when dealing with novel technology.

Gartner's 2020 Digital Friction Survey found that among a survey n size of 4,582 employees surveyed, only 31% experienced High-Quality UX that was productive, empowering and easy [23]. Why might system's need to have this level of quality? The same survey found that "employees with a high-quality UX are 1.8 times more likely to have a high intent to stay as indicated by a lack of interest in leaving their current employer or actively search for a new job." The Gartner research report further explained, "Employees with a high-quality UX are 1.5 times more likely to have high discretionary effort, which includes employees' willingness to help colleagues, take on additional responsibilities, put in extra effort and find better ways of working."

Augmented reality applications are a novel approach to a system's user experience. The user experience has opportunity to benefit users and organizations alike. The technology implementation requires quality for these benefits to be realized.

#### *4.1. Contextual Approaches to Quality*

The characteristics of the quality model are quickly identified in the following bulleted text. Further sections then develop the characteristics and apply the findings of the twitter text analysis to the previously established quality model.

- Presence
- Perspective
- Interaction
- Portability
- Persistence

#### *4.2. Twitter Analysis Applied to the Characteristic of Presence*

Accurately executed occlusion and collision behavior will make the AR application more realistic [1]. Together they form the characteristic of presence. Presence allows AR application users to experience both the physical and digital worlds accurately. Presence may be exemplified digital models obeying gravity or shadowing as physical objects would which meets the users' expectations. Presence must be tested for. Figure 8, a digital model of a dinosaur overlaid onto a physical landscape (the original tweet is a video of the skeletal dinosaur walking) maintains the existence of shadow beneath the digital model, an example of presence in AR.

According to Gartner, "realism, as well as usability, will be enhanced or undermined by the digital model's ability to have accurate presence within a physical landscape. There should therefore be a test for occlusion which occurs when a far object hides a near object, for example, and which also sometimes occurs when objects are floating above and closer than an original anchor point. Test for occlusion by placing multiple models or scenes and moving throughout the physical environment." [24]. Collision in AR applications exists when a digital object and a physical object appear to occupy the same physical place at the same time. Of the 211,269 tweets related to augmented and virtual reality in this research, only 70 tweets reference occlusion or collision. Average

sentiment for an AR tweet is .1176; however, for tweets mentioning collision the sentiment turns negative at -0.0157, and ventures more negative when mentioning occlusion -0.0475.

#### *4.3. Twitter Analysis Applied to the Characteristic of Perspective*

Perspective frequently refers to someone's point of view. In an augmented reality, perspective must encompass the physical and digital objects, combined behavior, and how well visual distinction exists among them.

Perspective may be affected by the physical environment, the digital model's attributes, and the interaction between them. This would include the inability to scale a digital model to the correct size when embedded into the physical view. Perspective fails when rendering white font in annotations or digital models given a physical environment containing bright light or predominately white backgrounds. Perspective within AR applications should provide visual distinction and improve the quality of the user experience.

Figure 12, the tweet within the scope of this research having the highest sentiment score, adequately illustrates the characteristic of perspective. This example uses a neon green annotation text color over a mostly grayscale physical scene. Quickly drawing attention to a physical device through digital annotation.

#### *4.4. Twitter Analysis Applied to the Characteristic of Interaction*

The characteristic of interaction is directly related to performance. Interaction directly affects the device, the costs, and the execution environment. The tweet text found in Figure 11 refers to AR applications used to model warehouses which may assist in finding specific freight. Warehouses can be quite large, requiring longer application runtime, as well as the rendering of many possible digital directional annotations before arriving at the correct location. If the AR application is utilizing commercial grade devices, we must concern how the interaction of such an application will consume WIFI network, the battery of the device, and whether other applications will run on the device at the same time. The greater the interaction, the more important the performance.

#### *4.5. Twitter Analysis Applied to the Characteristic of Portability*

There are many software development kits available to build AR applications. Table 2 references many examples of these technologies, and the runtime platforms they support. Not every AR application will run on your commercial or consumer system. It is not uncommon for some popular mobile AR apps to only run on iOS or Android, as example operating systems [25]. The value and availability of your application will depend upon the portability characteristic, which is a quality concern that must be considered and established during design and development.

Table 1 provides search terms that were utilized to scope the collection of tweets for this research. Many of those search terms are twitter handles or the names of SDK technologies. The top three mentioned development vendors are Apple (966), Unity (713), and Google (601). The popularity found in the mention analysis is not conducive with portability, as Apple is one of the more limited SDKs when concerned with runtime environment support as indicated within Table 2.

Table 2 Augmented reality development tools and supported runtime platform/s

Tool name	Supported Runtime Platform/s
Vuforia Studio [26]	Android, iOS, UWP, Unity
Spark AR [27]	Android, iOS
ARKit [28]	iOS
ARCore [29]	Android, iOS, Unity, Unreal
Wikitude [30]	Android, iOS, Microsoft Tablet
Maxst AR [31]	Android, iOS, Unity

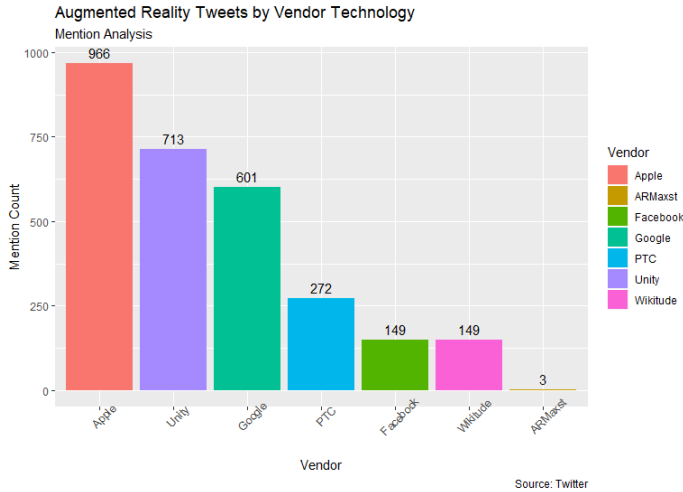


Figure 13 Example augmented reality development technology ordered by mention count

#### 4.6. Twitter Analysis Applied to the Characteristic of Persistence

Scaling augmented reality experiences across multiple users requires centralizing physical maps and virtual object anchors so that multiple users can share immersion [1]. Persistence allows an AR user experience to be shared among other application users from separate devices or application instances during runtime. Additionally, persistence allows the AR experience to be saved and resumed later by the original, or potentially other, application users.

Figure 4 is the most favored tweet within the scope of this research, and it adequately exemplifies many of the characteristics of this quality model. The tweet references the future of tabletop games, a physical board game with digital models rendered above. Persistence of the digital models allows multiplayer action in the game, so that competing players may see and experience the avatars and digital scenes that are being controlled by other’s play. It is possible that such a use case could leverage whichever phone device or headset might be handy to a player, which would encompass the portability characteristic. If the game is to be played for many hours, the interaction level will be high, perhaps enough to warrant recharging a device battery, or the requirement to plug an AR hardware device into a power source. Perspective is required so that the digital scenes are relevant to the location, size and position of the physical game board as seen by other’s points of view. Presence of the digital models, the prevention of

collision, will also provide an experience worth dedicating many hours towards.

While the previously mentioned and exemplified tweet represents the entertainment industry, certainly implementation across healthcare, manufacturing, real estate, education and many other industries would prioritize quality potential during investment in novel technology. Hence the AR quality model, which provides a basis for test case creation and categorization that promotes confidence in the application’s readiness and informs decisions towards the application’s release.

## 5. Conclusions and Future Works

### 5.1. Conclusions

The findings reported in this research study represent essential considerations for the investment and realization of value through augmented reality applications. Augmented reality applications have characteristics that are beyond traditional interfaces. Despite a novel technology, quality still matters. For these reasons, the AR quality model has been established and extended with input from social media and in some cases exemplified through such uses.

Among all tweets that mention either “augmented” or “virtual”, augmented reality mentions (135,319) are much more frequent than mentions of virtual reality (80,540). Top mentioned industries within the tweet collection scope were Education, Entertainment and Commerce. Top average tweet sentiment scores by mentioned industries within the study scope were Transportation, Hotel and Food.

AR quality model characteristics of Presence, Perspective, Interaction, Persistence and Portability all had representation among some of the highest retweeted or liked tweets, or among the tweets that drove industry spikes in the time series charts.

Still, limitations exist. Only one social media platform was utilized in this research, Twitter. Facebook, another such social media platform, is also a technology provider of AR software development. What information then may be more prevalent on their social platform? The AR quality model requires further research beyond social media and into specific teams and implementations within industry. Research towards failed AR implementations would also support the model or indicate where omissions of key characteristics exist.

### 5.2. Future Work

Future work on this quality model for augmented reality applications must extend into cyber security. 475 of the tweets within this research mentioned terms towards privacy or security. That is less than one quarter of one percent of tweets within the scope of this research. While the percentage is small, it is critical to consider augmented reality applications require access to the device camera. The device camera may capture personal or organizational information, potentially confidential or private in nature, that may become vulnerable to threat actors if exploited. Future work should engage in various threat and vulnerability research of AR applications and methods and technology which may help mitigate such risk or exposure.

Future works should also include case studies from many of the named industries within the paper with indication towards

whether the AR quality model adequately supports the users experience of if in fact key characteristics have been omitted, such as current lack of the security characteristic.

### Conflict of Interest

The authors declare no conflict of interest.

### Acknowledgment

The authors would like to appreciate Tyler Rinker, who has authored the sentimentr package and supported a question regarding sentiment scores during this research.

### References

- [1] J. Scheibmeir and Y. K. Malaiya, "Quality Model for Testing Augmented Reality Applications," 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York City, NY, USA, 0219-0226, 2019. doi: 10.1109/UEMCON47517.2019.8992974.
- [2] I. Rabbi, S. Ullah, A survey on augmented reality challenges and tracking, *Acta Graphica: J. Print. Sci. Graphic Commun*, 24(1-2), 2013, pp. 29-46, 2013.
- [3] X. Wang, A. Kotranza, J. Quarles, B. Lok, B.D. Allen, A pipeline for rapidly incorporating real objects into a mixed environment, in *The 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*, Washington, DC, USA, October 2005, 170-173, 2005.
- [4] D. Cearly, et al. "Top 10 Strategic Technology Trends for 2020: Multiexperience." Gartner, 10 Mar. 2020. [www.gartner.com/doc/3981950](http://www.gartner.com/doc/3981950).
- [5] S.P. Yun, H.J. Yap, R. Singh, S.W. Chang, K.L.R. Cheong, Augmented reality assisted factory layout planning and analysis for a flexible manufacturing cell, in *International Conference on Computer Science and Computational Mathematics*, 106-111, 2014.
- [6] S. J. Henderson, S. Feiner, Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret, in *8th IEEE International Symposium on Mixed and Augmented Reality*, Orlando, FL, USA, 135-144, 2009.
- [7] G. Westerfield, A. Mitrovic, M. Billingham, Intelligent augmented reality training for motherboard assembly, *Int. J. Artif. Intell. Educ.* 25(1), 157-172, 2015.
- [8] "ISO/IEC 25010." ISO 25000. Accessed May 22, 2020. <https://iso25000.com/index.php/en/iso-25000-standards/iso-25010>.
- [9] A.Dünser and M Billingham "Evaluating augmented reality systems," In: Furht B. (eds) *Handbook of Augmented Reality*. Springer, New York, NY, 289-307, 2011.
- [10] A. Sutcliffe and K. Kaur, "Evaluating the usability of virtual reality user interfaces," *Behaviour and Information Technology*, vol. 19, 2001.
- [11] J. Scheibmeir and Y. Malaiya, "An API Development Model for Digital Twins," 2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C), Sofia, Bulgaria, 2019, pp. 518-519, 2019. doi: 10.1109/QRS-C.2019.00103.
- [12] J. Scheibmeir, T. Murphy, and J. Herschmann. "Promote Continuous Quality for APIs to Support Digital Business." Gartner, February 21, 2019. <https://www.gartner.com/doc/3902567>.
- [13] A. Cirillo, *RStudio for R Statistical Computing Cookbook*. Birmingham (Gran Bretaña): Packt Publishing, 2016.
- [14] "How to Tweet." Twitter. Accessed May 23, 2020. <https://help.twitter.com/en/using-twitter/how-to-tweet>.
- [15] J. Wong, A. Leow, Gartner Survey Analysis: Insights to Kick-Start an Enterprise Multiexperience Development Strategy, 12 February 2019. <https://www.gartner.com/document/3901479>
- [16] R.S. Perdana, A. Pinandito, "Combining Likes-Retweet Analysis and Naive Bayes Classifier within Twitter for Sentiment Analysis", *Journal of Telecommunication, Electronic and Computer Engineering*, 10(1-8), 41-46, 2018.
- [17] Apple's ARKit: Cheat sheet, Cory Bohon, TechRepublic, June 7, 2019. Accessed September 28, 2019. <https://www.techrepublic.com/article/apples-arkit-everything-the-pros-need-to-know/>
- [18] International Labor Organization. "Industries and Sectors." Industries and sectors, n.d. <https://www.ilo.org/global/industries-and-sectors/lang-en/index.htm>.
- [19] H. Purohit, A. Hampton, V.L. Shalin, A.P. Sheth, J. Flach, S. Bhatt. What kind of #conversation is Twitter? Mining #psycholinguistic cues for emergency coordination *Computers in Human Behavior*, 29 (6), 2438-2447, 2013. DOI: 10.1016/j.chb.2013.05.007
- [20] Y. Tanulia, A.S. Girsang "Sentiment Analysis on Twitter for Predicting Stock Exchange Movement", *Advances in Science, Technology and Engineering Systems Journal*, 4(3), 244-250 2019.
- [21] C. Hung, "Word of Mouth Quality Classification Based on Contextual Sentiment Lexicons." *Information Processing & Management* 53(4), 51-63, 2013.
- [22] T. Rinker, "Sentimentr." README, n.d. <https://cran.r-project.org/web/packages/sentimentr/readme/README.html>
- [23] Application Research Team. "Proven Design Principles to Deliver a High-Value Employee User Experience." Gartner, May 12, 2020. <https://www.gartner.com/doc/3985028>.
- [24] J. Scheibmeir, J. Herschmann, T. Murphy, "Quality Is the Key to Avoiding 'Digital Distortion' With Your Augmented Reality Strategy." Gartner, April 16, 2020. <https://www.gartner.com/doc/3983581>.
- [25] M. Jansen, "Escape Reality with the Best Augmented Reality Apps for Android and IOS." *Digital Trends*. Digital Trends, June 9, 2019. <https://www.digitaltrends.com/mobile/best-augmented-reality-apps/>.
- [26] Use Vuforia for computer vision and augmented reality apps, 2019, <https://www.ptc.com/en/products/augmented-reality/vuforia>
- [27] Downloads, 2019. <https://developers.facebook.com/docs/arstudio/downloads/>
- [28] Get ready for the latest advances in augmented reality, 2019. <https://developer.apple.com/augmented-reality/>
- [29] ARCore overview, February 28, 2019. <https://developers.google.com/ar/discover/>
- [30] Wikitude Augmented Reality SDK, 2019, <https://www.wikitude.com/products/wikitude-sdk/>
- [31] AR SDK, 2019, Maxst. "Setup Guide." 3.4.x | MAXST Developer Site, n.d. <https://developer.maxst.com/MD/doc/unity/setup>.

## Cause of Time and Cost Overruns in the Construction Project in Nepal

Bishnu Prasad Khanal\*, Sateesh Kumar Ojha

Lincoln University College, Mayang Plaza, Taman Mayang Jaya, Petaling Jaya, 47301, Malaysia

### ARTICLE INFO

Article history:

Received: 14 January, 2020

Accepted: 27 June, 2020

Online: 22 July, 2020

Keywords:

Construction

Contractor

Bid

Nepal

Time Overrun

Cost Overrun

### ABSTRACT

*There are many works in construction in Nepal. To build the country, there must be different buildings, roads, bridges, etc. with rapid flights. But here it seems to be slowly delayed for various reasons. Many people have studied in this regard. But no one seems to have thought about reducing it. Therefore, I have prepared this article by studying the main reasons for construction delay and cost overruns and looking at how this cause can be reduced. This will enhance the delivery of project within the set time frame. Based on the literature available and their study the deviation of time taken, and cost incurred by the projects than estimated were examined. Construction projects such as road transport and highways, buildings, Hydropower were found to deviate more. Commitments to finish the projects in time was found as primary reason caused by easily postponing practices by the authority on the request of concerned contractors, was the reason for time overrun whereas capital cost rise due to price escalation was the major reason for cost overrun. This means exceptional provisions of laws were seen followed easily, actually they were supposed to use in extraordinary situation. Similarly planned schedules were found distorted from the beginning due to lack of seriousness to start timely due to budgetary provisions and adjustments.*

## 1. Background

The development business undertakes a significant job in doing completely created country status. To complete this a country needs to keep its development industry sound by acceptance most recent innovations and dealing with the activities appropriately. Finishing projects on time is a pointer of proficient development industry. Truth be told, a task is considered 'fruitful' on the off chance that it is finished on schedule, inside spending plan and to the predefined quality. Regularly, when the activities are delayed, they are either expanded or quickened and hence, obtain extra expense. To the dislike of proprietors, contractual workers and experts' numerous tasks experience broad postponements and along these lines exceed beginning time and cost measures. The advancement methodology is at risk to various variables and eccentric segments. Conveying a task on time does not happen by trusting that the required fulfillment date will be met. To design and deal with an effective task, the three parameters of time, cost and quality ought to be considered. The customers in the development business are fundamentally worried about quality, time and cost. However, larger part of development activities is secured based on the requirements time and cost. Cost acceleration and time overwhelms are regularly connected with

poor administration trains. In Nepal the development business is the most remarkable supporter of the financial development. The greatest customer of the development business is government open works projects. These projects are huge for nearby network and must be conveyed on schedule. Any deferral in conveyance won't encourage the full motivation behind the tasks' implementation a work in progress plan. Among all acquirement strategies the Design and Build (D&B) projects is noteworthy in executing improvement extends in Nepal and different nations.

The undertaking directors and the group need to comprehend the elements causing these postponements and cost overwhelm prompting project disappointments. The quantitative examination of poll review will rundown down the hazard causing factors by positioning. This will help the task administrators and the specialists to make proactive move to such disappointment circumstances. Next area manages time and cost related writing audit pursued by research procedure, exchange and examination by SPSS lastly the ends.

## 2. Statement of problems

There are many works in construction in Nepal. To build the country, there must be different buildings, roads, bridges, etc. with

\*Corresponding Author: Bishnu Prasad Khanal, Kathmandu, Nepal, [bpkhanal123@gmail.com](mailto:bpkhanal123@gmail.com)

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050423>



rapid flights. But here it seems to be slowly delayed for various reasons. Many people have studied in this regard. But no one seems to have thought about reducing it. Therefore, I have prepared this research proposal prepared by studying the main reasons of construction time and cost overrun and looking at how this can be reduced.

Hence urgently needed to study Cause and impact of delay completion of different construction project in Nepal. I have seen this study very important to complete the following tasks.

- Find out the main reason why the construction work is delayed.
- To speed up the construction work quickly, no reason to come.
- Addressing the elements of delay in construction and addressing timely manner.

In addition, there are numerous hazard factors inflicting cost within the Construction Project, afterward, analysts tend to bring this issue into to boot examine with the purpose of characteristic the elements and their positioning of significance. Within the current investigation, scientists have tried to seek out the needs for the underlying evaluated and last add, and completely different reasons for time overcome. In [1] it is discovered four key factors that were the little question affected on cost. These variables incorporate define changes, lack of wisdom, capricious climate condition and also the uncertainty in prices of building materials. it's important to possess a high to the underside understanding of the exceptional problems within the Construction Project. Consequently, the paper expects to offer high to bottom light-weight regarding this reason for project deferral and price overrun in each created and making nations, and framework the attainable proposals for dominant enterprise postponement and price overrun in future activities by dissecting the suitable investigations from numerous nations and Nepal.

Delays on Construction projects are a complete miracle. They are frequently joined by cost and time survives. Development venture postpones have a horrible effect on social occasions (owner, impermanent specialist, and master) to an understanding in regards to a Construction in opposing associations, question, suit, intercession, pay issues, and a general opinion of qualm towards one another [2].

Every project must be in comply with the cost and time estimated and the reasons of delay are to be spotted to assure the project efficiency with respect to time and costs. We can discover an excessive number of research articles about the reason for Construction delay in Nepal. In any case, we can't locate the real reason to postpone top to bottom and we are not the accomplishment to limit the accept Construction works. In Nepal. Henceforth earnestly expected to examine Cause and effect of delay inside and out the fulfillment of various Construction project in Nepal.

### **3. Significance of Study**

Development industry has been a bustling industry in the twentieth century. Immense huge numbers of tasks have jumped up, particularly since the Second World War. With the development rate going up continually, a savage rivalry has been set up among the manufacturers. A development project is

commonly a progression of exercises that have a particular goal to be finished inside specific determinations and the begin/end dates are all around characterized. The development projects are normally capital serious with a ton of obligation and premium parts and everyone needs to gain benefit on the speculation, as quickly as time permits. This has provoked the flood of quick track projects around the world.

The motivation behind this paper is to examine for all intents and purposes all viewpoints identified with postpone investigation on a development project, trailed by a contextual analysis. Despite the fact that the paper subtle elements for all intents and purposes each perspective related with the defer investigation, spotlight will be on the Lump Sum Contract from the perspective of a contractual worker, as it conveys the greatest hazard for a temporary worker.

As it is the agreement record which subtle elements the approaches if there should be an occurrence of postpone claims, the accompanying segment will examine different kinds of agreements and components for choosing the sort of agreement before formally tolerating it.

## **4. Methodology**

### *4.1 Data collection*

Information was collected from Construction project. It's time duration on, cost and Management process for Construction works, Procurement process, availability of Labor, materials, government policies, Decision making procedure etc, was studied. This study was carried out based on literature review, questionnaire survey, interview and case study. The research methodology was contained three phases. The first phase was included a literature review, second phase was questionnaire survey, interviews and third phase were carried out by case study. The writing audit was directed through books, meeting procedures, web and universal undertaking the executive's diaries. Questionnaire survey will include 75 respondents made up 10 Consulting firm, 20 Engineers, 15 Contractors, 10 Clients/ Owners, 20 Construction workers/employees. For Contextual investigation two task was chosen. Subsequently, will distinguish reasons for postponement and cost invade in each task independently increase better comprehension for future notice and proactive activities to be embraced in a future undertaking.

### *4.2 Literature Review*

The mission to facilitate by providing product and service timely and economically is bound with the constructions projects but they are deviated from estimated costs and time means deviated from, which is often practiced [3]. There is a need of in-depth investigation upon the deviations of constructions with respect to time, costs and quality as to minimize its frequent occurrence.

Author in [4] have identified the need of adequate planning to overcome the issues of delays in completion, over-costs and inferior quality of construction projects.

Researches on construction projects, like [1] found the extent of costs exceeded around 30 % in some developing countries. Similarly, [5] did mos detailed studies about these issues and found that 90% of the projects have cost overrun.



According to [6], behind Cost overruns were high transportation cost, change in material detail, and heightening of material cost, visit breakdown of development plants and equipment's and improve. Financial problems are the also major factor causing a delay in construction projects. Technical support is also necessary since the study shows that coordination problems are the second major factor causing delays in construction projects in Malaysia [7]. According to [8], Deferral additionally happens because of outer calculate like change government, guideline and area, inadequate of works and so on.

The mission of facilitating with good and services timely with reasonable costs is affected by the delays. Costly than estimated might have been common in Nepal too. Being developing country and research in this area is needed especially on these issues.

## **5. Finding out Causes of cost and Time overrun**

A number of methods applied to study these issues in the country. These methods include survey, observation, literature survey, interview with concerned authorities. These methods together have given insight about the reasons of delay, cost overruns, and quality is that few contractors are assigned many projects in favor. The easy reasons to prove justification for delay and cost overruns are climate change, inflation and natural calamities. All most all the projects have common reasons.

Other issues identified in the study are budget allocation, over ambitious contract duration to complete the project with the contractors merely to communication development efforts, faulty premises at the time of designing and starting the projects, weak monitoring and supervision system, lengthy litigation process to grant project approval, favoritism to the contractors based on the political and financial interests, power influence on the project rather than needs and cost benefits analysis, frequent changes in government and politically favored government officials, who works to please power rather than interest of the country and society. Similarly, studies have indicated some reasons like the interest of client group to earn money and take personal benefits, absent of awareness of common people towards what has been being constructed in their areas.

Some other issues are distributing the benefits of the construction to mass political cadres and relatives of political authority, donation of the contractors to run the elections, and win the political battle for which contractors are bound to raise their losses from minimizing the costs and compromising the quality of construction materials and mixtures of materials in a way to violate to standard and eleventh-hour pressure to complete the projects.

Construction delays can be grouped in many ways for their treatments. Some are not so damaging so they can be tolerated, and some are in tolerated to forgive. As the description they are excusable and inexcusable. When delays are there, some can be compensated because no one seems intentionally faulty but by act of god whereas some are caused due to negligence of human beings which cannot be compensated. Whether it is compensable or not also is subject of the terms and conditions of contracts between the contractors and proprietors or clients. Most of the faults and delays occurred due the reasons unspecified in the contracts are compensated. And some are termed as ongoing

project which can be improved if any faults found, for instance, schedule can be met with more additional employees, damaging parts of the works can be replaced, and so on. but some have been completed and handed over a in the past for which no other ways remain their rather than to accept it.

Whether a project is excusable or not is subject to reasonableness to excuse or not. For instance, if people involve in the projects make mistakes due to their inefficiency and ignorance then it is not excusable. But many times, it becomes the reason of beyond human ability to repair and maintain then it is called power of god.

The most effective methods of minimizing delay in construction projects are, sufficient financing all through the undertaking, able task administrator, granting offers to the correct originator/contractual worker, complete and appropriate structure at the opportune time, accessibility of assets, clear goal and degree, utilization of experienced subcontractors and providers [9]. The connection between progress nearby and 'solid' supervisory crews underlines the requirement for successful site the executives and oversight by temporary workers and consultants. Labor, at both the specialized and the administrative levels, ought to have their own insight refreshed by nonstop expert improvement plans [10].

In concurrent delays, it should be analyzed just one factors or many to make the project delay and costs. Necessary adjustments can be made through efforts and costs to mitigate the problems reasonably.

According to [11], Helpless site coordination has supposedly been one of the key disappointment boundaries for most Construction ventures. First characteristic under this factor, non-accessibility of drawing/structure on schedule. Second trait, slow end from proprietor is because of absence of legitimate coordination among proprietor and specialist or proprietor and temporary worker. Third property unworkable time plan given in contract is because of absence of coordination among customer and contractual worker about the pragmatic troubles at the site. Fourth quality helpless site the executives and oversight obviously feature the nonattendance of coordination between different legislative chains of command associated with development industry.

## **6. Conclusion**

Problems remain problems if thought efforts are not made to resolve the problems. Reasonably constructions projects are complicated problems inching lots of activities and tools and materials. More parts may call more complications so careful analysis and discussion are necessary to find the resolutions on delays, costs and quality. Construction projects have long impacts, huge investment and lengthy time consuming to complete so mistake also affects these three things. Timely resolving these issues certainly assure the people concerned that they are getting more facilities reasonably. The more expense, more time can be use in other new projects so fast development can be insured.

## **Reference**

- [1] R.K. Shah, "An Exploration of Causes for Delay and Cost Overruns in Construction Projects: Case Study of Australia, Malaysia and Ghana".

Journal of Advanced College of Engineering and Management, 2, 2016.  
<https://doi.org/10.3126/jacem.v2i0.16097>.

- [2] N.A. Ansari, R.M. Swamy, "Methodology to analyse Delay and its impact on construction project". International Research Journal of Engineering and Technology (IRJET), 5(2), 1823-1826, 2018. ISSN:2395-0056
- [3] A. Rauzana, "Analysis of Causes of Delay and Time Performance in Construction Projects" IOSR Journal of Mechanical and Civil Engineering, 13, 05, 2016. <https://doi.org/10.9790/1684-130503116121>
- [4] S.A. Hammadi, "Study of Delay Factors in Construction Projects". International Advanced Research Journal in Science, Engineering and Technology, 3(4), 2016. DOI 10.17148/IARJSET.2016.3420
- [5] S. Famiyeh, "Major causes of construction time and cost overruns". Business School, Ghana Institute of Management and Public Administration, Accra, Ghana, 2017. <https://doi.org/10.1108/JEDT-11-20150075>.
- [6] S.S. Mulla, "A Study of Factors Caused for Time & Cost Overruns in Construction Project & their Remedial Measures", 5(1), 48-53, 2015 [www.ijera.com](http://www.ijera.com).
- [7] W. Alaghbari, "The significant factors causing delay of building construction projects in Malaysia", 2005. <https://doi.org/10.1108/09699980710731308>.
- [8] L. Mali, A. Warudkar, "Analysis on Causes of Delay in Construction Industry in Pune City". International Journal of Innovative Research in Science, Engineering and Technology, 5(8), 14509-14519, 2016. <https://doi.org/10.15680/ijirset.2016.0508028>.
- [9] T. Pourrostan, A. Ismail, "Study of Methods for Minimizing Construction Delays: Evidences from a Developing Country". Advanced Material Research, 201-203, 2939-2942, 2011. <https://doi.org/10.4028/www.scientific.net/amr.201-203.2939>.
- [10] D.W.M Chan., M.M. Kumaraswamy, "A comparative study of causes of time overruns in Hong Kong construction projects". Internal Journal of Project Management, 15(1), 55-63, 1997. [https://doi.org/10.1016/s02637863\(96\)00039-7](https://doi.org/10.1016/s02637863(96)00039-7).
- [11] H. Doloi, A. Sawhney, K.C. Iyer, S. Rentala, "Analysing factors affecting delays in Indian construction projects". International Journal of Project Management, 30(2012), 479-489, 2012. <https://doi.org/10.1016/j.ijproman.2011.10.004>.

## High-Order Thinking Skills: The Educational Treasure Hunt Game

Yogi Udjaja<sup>1,\*</sup>, Sasmoko<sup>2</sup>, Jurike V. Moniaga<sup>1</sup>, Kevin Zulfian Bay<sup>1</sup>

<sup>1</sup>Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, 11480, Indonesia

<sup>2</sup>Primary Teacher Education Department, Faculty of Humanities, Bina Nusantara University, Jakarta, 11480, Indonesia

---

### ARTICLE INFO

Article history:

Received: 27 March, 2020

Accepted: 12 May, 2020

Online: 22 July, 2020

---

Keywords:

HOTS Game

Educational Video Games

Treasure Hunt

Puzzle

Game Development

---

---

### ABSTRACT

Many people who have no experience want to make Games. Among them are people who only like to play Games, people who want to make business through Games, and there are also people who really want to make Games. Assume making the game is very difficult and confusing, but forget the many examples that already exist, and can evolve if modified only slightly. In this case, there are several genres of games that are quite popular, which can be changed and modified into another sub-genre. One of them is the Treasure Hunt. It is one of the well-known genres in various video games and has a straight and simple design that is also easy to understand. Which is used in a game that aims to "Find and Work".

---

### 1. Introduction

Lack of perseverance and one's desire when pushed to work on a problem is very common among people. Especially if the individual is required to work on mathematical-themed problems. Indeed, when you're relaxed, people are lazy to do something, especially doing something that in general, or almost everyone, hates it.

There are several methods that can be done to encourage someone's desire or willingness to do and work on these problems. And some of them have their own different interests. A very effective way, at least in this digital age.

Which entertainment and media can be easily obtained and enjoyed only with a smartphone and one of them is video games, which are almost from all walks of life, and at least have played, really enjoyed video games and a genre that is very popular among young and old alike, namely A Series of Clues. Or rather Puzzle and Treasure Hunt [1].

And There is High Order Thinking Skills (HOTS). The game that has both Treasure Hunt and an Adventure Puzzle genre that aims to represent the 3T regions (Leading, Outermost, and

Disadvantaged) in Indonesia and train the ability to think through puzzle games [2].

HOTS Game was developed for mobile devices with the Android platform which will be easily accessible to most people [3].

HOTS Game is a game with adventure and puzzles in it. The puzzles in the game are math problems that contain some Creativity, Critical Thinking, and Problem Solving. The target players of the game are for all people but the main goal is elementary school children with the aim of training complex thinking skills. Not only that, but also aimed at encouraging the strength of one's thinking and one's skills at a large age.

Follow the story line starting from one point to another point. Get a few clues to interpret, then examine it for later to be used as a pointer. Look for a series of continuous problems. which are hidden somewhere in areas located in Indonesia.

Accompanied by music and an atmosphere that focuses on important points in the regions of Indonesia, to add to the impression of culture and nationality.

By covering various aspects given from most Video Games, such as Atmosphere, Music, Storyline, puzzles, etc. Where will it add to one's desire to get things done. Completion or game

---

\*Yogi Udjaja, Jakarta, Indonesia 11480, +62 898 268 3399, udjaja.yogi@gmail.com, yogi.udjaja@binus.ac.id

observation where completing all missions, getting all achievements and collectible is also the biggest motivation that is most liked. There is also Fantasy, where they prefer games where they can feel like someone else in another place. And the third is Design [4].

## 2. Related Research

### 2.1. Video Game

Is a type of play activity, which is carried out in the context of mock realities, where participants try to achieve predetermined goals and mediate in the form of digital media [5]. To play a game, one must play it on the right platform. Platform games can be categorized into three types: console, cellular, and cross-platform. Each platform has different SDK distribution characteristics and methods. The console SDK has a closed distribution method, while most mobile SDKs can be downloaded freely, although with some constraints. That is why the development of mobile games is becoming increasingly popular [6].

### 2.2. Treasure Hunt

A treasure hunt type game that uses positioning. Players are given clues or directions to proceed along with predetermined treasure hunt routes based on their location as determined by the clues. Based on these clues or directions, players change their locations. Based on these changed locations, as determined by the GPS, additional clues or directions are given to the players until one of the players completes the treasure hunt routes and arrives at the treasure to win the game. The game also incorporates other variables in determining the clues to provide to players, such as the location of other players as determined by their GPS devices, the previous locations of the players and other players and whether the player has correctly solved certain clues [7].

### 2.3. User Interface

An Interface Experience is information that is intended between the Player and the Device he uses. In Video Games, the Interface takes the form of a Controller, and much more to the Graphical User Interface.

As the name implies, the GUI contains Graphics-Visual information. With the lack of physical form and pleasantness, relying on Display to send information to players. The quality of the UI can be assessed from two aspects [8, 9]:

1. Usability: How well it facilitates the intended purpose of the tool, and how easy it is for users to absorb and operate.
2. Consistency: performing the same actions several times will produce the same results. However, that does not mean that consistently bad results are good.

### 2.4. User Experience

Different, but directly related to UI, User Experience concerns the quality of the user's time with the product [3, 10- 12], which can certainly affect the UI. Where the UI is about how well

a tool can convey the information it needs to users, UX is about how fun it is to use it. The interface can be very bad in providing all important information, but users might still consider it a good tool, for several reasons such as ease of use and aesthetically pleasing design.

## 3. Research Method

Like the Game Development Life Cycle [13-15], In the Initiation Phase, some problems have been thought by people to be lacking in the desire to do something. So, made a game that is intended to encourage the spirit of someone to do it. Which at least makes the individual want to try.

In the next Phase, Game System has been analyzed, and made a design. How this game will run and focus to help players increase their desires and also made in writing.

In the production phase, the Designer makes some basic assets needed to just try. Creating several data sets as test data such as character names and scores. Make a number of basic programming sets to focus the player looking for a hidden problem [16].

In the alpha testing phase, adjustments are made to the Design Document about the game. And do some bug testing and how the game runs. Collect some basic player data such as username and score and evaluate it.

Then in the Beta Testing phase, a number of surveys are conducted to the players which are focused on how the players react to how to find a set of puzzles and do the questions that have been made. A written test is also carried out.

Some feedback from players who have been tested, mostly including the difficulty of the questions presented. But this problem is ignored, because this game is intended to give ordinary people who are in the public area, a problem with a fairly high level

```
void Start()
{
    StageStatus = PlayerPrefs.GetInt("levellocation");
}

// Update is called once per frame
0 references
void Update()
{
    StageStatus = PlayerPrefs.GetInt("levellocation");

    if(StageStatus <= 9){
        Stage1.SetActive(true);
        Stage2.SetActive(false);
        Stage3.SetActive(false);
    }else if(StageStatus >= 10 && StageStatus <= 19){
        Stage1.SetActive(false);
        Stage2.SetActive(true);
        Stage3.SetActive(false);
    }else if(StageStatus >= 20){
        Stage1.SetActive(false);
        Stage2.SetActive(false);
        Stage3.SetActive(true);
    }else{};
}
```

Figure 1: Unlock Method



Almost all the methods used are basic and have logic that is easy to apply to anyone. Although many seem to be able to give a faster or more efficient effect, but here are some methods that use a little extra that can add effectiveness.

In the last phase, the release was being held for some times, then proceeded to release it via Google Play Store and also still do some repairs and maintenance at time.

### 3.1. Unlocks

In this game, many use the method that is locked and must be obtained. Use something simple, just by validating the status of the player's progression, and making it active.

### 3.2. Transition

Transition from one view to another, using the SceneManager which is divided into several functions and methods. And will be called in each particular situation.

```
0 references
public void LoadKalimantan()
{
    SceneManager.LoadScene("Stage Kalimantan");
}

0 references
public void LoadSulawesi()
{
    SceneManager.LoadScene("Stage Sulawesi");
}

0 references
public void LoadPapua()
{
    SceneManager.LoadScene("Stage Papua");
}
```

Figure 2: Loading Stages

```
void Start()
{
    Username = PlayerPrefs.GetString("username");
}

0 references
void Update()
{
    if (string.IsNullOrEmpty(Username)){
        Login.SetActive(true);
        Logoff.SetActive(false);
    }else{
        Login.SetActive(false);
        Logoff.SetActive(true);
    }
}
```

Figure 3: Logged Status

And in some cases, as in the Title Screen, only use two different canvases, that is, the display status Logged-In, and the status Logged-Off.

### 3.3. Data Management

For storing data and status obtained by the player from the results of solving the problem, as well as from the input requested in each game, as well as data taken from the value of an object, the Get Playerpref function is used.

```
0 references
void start()
{
    Username = name.GetComponent<InputField>().text;
}

0 references
public void SaveName(InputField name){
    PlayerPrefs.SetString("username", name.text);
}

0 references
public void LoadMenu(){
    SceneManager.LoadScene("Menu");
}
```

Figure 4: Setting Data

And to send and specify the desired data. Get Playerpref function will be used. The function is used in several different uses and different data types.

```
void Update()
{
    GetComponent<Text>().text = PlayerPrefs.GetString("username", "Default");
}
```

Figure 5: Loading Username

```
void Update()
{
    GetComponent<Text>().text = PlayerPrefs.GetInt("score").ToString();
}
```

Figure 6: Loading Score

```
void Update()
{
    GetComponent<Text>().text = PlayerPrefs.GetInt("skor1").ToString();
}
```

Figure 7: Loading Specific Score

```
void Update()
{
    GetComponent<Text>().text = PlayerPrefs.GetInt("skor2").ToString();
}
```

Figure 8: Loading Specific Score 2

```
void Update()
{
    GetComponent<Text>().text = PlayerPrefs.GetInt("skor3").ToString();
}
```

Figure 9: Loading Specific Score 3

```
void Update()
{
    GetComponent<Text>().text = PlayerPrefs.GetInt("levellocation").ToString();
}
```

Figure 10: Loading Location Status



#### 4. Analisis

We conducted a survey by giving 5 mathematical-themed problems, in writing. And gave the same problem again, but required to do it using a video game. Of the 30 people we surveyed, two statistics were obtained.

##### 4.1. Total questions answered

From 30 that has been surveyed, there are 2 people that can answer 1 Question and only 1 that can answer 2 Questions. Then the rest can't answer at all.

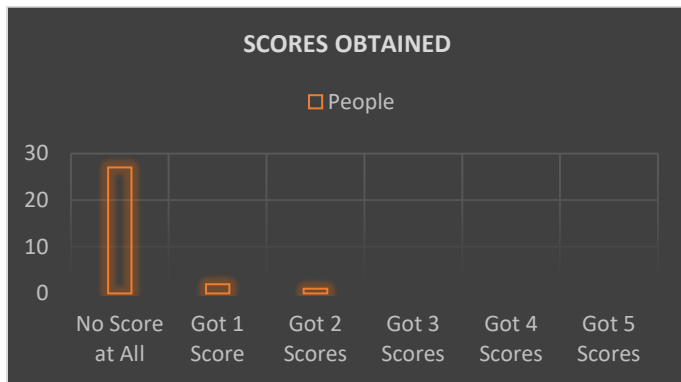


Figure 11: Total Answered Pre-Test

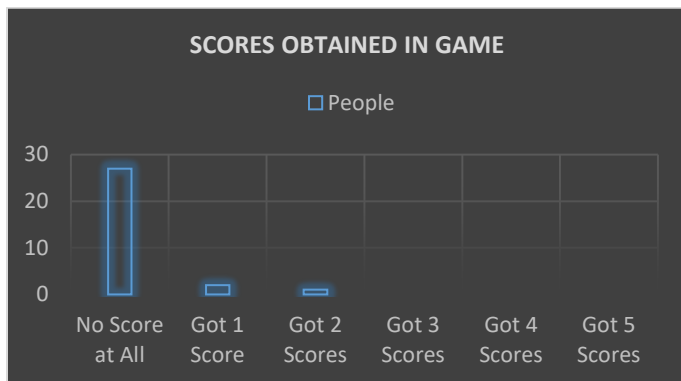


Figure 12: Total Answered Post Test (in-game)

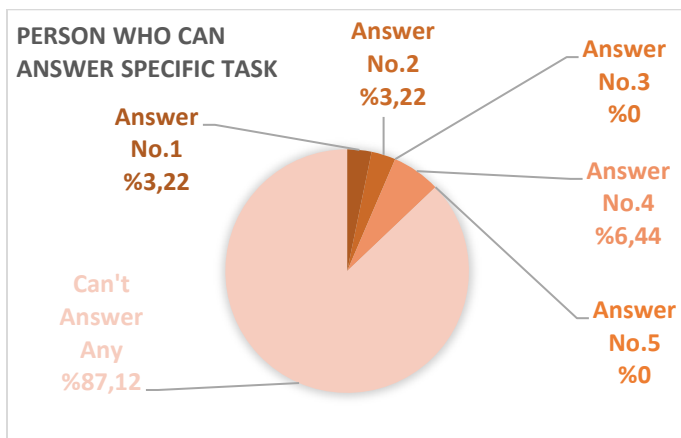


Figure 13: Task Answered

##### 4.2. Specific Number Answered

Only one person can answer Problem number 1, also one person can answer Problem number 2 and that one person earlier that can answer Problem number 2, can also answer Problem number 4, whereas no one can answer questions number 3 and number 5.

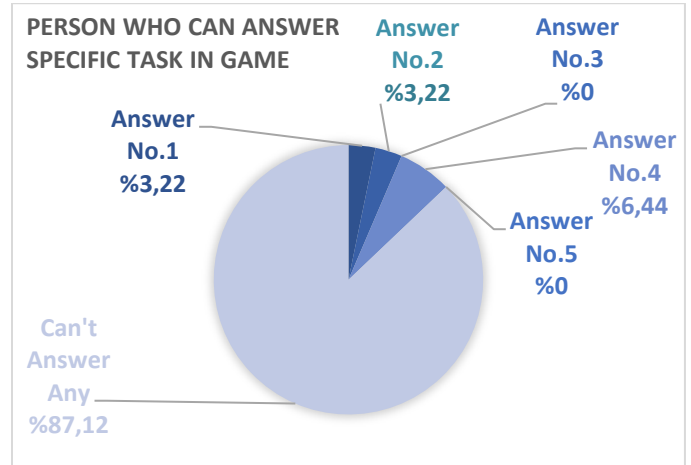


Figure 14: Task Answered in-game

##### 4.3. Difference

Among those who can answer their questions in writing, they will easily answer the questions in the Game. Yes, it is because the working period from written to game is quite short.

But there are no changes that occur to people who are unable to answer written questions, then to questions on the Game.

Assuming the difficulty of finding people, who are happy to work on math-related problems seriously in their spare time. Where most people tend to be reluctant to deal with things that require them to think hard when being relaxed.

#### 5. Result and Discussion

##### 5.1. Storyline

Indonesia is a beautiful country and rich in culture. The richness of Indonesian culture comes from the legacy of ancestors who always passed down to their future generations. Certainly, an ancestral inheritance has a very valuable heritage to guard. It is the duty of each region to safeguard the cultural treasures they possess. Saipul, an adventurous child, loves Indonesia very much and the cultural riches it possesses. Saipul is always passionate about knowing new cultures that have not been exposed to the outside world. Currently there are still 3 areas that he wants to explore. Namely East Kalimantan, South Sulawesi and West Papua. Saipul wants to know what kind of heritage the three regions have. To explore these areas, Saipul will be faced with a puzzle challenge that will direct him to the relics of each region. To achieve this, Saipul must complete the puzzle quickly and precisely to get clues to the next puzzle.

### 5.2. Title Screen

There are two types of Title Screen in this game, the first is the Normal version, and the other one is the Logged-In version. On the Title Screen the Normal player version will be shown the Game Logo, the Background depicts the Indonesian Archipelago, Input Field for Username and Button to login.



Figure 15: Title Screen

If the Player has filled in the Username and has already logged in, the Title Screen display will change to Logged-In. Then the Title Screen will display the phrase "Welcome Back" along with the Player Username. Two new buttons will also be displayed, namely the "Play" and "Exit". The "Play" button is used to proceed to the next Scene, the Main Menu. While the "Exit" button is used to exit the game, as well as log out from the player's username.

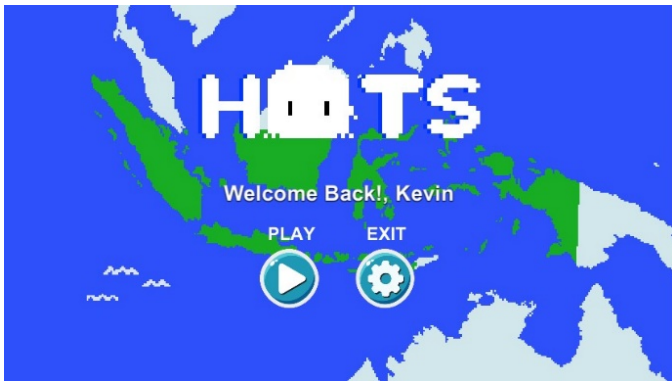


Figure 16: Alternate Title Screen

### 5.3. Main Menu

On the main menu, players will be shown a Map of Indonesian Island, which has a yellow area, which indicates that there are stages that can be played on the island. However, not all the available Stages are available at the start of the Game. The Stages are still locked and will open if the player has performed certain conditions. The playable stage will be indicated by the Exclamation Mark symbol in Orange. Each symbol of the Stage is also accompanied by a Panel containing the name of the Stage or Island and a Description presented by the Stage.



Figure 17: Stage Menu

Then there is also a panel at the bottom left of the Screen, called the Score Status Panel. This panel aims to show the Progress or Status Score that the player has obtained from each question on the existing Stages. The Panel has three indications that show the Player's Score and Total Score can be obtained from each Stage.

### 5.4. Gameplay

As both Treasure Hunt and Puzzle game HOTS Game has a fast and precise puzzle-solving gameplay. To proceed to the next puzzle, the player must complete it perfectly, that is, the correct time. This puzzle will not be displayed directly. In the Gameplay view, you will see a Joystick to move the Player Model. The player can move the Camera vertically by touching the layers on the left and right, and will move in accordance with the direction of the logistics that it touches.



Figure 18: Gameplay Scene

Players will see a button in the form of Scoll or Scroll, which serves to display a Panel that contains instructions where the location of hidden questions.

Then, the Player must move to find it by exploring an area to find the point where the puzzle is located. HOTS games are linear which means the next puzzle will not appear before completing the previous one.



Figure 19: Gameplay showing the clue

After completing a set of puzzles in one area, the player gets hidden treasures in that area. Then the player will shown a UI indication of a Task or Puzzle to complete.



Figure 21: Appearance of Task no.2

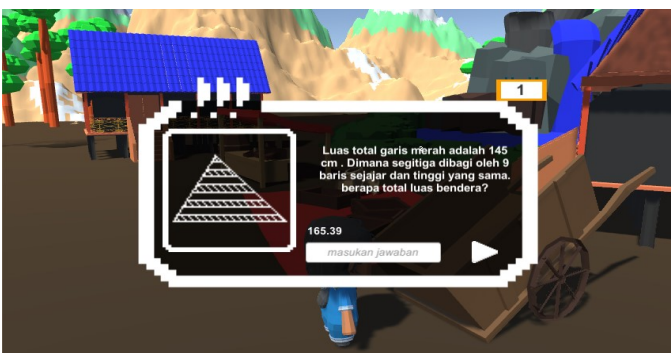


Figure 22: Task Display

To solve the puzzle in the form of mathematical problems, players are required to enter the correct number of digits after making calculations according to the existing problems. In addition to entering the correct numbers, players are required to interact with puzzles in the environment such as sliding, inserting, moving objects.

If the player answers correctly, the player will get a clue to the next puzzle and the player will get a score.

If the player answers incorrectly, the player will be given another chance to answer. If the second chance is still wrong, then the 'skip button' screen will appear to skip the puzzle problem and still be able to continue to the next puzzle.

The reward given is that the player will not get a score. In the exploration section, the gameplay display is third-person perspective. The player's character can explore the area in the game and can rotate the camera for different points of view.

Players must go around to find puzzles that have been marked and do interaction activities to start working on the puzzles found. When solving problems, in addition to getting instructions for the next question, there are questions that will get additional rewards, namely items that will be useful during gameplay exploration.

For example, a machete that will be used to cut branches to open roads covered with grass or branches. In each category, if you solve the problem until the last question, then you will get a reward in the form of a map of Indonesia. If it succeeds in collecting all the map fragments, it will display a map of Indonesia which shows famous points on Indonesia. For example, in the North Sumatra area there is the Lake Toba Icon, in the Yogyakarta Region there is the Borobudur Temple Icon.



Figure 23: New Clue Appear



Figure 24: Display of Pause Menu



### 5.5. Pause Menu

On the pause menu, there are several posts and buttons. There are two indications that show the username of the player and the score that has been obtained from the player.

Then there are several buttons. Namely Resume, Status, Main Menu, and Quit. On the Status menu, the player will be moved to a menu that shows details of player progressions when they play the game.

### 5.6. Score

Each question will give players 1 score and will accumulate on each Stage / island and will be shown to players via the pause menu and status menu. The status menu shows several indications, including the Score of the three Stages, the location of the player and the total score that has been obtained



Figure 25: Display of Status Menu

### References

- [1] M.D. Aguilera, A. Mendiz, "Video games and education. *Computers in Entertainment*, 1(1), 1-10, 2003. <https://doi.org/10.1145/950566.950583>.
- [2] Y. Udjaja, "Gamification Assisted Language Learning for Japanese Language Using Expert Point Cloud Recognizer", *International Journal of Computer Games Technology*, 2018. <https://doi.org/10.1155/2018/9085179>
- [3] A.W. Ruch, "Videogame Interface: Artefacts and Tropes. *Videogame Cultures and the Future of Interactive Entertainment*, 2010. [https://doi.org/10.1163/9781848880597\\_002](https://doi.org/10.1163/9781848880597_002)
- [4] M.D. Griffiths, "The educational benefits of videogames. *Education and health*, 20(3), 47-51, 2002.
- [5] M. Hassenzahl, N. Tractinsky, "User experience - A research agenda. *Behaviour and Information Technology*, 25, 91-97, 2006. <http://dx.doi.org/10.1080/01449290500330331>
- [6] D.P. Kristiadi et al., "The Effect Of Ui , Ux And Gx On Video", *The IEEE CyberneticsCom* 2017, 157-162, 2017. <https://doi.org/10.1109/CYBERNETICSCOM.2017.8311702>
- [7] P. Sporgis, "U.S. Patent No. 6,320,495. Washington, DC: U.S. Patent and Trademark Office, 2009.
- [8] M. Dondlinger, "Educational Video Game Design: A Review of the Literature. *Journal of Applied Educational Technology*, 4(1), 21-31, 2007.
- [9] D.P. Kristiadi et al., "The effect of UI, UX and GX on video games", In 2017 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom) 158-163, 2007. <https://doi.org/10.1109/CYBERNETICSCOM.2017.8311702>
- [10] Sasmoko, J. Harsono, Y. Udjaja, Y. Indrianti, J. Moniaga, "The Effect of Game Experience from Counter-Strike" *Global Offensive*. In 2019 International Conference of Artificial Intelligence and Information Technology (ICAIT), 374-378, 2019. <https://doi.org/10.1109/ICAIT.2019.8834521>
- [11] Y. Udjaja, "Ekspanpixel Bladsy Stranica: Performance Efficiency Improvement of Making Front-End Website Using Computer Aided

- Software Engineering Tool", *Procedia Computer Science*, 135, 292-301, 2018. <https://doi.org/10.1016/j.procs.2018.08.177>
- [12] Sasmoko, S.A. Halim, Y. Indrianti, Y. Udjaja, J. Moniaga, B.A. Makalew, "The Repercussions of Game Multiplayer Online Battle Arena. In 2019 International Conference of Artificial Intelligence and Information Technology (ICAIT), 443-447, 2019. <https://doi.org/10.1109/ICAIT.2019.8834518>
- [13] Y. Udjaja et al., "Gamification for Elementary Mathematics Learning in Indonesia. *International Journal of Electrical and Computer Engineering*, 8(5), 3859, 2018. <https://doi.org/10.11591/ijece.v8i5.pp3859-3865>
- [14] Y. Udjaja et al., "The Use of Role Playing Game for Japanese Language Learning", *Procedia Computer Science*, 157, 298-305, 2019. <https://doi.org/10.1016/j.procs.2019.08.170>
- [15] Meiryani, Y. Udjaja, J. Jeviro, S. Sabrina, "The Utilization of Games in Learning Calculation (Accounting) Based on Andriod. *International Journal of Psychosocial Rehabilitation*, 24(9), 7-11, 2020. <https://doi.org/10.37200/IJPR/V24I/PR290002>
- [16] I. Dart, M.J. Nelson, "Smart terrain causality chains for adventure-game puzzle generation", 2012 IEEE Conference on Computational Intelligence and Games, CIG, 328-334, 2012. <https://doi.org/10.1109/CIG.2012.6374173>.

## Entertainment Technology: Dynamic Game Production

Yogi Udjaja\*

Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, 11480, Indonesia

---

### ARTICLE INFO

*Article history:*

*Received: 27 March, 2020*

*Accepted: 06 June, 2020*

*Online: 22 July, 2020*

---

*Keywords:*

*Dynamic Game Production*

*Game Development*

*Software Engineering*

---

---

### ABSTRACT

*The times have made drastic changes in the field of technology, these changes have penetrated into various aspects, one of which is the game. Starting from the tactics of war in the real world that produces the game theory, then adopted in the form of games that exist today. Starting from without visuals to extraordinary visual elements, each of these changes always changes the way game developers in terms of making games, so that proposed methods that can adjust to the development era that can maintain quality and interest of users so that the game is made right on target, namely Dynamic Game Production.*

---

## 1. Introduction

Game theory was first discovered in 1944 by Neumann and Morgenstern, where the theory was adopted for war strategy [1]. On the other hand, in 1958 found the first computer-based game (digital game) made by Higinbotham and named Tennis for Two [2], this development changed history, so the way of making games also developed.

In general, the development of the game is directly proportional to the development of technology. It can be seen from the beginning of the game formed, starting from the model of the game that was played without visual elements until now the game has a visual element that is extraordinary. Developments that occur not only visualization but as a whole, so that the way of making it from time to time is also increasingly developing.

Then in 2005 the game development life cycle (GDLC) method was born which served to improve the quality of the game. According to Aleem et al., there are 5 continents in the spotlight on GDLC research, namely: Asia, Europe, North America, South America and Australia [3]. Total research on GDLC from 5 continents is around 296 studies both from pre-production, production to post-production stages. Aleem et al. said the quality of game development is still immature because there are still very few activities in game research especially in post-production [3]. Based on research that has been done, making a game requires

systematic investigation, where the games are made whether it is in accordance with the desires and needs of players [4].

This can be taken based on several factors that support the player's interest, including the most important is the psychological aspect. Based on this interaction with the user or player is very important to improve the quality of the game made.

## 2. Related Research

### 2.1. Game Production Pipeline

GameDev Jakarta is a game developer community located in Jakarta, Indonesia. Based on the results of discussions from the GameDev Jakarta Gathering (2016) obtained three main processes in making games, namely: pre-production, production and post-production.

- Pre-production: Stages before making a game, in the form of concepts, data analysis and game design.
- Production: The stages of making the following games are alpha and beta testing.
- Post-Production: Stages after the game is finished and marketed. When it has been marketed or released, the game's performance is checked again, is it feasible to maintain or make a new version or if the performance is not in accordance with what is desired, the game is not developed anymore (abandon).

### 2.2. Game Development Life Cycle

Ramdan and Widyaning analyzed several game development life cycle (GDLC) models performed by Blitz Game Studios, Arnold

---

\*Yogi Udjaja, Jakarta, Indonesia 11480, +62 898 268 3399, udjaja.yogi@gmail.com, yogi.udjaja@binus.ac.id



Hendrick, Doppler Interactive and Heacher Chandler [5]. Then after analyzing the GDLC model, the new GDLC model was proposed, in accordance at that time. After re-analysis [5, 6], there are still many things that need to be improved from this method, therefore this study was made.

### 2.3. Game Development Based Learning

Game development based learning (GDBL) is used to increase motivation and overall productivity of players [7]. In general it can be seen the relationship between serious games, game based learning, gamification and GDBL (see figure 1).

In figure 1 explained that the serious game is as a whole of a game that has a goal that involves all aspects of teaching, training and providing information. Game based learning (GBL) is a serious game created to improve teaching activities and initiatives based on role-playing, motivation, involvement and repetition. While gamification is the application of game elements in a non-game context. Then it can be concluded that GDBL is making games by combining all aspects contained in serious games, GBL and gamification.

According to Spieler and Slany, in stages GDBL was developed in 4 stages viz [7]:

- Research: This stage is carried out data collection, investigation and analysis of the story, title, genre and theme of the game, as well as the basic concepts of game structure and gameplay.
- Design: This stage creates a game design document that contains artwork, game content, and other game elements (assets, characters, sound, etc.).
- Development: This stage covers all the processes of making game elements, programming and followed by testing.
- Release: This stage is the last stage when the game is finished and is running well.

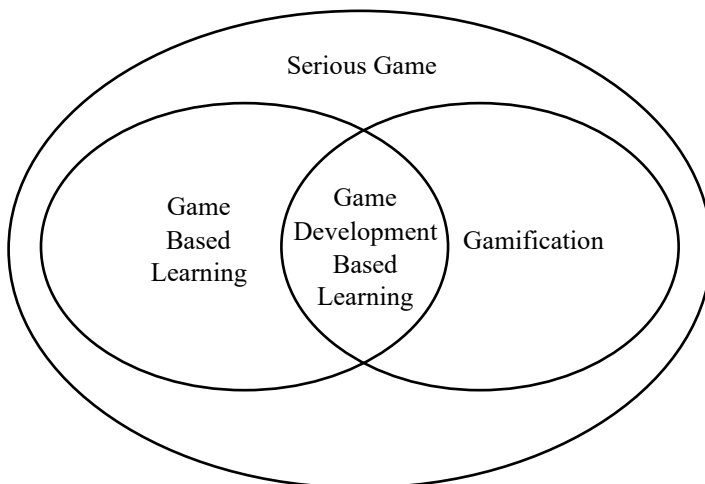


Figure 1: Relationship between Serious Game, GBL, gamification and GDBL [7]

### 2.4. Challenge Based Learning

Challenge Based Learning (CBL) is one way to learn so that someone can think critically [8, 9]. Where in this method begins

with a real problem and collaborates with the team to get the best solution.

### 2.5. Interpersonal Circumplex

According to Isbister, Interpersonal circumplex can be used as a model to recognize the player's personality or to make expressions, traits, culture, social activities and other personalities of characters in the game [10]. These models can be used to determine the target player and market research of the game to be made.

### 2.6. Game Experience

Based on the explanation from Udjaja, game experience comes from a combination of user interface (UI), user experience (UX), gameplay experience (GX) and game balancing [11]. The model has been proven to see conformity to the comfort, interests, needs and motivation of users to play.

## 3. Propose Method

Based on previous research [4-16], game development methods always change dynamically as needed. Based on this, the proposed game development method was developed according to the needs and changing times of the dynamic game production (DGP). This method looks at the emotional engagement aspect of the player so that the game is made right on target. The following are the DGP stages (see figure 2):

### A. Initiation

This stage is a determinant of the suitability of making a game to suit the needs of the player. In this stage, the analysis phase can be done simultaneously or alternately, so as to produce the required system requirements. The stages are:

- Player Analysis

At this stage an analysis of player needs and payer motivation to play the game, as well as the determination of the game concept.

- Competitor Analysis

At this stage an analysis of existing games is performed. Can be seen from the manufacturing process, marketing to game maintenance.

- Interaction Analysis

At this stage an analysis of the interaction will be performed by the player. Can be seen from the device and controller used, and how to use it.

- System Requirement

After conducting the analysis, the results of the analysis are formed into a system requirement. This stage is the required requirements. Can be made using mind maps or other methods that are easily understood by developers.

Stages of analysis carried out can use the CBL method so that the results obtained are right on target. For workmanship guidelines can be seen in ref. [16].

## B. Alpha Version

This stage is carried out alternately, starting from preproduction, then proceed with production and prototype testing, the prototype is evaluated through performance analysis, when the results of the prototype testing are not in accordance with the requirements then the preproduction stage will be carried out again or can return to the initiation stage.

- Preproduction

Based on the system requirements that have been made, an estimated time of manufacture is made. Then the game design document (GDD) is made, which contains the design of the game to be made, as has been made by Adams [17]: high concept document, game treatment document, character design document, world design document, user interface design document, flowboard, story and level progression, on-screen text and dialog script, game script. In GDD can also be added specific design in the form of game experience, in-game algorithm, artificial intelligence, debug system, system design using object oriented or procedural and so on. The more detailed design that is made, the clearer the flow of the game to be made.

- Production

At this stage the game is made based on the existing GDD. Making a game is divided into several modules, according to the system requirements and estimated time that has been made. Every single title was completed, a play was tested and a work checklist was made.

- Prototype Testing

After the prototype game is finished making automatic testing using Selenium or other applications to see if there are errors in the code that has been made. Besides that, playtesting is also done internally and added by some people from external (can be from players or people who are experts in their field) to get feedback on the interest in game and game experience. Then a performance analysis is performed, where the results of automatic testing and feedback from people who have done the testing are analyzed and compared with the GDD that is made, so that the quality of the game made can be increased again. If it is not in accordance with the expectations or the GDD made there is an error done again initiation stage.

## C. Beta Version

This stage was performed soft-launch playtesting to several players, then performance analysis was re-conducted. When there is a mismatch, changes will be made at the production stage, otherwise it will be returned to the initiation stage.

## D. Release

This stage is the stage of launching the game to the public. player research analitic is done to see the player's interest, if there is a monetize can also be seen from items that are often purchased or from the revenue obtained.

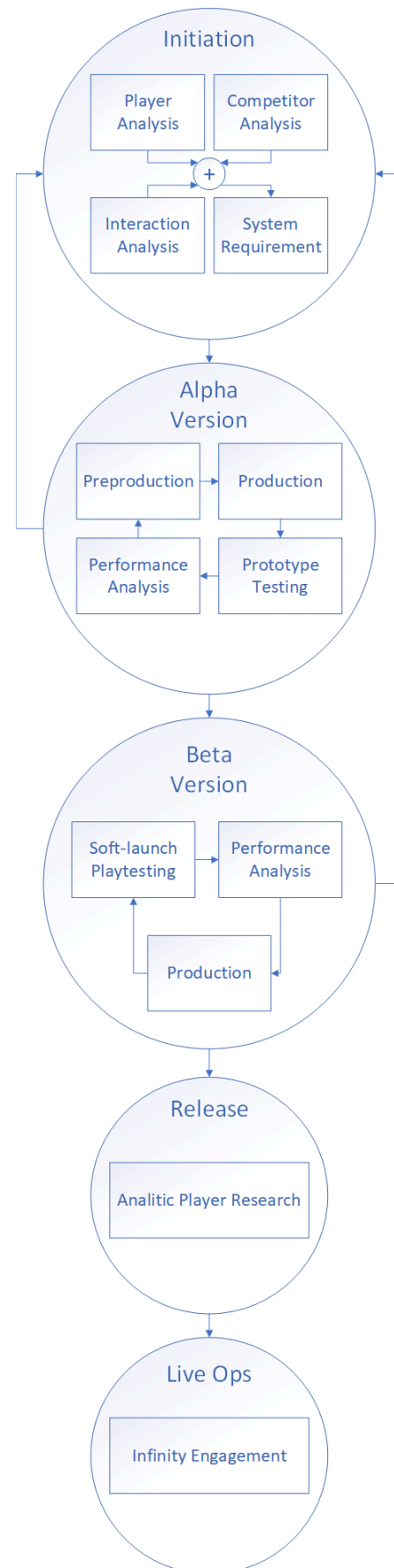


Figure 2: Dynamic Game Production

#### E. Live Ops

This stage is the stage of maintaining player interest in continuing to play games that have been live by making changes in the game. In general, this change does not change the code that has been made, if you change the code, changes made in a small scope or add some new functions. Changes that occur can be made such as adding cosmetic items, limited offers, new mechanics, camapign to acquire paid users or others.

#### 4. Conclusion and Discussion

Every making something should begin with research so that what is made right on target. One simple concept like 5W1H (What, Where, When, Who, Why, How) can be used to understand something, or combined with challange based learning.

This DGP was created to maintain the quality of the game and the interests of the players so that the game was made right on target. In practice, this method is a method of making a dynamic game, when there is a change from a condition, this method is able to adjust and solve it—such as the alpha version or beta version can return to initiation (See figure 2), so that the results of the game are made better.

On the other hand, emotional engagement in games has positive and negative impacts, when getting someone's emotional engagement, playing games can change one's behavior. It also must be considered, whether this change is good or not. The developer must also consider this so that undesirable things do not happen.

#### References

- [1] J.V. Neumann, O. Morgenstern, “*Theory of games and economic behavior (commemorative edition)*” Princeton university press, 2007.
- [2] W. Higinbotham, “Tennis for two, 1958. <https://www.sunysb.edu/libspecial/videogames/tennis.html>
- [3] S. Aleem, L.F. Capretz, F. Ahmed, “Game development software engineering process life cycle: a systematic review. *Journal of Software Engineering Research and Development*, 4(1), 6, 2016. <https://doi.org/10.1186/s40411-016-0032-7>
- [4] Y. Udjaja, V.S. Guizot, N. Chandra, “Gamification for Elementary Mathematics Learning in Indonesia. *International Journal of Electrical and Computer Engineering (IJECE)*, 8(6), 2018. <http://doi.org/10.11591/ijece.v8i5.pp3860-3865>
- [5] R. Ramadan, Y. Widyani, “Game development life cycle guidelines”, In *Advanced Computer Science and Information Systems (ICACSIS), 2013 International Conference*, 95-100, 2013. <https://doi.org/10.1109/ICACSIS.2013.6761558>
- [6] Y. Udjaja, A.C. Sari, “A Gamification Interactive Typing for Primary School Visually Impaired Children in Indonesia”, *Procedia computer science*, 116, 638-644 2017. <https://doi.org/10.1016/j.procs.2017.10.032>
- [7] B. Spieler, W. Slany, “Game Development-Based Learning Experience: Gender Differences in Game Design, 2018. *arXiv preprint arXiv:1805.04457*.
- [8] L. Johnson, S. Brown, “*Challenge based learning: The report from the implementation project*, 1-36, 2011.
- [9] The Challenge Institute “Challenge Based Learning Viewed on 20 October 2019 18:56 WIB, 2018. Retrived on: <https://www.challengebasedlearning.org>
- [10] K. Isbister, “*Better game characters by design: A psychological approach*. San Francisco: Elsevier, 2006.
- [11] Y. Udjaja, “Gamification Assisted Language Learning for Japanese Language Using Expert Point Cloud Recognizer. *International Journal of Computer Games Technology*, 2018. <https://doi.org/10.1155/2018/9085179>
- [12] D.P. Kristiadi et al., “The effect of UI, UX and GX on video games”, In *Cybernetics and Computational Intelligence (CyberneticsCom), 2017 IEEE International Conference*, 158-163, 2017. <https://doi.org/10.1109/CYBERNETICSCOM.2017.8311702>
- [13] Y. Udjaja et al., "The Use of Role Playing Game for Japanese Language Learning." *Procedia Computer Science* 157, 298-305, 2019. <https://doi.org/10.1016/j.procs.2019.08.170>
- [14] S.A. Halim et al., “The Repercussions of Game Multiplayer Online Battle Arena. In *2019 International Conference of Artificial Intelligence and Information Technology (ICAIIIT)*, 443-447, 2019. <https://doi.org/10.1109/ICAIIIT.2019.8834518>
- [15] S. Harsono et al., “The Effect of Game Experience from Counter-Strike: Global Offensive. In *2019 International Conference of Artificial Intelligence and Information Technology (ICAIIIT)*, 374-378, 2019. <https://doi.org/10.1109/ICAIIIT.2019.8834521>
- [16] Apple, Challenge Based Learning: A Classroom Guide. Viewed on 20 October 2019 18:29 WIB, 2010. Retrived on: [https://images.apple.com/education/docs/CBL\\_Classroom\\_Guide\\_Jan\\_20\\_11.pdf](https://images.apple.com/education/docs/CBL_Classroom_Guide_Jan_20_11.pdf)
- [17] E. Adams, "*Fundamentals of game design*. Pearson Education, 2014.

## Malware Classification Based on System Call Sequences Using Deep Learning

Rizki Jaka Maulana, Gede Putra Kusuma\*

Computer Science Department, BINUS Graduate Program, Bina Nusantara University, Jakarta, 11480, Indonesia

---

### ARTICLE INFO

*Article history:*

*Received: 27 March, 2020*

*Accepted: 06 June, 2020*

*Online: 22 July, 2020*

---

*Keywords:*

*Malware Classification*

*Malware Detection*

*System Call Sequence*

*Deep Learning*

*LSTM Model*

---

### ABSTRACT

*Malware has always been a big problem for companies, government agencies, and individuals because people still use it as a primary tool to influence networks, applications, and computer operating systems to gain unilateral benefits. Until now, malware detection with heuristic and signature-based methods are still struggling to keep up with the evolution of malware. Machine learning is known to be able to automate the work needed to detect families of existing and newly discovered malware. Unfortunately, the machine learning method using Support Vector Machine (SVM) for detecting malware can only reach a low level of accuracy. In this work, we propose a dynamic analysis method and uses a system call sequence to monitor malware behavior. It uses the word2vec technique as word embedding and implements deep learning models, namely Long Short-Term Memory (LSTM) and Nested LSTM, as classifiers. To compare with existing machine learning approach, we also apply the Support Vector Machine (SVM) as a benchmark method. The Nested LSTM gets an accuracy of 93.11%, while the LSTM gets the best accuracy of 98.61%. The LSTM also achieved the best performance in terms of average precision at 97.57%, the average recall at 97.29%, and the average score of f1 at 97.43%. We have found that our model is lightweight but powerful for detecting malware with significant accuracy.*

---

### 1. Introduction

The high use of the internet increases the level of connectivity of electronic devices, making questions about the integrity of the system. Conventionally, software and computer systems are developed for good purposes. However, some software was developed to produce crime (malware). Malware is a common word used for programs that have malicious code snippets that can cause significant threats to computer users or any digital device. Malware can contain malicious code viruses, worms, Trojan horses, can also make a back door to divulge personal information or control a person's system. Through malware, serious crimes can be done; This is why malware detection is needed [1]. To detect malware definitions must be made for analysis of which malware is essential. Malware analysis consists of analyzing various aspects of malware so that malware can be detected [2]. The definition of malware is also known as a signature/signature. This signature is used by virus scanners known as anti-viruses to detect malware. The research will experiment on seven types of malware, which are adware, backdoor, packed, riskware, trojan, virus, and worm.

Traditional malware detection is done on susceptible files that are not processed. This is mostly done with a signature, heuristic, and behavioral approach. The signature approach looks for static patterns of malware known in suspicious files [3]. Research has shown that the signature approach is very weak in dealing with polymorphic and metamorphic malware. The heuristic approach checks the characteristics of suspicious malware from suspicious files. Despite being able to detect unknown malware, they are very high at the false-positive level.

The Behavioral Approach monitors the implementation of programs to monitor suspicious behavior. Although this approach can detect different malware variants, this approach also has a high false-positive [4]. To help malware analysts retrieve useful information from large malware samples, the need for automatic classification in statistical variants is needed. Malware detection based on a signature cannot overcome this variant because it does not take polymorphic malware into account. Polymorphic is a form of malware that frequently always changes its identifiable features to evade detection. Furthermore, such a system can be easily avoided.

---

\*Corresponding Author: Gede Putra Kusuma, [inegara@binus.edu](mailto:inegara@binus.edu)



The most important event that can be tracked to determine malware behavior is the system call. Before malware performs a malicious action, malware needs to use the operating system (OS) service of the target. For each activity that is carried out, such as opening a file, running a thread, writing a command to the administrator, or opening a network connection, interaction with the operating system is required. This interaction is carried out via the API call system of the target OS. Therefore, monitoring the behavior of malware is very important to monitor the order of system calls during malware execution. Different malware families certainly have different goals.

Detected malware is easily handled mainly by elimination. However, the current nature of malware is polymorphic and metamorphic, making them difficult to detect in traditional ways. They disguise their structure but not their operations. Because all malware must be executed to carry out its malicious actions successfully, some studies [5], analyze API calls to detect malware in high accuracy execution. However, this detection ends by marking malware or not malware [6]. It does not classify malware into its type (viruses, worms, Trojans, etc.). Classification is important because it helps simplify the course of action to neutralize it.

Research on malware classification has been done before. However, these studies do not use the Word2vec method. One example is a study of classification in system call sequences conducted in 2019, wherein that study the classification contained nine types of malware, namely kelihos\_v3, vundo, rammit, lolipop, simda, tracur, obfuscator.ACY and gatak. The methods used are text and hex commands and LSTM [7]. Also, in 2016 there was research on evaluating machine learning methods such as the Hidden Markov Model [8] and SVM [9] in determining malware classification.

We use word embedding techniques in processing to convert malware system call sequences into vectors to achieve an increase in capturing the relationship between n-grams in the system call sequence and then proceeding to LSTM for the classification process. In essence, this approach expected to improve accuracy and precision for most families of malware, which brings a significant improvement from the methods used by previous researchers and thus can help to classify malware more accurately.

All sections of the paper have been organized as follows. Section 2 discusses related works on the detection and classification of malware. In section 3 provides details of the background Theories, word embedding, and deep learning methods. Section 4 discusses the details of the dataset, the methodology used, and the evaluation design. Section 5 discusses the details of all experimental results, including training and testing results. The conclusion and future work are places in section 6.

## 2. Related Works

Previous researchers have shown that using program behavior features such as API calls can detect malware, including metamorphic and polymorphic malware, with high accuracy. This is because, at a higher level, malware disguises itself by changing their behavior or continuously changing their signatures. However, to cause damage, they must execute and change execution

behavior more difficult. This can make them harmless. Therefore, this approach targets malware at the execution level.

The first researcher who used a deep learning-based malware detection (DLMD) approach relied on static methods to predict behaviors that can be executed using system call sequences that provide sequences taken from running processes. Using SVM and CNN the results show that this method is quite effective in detecting polymorphic and metamorphic malware with an accuracy and detection rate of 89% to 96% [10]. In the proposed DLMD technique, SVM is used as a feature selector and CNN autoencoder is used as a feature extractor. After that, a Multilayer perceptron is used as a classifier.

Other researchers develop an 18-layer deep residual network to be issued bytecode to a 3-channel RGB image and then apply deep learning to classify malware. To convert malware to images, they first convert malware binaries to 8-bit vectors (bytecodes) [11]. After that, the bytecodes are converted into grayscale images with contribution values from 0 to 255, each vector that turns into pixels with added values from 0 to 255. In the next step, they then convert the grayscale images to 3-channel RGB images with duplicate the grayscale channel three times and then collect all three channels to create an RGB image. Their experimental results show that the network residual model achieves an average accuracy of 86.54% with 5-fold cross-validation.

In [12], the author proposed a new malware detection method based on Deep Graph Convolutional Neural Networks (DGCNNs) to learn directly from the sequence of API calls and related behavior graphs. The experimental results show that the model reaches a similar area under the ROC curve (AUC-ROC) and F1-Score of Long-Short Term Memory (LSTM) networks that produce up to 96%.

In [13], the author proposed a method for detecting malware variants that are packaged based on sensitive system calls and the Deep Belief Network. Different experimental groups and different data samples were used for analysis. The 10-fold cross-validation method is used for classification. Theoretical analysis and experimental results show that the proposed method can detect packed malware which reaches an accuracy of 92% and requires a detection time of fewer than 0.001 seconds.

In [14], the author proposed a conventional approach with deep learning-based using Recurrent Neural Networks (RNN) that are vulnerable to redundant API injection. They investigated the effectiveness of Convolutional Neural Networks (CNN) against injection of redundant APIs. Their malware detection system converts malware files into image representations and classifies image representations with CNN. CNN is implemented with spatial pyramid pooling layers (SPP) to handle various input sizes. They also evaluated the effectiveness of SPP and image color space (greyscale / RGB) by measuring system performance on unaltered data and adversarial data with the injected redundant API. The results show that Naive SPP implementation is not impractical due to memory constraints and effective greyscale imaging against redundant API injection.

The last researcher proposed an approach of how deep learning architecture using the stacked AutoEncoders (SAEs) model can be designed for intelligent malware detection. The SAEs model



functions as a greedy layerwise training operation for unsupervised feature learning, followed by supervised fine-tuning parameters (eg Weights and offset vectors). Based on the representation of different features, various types of classification methods, such as Artificial Neural Networks (ANNs), Support Vector Machines (SVM), Naïve Bayes (NB), and Decision Tree (DT) are used as a model construction to detect malware. Most of these methods are built on shallow learning architectures. Even though they have succeeded in isolating malware detection but shallow learning architectures are still unsatisfactory for malware detection problems [15]. The experimental results of the method showed that the proposed method achieves 96% accuracy. The bibliography comparison of previous works are summarised in Table 1.

Table 1: Bibliography comparison

Author	Dataset	Method	Accuracy	Class Malware
Rafique, Ali, Qureshi, Khan, & Mirza, 2019 [10]	10.868 Portable Execution (PE) files	SVM & CNN	89% - 96%	6 Class Malware
Yan Lu, Jonathan Graham, Jiang Li, 2019 [11]	2949 Portable Execution (PE) files	Deep Residual Network	86.54%	5 Class Malware
Oliveira, Julho, & Julho, 2019 [12]	42.797 Portable Executable (PE) files	LSTM & DGCNN	92.7 - 96.8%	2 Class Malware
Zhang, Chang, Han, & Zhang, 2020 [13]	7.195 Portable Executable (PE) files	SVM & Deep Belief Network	86.3% - 92.6%	4 Class Malware
Ke He, Dong-Seong Kim, 2018 [14]	2.413 Portable Executable (PE) files	CNN	95%	2 Class Malware
Hardy, Chen, Hou, Ye, & Li, 2016 [15]	22.500 Portable Executable (PE) files	SVM, NB, DT, ANN & DL4MD	92% - 95%	2 Class Malware

Based on the results of the literature review, there have been previous studies that have tried to classify malware based on system call sequences data. But, the methods used before did not achieve high classification accuracy. In other fields, many methods

of deep learning have proven to be more accurate and therefore we use deep learning and the word2vec as a word embedding to improve accuracy. Since deep learning models are used, therefore we do not use feature extraction specifically like the study above. However, we use a word embedding, which convert the input text into numeric data as input to the LSTM model. As a result, it will increase classification accuracy.

### 3. Background Theories

#### 3.1. Word2Vect

Word2vec is a two-layer neural network that can process text by converting words into vectors or can also be called "vectorization." Input from word2vec is a collection of text, and the output is a collection of vectors. Feature vector representing words in a corpus. Word2vec is not a deep neural network. Word2vec works by converting text into numerical forms which can then be translated by deep neural networks. Word2Vec is a word embedding technique that is quite popular and was developed by[16] at Google.

Word2vec can also be applied to codes, likes, playlists, social media graphics, sentiment sentences, and other verbal or symbolic series where patterns can be seen. The purpose of word vectorization is to group word vectors that are similar in vector space, which can later detect mathematical equations. Word2vec functions by making a distributed numerical vector representation of a word. For example like in the context of an individual word.

Word2vec works automatically. With enough data usage and context. Word2vec can make very accurate guesses about the meaning of the words based on previous appearance or interpretation. These guesses are used to build the association of words with other words (e.g. "Male" means "boy" and "woman" means "girl"), or classify a document and then group them according to their topic. Clusters can form the basis of sentiment analysis, e-commerce, search, malware analysis, and recommendations in areas such as scientific research and legal discovery. The output of word2vec is in the form of vocabulary where each item has a vector, which can be entered into further processes such as machine learning or deep learning. Also, it can be used just to detect the relationship between these words.

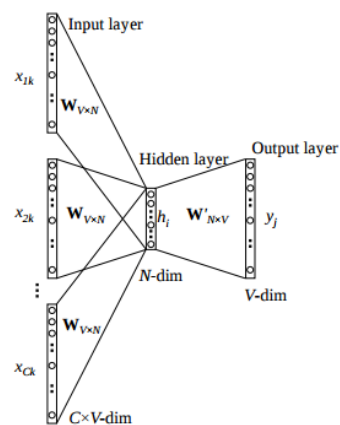


Figure 1: Continuous bag-of-words architecture

Figure 1 is a Word2vec Continuous-bag-of-words (CBOW) model. The way CBOW works is to take the context of each word and then make it as input and try to predict words that fit the context. As an example, When trying to predict the current target word (the center word) based on the source context words (surrounding words) [17]. If we make a simple sentence like “the black cat jump over the very big goat” this can be pairs of (context\_window, target\_word) where if we consider a context window of size 2, we have examples like ([the, cat], black), ([cat, over], jump), ([very, goat], big) and so on. This model tries to predict target\_word based on context\_window words.

### 3.2. Long Short-Term Memory (LSTM)

LSTM was first introduced by Sepp Hochreiter and Jurgen Schmidhuber in 1997 [18]. LSTM is a type of repetitive neuron that has been shown to increase the ability of RNN. LSTM can remove the effects of the problem by vanishing and bursting gradients, and is better to data-sensitivity relationships [19]. LSTM launched the forget gate inside the LSTM neuron, which allows accessing the information requested by the neuron allowing its access to focus on the critical parts and discard the information that is not useful.

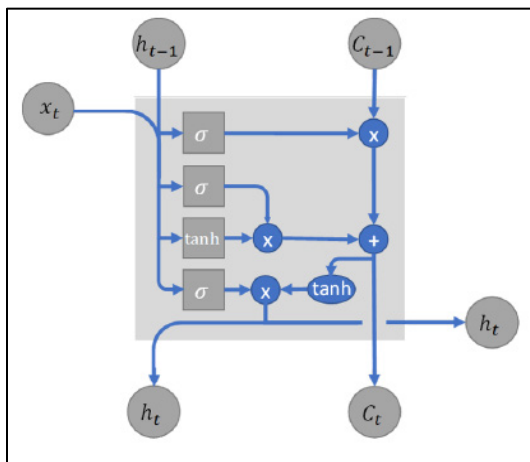


Figure 2: LSTM architecture

Figure 2 shows the structure of the LSTM. The key to LSTM architecture is its cell state. Cell state can be interpreted as a memory of a network and can delete or add information to a structure called a gate. For each “t” time-step in LSTM can be described by using this formula [20]:

$$f_t = \sigma(W_f * x_t + U_f * h_{t-1} + b_f) \quad (1)$$

$$i_t = \sigma(W_i * x_t + U_i * h_{t-1} + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_C * x_t + U_C * h_{t-1} + b_C) \quad (3)$$

$$C_t = i_t * \tilde{C}_t + f_t * C_{t-1} \quad (4)$$

$$O_t = \sigma(W_o * x_t + U_o * h_{t-1} + b_o) \quad (5)$$

$$h_t = o_t * \tanh(C_t) \quad (6)$$

while  $f_t$  is forget gate,  $i_t$  is the input gate,  $O_t$  is output gate  $C_t$  is a memory cell,  $h_t$  is a hidden layer,  $x_t$  is input when time “t”,  $\sigma$  is sigmoid activation function,  $\tanh$  is hyperbolic tangent activation function,  $W_t W_i W_C W_o U_f U_i U_C U_o$  are weight matrices for controlling the input and  $b_f b_i b_C b_o$  are bias vector.

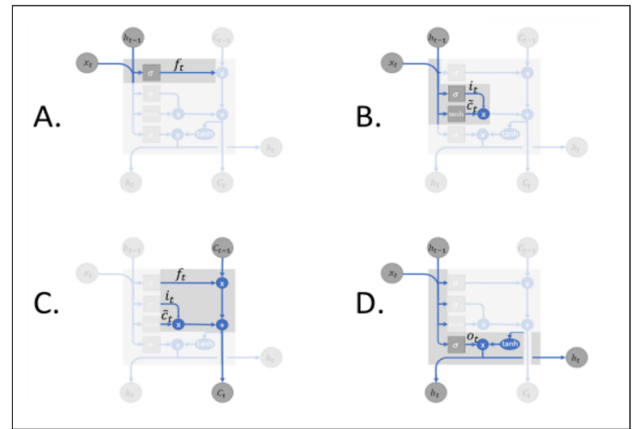


Figure 3: LSTM steps

Figure 3 shows the steps contained in the LSTM model architecture. There are four steps in LSTM namely: Step A. First, the model needs to determine what needs to be changed from the state of the cell. Figure 3 (A) will have a value called forget gate  $f_t$ . The input of this step is the output of the previous step, which written by  $h_{t-1}$  and  $x_t$  input. The activation function will give a result of “0” or “1”, where “0” means "not let anything pass" and “1” means "remember everything".

The next step is to determine what information will be added to the state of the cell. Shown by Figure 3 (B), Equations (2), and (3). At this stage, the input is  $h_{t-1}$  and  $x_t$ . The first layer is called the sigmoid layer, which serves to determine which part to be updated. And the tanh layer is to create a new candidate value  $C_t$ . In the next step, the two layers will be combined to update the status of  $C_t$  cell.

In step C, the old cell will be multiplied by  $f_t$  so that it can forget things that are no longer needed, so new information that will enter can be easily added to the cell's memory status. This section is shown in Figure 3 (C) and Equation (4). In the final step, the output of  $h_t$  is shown in figure 3 (D), Equation (5), and (6). Output results are based on the state of the cell but in the state that is being filtered. Initially, the sigmoid layer was applied to the previous output  $h_{t-1}$  and  $x_t$  input to determine the  $O_t$  gate output value. The resulting value is between “0” and “1”, which indicates part of the cell state is output. Then the state of a cell  $C_t$  is changed by the tanh function to get the value between “-1” and “1”. The value of the changed cell status is then multiplied by the output value at the  $O_t$  gate, which ends with  $h_t$  output and this output will be used for the next step in the model.

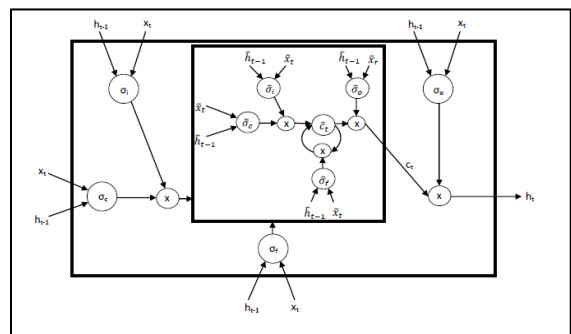


Figure 4: Nested LSTM architecture

### 3.3. Nested Long Short-Term Memory (NLSTM)

Figure 4 is an architectural drawing of a Nested LSTM [21]. Nested LSTM is a simple extension of the LSTM model that adding depth through nesting into the model. Inside Nested LSTM there are memory cells that make up internal memory and can only be accessed through external memory cells by applying a temporal hierarchy. The gate output in LSTM encodes the intuition that irrelevant memories at the current time step may still need to be remembered. Nested LSTM uses this intuition to create a temporal memory hierarchy. In Nested LSTM, access to internal memories is maintained in the same way, so that long-term information that is only situationally relevant can be selectively accessed. The equation in Nested LSTM can be described as follows:

$$\tilde{h}_{t-1} = f_t * C_{t-1} \quad (7)$$

$$\tilde{x}_t = i_t * \tanh(x_t W_{xc} + h_{t-1} W_{hc} + b_c) \quad (8)$$

$$C_t = \tilde{h}_{t-1} + \tilde{x}_t \quad (9)$$

$$\tilde{i}_t = \tanh(\tilde{x}_t \tilde{W}_{xi} + \tilde{h}_{t-1} \tilde{W}_{hi} + \tilde{b}_i) \quad (10)$$

$$\tilde{f}_t = \tanh(\tilde{x}_t \tilde{W}_{xf} + \tilde{h}_{t-1} \tilde{W}_{hf} + \tilde{b}_f) \quad (11)$$

$$\tilde{C}_t = \tilde{f}_t * \tilde{C}_{t-1} + \tilde{i}_t * \tanh(\tilde{x}_t \tilde{W}_{xc} + \tilde{h}_{t-1} \tilde{W}_{hc} + \tilde{b}_c) \quad (12)$$

$$\tilde{o}_t = \tanh(\tilde{x}_t \tilde{W}_{xo} + \tilde{h}_{t-1} \tilde{W}_{ho} + \tilde{b}_o) \quad (13)$$

$$\tilde{h}_t = \tilde{o}_t * \tanh(\tilde{C}_t) \quad (14)$$

$$C_t = \tilde{h}_t \quad (15)$$

Where,  $f_t$  is forget gate,  $\tilde{f}_t$  is inner forget gate,  $i_t$  is the input gate,  $\tilde{i}_t$  is inner input gate,  $O_t$  is the output gate,  $\tilde{o}_t$  is inner output gate,  $C_t$  is a memory cell,  $\tilde{C}_t$  is an inner memory cell,  $h_t$  is a hidden layer,  $\tilde{h}_t$  is an inner hidden layer,  $x_t$  is input when time "t,"  $\tilde{x}_t$  is inner input when time "t,"  $\sigma$  is sigmoid activation function,  $\tanh$  is hyperbolic tangent activation function,  $W_{xc} W_{hc}$  are weight matrices,  $\tilde{W}_{xi} \tilde{W}_{hi} \tilde{W}_{xf} \tilde{W}_{hf} \tilde{W}_{xc} \tilde{W}_{hc} \tilde{W}_{xo} \tilde{W}_{ho}$

are inner weight matrices  $b_c$  is bias vector and  $\tilde{b}_i \tilde{b}_f \tilde{b}_c \tilde{b}_o$  are the inner bias vector.

### 3.4. Support Vector Machine (SVM)

SVM is machine learning that is usually used for classification or regression. SVM is also a type of supervised learning. The main purpose of SVM is to determine data with decision boundaries and extend to non-linear boundaries using kernel tricks [22]. SVM is used in many applications such as word sentiment, categorization of text and documents, pattern recognition, face recognition, handwriting analysis, and binary classification. the idea behind SVM is to share data with the best method. The binary classification used to compile we need to classify 2 data sets. In multi-classification, the most frequent method is to create a one-versus-rest classifier (OVA) where each category is divided, and all other categories are combined and to choose the class that classifies collecting data with the largest margins. Divide the class into binary problems. The classifier learning step is carried out by all training data, taking certain class patterns as positive and all other examples as negative. Support Vector Machine has three main parameters, namely, C, gamma, and kernel. The kernel is always used as the Radial Base Function (RBF) because of its best performance [23]. While C and gamma are hyperparameters that have different values and produce different accuracy and results.

## 4. Research Methodology

### 4.1. Dataset Generation

We collect malware samples and track the behavior of malware using Cuckoo malware analysis [24]. The malware collection consists of samples collected from two primary sources: Virus Share [25] and GitHub / TheZoo [26]. We chose this source because it provided a large and varied sample Portable Executable (PE) file for evaluation. Because malware authors can use obfuscation and packers code for sub-vertical static analysis, we use dynamic malware analysis to collect data about malware behavior. Then, several tools allow tracking malware execution and gathering logs from the order of execution [27]. We use Cuckoo Sandbox, which is open-source and provides a controlled environment for executing malware. In the dataset experiment, that will be used as many as 13356 data, where the data is divided into three groups, namely training, validation, and testing.

Table 2: Description of malware dataset

Malware	Training Data	Validation Data	Testing data	Total
Adware	2159	719	719	3597
Backdoor	504	167	167	838
Packed	664	220	220	1104
Riskware	733	243	243	1219
Trojan	2484	827	827	4138
Virus	592	197	197	986
Worm	886	294	294	1474
Total	8012	2672	2672	13356

Table 2 shows the distribution of the amount of training and testing data used in this research. At the training stage, the model will be trained using 8012 data, while at the data testing stage will be tested using 2672 data. Experiments will be conducted on both models. Prediction of testing data will be an experimental result where the results will be described through a confusion matrix so that the accuracy of each model is obtained.

### 4.2. Word Embedding

We extracted the PE file by preprocessing the PE Headers and opcodes from the code section. To use this data in the classification process, we need to make numerical vectors with word embedding. The PE file is run in the Cuckoo sandbox which is a malware analysis tool. Can extract API calls from PE files during execution. The sandbox tool is configured on Ubuntu 18.04.2 LTS along with the Windows 7 virtual environment using the Oracle virtual box where the PE files are executed. Virtual environments help in such a way that malicious files are executed and behave in the same way as in a conventional system [6]. This is very helpful in understanding malware behavior when trying to infect a system.

During PE file execution, the Cuckoo sandbox generates log files. The log file contains snapshots taken during execution (behavior profile) [28]. This is done for every sample that is executed. Each sequence of API calls is recorded according to the

class label specified by Kaspersky [29] and VirusTotal [30]. We determined seven classes of malware (Adware, Backdoor, Packed, Riskware, Trojan, Virus, and Worm). The API call log that has been collected is always long and continuous. We will apply text mining with word2vec techniques. To select API calls that are relevant for classification. Word2vec helps identify a set of API calls that are more common in the malware class. This works in a way that if there is a word API call, it often appears in a class. But when it appears in many other classes, it is not a unique identifier and must be given a lower score. Only the words API calls with high scores or frequently appearing words are considered as PE file profile behavior.

Word2vec has two techniques, namely Skip-gram and Continuous Bag of Words (CBOW). This CBOW method takes the context of each word from the whole sentence or paragraph as input and tries to predict the word for word that fits the context. In contrast, the skip-gram model predicts the meaning of words after searching for their target words, and the author uses CBOW for this research. First, we did a mapping for seven labels and turned it into one-hot encoding. Then, the writer converts the whole sentence to the lower case and removes the punctuations. The next step is to create a word2vec embedding model generator to convert words to vectors with the specified model size.

```

p> M:
counts=Counter(word_bag).most_common(1200) # selecting only top 1200 most common words
print(counts)

words_list=[] # making words list out of tuples present in counts
for c in counts:
    words_list.append(c[0])

[('loadlibraryexw', 136606), ('localalloc', 123640), ('getprocaddress', 98383), ('getmodulehandlew', 79815),
('closehandle', 69078), ('createfilew', 68124), ('localfree', 59950), ('getsystemmetrics', 59885),
('mapiofileex', 52930), ('getmodulefilenamew', 52725), ('istolow', 52347), ('getcurrentthreadid', 52075),
('getmodulehandlew', 51668), ('getthreadlocale', 49396), ('getversionex', 48844), ('listriena', 47563),
('loadlibrary', 44979), ('freelibrary', 43570), ('loadresource', 43324), ('createfilemappingw', 41745),
('waitforsingleobjectex', 40288), ('waitforsingleobject', 39071), ('getversionexa', 39035), ('seterrormode',
37751), ('releasedc', 37417), ('comparestringw', 36537), ('gettickcount', 35830), ('disablethreadlibrarycalls',
35516), ('lsdbcsleadbyte', 34347), ('searchpathw', 33961), ('lstrcpw', 33138), ('loadcursorsw', 32191),
('lstrcpyna', 32089), ('loadlibraryexa', 30218), ('findresourcew', 29998), ('lockresource', 29646),

```

Figure 5: Code snippet for Word Bag

The next step is to create a Word Bag with the same number of words counted in various types of malware that is calculated to help determine how relevant a word is to a specific class or how often the word appears in the word bag. The code snippet for Word Bag is shown in Figure 5.

```

word='getversion'
print("Vecotr: ",model1.wv[word])

Vecotr: [ 2.59232450e+00  6.84048295e-01  6.26165152e-01 -1.40236092e+00
 1.06598151e+00  6.68558717e-01 -1.79842436e+00  3.80435050e-01
 1.46896768e+00  8.00294206e-02 -1.01184219e-01  1.30136698e-01
 9.99615102e-02  1.23285269e-02 -5.94269335e-01 -1.94395602e+00
 1.35559547e+00  1.02511263e+00  1.47519684e+00  1.61271095e-01
 -3.92647773e-01 -6.11079276e-01 -1.14168262e+00 -1.34283960e+00
 -4.08858657e-01 -1.46569645e+00 -1.60189641e+00 -1.31877005e+00
 7.46764421e-01  3.13887417e-01 -2.35597193e-01 -4.39115703e-01
 -1.52561033e+00  4.51700360e-01 -4.99426723e-01 -3.69457185e-01
 -1.37319434e+00  1.59534901e-01 -5.51769376e-01  3.89835179e-01
 6.48715258e-01  1.29709542e+00 -1.66487455e-01 -2.22325325e+00
 5.04632413e-01  6.96626082e-02 -4.84076947e-01 -6.93618238e-01
 2.18451515e-01 -3.01782936e-02  1.77319932e+00  7.37498820e-01
 -1.01685035e+00  5.78746140e-01 -7.53057301e-01  4.32285607e-01

```

Figure 6: Code snippet from changing word to vector

The word was changed to vector using the word2vec embedding model that was created, as shown in Figure 6. After getting the vector for the word, the average value of the vector (mean) is taken and multiplied (multiplied) by the frequency of

words in the class and label. The following entire preprocess process is summarized below:

- Enter a sentence and repeat each word
- For each word, it will be changed to represent a numeric / vector.
- Take the mean vector and multiply with the number of classes and add them as features.
- Pad the sentence to fixed-length 128 then move to the next sentence.

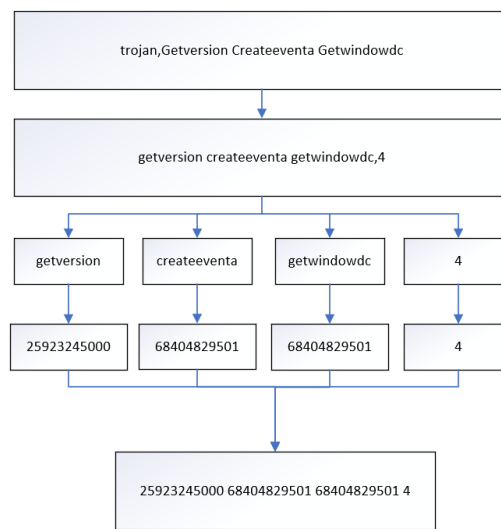


Figure 7: Illustration of preprocessing stages

After going through this process, as illustrated in figure 7, a fixed length of 128 vectors is obtained as a feature for each sentence. If a sentence has more than 128 words, the word will be truncated, and if it has less than 128 words, then padding "0" will be added so that each sentence has the same length.

After the feature making process, data mining classification is applied using a classification approach. We use Long Short-Term Memory (LSTM). Based on the type of API call chosen to describe a particular class of malware, the classification approach helps in concluding whether the file is malicious by determining the class in which the malware is. Because the process ends with the accuracy of determining the class in which the file is located after behavioral detection. All PEs have a direct relationship with the Operating System (OS) via the system calls API. This shows that API calls can easily notify malware behavior when attempting to execute.

### 4.3. Deep Learning Model

Deep Learning is one area of artificial neural networks to deal with problems on more large datasets. Deep Learning provides a very compelling architecture for supervised learning. By adding more layers to the deep learning model, it can do better at represent labeled malware data. To implement Deep Learning techniques for malware classification, a computer-based program is needed that can do computing. Therefore it is necessary to design an algorithm that can support the development of programs for this research. The algorithm used in this study is divided into three main parts, namely the training algorithm, the testing



algorithm, and the classification algorithm. These three algorithms follow the concept of writing code with API and the basic theory of Machine Learning for feature learning. In the training algorithm, five main stages will be carried out, namely the stages of Data Augmentation, Load Training Data, Modeling Long Short-Term Memory (LSTM), Training Model, and Final Weight Storage.

The LSTM model has several layers, including the embedding layer, LSTM layer, and Output layer. The input of the model is the preprocessing text that has been transformed into numeric where the input length is 128, where each number or vector represents a word, at the embedding layer, the input will be transformed into a vector that has a length of 128 vectors. Furthermore, LSTM consists of 3 gates, which will process each input vector to produce 128 vectors and where each output is connected to the output layer. At the output layer, there are seven neurons. each of these neurons has softmax activation to make a value in each classification. The classification prediction results are the highest output value.

The Nested LSTM model made consists of several layers, including the Embedding layer, Nested LSTM layer, and Output layer. Similarly, the LSTM input model of the Nested LSTM model is the preprocessing text that has been changed to numeric, where the length of the input is different where each number represents a word. At the embedding layer, the input will be transformed into a vector that has a length of 128. Furthermore, the Nested LSTM cell consisting of 3 gates (depth = 2) will process each input vector to produce 128 output vectors where each output is connected to the output layer. At the output layer, the same number neurons like LSTM and each of these neurons have softmax activation, which results in a value for each classification.

Support Vector Machine Model has three main parameters, namely C, Gamma, and Kernel. The kernel is always used as the Radial Base Function (RBF) because of its best performance. C and gamma are hyperparameters that have different values between the two and produce different accuracy and results. We need to find the best C and gamma values. That is why we use GridSearch. In GridSearch, we make all possible C and gamma combinations and then choose the one that has the best. Sklearn has a GridSearch Cross-Validation (CV) function that takes the SVM model, the Cs and Gammas grid parameters, and the number of folds. The number of folds means that the data will be divided into that many folds. In this case, it is three and then is trained on two and tested on one.

#### 4.4. Evaluation Design

In this research, the dataset will be used as many as 13356 data, where the data is divided into three groups, namely training, validation, and testing. Data need to be converted in numerical value before going into the deep learning model. So the first step is to convert labels to one-hot encoding. After that, sentences are being converted into lowercase and remove punctuations to create a clean word2vec model using CBOW. The deep learning model will be trained using 8012 data, while the data testing stage will be tested using 2672 data. The LSTM model uses Adam optimizer using 64 batch sizes and 30 epochs because after several test we

found that this combination works best on accuracy and added with 512 dense layers with 20% dropping units rate to prevent overfitting and also using softmax for classification. Whereas the Nested LSTM model uses Adam optimizer using 64 batch sizes and 50 epochs added with 1024, 2048, and 7 dense layers with recurrent dropping to prevent overfitting and also using softmax for classification. The Support vector machine uses the RBF kernel and Grid Search Cross-Validation for hyperparameter tuning to find the best value for the C parameter and gammas for the training model. In this research, performance will be measured based on the level of accuracy, recall, precision, and f1-score achieved to measure the performance, the results of the evaluation will be set forth in the form of a confusion matrix. The confusion matrix contains information from actual classifications and predicted classifications [31]. All methods were implemented on Python 3.6, Jupyter Notebook 6.0.3, Tensorflow 1.15.0 version, Intel Core i5-6400 (with 16 GB RAM), and Nvidia GeForce GTX 1050 Ti GPU.

## 5. Experimental Results

### 5.1. Training Result

After the two models are made, the model is trained using 13356 training data. After several trials, we decided to use 30 epochs with 64 batch sizes for the LSTM model and 50 epochs with 64 batch sizes for the nested LSTM model because they produce high accuracy in training.

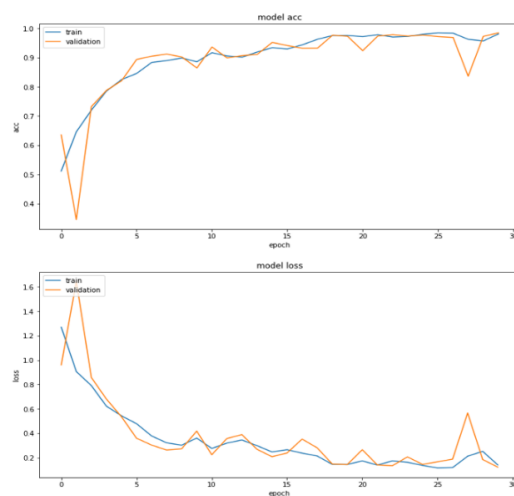


Figure 8: Accuracy and loss of the LSTM model

Figure 8 shows the training process of the LSTM model. The accuracy and loss of the model during training are indicated by the line above, the blue line shows the data in training, and the yellow line shows the validation data. The accuracy and loss of the LSTM model always increase from the start. At the end of the epoch, the accuracy reached 98%, and the loss was 0.14%.

Figure 9 shows the results of the training and validation process of the Nested LSTM model. The results of the accuracy and loss accuracy of this model can be seen in the line above, the blue line shows the data in training and the yellow line shows the validation data. the level of accuracy and loss in this model is quite good with an increase from the beginning to the end of the



test. At the end of the epoch, the accuracy reached 93.1%, and the loss was 0.18%.

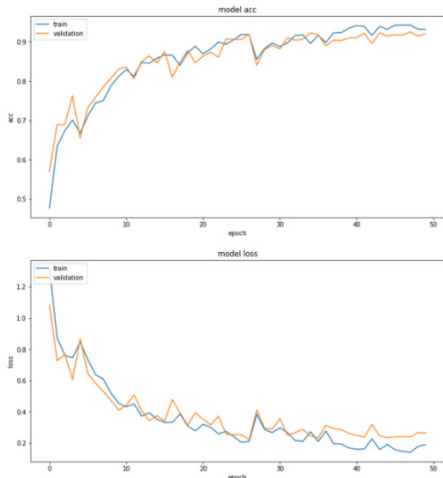


Figure 9: Accuracy and loss of the Nested LSTM model

### 5.2. Testing Result of LSTM Model

After the LSTM and Nested LSTM models are created and trained, the SVM is built and tested only as a benchmark to compare the two models above. All three models were tested using 2672 test data.

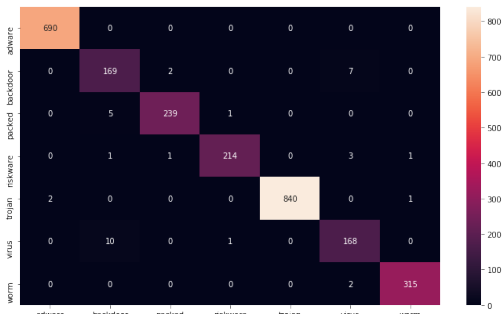


Figure 10: LSTM confusion matrix

Figure 10 shows the results of LSTM model testing, where the test results are shown using the confusion matrix. As we can see in every malware label, there is not much miss. This is because the loss rate of testing data is only 0.18%. The biggest label miss here is a virus where 30 labels are considered backdoor. The rest showed outstanding results with miss under 10. Thus, the testing accuracy obtained from the LSTM model is 98.61%.

Table 3: Details of the LSTM model results

Classification	Precision	Recall	F1-score
Adware	100%	100%	100%
Backdoor	95%	91%	91%
Packed	98%	99%	98%
Riskware	97%	99%	99%
Trojan	100%	100%	100%
Virus	94%	93%	95%
Worm	99%	99%	99%
<b>Average</b>	<b>97.57%</b>	<b>97.29%</b>	<b>97.43%</b>
<b>Accuracy</b>	<b>98.61%</b>		

Table 3 shows all the precision, recall, and f1-scores of the LSTM model from each malware. The LSTM method obtained an average precision of 97.57%, a recall of 97.29%, and an f1-score of 97.43%.

### 5.3. Testing Result of Nested LSTM Model

Figure 11 shows the results of Nested LSTM model testing, where the test results are shown using the confusion matrix. As we can see in every malware label, there is not much miss. It is because the loss rate of testing data is only 0.18%. The biggest label miss here is a backdoor where 66 labels are considered as a virus. The rest showed outstanding results with miss under 20. The testing accuracy obtained is 93.11%.

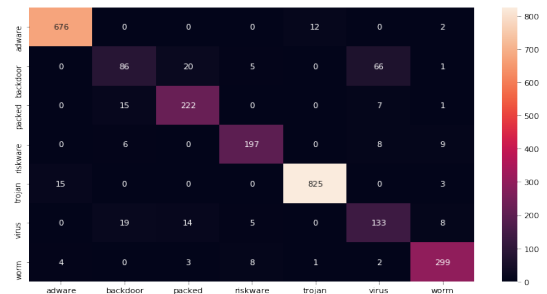


Figure 11: Nested LSTM confusion matrix

Table 4: Details of the Nested LSTM model results

Classification	Precision	Recall	F1-score
Adware	98%	100%	97%
Backdoor	55%	72%	67%
Packed	93%	71%	83%
Riskware	88%	87%	89%
Trojan	100%	99%	97%
Virus	65%	75%	76%
Worm	92%	92%	93%
<b>Average</b>	<b>84.43%</b>	<b>85.14%</b>	<b>86.00%</b>
<b>Accuracy</b>	<b>93.11%</b>		

Table 4 shows all the Precision, Recall, and F1-scores of the model from each classification. The Nested LSTM method obtained an average precision of 84.43%, a recall of 85.14%, and an f1-score of 86.00%.

### 5.4. Testing Result of Support Vector Machine Model

We also developed Support Vector Machine (SVM) so that it can be used as a comparison or benchmark. To see which ones perform better using the same word embedding method.

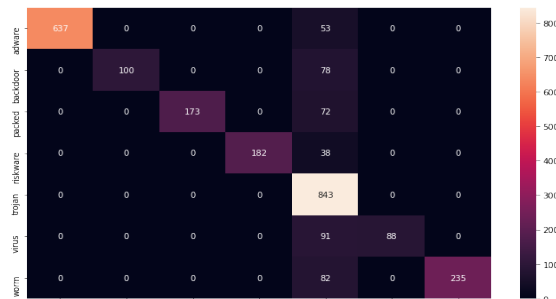


Figure 12: SVM confusion matrix

Figure 12 shows the results of the SVM testing model, where the test results are shown using the confusion matrix. As we can see in every malware label, there is much miss. For example, the Virus miscalculated up to 91 labels. Followed by almost every label more than 30 miss labels because of a low level of accuracy, and this is why we propose to use deep learning methods because this type of data is not suitable for SVM. The testing accuracy obtained is only 84.50%

Table 5: Details of the SVM model results

Classification	Precision	Recall	F1-score
Adware	94%	100%	97%
Backdoor	50%	100%	67%
Packed	70%	100%	83%
Riskware	85%	100%	92%
Trojan	100%	66%	80%
Virus	40%	100%	57%
Worm	78%	100%	87%
<b>Average</b>	<b>73.86%</b>	<b>95.14%</b>	<b>80.43%</b>
<b>Accuracy</b>	<b>84.50%</b>		

Table 5 shows all the results of the classification label, precision, recall, and f1-scores of the SVM model from each classification. The SVM method obtained an average precision of 73.86%, a recall of 95.14%, and an f1-score of 80.43%.

### 5.5. Summary of Testing Results

This section presented the result of all methods, including our proposed method, and compared it to other existing methods, which is SVM. Our proposed methods can overcome others. The results are shown in Table 6.

Table 6: Results comparison of all models

Method	Accuracy	Precision	Recall	F1-score
LSTM	<b>98.61%</b>	<b>97.57%</b>	<b>97.29%</b>	<b>97.43%</b>
Nested LSTM	93.11%	84.43%	85.14%	86.00%
SVM	84.50%	73.86%	95.14%	80.43%
CNN	96.60%	95.71%	95.12%	95.62%
DRN	86.54%	84.97%	84.23%	84.67%
DBN	92.60%	96.30%	89.60%	92.80%
DGCNN	96.87%	88.79%	92.83%	90.76%
DL4MD	95.64%	93.06%	94.60%	94.52%

Table 6 shows the overall comparison of all methods. We also make plot early stopping in LSTM and Nested LSTM train process so that we can take the best accuracy model when the training process happens. Overall, the table above shows that the LSTM model produces the best accuracy of 98.61% among the three methods, although the difference in accuracy does not differ significantly from the Nested LSTM. Both LSTM and Nested LSTM methods are still better than SVM methods. It shows that the deep learning method is far more accurate compared to ordinary machine learning methods.

## 6. Conclusion and Future Work

In this paper, we investigate the effectiveness of malware system call sequences that transformed into vectors and use word2vec as word embedding and then enter the LSTM layer repeatedly for the classification process with non-linear activation

functions like Softmax. We have also carried out various experiments with different parameters, network structures, and added early stopping plots to get the best model accuracy in the training process. The design of the model is also evaluated using different methods such as Nested LSTM and SVM as benchmarks. From the three models, it can be concluded that the LSTM method gets the highest accuracy reaching 98.61% in the real-world data set. Overall, LSTM is included as the most effective approach to learning long-range dependencies in cybersecurity tasks and more appropriate methods for detecting malware through system call sequences.

From this research, it can be concluded that the experimental results with the LSTM network are straight forward. Still, we have not tried to use more complex LSTM networks, such as use many different layers, use more automatic decoders or use word embedding techniques other than word2vec. It because such a network architecture will cost us more and more complex preprocessing, network architecture, and a clean dataset probably will improve the results.

## References

- [1] N. Aziz, Z. Yunos, and R. Ahmad, "A management framework for developing a malware eradication and remediation system to mitigate cyberattacks," in *Lecture Notes in Electrical Engineering*, 481, 513–521, 2019.
- [2] R. Bavishi, M. Pradel, and K. Sen, "Context2Name: A Deep Learning-Based Approach to Infer Natural Variable Names from Usage Contexts," 2018.
- [3] C. Raghuraman, S. Suresh, S. Shivshankar, and R. Chapaneri, "Static and dynamic malware analysis using machine learning," in *Advances in Intelligent Systems and Computing*, 1045, 793–806, 2020.
- [4] Abbasi, "Leveraging behavior-based rules for malware family classification," Dec. 2019.
- [5] H. Lim, "Detecting Malicious Behaviors of Software through Analysis of API Sequence k-grams," *Comput. Sci. Inf. Technol.*, 4, no. 3, 85–91, 2016.
- [6] Y. Ki, E. Kim, and H. K. Kim, "A novel approach to detect malware based on API call sequence analysis," *Int. J. Distrib. Sens. Networks*, 2015.
- [7] V. Zenkov and J. Laska, "Dynamic data fusion using multi-input models for malware classification," 2019.
- [8] M. Imran, M. T. Afzal, and M. A. Qadir, "Malware classification using dynamic features and Hidden Markov Model," in *Journal of Intelligent and Fuzzy Systems*, 31(2), 837–847, 2016.
- [9] A. F. Agarap, "Towards Building an Intelligent Anti-Malware System: A Deep Learning Approach using Support Vector Machine (SVM) for Malware Classification," 2017.
- [10] M. F. Rafique, M. Ali, A. S. Qureshi, A. Khan, and A. M. Mirza, "Malware Classification using Deep Learning based Feature Extraction and Wrapper based Feature Selection Technique," 1–20, 2019.
- [11] Y. Lu, G. Jonathan, and L. Jiang, "Deep Learning Based Malware Classification Using Deep Residual Network," 2019.
- [12] A. Oliveira, U. N. De Julho, and U. N. De Julho, "Behavioral Malware Detection Using Deep Graph Convolutional Neural Networks," 1–17, 2019.
- [13] Z. Zhang, C. Chang, P. Han, and H. Zhang, "Packed malware variants detection using deep belief networks," *MATEC Web Conf.*, 309, 02002, 2020.
- [14] K. HE and D.-S. KIM, "Malware Detection with Malware Images using Deep Learning Techniques," 2018.
- [15] W. Hardy, L. Chen, S. Hou, Y. Ye, and X. Li, "DL4MD: A Deep Learning Framework for Intelligent Malware Detection," *Proc. Int. Conf. Data Min.*, 61–67, 2016.
- [16] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," *Adv. Neural Inf. Process. Syst.*, 3111–3119, 2013.
- [17] D. Meyer, "How exactly does word2vec work?," *Uoregon.Edu, Brocade.Com*, 1–18, 2016.
- [18] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, 9, no. 8, 1735–1780, 1997.
- [19] A. Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and

- Long Short-Term Memory (LSTM) network,” *Phys. D Nonlinear Phenom.*, 404, p. 132306, Mar. 2020.
- [20] F. Miedema, “Sentiment Analysis with Long Short-Term Memory networks,” 1–17, 2018.
- [21] J. R. A. Moniz and D. Krueger, “Nested LSTMs,” *J. Mach. Learn. Res.*, 77, 530–544, 2017.
- [22] Y. Ahuja and S. Kumar Yadav, “Multiclass Classification and Support Vector Machine,” *Global Journal of Computer Science and Technology Interdisciplinary*, 12(11), 14–19, 2012.
- [23] C. Brew, “Classifying ReachOut posts with a radial basis function SVM,” 2016.
- [24] L. Wang, B. Wang, J. Liu, Q. Miao, and J. Zhang, “Cuckoo-based malware dynamic analysis,” *Int. J. Performability Eng.*, 15(3), 772–781, 2019.
- [25] “VirusShare.com.” [Online]. Available: <https://virusshare.com/>. [Accessed: 18-Apr-2020].
- [26] G. D. Webster, Z. D. Hanif, A. L. P. Ludwig, T. K. Lengyel, A. Zarras, and C. Eckert, “SKALD: A scalable architecture for feature extraction, multi-user analysis, and real-time information sharing,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 9866 LNCS, 231–249, 2016.
- [27] T. K. Lengyel, S. Maresca, B. D. Payne, G. D. Webster, S. Vogl, and A. Kiayias, “Scalability, fidelity and stealth in the DRAKVUF dynamic malware analysis system,” *ACM Int. Conf. Proceeding Ser.*, 2014-Decem, no. December, 386–395, 2014.
- [28] S. Jamalpur, Y. S. Navya, P. Raja, G. Tagore, and G. R. K. Rao, “Dynamic Malware Analysis Using Cuckoo Sandbox,” in *Proceedings of the International Conference on Inventive Communication and Computational Technologies, ICICCT 2018*, 2018, 1056–1060, 2018.
- [29] “Kaspersky Cyber Security Solutions for Home & Business | Kaspersky.” [Online]. Available: <https://www.kaspersky.com/>. [Accessed: 18-Apr-2020].
- [30] “VirusTotal.” [Online]. Available: <https://www.virustotal.com/gui/home/upload>. [Accessed: 26-Apr-2019].
- [31] A. K. Santra and C. J. Christy, “Genetic Algorithm and Confusion Matrix for Document Clustering.” 2012.

## The Importance of Sustainability Audit Report in Go Public Companies Sector, in Indonesia

Bambang Leo Handoko\*, Ang Swat Lin Lindawati

*Accounting Department, Faculty of Economics and Communication, Bina Nusantara University, Jakarta, 11480, Indonesia*

---

### ARTICLE INFO

*Article history:*

*Received: 09 April, 2020*

*Accepted: 07 July, 2020*

*Online: 22 July, 2020*

---

*Keywords:*

*Sustainability*

*Reporting*

*Audit*

*Indonesia*

*Corporation*

---

---

### ABSTRACT

*Sustainability is considered new for Indonesian state-owned enterprises, also has become an obligation since two years ago. More and more people are aware of the environment for the next generation, the survival of mankind in the future. Sustainability issues are no exception in Indonesia. Companies publish sustainability reports as well as financial and annual reports. This sustainability report creates a need, namely the need for audits for sustainability reports. This is often referred to as a sustainability audit or assurance report. It is a qualitative descriptive survey. Researchers collect data from the literature and study observations as well as from interviews with stakeholders. The results of our study were obtained because the increasing number of public companies reporting on sustainability requires sustainability reporting audits in Indonesian public companies. The advantage of a sustainability audit report is that, in addition to its reputation or image, it can also add value to the company's planning, structure and accountability. Brand name of the listed company. However, there are weaknesses, the weakness lies in comparability and costs.*

---

### 1. Introduction

The issue of sustainability in accountability is an absolute mandatory when relying on business decisions to participate in nature conservation. There is an important reporting requirement to support this accounting report, because it requires hard work and the role of companies that integrate accounting and environmental assessment into their work for the sustainability of their profession. Sustainability accounting was an approach or method used in organizations to increase sustainability. Sustainable development was defined in 1987 by the United Nations Commission on Brundtland [1].

One of the main challenges in the field of accounting is sustainability, which began with a lack of understanding of the definition of sustainable development [2]. Another challenge is that it is difficult to combine control and the enterprise as the ultimate goal, since, as it has always been, usually as a traditional enterprise, it should achieve profit maximization, so that it can be seen as an investment in modal capital for the benefit of understanding. While some research has shown that sustainable business practices can catch up to promote greater long-term profitability, maintain the apparent real relationship, and that approval is important, both organizations need to consider factors that are considered increasingly important.

To report on their efforts, companies can demonstrate their commitment to corporate social responsibility (CSR) by the following means: High-level participation (CEO, Board of Directors), investment policy, program, signatories to voluntary standards, principles (UN Global Compact - Ceres Principles), reports (Global Reporting Initiatives) [3].

The triple bottom line concept requires that corporate responsibility is based on the importance of stakeholders in understanding. In this case, the term "stakeholders" refers to those who are directly or indirectly affected by the actions of the company [4]. Employees, customers, suppliers, local residents, government agencies and creditors are examples of key stakeholders. The audit should also be able to provide stakeholders with a post-implementation assessment, as in the research conducted by [5].

One of the analytical works is the sustainable balanced scorecard model. In a study carried out by [6], based on both popular BSC practices and the latest BSC developments, the BSC strategy was implemented to measure and support sustainable practices in a company, which is why the concept is called BSC sustainability. This means that not only finances are at the core of an organization's central BSC strategy, but also social and environmental objectives are integrated (like the former GRI concept, but their measurement and strategy through the BSC approach), or in other words, BSC sustainability is integrated.

---

\*Corresponding Author: Bambang Leo Handoko, +6289522675001, bambang.handoko@binus.edu



Relationships based on the creation of economic value and environmental impacts added by the assessment form: Life Cycle Investments and Life Cycle Impacts. This evaluation links the Balanced Scorecard to the environmental systems of the companies' environment by integrating them in different ways. In Indonesia, the form and reports are influenced by accounting standards established by various global experts. In Indonesia, the form and format of the reports are influenced by accounting standards established by various global experts [7].

This national and international phenomenon clearly implies that today's companies can no longer focus solely on profits. In 1988, the concept of Triple Bottom Line (TBL or 3BL) or 3P - People, Planet and Profit - was introduced. In short, the three pillars that measure the value of business success using three criteria: economic, environmental and social.

In fact, this approach has been widely used since the beginning of 2007, as the development of the full cost accounting method has become widespread in public sector companies. In private sector companies, the application of corporate social responsibility (CSR) is also a form of implementation of the CTA. The concept of TBL implies that companies give priority to the interests of stakeholders (all parties affected and concerned by the company's activities) over the interests of shareholders (the shareholders [4]. It was designed in the minds of companies to make them aware of the need for non-market valuation factors, i.e. the integration of natural capital into their business capital [8].

### 1.1. Research Problem

Based on the context, the following problems can be formulated in this study

- What are the obligations of undertakings to make public offers in to report on sustainability reporting?
- Why is a review of the sustainability report necessary?
- What does the verification of a sustainability report look like?
- What are strengths and weaknesses of the sustainability audit?

### 1.2. Purpose and Benefit

The objective of our study was to provide information on the reports of operating companies, and thus to determine why an audit is necessary to test the quality of sustainability reporting. The benefits of this research are as follows:

Theoretical benefits; the document should increase knowledge and references concerning the image of sustainability reports.

Practical benefits: The document should be useful to companies in improving the quality of sustainability reports and can be considered.

## 2. Theoretical Framework

### 2.1. Sustainability Reporting

The quality of the presentation of the information compiled by management is a crucial factor in building trust and relationships with stakeholders. The sustainability report [9] has a positive influence, not only on the stakeholders, but also on the [www.astesj.com](http://www.astesj.com)

management and the company internally. [10] states that sustainability reports are media and a means for management to show stakeholders the company's performance and results.

An audit on sustainability reports is carried out to show the different perspectives of companies by showing that companies are involved and participate also in the context of sustainability. Auditing is an important factor in ensuring that the information presented is relevant and responsible both financially and non-financially. According to [11], management should not only focus on priorities financial activities, but should also be able to focus on the social and environmental activities of the company. In this way, the scope of the audit becomes broader, which has an impact on increasing the added value of the company.

The attention of public entities on the importance of sustainability reporting should be done through the GRI assessment. Companies must be able to assess the implementation and validity of sustainability reports. [12] which argue that honest and transparent disclosure of information on social and environmental activities impacts on the value of the company so that it influences business decisions made by stakeholders.

### 2.2. Triple Bottom Line Concept

TBL focuses on all aspects of business, not only profit, but also on preparing and preserving the earth's ecosystems for the future [4]. The description of TBL includes:

1. People. Focus on human resources through the protection of labor rights and duties and compliance with applicable labor regulations. Companies must be able to offer decent wages and working hours and provide health and education services for the workforce.

2. Planet. Problem area related to energy saving. It is also a matter of minimizing waste by treating it in advance in such a way that it does not harm the environment and ecosystems that can affect the quality of human life and flora and fauna.

3. Profit. It is emphasis that these companies should not only strive for profit, but also create fair and ethical trade in their business activities.

### 2.3. Sustainability Auditing

The world of auditing has, of course, been challenged by the evolution of processes in the industry; auditors must begin to understand the use of technology in order to be able to verify all the claims made in the company. New technologies that can defuse the audit work are beginning to evolve. One of them is continuous auditing, in which audits are conducted in real time without disrupting business operations and by reducing the time between initial reports and audit results [13].

However, the audit continued to identify various problems in its performance, with samples being taken for comparison in Brazil, South Africa, China and the United States for some of the studies. It has given the constraints that emerged in the conduct of the ongoing audits, the need for equal treatment in several countries [14].

Sustainability auditing involves the collection of automated audit indicators and evidence of company transactions, processes,

controls and IT systems by internal or external auditors at specified intervals or to ensure continuity. The information obtained makes it easier for auditors to ensure that they are not in breach of the law [15]. Sustainability auditing can be used as an early warning system to detect control deficiencies earlier than audits using traditional approaches.

Sustainability auditing differs from continuous monitoring. Continuous monitoring is the automatic feedback mechanism for management to ensure that controls and systems are functioning properly in accordance with design and transaction processes [16]. Management can use this information to make new arrangements for new audit activity by using analysis to identify performance gaps or unusual transactions that may indicate a control failure. Ongoing monitoring is beneficial in providing benefits to management and in developing the business by increasing the ability to monitor and manage the environment without compromising the optimal performance of the business.

#### 2.4. Sustainability Auditing in Indonesia

In Indonesia, there are regulations for business activities related to environmental and social issues, including the Partnership for Environmental Development. To this end, the company has begun to produce a sustainability report to demonstrate its responsibility in the development of social and environmental activities and its ability to be accountable to stakeholders [17].

According to [18], the Certified Sustainability Reporting Specialist (CSRS) is also used as an additional infrastructure for sustainability reporting. The CSRS is responsible for preparing a sustainability report by presenting the actual activities of the company.

In [19], the assessment of the quality of sustainability reporting in Indonesia shows that investors feel that the company's efforts to disclose non-financial information have not been fully implemented.

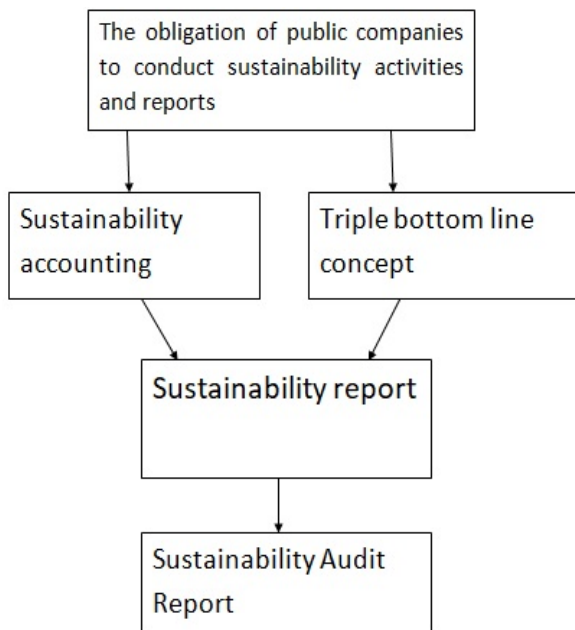


Figure 1: Research Framework

In Indonesia, there is one organization that is active in the field of sustainability reporting, namely the NCSR or the National Center for Sustainability Reporting. The main activity of NCSR is to advocate for reporting in accordance with the guidelines established by the Certified Public Accountants Association (ACCA) by working with the Indonesian Ministry of Environment.

The novelty of our research is the review of an audit of sustainability reporting in Indonesia. This is a new concept for developing countries that, like Indonesia, thus we connected our sustainability report with point of view of OJK. Here is an overview of the theoretical framework of this research

### 3. Research Methodology

This research is qualitative research, qualitative descriptive research. Qualitative because the researchers want to describe or paint facts or a visible condition or symptoms. Qualitative descriptive research is research that describes or depicts the objects of the researchers based on facts that appear or are as they are.

Our method is to conduct a literature research on sustainability reporting, sustainability auditing and their development in Indonesia. We try to illustrate the phenomenon of auditing sustainability reports in Indonesia. The method of data collection consists of literature research, observations and interviews with parties involved in sustainability auditing. We conducted in-depth interviews with various departments; public enterprises, state supervisory authorities and also auditors of auditing companies.

### 4. Research Analysis

#### 4.1. Reporting Obligations of Go-Public Companies Sustainability Report

The business is the place where the activities of producing goods or services are carried out. In principle, the main objective of the enterprise is to achieve maximum profit. However, in order to achieve these objectives, companies often also ignore the environmental and social impacts of environmental or economic activities.

Some recent companies in Indonesia are beginning to balance the benefits of environmental orientation and improvement. The 3P (Triple Bottom Line) concept is seen as a pillar in the development of sustainable businesses and the achievement of sustainable development. Growing with these 3P concepts by encouraging companies to adopt CSR and thus become an integral part of the future survival of business [20].

With the development of CSR, companies began to realize that they had to publish a report based not only on financial conditions, but also on the provision of social and environmental information, and this report was then called a sustainability report. This sustainability report is based on the guidelines of the Global Reporting Initiative (GRI) and is prepared separately from financial reports or annual reports.

In Indonesia, the sustainability report is still voluntary and is in the process of being introduced; there are still a few companies that are interested in disclosing the sustainability report. In 2016, the number of listed companies that disclosed their sustainability report to ISRA was only 31, representing 7.35% of all IDX-listed

companies [20]. This is due to the fact that the company is not aware of the benefits of sustainability reporting disclosure.

The company is still not aware of the disclosure of sustainability reports. Out of the 438 companies, only 25 have produced such reports. Several factors make companies reluctant to produce sustainability reports. The first company is not transparent in the way it conducts its business and has no obligation to be a company with good corporate governance.

#### 4.2. Need of Sustainability Audit Report

The Sustainability Report is a report prepared by a company to measure, disclose and report on the company's efforts to measure, disclose and report on how it is accountable to all stakeholders in the company's work for sustainable development. Bapepam-LK (OJK/Similaires predecessor of the SEC in Indonesia) has published regulatory activities in its reports. By implementing the sustainability report, it is expected that companies will be able to develop sustainable growth based on business ethics.

The process of sustainability reporting is implemented through 5 mechanisms [21], namely

- Compilation of corporate guidelines. In this case, the company creates policies related to sustainable development and then publishes the policies as they appear.
- Pressure on the supply chain. The public's expectation that the company will provide environmentally friendly products and services also puts pressure on the company to produce performance standards and sustainability reports for its suppliers and their livelihoods.
- Stakeholder involvement. The need for the involvement and support of all relevant stakeholders
- Voluntary codes. In this mechanism, the Community encourages companies to develop aspects of sustainability performance and requires companies to report on the implementation of sustainability. If the company has not yet implemented, it must make a statement.
- 5 Other mechanisms include rating and benchmarking, taxes and subsidies, tradable permits, and commitments and limitations.

Sustainability reports can be published separately from the integration into annual reports. Below are some of the reasons why companies report on sustainable development separately from annual reports:

- a. The sustainability report is a means of communicating with stakeholders to convey the message that the company has achieved sustainable development.
- b. To maintain a good (positive) image of the stakeholders.
- c. Strive for legitimacy from stakeholders.

The audit of the sustainability report aims to provide a comprehensive picture of corporate governance, focusing on corporate management, performance and reporting in the context of sustainability [13]. The corporate sustainability audit report has the following specific objectives

- a. Assess how the organization identifies sustainability issues that are consistent with the organization's activities,

- b. To see how companies, respond to management on other issues related to the company's operations
- c. Conduct assessments of the sustainability information and disclosures provided by management.

The subsequent audit report is not only used to assess the sustainability information provided by management, but may also be useful for internal company processes [13]. Advantages of auditing corporate sustainability reports for internal company processes, namely

- a. Improving sustainability and business performance
- b. Evaluation of the commercial activities of enterprises in relation to the effects caused by commercial activities.

#### 4.3. Advantage of Sustainability Auditing

1. Compare your practices with industry best practices

This is perhaps the most obvious benefit of a sustainability audit. There is no chief sustainability officer or sustainability manager who has the time to keep up with the latest developments in the world of sustainability [20]. Proper sustainability audits should provide this benefit. It is very important that the working framework and categories of sustainability are comprehensive. If this is the case, your organization should be able to compare them with industry leaders in sustainability.

2. Planning, structure and accountability

The big challenge facing many organizations is that they want to implement a comprehensive sustainability plan, but ongoing initiatives are often integrated into operational initiatives, facilities or information technology. The sustainability audit identifies these initiatives and provides an opportunity to explain them in the sustainability program. This enables effective planning, a structured approach and accountability that extends beyond organizational boundaries.

3. Identifying the new sustainability initiative

One of the most outstanding benefits of SIAs is that it allows a tactical organization to identify tithing initiatives in areas that were not previously considered. This increased awareness is translated into enthusiasm, innovation and unexpected benefits throughout the organization

4. Time periods and business-to-business tracking

One complaint about sustainability management is the lack of "time frame" and "tracking and accountability" of companies. After an ongoing review, your organization will have a comprehensive picture of all your green activities. If your organization has several awards, you now have the opportunity to compare them using the same criteria. The same applies to comparing operations from period periods.

5. Improved sustainability performance

An ongoing audit can enable your organization to improve its continuous overall performance. You have all the information you need to improve the planning, implementation and follow-up of various initiatives. In this way, you will be prepared to realize the key benefits of sustainability listed below.

- Cost - The natural outcome of most sustainability initiatives is cost reduction. Look for ways to reduce consumption (raw materials, energy), increase the efficiency and effectiveness of business practices, and explore new and innovative products and services that all lead to operational performance.
- Compliance - Compliance risk is a growing challenge for companies. The number and complexity of international, national and regional regulations that address the environmental impact of products, services and business activities is increasing. Companies can make them take enough risk by not managing these issues effectively.
- Reputational risk - Reputational risk can be defined as an action, event or situation that may have a negative or positive impact on a company's reputation. Non-compliance, unethical conduct, and the perception of having to address all these issues are representative of the ways in which the organization can have a negative impact on its reputation. Apart from financial risks and litigation, reputational risk in the global marketplace is the biggest risk for the long-term organization.
- Competitive advantage - Sustainable practice can provide organizations with significant competitive advantages in the form of reputation, lower costs, better compliance, higher levels of innovation and many others.
- Innovation - One of the key benefits of embarking on sustainability activities is innovation. At the heart of the process leading to better sustainability, organizations need to see that newcomers are running their businesses, what goes into their products and how their products impact the environment and the society they sell.

#### 4.4. Disadvantage of Sustainability Auditing

##### 1. Comparability

Comparative challenges include assessing and measuring the appropriate environmental and social activities of the company, how to integrate the information into financial measurements, and then effectively reporting to investors and other stakeholders.

##### 2. Costs

The weakness of both is the additional time and cost involved in collecting information, organizing and reporting. The company may need to use additional equipment to meet CSR reporting requirements.

##### 3. Contradiction of the critics

These challenges may be small, as the trend continues to grow in popularity, and anyone who makes sense should recognize the potential ethical problems of the past that have been identified by the company and the corrective actions taken.

#### 4.5. Proposed sustainability report audit form

One of the forms is to use the form of an independent declaration of assurance where the declarations containing the assurance provider's statement on the content of the sustainability report of the audited company. This declaration is independent and can be justified [21]. The declaration should contain related information, some of which is

The intended stakeholder classifies this report as a decision support tool; the first information to be included in the independent statement of assurance is usually an independent statement of assurance addressed to the management of the audited undertaking.

The client depends on the presentation and the auditors on the audit; the next information that has to be included in the independent statement of assurance is the obligation of the audited company and the obligations of the assurance provider (auditor).

An explanation of everything, such as the attestation service regarding the sustainability report; free of bias, the scope should be indicated, including the type and level of audit that has been carried out. The determination of the scope, type and level of the audit should be made prior to the audit activities and constitutes an agreement between the audited.

Description of disclosures covered; the following information is a description of the scope of the disclosure, methodology; the information that must also be included in the independent assurance is the technique or methodology used by the assurance provider in performing the audit activities.

Any limitations; assurance providers must disclose any limitations encountered in performing the audit activities. It may also be explained how assurance providers can be assured that the impact of any limitations encountered will be minimized.

## 5. Conclusion and Suggestion

### 5.1. Conclusion

Sustainability accounting is the result of the process and measurement of the accounting system to communicate how the financial system is managed and takes environmental sustainability into account. However, in order for the report to be considered, it must be subject to a sustainability audit, which examines and assesses whether the report presented is consistent with the actual activities of the company.

Continuous reviews have many benefits in terms of reducing the time it takes to complete financial and reporting in different countries, speeding up the review process and facilitating the work of the auditors, but conducting continuous reviews requires an awareness of the challenges and their own challenges, based on different research reports, speeding up the review process and facilitating the review process, and the highest priority in the system.

### 5.2. Proposal

The accounting sustainability report is very important for the company, because with its introduction, the information of a company or an audit will be really transparent in the current *Otoritas Jasa Keuangan* (OJK), the OJK; similar to the SEC, which has required a sustainability report from every company to avoid undesirable things like fraud etc. So if every company in Indonesia, especially those that go public, is required to submit sustainability reports, there is no reason why this should not be the case.

Disclosure of sustainability reports is still voluntary in most countries, including Indonesia, and last year a few companies still wanted to disclose sustainability reports. Sustainability reports are becoming increasingly important in global business practice and are one of the criteria for assessing a company's social



responsibility. Managers in the global economy are increasingly aware that more comprehensive reporting (not just financial reporting) will support corporate strategy. They can also demonstrate their commitment to sustainable development. Sustainability reporting can also improve the financial performance and legitimacy of companies. We suggest that public companies should be able to publish sustainability reports in accordance with the applicable standards and that auditors should be able to conduct audits in accordance with the applicable auditing standards, so that the process of sustainability reporting is not just a formality but actually benefits the industry in Indonesia.

## References

- [1] World Commission on Environment and Development, "Brundtland Report - Our common future," *Our Common Future*, 2017.
- [2] A. Masiulevičius and V. Lakis, "Differentiation of performance materiality in audit based on business needs," *Entrep. Sustain. Issues*, 2018.
- [3] GRI, "Consolidated Set of GRI Sustainability Reporting Standards," 2018.
- [4] H. Alhaddi, "Triple Bottom Line and Sustainability: A Literature Review," *Bus. Manag. Stud.*, 2015.
- [5] E. Symeonaki, M. Papoutsidakis, D. Tseles, and M. Sigala, "Post-Implementation Evaluation of a University Management Information System (UMIS)," in *Third International Conference on Mathematics and Computers in Sciences and in Industry (MCSI)*, 14–19, 2016.
- [6] F. Aliakbari Nouri, M. Shafiei Nikabadi, and L. Olfat, "Developing the framework of sustainable service supply chain balanced scorecard (SSSC BSC)," *Int. J. Product. Perform. Manag.*, 2019.
- [7] Krismiaji, Y. Anni Aryani, and D. Suhardjanto, "International Financial Reporting Standards, board governance, and accounting quality A preliminary Indonesian evidence," *Asian Rev. Account.*, 2016.
- [8] J. Elkington, "Enter the triple bottom line," in *The Triple Bottom Line: Does it All Add Up*, 2013.
- [9] R. G. Eccles, I. Ioannou, and G. Serafeim, "The impact of corporate sustainability on organizational processes and performance," *Manage. Sci.*, 2014.
- [10] T. Rusmanto and C. Williams, "Compliance Evaluation on CSR Activities Disclosure in Indonesian Publicly Listed Companies," *Procedia - Soc. Behav. Sci.*, 2015.
- [11] D. Dienes, R. Sassen, and J. Fischer, "What are the drivers of sustainability reporting? A systematic review," *Sustainability Accounting, Management and Policy Journal*. 2016.
- [12] I. Ioannou and G. Serafeim, "Corporate Sustainability: A Strategy?," *SSRN Electron. J.*, 2019.
- [13] E. Goicoechea, F. Gómez-Bezares, and J. V. Ugarte, "Integrated reporting assurance: Perceptions of auditors and users in Spain," *Sustain.*, 2019.
- [14] W. Ahmad, A. V. Mishra, and K. J. Daly, "Financial connectedness of BRICS and global sovereign bond markets," *Emerg. Mark. Rev.*, 2018.
- [15] J. Hammer and G. Pivo, "The Triple Bottom Line and Sustainable Economic Development Theory and Practice," *Econ. Dev. Q.*, 2017.
- [16] H. Al-Shaer and M. Zaman, "Credibility of sustainability reports: The contribution of audit committees," *Bus. Strateg. Environ.*, 2018.
- [17] S. A. Sarkodie and V. Strezov, "Effect of foreign direct investments, economic development and energy consumption on greenhouse gas emissions in developing countries," *Sci. Total Environ.*, 2019.
- [18] P. S. Kurniawan, "Profesi Certified Sustainability Reporting Specialist, Pelaporan Keberlanjutan, Dan Teori Enterprise (Suatu Tinjauan Mengenai Profesi CSRS dalam Pelaporan Keberlanjutan)," *J. Ilmu Sos. dan Hum.*, 2017.
- [19] Ernst & Young and Boston College Centre, "Value of sustainability reporting - A study by EY and Boston College Center for Corporate Citizenship," *EYGM Ltd.*, 2014.
- [20] S. Withers and P. Demediuk, "Sustainability Reporting Guidelines," *Int. J. Sustain. Econ. Soc. Cult. Context*, 2014.
- [21] M. C. Wang, "The relationship between firm characteristics and the disclosure of sustainability reporting," *Sustain.*, 2017.

# Dynamic Decision-Making Process in the Opportunistic Spectrum Access

Mahmoud Almasri<sup>\*1</sup>, Ali Mansour<sup>1</sup>, Christophe Moy<sup>2</sup>, Ammar Assoum<sup>3</sup>, Denis Lejeune<sup>1</sup>, Christophe Osswald<sup>1</sup>

<sup>1</sup>LABSTICC, UMR 6285 CNRS, ENSTA Bretagne, 2 rue F. Verny, Brest, 29806, France

<sup>2</sup>Univ Rennes, CNRS, IETR - UMR 6164, Rennes, F-35000, France

<sup>3</sup> Faculty of Science, Lebanese University, Tripoli, Lebanon

## ARTICLE INFO

Article history:

Received: 08 April, 2020

Accepted: 08 July, 2020

Online: 28 July, 2020

Keywords:

Opportunistic Spectrum Access

Cognitive Networks

Multi-Armed Bandit

Quality of Service

Priority Access

## ABSTRACT

We investigate in this paper many problems related to the decision-making process in the Cognitive Radio (CR), where a Secondary User (SU) tries to maximize its opportunities by finding the most vacant channel. Recently, Multi-Armed Bandit (MAB) problems attracted the attention to help a single SU, in the context of CR, makes an optimal decision using the well-known MAB algorithms, such as: Thompson Sampling, Upper Confidence Bound,  $\epsilon$ -greedy, etc. However, the big challenge for multiple SUs remains to learn collectively or separately the vacancy of channels and decrease the number of collisions among users. To solve the latter issue for multiple users, the All-Powerful Learning (APL) policy is proposed; this new policy considers the priority access and the dynamic multi-user access, where the number of SUs may change over time. Based on our APL policy, we consider as well as the Quality of Service (QoS), where SUs should estimate and then access best channels in terms of both quality and availability. The experimental results show the superiority of APL compared to existing algorithms, and it has also been shown that the SUs are able to learn channels qualities and availabilities and further enhance the QoS.

## 1 Introduction

Game theory represents a decision-making mathematical tool that attracts much attention, when it comes to networks for resource sharing, congestion control, transmission-rate adaptation, etc. This theory was originally and exclusively proposed for economics before being applied to many other topics, such as: financial, regulation, military, political science and also biology. The main objective for using the game theory is to study and analyze cooperative or competitive situations for rational players in order to find an equilibrium among them. When, players reach the equilibrium point, then none of them can gain more by changing its action.

Game theory is widely applied in Cognitive Radio (CR) in order to enhance the spectrum efficiency of the licensed frequency bands. Indeed, according to many recent studies, the frequency bands are not well used. On the one hand, the demands on high data rate applications and wireless devices have experienced unprecedented advancement since 1990s

which makes the frequency bands more and more crowded. On the other hand, several simulations have been conducted in the United States and showed that 60 % of the frequency bands are not used [1]. Several solutions have been recommended by the Federal Communications Commission (FCC) in order to enhance the usage of the spectrum. Opportunistic Spectrum Access (OSA) in CR, represents one of the proposed solutions, where users are categorized into two groups namely: Licensed users (Primary Users: PUs) who have the right to access the frequency bands at any time, and unlicensed users (Secondary Users: SUs) that can access the frequency bands in an opportunistic manner. Usually, SUs can coexist with PUs in the same frequency bands as far as they don't cause any harmful interference to these latter. Indeed, SUs are able to access the frequency bands currently unused by PUs. SUs in OSA have many challenges in order to reduce the interference with PUs:

- **Spectrum Sensing:** A SU should sense the frequency bands and identify the available spectrum holes before

\* Corresponding Author: Mahmoud Almasri. 2 rue F. Verny, 29806 Brest, France. email: mahmoud.almasri@ensta-bretagne.fr

making any decision. The main challenge is to gather an accurate information about the status of the spectrum (free or busy) in order to access only the unused channels without causing any harmful interference to PUs. Due to hardware constraints, delay and high energy consumption, a SU may be able to sense a portion of the frequency bands (e.g. one channel at each time slot) and decides whether the selected channel is free to transmit.

- **Spectrum Decision:** At each time slot, a SU should decide which channel to access based on past success or failure decisions. As a result, a SU can gather some information about the availability and quality of channels and build a database of the spectrum access environment. This database is used in order to make a good decision and enhance the future actions of the SU.
- **Spectrum Sharing:** In order to share the available spectrum among SUs, two main models exist: Cooperative or competitive access. In the cooperative behaviors, the users need to exchange information with each other in order to maximize their opportunities and thus decrease the interference among themselves. Despite the latter benefits of the cooperative access, each user should be informed about others decisions before making any action which may increase the complexity of the secondary network. While, in the competitive access, each SU makes an action based on its local observation. However, this lack of information exchange can increase the number of collisions among users. To solve this issue, a specific policy is required to learn the vacancy probabilities of available channels and decrease the number of collisions among users.

ate the selected channel when a PU reappears. Moreover, a SU may badly identify its dedicated channel and then access a channel that does not correspond to its prior rank<sup>1</sup>. Therefore, the user should evacuate its current channel when he identifies its targeted channel.

This paper is an extension of our original work presented in [2] with a novel policy called All-Powerful Learning (APL) is proposed in order to maximize the opportunities of SUs, share the available spectrum among them, and limit the interference among PUs and SUs. Instead of only considering the availability, this paper takes into account a quality information metric, where the priority users should access only best channels with the highest availability and quality.

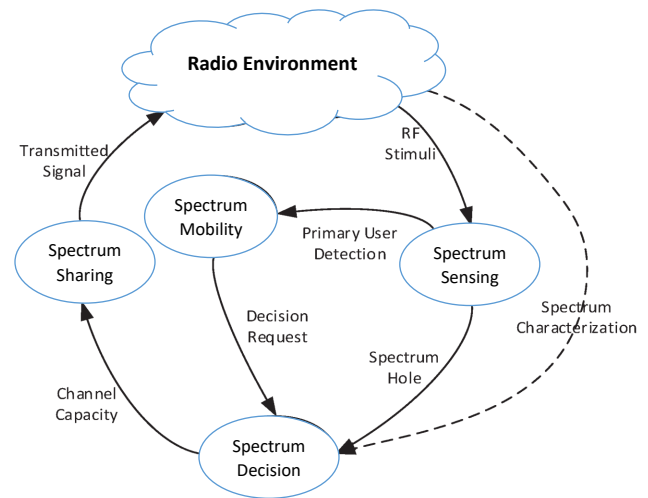


Figure 1: Cognitive cycle as introduced in [3].

Table 1: List of acronyms

APL	All-Powerful Learning
CR	Cognitive Radio
DMC	Dynamic Musical Chairs
EXP3	Exponential weights for Exploration and Exploitation
FCC	Federal Communications Commission
MAB	Multi-Armed Bandit
MEGA	Multi-user $\epsilon$ -greedy collision Avoiding
OSA	Opportunistic Spectrum Access
PU	Primary User
QoS	Quality of Service
SU	Secondary User
SLK	Selective Learning of the $k^{th}$ largest expected rewards
TS	Thompson Sampling
UCB	Upper Confidence Bound

## 2 Multi-Armed Bandit Problem

Multi-Armed Bandit (MAB) model represents one of the famous models, in game theory, that is adopted to enhance the efficiency of the licensed frequency bands. Moreover, MAB problem represents a simple case of the Reinforcement Learning (RL).

In the RL, the agent should enhance his behavior from the feedback (e.g. reward). Indeed, the RL may allow an agent to adapt to his environment by finding a suitable action to reach the best reward. The agent can maximize his reward without any prior information about his environment. However, by memorizing the states of an environment or the actions he took, the agent can make a better decision in the future. The reward feedback, also called reinforcement signal, has an important role to help an agent to learn from its environment. The RL is widely used in several domains: Robotics, Aircraft control, self-driving cars, Business strategy planning, etc. It was first developed for a single agent who should find an optimal policy that maximizes his expected reward knowing that the optimal policy depends on the environment. Unlike the case of a single agent, for multiple agents, the optimal policy

- Finally, in the **Spectrum Mobility**, a SU should evacu-

<sup>1</sup>Based on our APL policy, each user has a prior rank and should access the channel corresponding to its rank.

depends not only on the environment but also on the policies selected by other agents. Moreover, when multiple agents apply the same policy their approaches in such systems often fail because each agent tries individually to reach a desired result. In other words, it is impossible for all agents in a certain system to maximize simultaneously their personal reward, although find an equilibrium for the system representing a point of interest. Subsequently, it is important to find a policy for each agent in order to guarantee the convergence to an equilibrium state in which no agent can gain more when modifying its own action. In RL, Exploitation-Exploration dilemma represents an attractive problem. In order to maximize his performance (exploitation), the agent should gather some information about his environment (exploration). This is known as the Exploration-Exploitation dilemma in the reinforcement learning. If the agent spends a lot of time on the exploration phase, then he cannot maximize his reward. Similarly, when the agent focuses on the exploitation phase by exploiting his current information, then he may miss the best action that leads to the highest reward. Thus, the agent needs to balance the tradeoff between Exploration and Exploitation in order to obtain an appropriate result.

Due to its generic nature, the MAB model is widely adopted in many fields, such as: wireless channel access, jamming communication or object tracking. In such model, an agent can play a single arm at each time trying to maximize its long-term reward. To reach its goal, the agent needs to find the best arm in terms of expected reward. At each time slot, the agent can choose the current best arm (exploitation) or play other arms trying to obtain a robust estimation of their reward (exploration). Generally, an optimal policy, used by the agent, should balance between the exploitation and the exploration phases while pulling the arms.

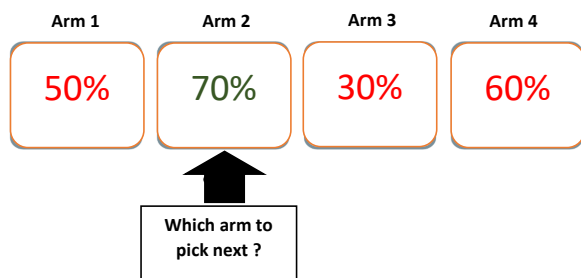


Figure 2: Several Arms with different expected reward. After a finite period of time the agent has a perception about the reward obtained from each arm.

Like most RL frameworks, the agent starts the game without any priori knowledge about the expected reward of the arms. The main goal of the agent is to find the arm with the highest expected reward. Here, we should define two classes of arms:

**Optimal arm:** This arm has the highest expected reward and is represented by the arm 2 in Fig. 2. The agent tries to reach this arm in order to maximize his expected reward.

**Suboptimal arms:** Include all other arms considered as

non-optimal. Efficient MAB algorithms should be able to limit playing with suboptimal arms.

To solve the MAB problem, several algorithms have been proposed, such as: Thompson Sampling [4], Upper Confidence Bound (UCB) [5],  $\epsilon$ -greedy [6], Exponential weights for Exploration and Exploitation (EXP3) [7], etc. The performance of a given MAB algorithm is usually measured by a regret that represents the gap between the reward obtained in the ideal scenario, where the user know the expected reward of each arm and often pulls the best one, and that obtained using a given MAB algorithm.

It is worth mentioning that these algorithms have been suggested for a single SU in the context of OSA where the SU is considered as an agent and the channels become equivalent to the different arms. Then, it is assumed that each channel is associated with a distinct availability probability and the SU should estimate this latter after a finite number of time slots. In this work, we first start to formulate the classical OSA as a MAB problem, in which, we consider a single Secondary User (SU) that needs to access opportunistically the frequency band. Later on, we will consider more realistic conditions that deal with the OSA (e.g. multiple users, Quality of Service, collision among users, dynamic access).

## 2.1 Thompson Sampling

Thompson Sampling (TS), a randomized algorithm with a bayesian spirit, represents one of the earliest algorithms proposed to tackle the MAB problem. In TS, each arm has assigned an index  $B_i(t, T_i(t))$  that contains information based on the past success and failure observations. After a finite number of time slots, the index  $B_i(t, T_i(t))$  will be very close to the mean reward of each arm. By selecting the arm with the highest index at each time slot, the agent often selects the best arm with the highest reward. This index achieves a trade-off between the exploration and the exploitation phases and can be defined as follows:

$$B_i(t, T_i(t)) = \frac{W_i(t, T_i(t)) + a}{W_i(t, T_i(t)) + Z_i(t, T_i(t)) + a + b} \quad (1)$$

where  $W_i(t, T_i(t))$  and  $Z_i(t, T_i(t))$  represent respectively the success and failure access;  $a$  and  $b$  are constant numbers.

Despite its excellent performance that can exceed the state-of-the-art MAB algorithms [8, 9, 10], TS is widely ignored in the literature. This ignorance is due to the fact that this algorithm is proposed with a lack of proof and a slight mathematical background unlike other MAB algorithms, such as: UCB or  $\epsilon$ -greedy. Recently, TS has attracted more attention and is being used in several fields [11, 12, 13]. Recent studies have found a theoretical upper bound for its convergence to the best choice [14, 15, 16].

## 2.2 Upper Confidence Bound

Upper Confidence Bound (UCB) represents one of the famous MAB algorithms firstly proposed in [5]. Like TS, the index  $B_i(t, T_i(t))$  of UCB contains two phases, the exploration



and the exploitation phases, in order to estimate the vacancy probabilities of channels and then access the best one. In the literature, several variants of UCB have been proposed to enhance the performance of the classical UCB, such as: UCB1, UCB2, UCB-tuned, Bayes-UCB, KL-UCB [8, 17, 18, 19]. UCB1 [17] represents the simplest version that balances between the complexity and the optimality.

---

**Algorithm 1: Thompson Sampling Algorithm**


---

**Input:**  $C, n,$   
1  $C$ : number of channels,  
2  $n$ : total number of slots,  
3 **Parameters:**  $S_i(t), T_i(t), W_i(t, T_i(t)), Z_i(t, T_i(t)),$   
4  $S_i(t)$ : the state of the selected channel, equals one if the channel is free and 0 otherwise,  
5  $T_i(t)$ : number of times the  $i^{th}$  channel is sensed by SU,  
6  $W_i(t, T_i(t))$ : the success access of the  $i^{th}$  channel,  
7  $Z_i(t, T_i(t))$ : the failure access of the  $i^{th}$  channel,  
**Output:**  $B_i(t, T_i(t)),$   
8  $B_i(t, T_i(t))$ : the index assigned for the  $i^{th}$  channel,  
9 **foreach**  $t = 1$  to  $n$  **do**  
10      $a_t = \arg \max_i B_i(t, T_i(t)),$   
11     Observe the State  $S_i(t),$   
12      $W_i(t, T_i(t)) = \sum_{i=0}^n S_i(t)1_{a_t=i},$   
13     %  $1_{a_t=i}$ : equal 1 if the user selects the  $i^{th}$  channel and 0 otherwise,  
14      $Z_i(t, T_i(t)) = T_i(t) - W_i(t, T_i(t)),$   
15      $B_i(t, T_i(t)) = \frac{W_i(t, T_i(t)) + a}{W_i(t, T_i(t)) + Z_i(t, T_i(t)) + a + b}$

---



---

**Algorithm 2: UCB1 Algorithm**


---

**Input:**  $\alpha, C, n,$   
1  $\alpha$ : exploration-exploitation factor,  
2  $C$ : number of channels,  
3  $n$ : total number of slots,  
4 **Parameters:**  $T_i(t), X_i(T_i(t)), A_i(t, T_i(t)),$   
5  $T_i(t)$ : number of times the  $i^{th}$  channel is sensed up to  $t,$   
6  $X_i(T_i(t))$ : the exploitation contribution of  $i^{th}$  channel,  
7  $A_i(t, T_i(t))$ : the exploration contribution of  $i^{th}$  channel,  
**Output:**  $B_i(t, T_i(t)),$   
8  $B_i(t, T_i(t))$ : the index assigned for  $i^{th}$  channel,  
9 **foreach**  $t = 1$  to  $C$  **do**  
10     SU senses each channel once,  
11     SU updates its index  $B_i(t, T_i(t)),$   
12 **foreach**  $t = C + 1$  to  $n$  **do**  
13      $a_t = \arg \max_i B_i(t - 1, T(t - 1)),$   
14      $T_i(t) ++,$   
15      $X_i(T_i(t)) = \frac{1}{T_i(t)} \sum_{\tau=1}^t S_i(\tau),$   
16     %  $S_i(\tau)$  is the observed state from channel  $i$  at  $\tau,$   
17     %  $S_i(\tau) = 1$  if the channel  $i$  is vacant and 0 otherwise,  
18      $A_i(t, T_i(t)) = \sqrt{\frac{\alpha \ln(t)}{T_i(t)}},$   
19      $B_i(t, T_i(t)) = X_i(T_i(t)) + A_i(t, T_i(t)),$

---

For this reason, UCB1 is the widely adopted version

scheme in the context of CR to help a SU make an optimal decision [20, 21, 22, 23, 24, 25]. In UCB1, the index  $B_i(t, T_i(t))$  essentially comprises two important factors:  $X_i(T_i(t))$  and  $A_i(t, T_i(t))$  that represent respectively the exploitation (or the expected reward) and the exploration phases:

$$B_i(t, T_i(t)) = X_i(T_i(t)) + A_i(t, T_i(t)) \quad (2)$$

where the exploitation and the exploration factors can be expressed as:

$$X_i(T_i(t)) = \frac{1}{T_i(t)} \sum_{j=1}^t r_i(j) \quad (3)$$

$$A_i(t, T_i(t)) = \sqrt{\frac{\alpha \ln(t)}{T_i(t)}} \quad (4)$$

The factor  $A_i(t, T_i(t))$  has an important role in learning the availability probabilities of channels by pushing the algorithm to examine the state of all available channels. Thus, after a finite time  $t$ ,  $X_i(T_i(t))$  of the  $i^{th}$  channel will approximately equal to its availability probability  $\mu_i$ .

In [17], the authors found an upper bound of the sum of regret (i.e. the loss of reward by selecting the worst channels) for a single agent and  $C$  arms. It has shown that the upper bound of the regret achieves a logarithmic asymptotic behavior, which means that after a finite number of time slots, the agent will be able to identify the best arm and always select it.

### 2.3 $\epsilon$ -greedy

One of the simplest MAB algorithms to tackle the MAB problem is referred to  $\epsilon$ -greedy that was firstly proposed in [6]. A recent version of this algorithm is proposed in [17] in order to achieve a better performance compared to several previous versions (see algorithm 1). Like several MAB algorithms,  $\epsilon$ -greedy contains two phases completely separated: exploration and exploitation. During the exploration phase, the user chooses a random channel in order to learn the vacancy probability of channels.

While in the exploitation phase, the user usually selects the channel with the highest expected reward  $X_i(T_i(t))$ . The authors of [17] have also investigated the analytical convergence of the  $\epsilon$ -greedy and proved that the regret (i.e. the loss of reward by selection the worst channel) achieves a logarithmic asymptotic behavior.

## 3 Problem Formulation

In the previous section, we introduced the well-known MAB algorithms that help a MAB agent makes a good decision. In this section, we present the classical OSA for a single SU in order to formulate it as a MAB problem. However, MAB algorithms can represent an optimal solution for the classical OSA, as it can be seen in section 5. On the other hand, we

consider more developed scenarios compared to the classical OSA such as multiple SUs, decreasing the collisions among users and also estimating the quality of the available channels. We first present the OSA for multiple SUs in the next section and, hereinafter, we propose the new APL policy to manage a secondary network.

---

**Algorithm 3:**  $\epsilon$ -greedy Algorithm

---

**Input:**  $C, H, n,$

- 1  $C$ : number of channels,
- 2  $H$ : exploration constant,
- 3  $n$ : total number of slots,
- 4 **Parameters:**  $T_i(t),$
- 5  $T_i(t)$ : number of times the channel is sensed up to time  $t,$
- 6  $\chi$ : a uniform random variable in  $[0,1],$

**Output:**  $X_i(T_i(t)),$

- 7  $X_i(T_i(t))$ : the expected reward that depends on  $T_i(t),$

8 **foreach**  $t = 1$  to  $n$  **do**

9     **if**  $\chi < \min\{1, \frac{H}{t}\}$  **then**

10         SU makes a random action  $a_t,$

11     **else**

12          $a_t = \max_i X_i(T_i(t)),$

13          $T_i(t) ++,$

14          $X_i(T_i(t)) = \frac{1}{T_i(t)} \sum_{\tau=1}^t S_i(\tau),$

15         %  $S_i(\tau)$  is the observed state from channel  $i$  at  $\tau,$

16         %  $S_i(\tau) = 1$  if the  $i^{th}$  channel is vacant and 0 otherwise,

---

### 3.1 Single User Case

Let us consider a SU accesses  $C$  channels, each of which associated with a vacancy probability  $\mu_i \in [0,1]$ . Let the vacancy probabilities be ordered by their availability probabilities,  $\mu_1 > \mu_2 > \dots > \mu_C$ , which are initially unknown for the secondary user. A most important objective of the SU is to estimate the vacancy probabilities of channels after a finite time in order to access the best channel that has  $\mu_1$  as vacancy probability. At each time slot, the user can select one channel and transmit its data if available; otherwise, it should wait the next slot to sense another channel. Let the state of the  $i^{th}$  channel at slot  $t$  be referred to  $S_i(t)$ :  $S_i(t)$  equals 1 if the  $i^{th}$  channel is free and 0 otherwise. Hereinafter, we consider that the obtained reward from the  $i^{th}$  channel  $r_i(t)$ , at slot  $t$  is equal to its state:  $r_i(t) = S_i(t)$ . Let  $T_i(t)$  represent the number of times to access the  $i^{th}$  channel up to the slot  $t$ . The user should be rational by adopting a given policy in order to quickly identify the best channel. A policy selected by the SU may not be considered as optimal in term of the accuracy of the channels' vacancy estimation or the convergence speed towards the best channel. Finally, let us introduce the regret that rerepresents the gap between the reward obtained in an ideal scenario and that can be obtained using a given policy as follows:

$$R(n, \beta) = n\mu_1 - E \left[ \sum_{t=1}^n \mu_i^\beta(t) \right] \quad (5)$$

where  $n$  represents the total number of time slots and  $\mu_i^\beta(t)$  stands for the vacancy probability of the selected channel at slot  $t$  under the policy  $\beta$ , and  $E(\cdot)$  is the mathematical expectation.

### 3.2 Multi-User Case

In this section, we consider  $U$  SUs trying to learn the vacancy probabilities of the  $C$  channels and then access only the  $U$  best ones ( $C > U$ ). When several SUs existing in the spectrum, their main challenge is to learn collectively or separately the vacant probability of channels as much as possible in order to access the best ones. Therefore, a policy selected by users should estimate the vacancy of channels as much as possible, and should also be able to decrease the collisions number among users. Therefore, let us define the regret for multiple users that takes into account both the convergence speed to the  $U$  best channels and the collision number among users as follows:

$$R(n, U, \beta) = n \sum_{k=1}^U \mu_k - \sum_{t=1}^n E \left[ S^\beta(t) \right] \quad (6)$$

where  $\mu_k$  stands for the vacancy probability of the  $k^{th}$  best channel;  $S^\beta(t)$  represents the global reward obtained by all users at time  $t$  using the policy  $\beta$  and is defined as follows:

$$S^\beta(t) = \sum_{j=1}^U \sum_{i=1}^C S_i(t) I_{i,j}(t) \quad (7)$$

where  $S_i(t)$  represents the state of the  $i^{th}$  channel at time  $t$ :  $S_i(t) = 1$  if the  $i^{th}$  channel is available and 0 otherwise;  $I_{i,j}(t)$  indicates that no collisions have appeared in the  $i^{th}$  channel by the  $j^{th}$  user at slot  $t$ :  $I_{i,j}(t) = 1$  if the  $j^{th}$  user is the sole occupant of the channel  $i$  and 0 otherwise. Finally, the regret that takes into consideration the channels' occupancy and the collisions number among users can be expressed by:

$$R(n, U, \beta) = n \sum_{k=1}^U \mu_k - \sum_{j=1}^U \sum_{i=1}^C P_{i,j}(n) \mu_i \quad (8)$$

where  $P_{i,j}(n) = \sum_{t=1}^n E [I_{i,j}(t)]$  represents the expectation of times that the  $j^{th}$  user is the only occupant of the  $i^{th}$  channel up to  $n$ , and the mean of reward can be given by:

$$\mu_i \approx \frac{1}{n} \sum_{t=1}^n S_i(t)$$

## 4 Multi-Priority Access

In the existing models of OSA where several SUs exist in the network, the main challenge is to learn collectively (via a cooperative learning) or separately (via a competitive learning) the

available channels while decreasing the number of collisions with each other. In our work, we focus on the competitive priority access, where the  $k^{th}$  user should selfishly estimate the vacancy probabilities of channels in order to access the  $k^{th}$  best one. Our proposed policy for the priority access takes into account the dynamic access where the priority users can enter/leave the network at any time. To the best of our knowledge, only the priority or the random access are considered without the dynamic access in several proposed MAB policies [24, 25, 26, 27] (a simple example for the priority dynamic access is shown in Fig. 3).

To formulate the OSA as a MAB problem, recent works extend the simple case of MAB (i.e. the case of a single agent) to consider several agents [20, 25, 26, 28, 29]. In our work, we are interested in the OSA for multiple priority access in which SUs should access the spectrum according to their ranks. Moreover, decreasing the number of collisions among SUs represents a point of interest to enhance the global performance of the secondary network. In general, when two SUs access the same channel to transmit, their data cannot be correctly received because of the interference between them. When a collision occurs among users, several proposals can be found in the literature in order to enhance their behavior in the next slots. We present below two well-known collision models in the literature that are widely used in OSA:

- ALOHA-like model: If a collision occurs between two or more users, then none of them receives a reward, despite the selected channels is free. This model may ensure the fairness among users, and no collision avoidance mechanism is used.
- Reward sharing model: If two or more users select the same channel at the same time, the colliding users share the obtained reward from the selected channel (each of them receives the same reward).

The above models can affect the methodologies used to collect the reward from the target channel while the learning phase is not affected. In our work, we consider the most widely used, ALOHA-like.

Based on the ALOHA-like, the works of [2, 20, 21, 25, 26, 27, 28, 30] proposed semi-distributed and distributed algorithms in which users cannot exchange information with each other. Liu and Zhao in [28], proposed Time-Division Fair Share (TDFS) policy and showed that the proposed algorithm may achieve an asymptotic logarithmic behavior. In such algorithm, the users access the channels with different offsets. TDFS also ensures the fairness among users; while in our work we are interested in the priority access where users access the channels based on their prior rank. In [28], TDFS policy was been used to extend UCB1 algorithm to consider multiple users. Beside TDFS, the authors of [20] proposed Random Rank policy, based on UCB1, to manage the secondary network. Random Rank represents a distributed policy (i.e. no-information exchange among users) in which the user achieves a different throughput.

The authors of [24] proposed the Selective Learning of the  $k^{th}$  largest expected rewards (SLK) policy, based on UCB1, that represents an efficient policy for the priority access. However, SLK allows only a fixed number of users to access the available channels. So that, the dynamic access under SLK cannot be considered since this latter restricts the access. Similarly to SLK, the authors of [25] proposed the  $k^{th}$  – MAB for the priority access which is based on UCB1 and  $\epsilon$ -greedy. In  $k^{th}$  – MAB, the time is slotted and each slot is divided into multi sub-slots depending on the users priority ranks. For instance, the slot of  $SU_U$  is divided into  $U$  sub-slots in order to find the  $U^{th}$  best channel and transmit data via this channel. Therefore, the main limitation of this policy remains in the dissatisfaction of transmission time of high ranked users.

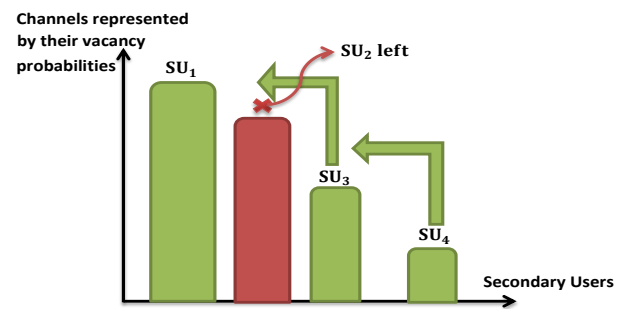


Figure 3: Priority access after a user left its dedicated channel

For the random access, several learning policies can be found in the literature, where the SU selects randomly its channel. The authors of [26] proposed the Musical Chairs policy as well the Dynamic Musical Chairs (DMC) policy for a dynamic access. In both policies, the SU selects a random channel up to time  $T_0$  in order to estimate the vacancy of channels and the number of users,  $U$ , in the network. After  $T_0$ , the SU chooses a random channel between  $\{1, \dots, U\}$ . The main drawback of the Musical Chairs and DMC is that the users should know the total number of transmission time slots as well as the number of available channels. Moreover, in DMC a restrict access is considered, where the users cannot leave the network during the time  $T_0$ . To find the  $U$ -best channels, the authors of [27] proposed the Multi-user  $\epsilon$ -greedy collision Avoiding (MEGA) algorithm based on the  $\epsilon$ -greedy algorithm proposed in [17]. However, their algorithm suffers the same drawbacks of the Musical Chairs and the Dynamic Musical Chairs and it does not consider the priority access. In order to solve all these limitations, we propose in section (4.1) a novel policy called APL for the priority dynamic access.

#### 4.1 APL for the Priority Access

In this section, we propose a new policy for the priority access. This policy enables a secondary user to learn the vacant probabilities of channels and ensures the convergence to his dedicated channel. Moreover, it can be used with all learning MAB algorithms such as: Thompson Sampling (TS), Upper

Confidence Bound (UCB), AUCB,  $e$ -UCB,  $e$ -greedy, etc. We should highlight that our proposed policy does not require prior knowledge about the channels as in the case for other policies, such as: Musical Chair [26], SLK [24],  $k$ -th MAB [25], MEGA [27], etc. Indeed, existing policies to manage a secondary network suffer from one or more of the following disadvantages:

1. The number of users should be fixed and known to all users.
2. SUs should have a prior information about the number of channels.
3. Expected transmission time should be known.
4. The dynamic access is not suggested. To recall, in a dynamic access, the users can at any given time enter or leave the network.
5. Some algorithms consider a restricted dynamic access, where a SU can't leave the network during the learning or the exploration phases.
6. The vacant probabilities of channels should be static; otherwise, users cannot adapt to their environment.
7. The priority access is seldomly suggested in the literature, while the random access represents the most used model.

Unlike SLK and  $k$ -th MAB, our proposed policy for the priority access, called All-Powerful Learning algorithm (APL), doesn't suffer from the above mentioned drawbacks. As a matter of fact, SLK and  $k$ -th MAB policies suffer from the 1<sup>st</sup>, 2<sup>nd</sup> and 4<sup>th</sup> mentioned drawbacks.

In a classical priority access, each channel has assigned an index  $B_i(t)$  and the highest priority user  $SU_1$  should sense and access the channel with the highest index  $B_i(t)$  at each time slot. Indeed, the best channel, after a finite number of time slots, will have the highest index  $B_i(t)$ .

As the second priority user  $SU_2$  should avoid the first best channel and try to access the second best one. To reach his goal,  $SU_2$  should sense the first and second best channels at each time slot in order to estimate their vacant probabilities and then access the second best channel if available. In this case, the complexity of the hardware is increased, and we conclude that a classical priority access represents a costly and impractical method to settle down each user to his dedicated channel. In the case of APL, at each time slot, the user senses a channel and transmits his data if the channel is available (see algorithm 4). In our policy, each  $SU_k$  has a prior rank,  $k \in \{1, \dots, U\}$ , and his target is to access the  $k$ -th best channel. The major problem of the competitive priority access is that each user should selfishly estimate the vacant probabilities of the available channels. Our policy can intelligently solve this issue by making each user generate a rank around his prior rank to get information about the channels availability. For instance, if the rank generated by the  $k$ -th user equals 3 (considering that  $k > 3$ ), then he should access the channel that

has the third index, i.e.  $B_3(t)$ . In this case,  $SU_k$  can examine the states of the  $k$  best channels and his target is the  $k$ -th best one.

---

**Algorithm 4:** APL for the priority dynamic access

---

**Input:**  $k, \zeta_k(t), r_i(t)$ ,  
 1  $k$ : indicates the  $k$ -th user or  $k$ -th best channel,  
 2  $\zeta_k(t)$ : indicates a presence of collision for the  $k$ -th user at instant  $t$ ,  
 3  $r_i(t)$ : indicates the state of the  $i$ -th channel at instant  $t$ ,  $r_i(t) = 1$  if the channel is free and 0 otherwise,  
 4 **Initialization**  
 5  $k = 1$ ,  
 6 **for**  $t = 1$  to  $C$  **do**  
 7      $SU_k$  senses each channel once,  
 8      $SU_k$  updates his index  $B_i(t)$ ,  
 9      $SU_k$  generates a rank of the set  $\{1, \dots, k\}$ ,  
 10      $k + 1$ ,  
 11 **for**  $t = K+1$  to  $n$  **do**  
 12      $SU_k$  senses a channel in his index  $B_i(t)$  according to his rank,  
 13     **if**  $r_i(t)=1$  **then**  
 14          $SU_k$  transmits his data,  
 15         **if**  $\zeta_k(t)=1$  **then**  
 16              $SU_k$  regenerates his rank of the set  $\{1, \dots, k\}$ ,  
 17         **else**  
 18              $SU_k$  keeps his previous rank,  
 19     **else**  
 20          $SU_k$  refrains from transmitting at instant  $t$ ,  
 21      $SU_k$  updates his index  $B_i(t)$

---

However, if the rank created by  $SU_k$  is different than  $k$ , then he selects a channel with one the following probabilities:  $\{\mu_1, \mu_2, \dots, \mu_{k-1}\}$  and he may collide with a priority user, i.e.  $SU_1, SU_2, \dots, SU_{k-1}$ . Therefore,  $SU_k$  should avoid regenerating his rank at each time slot; otherwise, a large number of collisions may occur among users and transmitted data can be lost. So, after each collision,  $SU_k$  should regenerate his rank from the set  $\{1, \dots, k\}$ . Thus, after a finite number of slots, each user settles down to his dedicated channel. It remains to investigate the analytical convergence of APL to verify its performance in a real radio environment.

## 4.2 Quality of Service

As mentioned before, UCB represents one of the popular MAB algorithms that is widely suggested in the literature, where several variants have been proposed. In [23], we proposed a new variant of UCB called the Quality of Service UCB1 (QoS-UCB1) for a single SU, where this latter is able to learn channels' vacancy and quality. To consider multiple SUs, this version of UCB is extended using the Random Rank policy proposed in [20] to manage a secondary network. It has been shown that the Random Rank policy with the QoS-UCB1 represents an optimal solution to allow users to learn separately channels' vacancy and quality. However, in this paper, we



evaluate the performance of our APL policy with QoS-UCB1 for the priority access.

Supposing that each channel has a binary quality represented by  $q_i(t)$  at slot  $t$ :  $q_i(t) = 1$  if the channel has a good quality and 0 otherwise. Then, the expected quality collected from the channel  $i$  up to time  $n$  is given by:

$$G_i(T_i(n)) = \frac{1}{T_i(n)} \sum_{\tau=1}^{T_i(n)} q_i(\tau) \quad (9)$$

The global mean reward, that takes into account channels' vacancy and quality, can be expressed as follows [23]:

$$\mu_i^Q = G_i(T_i(n)) \cdot \mu_i \quad (10)$$

The index assigned to the  $i^{\text{th}}$  channel that considers both vacancy and quality  $B_i^Q(t, T_i(t))$  can be defined by:

$$B_i^Q(t, T_i(t)) = X_i(T_i(t)) - Q_i(t, T_i(t)) + A_i(t, T_i(t)) \quad (11)$$

According to [23], the term  $Q_i(t, T_i(t))$  of the quality factor is given by the following equation:

$$Q_i(t, T_i(t)) = \frac{\gamma M_i(t, T_i(t)) \ln(t)}{T_i(t)}$$

where the parameter  $\gamma$  stands for the weight of the quality factor;  $M_i(t, T_i(t)) = G_{\max}(t) - G_i(T_i(t))$  being the difference between the maximum expected quality over channels at time  $t$ , i.e.  $G_{\max}(t)$ , and the one collected from channel  $i$  up to time slot  $t$ , i.e.  $G_i(T_i(t))$ . However, when the  $i^{\text{th}}$  channel has a good quality  $G_i(T_i(t))$  as well as a good availability  $X_i(T_i(t))$  at time  $t$ . The quality factor  $Q_i(t, T_i(t))$  decreases while  $X_i(T_i(t))$  increases. Subsequently, by selecting the maximum of its index  $B_i^Q(t, T_i(t))$ , the user has a large chance to access the  $i^{\text{th}}$  channel with a high quality and availability.

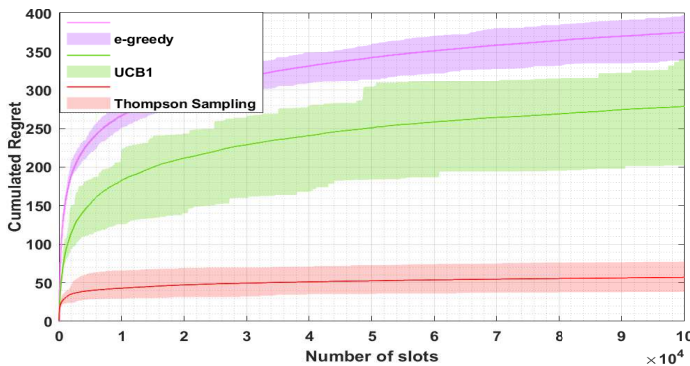


Figure 4: Evaluate the performance of TS, UCB1 and  $\epsilon$ -greedy in OSA

## 5 Simulations and Results

In our simulations, we consider three main scenarios: In the first one, a SU tries to learn the vacancy of channels using the MAB algorithms: TS, UCB1 and  $\epsilon$ -greedy in order to access

the best one with the highest vacancy probability. We also compare the performance of these MAB algorithms to show which one can offer more opportunities for the SU. In a second scenario, we considered 4 SUs trying to learn the vacancy of channels with a low number of collisions. In this scenario, we show that, based on our policy APL, users reach their dedicated channel faster than several existing policies. In the last scenario, using APL with the QoS-UCB1, users should learn both vacancy and quality of channels and then converge towards channels that have a good vacancy and quality.

In our algorithm, two factors can affect the convergence:  $\alpha$  or  $H$  while the convergence of UCB1 and  $\epsilon$ -greedy are affected by  $\alpha$  and  $H$  respectively. We consider the value of  $\alpha$  and  $H$  for which UCB1 and  $\epsilon$ -greedy achieve their best performance. According to [17], the best value of  $H = \frac{c \times K}{d^2}$  (i.e.  $c = 0.1$  is a constant number,  $K = 9$  and  $d = \min_i(\mu_1 - \mu_i) = 0.1$ ) and  $\alpha$  are 90 and 2 respectively in order to ensure a balance between the exploration and exploitation phases.

Let us initially consider a SU trying to access 9 channels associated with the following vacancy probabilities:

$$\Gamma = [0.9 \ 0.8 \ 0.7 \ 0.6 \ 0.5 \ 0.4 \ 0.3 \ 0.2 \ 0.1]$$

Fig. 4 compares the regret of the SU using the three MAB algorithms: TS, UCB1 and  $\epsilon$ -greedy over 1000 Monte Carlo runs. The simulation outcomes are presented with a shaded region enveloping the average regret. As we can see, the regrets of the 3 MAB algorithms have a logarithmic asymptotic behavior with respect to the number of slots, while TS produces a lower regret for all simulations. That means that the SU can quickly reach the best channel that offers more opportunities for the user compared to other channels. In the second scenario of our simulation, we evaluate the performance of APL and its ability to make each user selects his dedicated channel after a finite number of time slots. We evaluate the performance of APL compared to the existing learning policies such as Musical Chair and SLK.

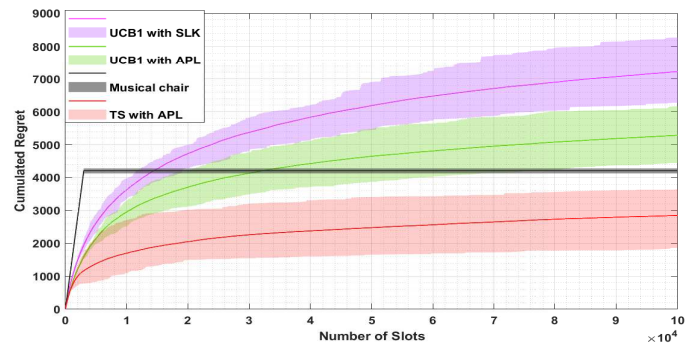


Figure 5: TS, UCB1 with APL compared to SLK and Musical Chairs

To make this comparison, we use two main performance indexes: the regret related to the access of worst channels and the percentage of times to access best channels by each user. A collision may occur when two or more users try to access the same channel. We adopt in our simulations the

ALOHA model, widely used one in OSA, in which none of the collided users receives a reward. After each collision, and based on our policy APL, the collided users should regenerate their rank. First, we consider a static setting of users, then we investigate the dynamic access in which the priority users can enter or leave the network.

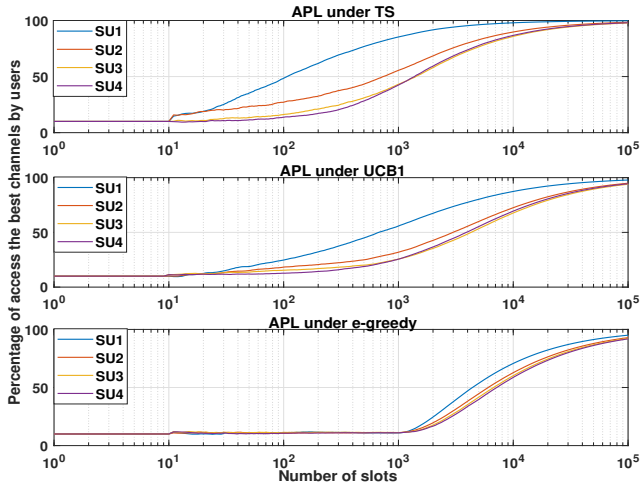


Figure 6: The percentage of times where each  $SU_k$  selects its optimal channel using the proposed approach

In Fig. 5, we compare the regret of APL to SLK and Musical Chair. APL and SLK take into consideration the priority access while Musical Chair is proposed for the random access. Despite the regret of APL and SLK has a logarithmic asymptotic behavior, the regret of Musical Chair has two parts:

- A linear part at the beginning, during the learning period, due to the large number of collisions resulting from the random selection.

- A constant part in which the users exploit the U best channels.

As we can see from Fig. 5, APL using TS outperforms Musical Chair and SLK by achieving the lower regret.

Fig. 6 shows the percentage of times that the  $k$ -th user accesses his dedicated channel based on our policy APL up to  $n$ ,  $P_k(n)$ . This latter is given by:

$$P_k(n) = \frac{1}{n} \sum_{t=1}^n 1_{(\text{if } \beta_{APL}^l(t)=k)} \quad (12)$$

where  $\beta_{APL}^l(t)$  represents the channel selected at time  $t$  under APL using the learning algorithm  $l$ , such as: TS, UCB1 or  $\epsilon$ -greedy. As we can see, based on our policy APL, the users are able to converge to their targeted channels:  $SU_1$  converges to the best channel  $\mu_1$ , followed by  $SU_2$ ,  $SU_3$  and  $SU_4$  to the channels  $\mu_2$ ,  $\mu_3$  and  $\mu_4$  respectively. In addition, we can observe a fast converges of APL using TS compared to TS.

This figure clearly shows that, based on APL, the users converge to their dedicated channels: the first priority user  $SU_1$  converges towards the best channel  $\mu_1 = 0.9$ , followed by  $SU_2$ ,  $SU_3$  and  $SU_4$  towards channels  $\mu_2 = 0.8$ ,  $\mu_3 = 0.7$  and  $\mu_4 = 0.6$  respectively. In addition, we can see that the users quickly reach their dedicated channels using TS and a slow one under UCB1 and  $\epsilon$ -greedy.

Fig. 7 compares the regret of APL and DMC for the dynamic access where the dotted line indicates the entering and leaving of users on the network. Figures (6a) and (6b) represent respectively the cumulative and average regrets of APL, where at each entering or leaving of users, a significant increase in the regret is observed. It is worth mentioning that, in the dynamic scenario and based on APL, the user can change its current channel for two reasons:

1. When a collision occurs,  $SU_k$  should generate a random

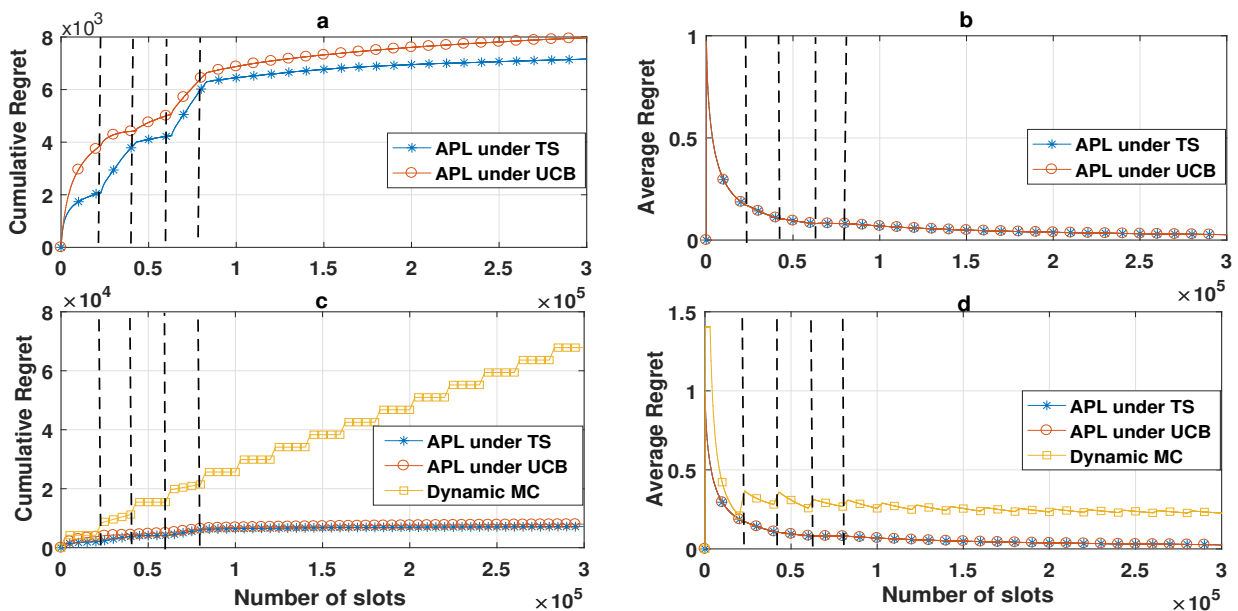


Figure 7: APL and DMC for dynamic access

rank from the set  $\{1, \dots, k\}$ .

- When a PU reappears in the network and accesses the current channel used by  $SU_k$ , the index of this channel decreases, and it may be overwhelmed by another channel that has a low index.

To the best of our knowledge, two policies exist in the literature that consider the dynamic access but without considering priority access: DMC [26] and MEGA [27]. The authors of [26] show that the DMC achieves better performance compared to MEGA policy. In Figures (6c) and (6d), we can see that the performance of APL outperforms the one of DMC and achieves a lower regret. However, after the dynamic access interval, our algorithm achieves a logarithmic regret although the regret of DMC keeps growing with time. Thus, the access under DMC algorithm is realized in epochs, where each one is composed of a learning phase with enough rounds of random exploration to learn the  $U$  best channels and the number of users under the dynamic access. The length of an epoch and the learning phase are  $T_1$  and  $T_0$  respectively. These two parameters depend on the number of channels  $C$  and the total number of slots  $n$ .

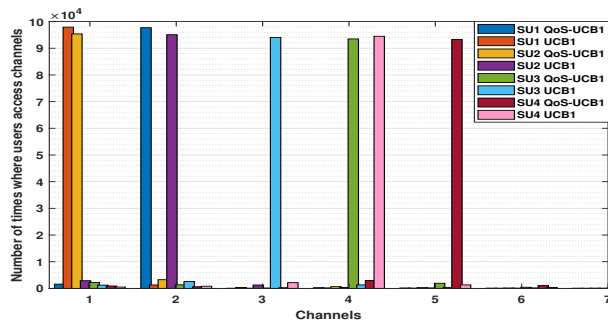


Figure 8: Access channels by priority users using APL

Let us start with the last scenario in which users are able to learn both channels' vacancy and quality using our APL policy where the empirical mean of the quality collected from channels as follows:  $G = [0.7 \ 0.9 \ 0.2 \ 0.8 \ 0.8 \ 0.7 \ 0.7 \ 0.8 \ 0.8]$ . Thus, the global mean reward that takes into consideration both quality and vacancy  $\mu_Q$  is given by:  $\mu_Q = [0.63 \ 0.72 \ 0.14 \ 0.48 \ 0.4 \ 0.28 \ 0.21 \ 0.16 \ 0.08]$ . After estimating the channels' availability and quality (i.e.  $\mu_Q$ ) and based on our APL policy with QoS-UCB1, the first priority user  $SU_1$  should converge towards the channel that has the highest global mean, i.e. channel 2, while the target of  $SU_2$ ,  $SU_3$  and  $SU_4$  should be respectively channels 1, 4 and 5. On the other hand, in the case of APL with UCB1, the target of the priority users  $SU_1$ ,  $SU_2$ ,  $SU_3$ , and  $SU_4$  should be respectively the channels 1, 2, 3 and 4. This result can be confirmed in Fig. 8, where the priority users access their dedicated channels using APL with QoS-UCB1 or UCB1. Fig. 9 displays the achievable regret of APL with QoS-UCB1 and UCB1 in the multi-user case. Despite the fact that the two curves have a logarithmic asymptotic behavior, we notice an improvement regret of APL with QoS-UCB1 compared to UCB1.

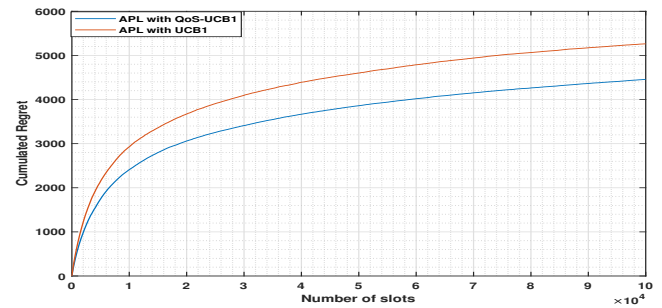


Figure 9: The regret of APL with QoS-UCB1 and UCB1

## 6 Conclusion

This paper deals with the Opportunistic Spectrum Access (OSA) problem in the context of Cognitive Radio (CR) for a single or multiple Secondary Users (SUs). Recently, several Multi-Armed Bandit (MAB) algorithms have been suggested to help a single SU make a good decision. To tackle the problem of OSA with several SUs, we proposed a novel policy for the priority access called All-Powerful Learning (APL) that allows several SUs to learn separately the channels' vacancy without any cooperation or a prior knowledge about the available channels. Moreover, APL considers the priority dynamic access while only the priority or the dynamic access are separately considered in several recent works, such as Selective Learning of the  $k^{th}$  largest expected rewards (SLK), Musical Chairs, Multi-user  $\epsilon$ -greedy collision Avoiding (MEGA) and  $k^{th}$  - MAB. In our work, the Quality of Service (QoS) have been also investigated where SU is able to learn both quality and availability of channels and then make an optimal decision with respect to its prior rank. Like most important works in OSA, this work focuses on the Independent Identical Distributed (IID) model in which the state of each channel is supposed to be drawn from an IID process. In future work, we will consider the Markov process as a dynamic memory model to describe the state of available channels, although it is a more complex process compared to IID.

## References

- [1] M. Marcus, C. Burtle, B. Franca, A. Lahjouji, N. McNeil, "Federal Communications Commission (FCC): Spectrum Policy Task Force," in ET Docket no. 02-135, November 2002.
- [2] M. Almasri, A. Mansour, C. Moy, A. Assoum, C. Osswald, D. Lejeune, "All-Powerful Learning Algorithm for the Priority Access in Cognitive Network," in EUSIPCO, A Corua, Spain, September 2019.
- [3] S. Haykin, "Brain-Empowered Wireless Communications," IEEE Journal on Selected Areas in Commun.
- [4] W. Thompson, "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples," Biometrika, 25(3), 285-294, 1933.
- [5] T. Lai, H. Robbins, "Asymptotically efficient adaptive allocation rules," Advances in Applied Mathematics, 6(1), 4-22, 1985.
- [6] C. Watkins, Learning from delayed rewards, Ph.D. thesis, University of Cambridge, 1989.

- [7] P. Auer, N. Cesa-Bianchi, Y. Freund, R. E. Schapire, "The nonstochastic multiarmed bandit problem," *SIAM journal on computing*, **32**(1), 48–77, 2002.
- [8] G. Burtini, J. Loepky, R. Lawrence, "A survey of online experiment design with the stochastic multi-armed bandit," arXiv preprint arXiv:1510.00757, 2015.
- [9] S. Scott, "A modern Bayesian look at the multi-armed bandit," *Applied Stochastic Models in Business and Industry*, **26**(6), 639–658, 2010.
- [10] O. Chapelle, L. Li, "An empirical evaluation of thompson sampling," in *Advances in neural information processing systems*, Granada, Spain, December 2011.
- [11] S. Guha, K. Munagala, "Stochastic regret minimization via Thompson sampling," in *Conference on Learning Theory*, 317–338, 2014.
- [12] T. Kocák, M. Valko, R. Munos, S. Agrawal, "Spectral thompson sampling," in *Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014.
- [13] I. Osband, D. Russo, B. V. Roy, "(More) efficient reinforcement learning via posterior sampling," in *Advances in Neural Information Processing Systems*, 3003–3011, 2013.
- [14] S. Agrawal, N. Goyal, "Analysis of thompson sampling for the multi-armed bandit problem," in *conf. on Learning Theory*, Edinburgh, Scotland, June 2012.
- [15] E. Kaufmann, N. Korda, R. Munos, "Thompson sampling: An asymptotically optimal finite-time analysis," in *International conf. on Algorithmic Learning Theory*, Lyon, France, October 2012.
- [16] S. Agrawal, N. Goyal, "Further optimal regret bounds for thompson sampling," in *Artificial intelligence and statistics*, Scottsdale, USA, April 2013.
- [17] P. Auer, N. Cesa-Bianchi, P. Fischer, "Finite-time Analysis of the Multi-armed Bandit Problem," *Machine Learning*, **47**(2), 235–256, 2002.
- [18] E. Kaufmann, O. Cappé, A. Garivier, "On Bayesian upper confidence bounds for bandit problems," in *Artificial intelligence and statistics*, La Palma, Canary Islands, April 2012.
- [19] O. Maillard, R. Munos, G. Stoltz, "A finite-time analysis of multi-armed bandits problems with kullback-leibler divergences," in *Annual conf. On Learning Theory*, Budapest, Hungary, July 2011.
- [20] A. Anandkumar, N. Michael, A. Tang, A. Swami, "Distributed Algorithms for Learning and Cognitive Medium Access with Logarithmic Regret," *IEEE Journal on Sel. Areas in Com.*, **29**(4), 731–745, 2011.
- [21] M. Almasri, A. Mansour, C. Moy, A. Assoum, C. Osswald, D. Lejeune, "Distributed Algorithm to Learn OSA Channels Availability and Enhance the Transmission Rate of Secondary Users," in *ISCIT, HoChiMinh, Vietnam*, September 2019.
- [22] M. Almasri, A. Mansour, C. Moy, A. Assoum, C. Osswald, D. Lejeune, "Distributed algorithm under cooperative or competitive priority users in cognitive networks," *EURASIP Journal on Wireless Communications and Networking*, **2020**(1), 1–31, 2020.
- [23] N. Modi, P. Mary, C. Moy, "QoS driven Channel Selection Algorithm for Cognitive Radio Network: Multi-User Multi-armed Bandit Approach," *IEEE Trans. on Cog. Com. & Networking*, **3**(1), 1–6, 2017.
- [24] Y. Gai, B. Krishnamachari, "Decentralized Online Learning Algorithms for Opportunistic Spectrum Access," in *GLOBECOM, Texas, USA*, December 2011.
- [25] N. Torabi, K. Rostamzadeh, V. C. Leung, "Rank-optimal channel selection strategy in cognitive networks," in *GLOBECOM, California, USA*, December 2012.
- [26] J. Rosenski, O. Shamir, L. Szlak, "Multi-player bandits-a musical chairs approach," in *ICML, New York, USA*, June 2016.
- [27] O. Avner, S. Mannor, "Concurrent bandit and cognitive radio networks," in *European Conf. on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, Nancy, France, September 2014.
- [28] K. Liu, Q. Zhao, B. Krishnamachari, "Decentralized multi-armed bandit with imperfect observations," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, USA, October 2010.
- [29] Y. Gai, B. Krishnamachari, "Decentralized Online Learning Algorithms for Opportunistic Spectrum Access," in *GLOBECOM, Texas, USA*, December 2011.
- [30] M. Almasri, A. Mansour, C. Moy, A. Assoum, C. Osswald, D. Lejeune, "Distributed Algorithm under Cooperative or Competitive Users with Priority Access in Cognitive Networks," *EURASIP journal on wireless communications and networking*, (Accepted).



# Nonlinear $\ell_{2,p}$ -norm based PCA for Anomaly Network Detection

Amal Hadri<sup>\*1</sup>, Khalid Chougali<sup>2</sup>, Raja Touahni<sup>1</sup>

<sup>1</sup>LASTID Laboratory, Faculty of Science, Ibn tofail University, 14000, Morocco

<sup>2</sup>GREST Research Group, National School of Applied Sciences (ENSA), 14000, Morocco

## ARTICLE INFO

Article history:

Received: 18 May, 2020

Accepted: 17 July, 2020

Online: 28 July, 2020

Keywords:

PCA

$\ell_{2,p}$ -norm PCA

Nonlinear  $\ell_{2,p}$ -norm PCA

Network Anomaly Detection

Feature extraction methods

KDDCup99

NSL-KDD

## ABSTRACT

Intrusion detection systems are well known for their ability to detect internal and external intrusions, it usually recognizes intrusions through learning the normal behaviour of users or the normal traffic of activities in the network. So, if any suspicious activity or behaviour is detected, it informs the users of the network. Nonetheless, intrusion detection system is usually prone to a high false positive rate & a low detection rate as a consequence of the tremendous amount of meaningless information used in the network traffic utilized to create the intrusion detection system. To overcome that, many techniques like Principal Component Analysis (PCA),  $L_1$ -PCA and  $\ell_{2,p}$ -norm based PCA were suggested. However, these methods are linear and not robust to outliers. This paper introduces the nonlinear variant of the  $\ell_{2,p}$ -norm principal component analysis. Namely, the nonlinear  $\ell_{2,p}$ -norm principal component analysis intends to project the data sets into a more feasible form so that the meaning of the data is damaged as less as possible. The proposed technique is not uniquely robust to outliers but keeps PCA's positive properties as well. Experimental results on the datasets KDDCup99 and NSL-KDD show that the proposed technique is extra effective, robust and outperform PCA,  $L_1$ -PCA and  $\ell_{2,p}$ -norm based PCA algorithms.

## 1 Introduction

Substantial shift in the proliferation of network security tools as well as the highly sophisticated attacks and intrusions are occurring in world during this information age. The conventional techniques in network security landscape such as data encryption, firewalls & user authentication are not sufficient to protect against existing threats. In order to detect any damages caused by attackers in a particular network. Denning [1] has introduced the Intrusion Detection System (IDS). IDSs are used to analyse network packets and determine if the intrusion is threat or not. Two types for IDSs have been developed network-based and host-based IDSs.

Generally, IDS are either, anomaly based or misuse based techniques (equally called knowledge or signature-based) [2]. These techniques have their advantages and limits. In misuse-based method, a database of known attacks signatures (also known as patterns) compares attacks signatures to the data (packets) existing in the network. When a signature is detected, this method produces an alarm signalling a known intrusion. However, the weakness of this method is in its inability to detect new intrusions (or attacks). In the anomaly-based IDS, behavioural reference of system or network is built based on the use of the data that represent the normal

behaviour. Consequently, the zero day attacks are well managed and attacks or intrusions in this case are simply any action that deviates from the pre-defined reference. Nonetheless, due to the noisy and redundant traffic data that contain many irrelevant features, anomaly-based technique may produce a significant amount of false alarms leading to unsatisfactory detection rate.

To address the issue of high dimensionality, some feature reduction and feature selection techniques have been used such as the principal component analysis (PCA) [3, 4] & the linear discriminant analysis (LDA) [5]. Additionally in the context of feature extraction, regularized discriminant analysis RDA [6], the quadratic discriminant analysis (QDA), with maximum margin criterion(MMC) [7] have been used.

Other techniques exist which are extended variants of PCA like for instance: (i) Kernel PCA [8] that maps nonlinearly the original data into a higher-dimensional space & after that it applies the PCA algorithm to extract the features (ii) the weighted PCA (WPCA) [9] it employs a weighted distance to address the impact of outliers onto the directions, (iii) the popular Fuzzy PCA [10–13] that fuzzify the original data to get the fuzzy membership for every data & transform PCA into Fuzzy PCA, (iv) Sparse PCA [14] that extends the classical PCA by introducing sparsity structures to the input variables.

\*Corresponding Author Amal Hadri, LASTID Laboratory, Faculty of Science, Ibn tofail University, 242, Morocco, amal.hadri@uit.ac.ma

The precedent algorithms are generally based on a global Euclidean structure. Unlike the manifold learning algorithms, which are well designed to maintain the local geometric structure of data and captivate the attention of many researchers in machine learning and the recognition of patterns fields. The most-known manifold learning techniques are: marginal fisher analysis (MFA) [15], locality preserving projection (LPP) [16] & Neighbourhood Preserving Embedding (NPE) [17]. MFA is a supervised manifold learning technique that endures the local manifold information, LPP is mainly a linear approximation of Laplacian embedding (LE) [18] & NPE is principally a linear approximation of locally linear embedding (LLE) [19]. Many interesting methods for dimensionality reduction have been made based on these techniques (MFA, LPP and NPE).

Lately, and to enhance the effectiveness to outliers for feature extraction many approaches use several criterion functions like  $L_1$ -norm maximization or minimization and nuclear norm [20–25]. Nuclear norm obtains clean data with low-rank structure, however it remains out of sample issue.  $L_1$ -norm based subspace learning method is among them & it turn into a very attractive subject in dimensionality reduction & machine learning fields. For instance, [22] developed a technique called  $L_1$ -PCA, where the projection matrix is obtained through minimizing  $L_1$ -norm-based reconstruction error in the objective function of PCA. Solving  $L_1$ -PCA is computationally costly, to tackle this problem, Kwak [21] proposed a method called PCA-L1 that solves the principal components through maximizing the variance, that is computed via  $L_1$ -norm. To better demonstrate the efficiency of subspace learning techniques, Kwak et al. extended  $L_1$ -norm into  $L_p$ -norm and introduced  $L_p$ -norm-based LDA and PCA [26, 27].

However, almost all the  $L_1$ -norm based PCA techniques can not optimally minimize the reconstruction error, that is the main point of PCA. Additionally, these techniques are not invariant to rotation, which is significant property in learning algorithms [26, 27]. To address these issues, R1-PCA [27] was introduced to minimize the reconstruction error by way of putting  $l_2$ -norm on the spatial dimension and the  $L_1$ -norm on the data. Optimal mean R1-PCA [28] was proposed as well, this algorithm uses the optimal mean in R1-norm instead of the fixed mean utilized in R1-PCA. Inspired by this, [29] proposed the  $\ell_{2,p}$ -norm based PCA and extends R1-PCA into a generalized robust distance metric learning formulation for PCA. The idea behind this algorithm is to utilize  $\ell_{2,p}$ -norm as a distance metric for the reconstruction error, and uses a non-greedy algorithm as an optimal solution, that has a closed-form solution in every iteration.  $\ell_{2,p}$ -norm PCA keeps all PCA's advantages such as rotational invariance. The optimal solution involves the covariance matrix, and it is robust against outliers.

$\ell_{2,p}$ -norm PCA [29] has its own weaknesses. In fact, it cannot be efficient against noise and outliers if the data that we are dealing with have nonlinear structures, which give rise to false results. To address one of the weaknesses in the area of intrusion detection system, we suggest a nonlinear version of  $\ell_{2,p}$ -norm based PCA less prone to outliers. Our nonlinear  $\ell_{2,p}$ -norm based PCA technique is extra robust than the conventional  $\ell_{2,p}$ -norm based PCA as demonstrated by the experiments we performed using two well-known data sets namely KDDcup99 [30–32] and NSL-KDD [33, 34].

The rest of this paper is organized as follows. Section II reviews  $\ell_{2,p}$ -norm based PCA. Nonlinear  $\ell_{2,p}$ -norm based PCA is suggested

in Section III. In Section IV, we present the simulated datasets. Section V reports the experiments and discussion of the results. The conclusions are presented in Section VI.

## 2 $\ell_{2,p}$ -norm PCA Algorithm

The  $\ell_{2,p}$ -norm based PCA utilized here and from where the nonlinear case is originated was proposed in [29]. The  $\ell_{2,p}$ -norm based principal component algorithm which Wang & al. suggested in [29] is mainly based on the principal component algorithm (PCA)[3, 4] where the large reconstruction errors, dominate the objective function. In Wang & al. proposed algorithm, the objective function was extended to reduce the impact of large distance and to include the rotational invariance. Following, the algorithm is briefly presented. A more detailed description can be found in [29].

Wang & al. proposed a generalized robust PCA where:

$$\min_W \sum_{i=1}^N \|x_i - WW^T x_i\|_2^p \quad (1)$$

$$\text{subject to } W^T W = I_k.$$

where  $0 < p \leq 2$ .

Notice here that the optimization process of the objective function (1) is very hard, so they simplified the objective function by using simple algebra as follows:

$$\sum_{i=1}^N \|x_i - WW^T x_i\|_2^2 \|x_i - WW^T x_i\|_2^{p-2} \quad (2)$$

$$= \sum_{i=1}^N \text{tr} \left\{ (x_i - WW^T x_i)^T * (x_i - WW^T x_i) \right\} d_i \quad (3)$$

$$= \sum_{i=1}^N \text{tr} \left\{ x_i^T x_i - x_i WW^T x_i - x_i WW^T x_i + x_i WW^T WW^T x_i \right\} d_i \quad (4)$$

$$= \sum_{i=1}^N \text{tr} \left\{ x_i^T x_i - x_i WW^T x_i \right\} d_i \quad (5)$$

$$\text{where } d_i = \|x_i - WW^T x_i\|_2^{p-2}.$$

Replacing Eq. (5) within the objective function (1), and through using simple algebra, the function (1) turns into :

$$\min_W \sum_{i=1}^N \text{tr} \left\{ x_i^T x_i \right\} d_i - \sum_{i=1}^N \text{tr} \left\{ W^T x_i x_i^T W \right\} d_i \quad (6)$$

The main goal now, is how to solve the optimal projection matrix  $W$  of the objective function (6). The aim is to obtain a projection matrix  $W$  that will minimize the value of the objective function (6). The objective function (6) has unknown variables  $W$  &  $d_i$  that is connected with  $W$ . Therefore, it is very hard to straightforwardly solve the objective function (6) since it does not have a closed-form solution. Therefore, an algorithm will be elaborated now for alternately updating  $W$  (while keeping  $d_i$  fixed) and  $d_i$  (while keeping  $W$  fixed). To get extra precise, in the  $(t + 1)^{\text{th}}$  iteration, when  $d_i^{(t)}$

is known, accordingly we will minimize the objective function (6) in order to update  $W$ . In this particular case, the first term in the objective function (6) turns into a constant. Consequently, Eq. (6) becomes:

$$W^* = \operatorname{argmax} \operatorname{tr}(W^T XDX^T W) \quad (7)$$

subject to  $W^T W = I_k$ .

where  $D$  is a diagonal matrix &  $d_i$  are its elements on diagonal, and where the column vectors in  $W$  of the objective function (7) which contains the eigenvectors of  $XDX^T$  matching to the  $k$  largest eigenvalues. Then, the diagonal element  $d_i$  of the matrix  $D$  is updated. Until the algorithm is converged, the prior iterative procedure will be repeated. The pseudocode of solving the objective function (1) is summarized in Algorithm 1.

**Algorithm 1**  $\ell_{2,p}$ -norm based PCA

Input:  $X = [x_1, x_2, x_3, \dots, x_N] \in R^{m \times N}$ ,  $k, p$ , where  $X$  is centralized.

Initialize:  $W_1 \in R^{m \times k}$  which satisfies the equation  $W^T W = I_k$ ,  $t = 1$ . While not converge do

1. Calculate diagonal matrix  $D$  whose diagonal elements are  $d_i = \|x_i - WW^T x_i\|_2^{p-2}$ .
2. Compute the weight covariance matrix  $XDX^T$ .
3. Solve  $W^* = \operatorname{argmax} \operatorname{tr}(W^T XDXW)$   
The columns vectors of optimal projection matrix  $W_t$  which contains the first  $k$  eigenvectors of  $XDX^T$  matching to the  $k$  largest eigenvalues.
4.  $t \leftarrow t + 1$ ;

end while

Output:  $W_t \in R^{m \times k}$

### 3 The proposed method

$\ell_{2,p}$ -norm based PCA [29], like all the linear variants of PCA fails sometimes in producing the optimal projection vectors since it permits uniquely a linear dimensionality reduction [35]. Therefore, if we are dealing with complex nonlinear structures of data, that can be presented differently in a linear space, linear variants of PCA will skew the results. To address this problem, this section introduces a new nonlinear version of  $\ell_{2,p}$ -norm based PCA namely nonlinear  $\ell_{2,p}$ -norm based PCA.

#### 3.1 Nonlinear $\ell_{2,p}$ -norm PCA Algorithm

We suggest a generalized nonlinear robust PCA where:

$$\min_W \sum_{i=1}^N \|x_i - W^T g(y)\|_2^p \quad (8)$$

subject to  $W^T W = I_k$ .

Where  $y = x_i * w$ ,  $0 < p \leq 2$  and  $g$  can be chosen as nonlinear function. In this article, the function  $g$  was chosen to be sigmoid function like:

- Gudermannian function  
 $g(y) = \int_0^y \frac{1}{\cosh t} dt = 2\arctan(\tanh(\frac{y}{2}))$
- Generalised logistic function  
 $g(y) = (1 + e^{-x})^{-\alpha}, \alpha > 0$
- Arctangent function  
 $g(y) = \operatorname{artan} y$
- Hyperbolic tangent  
 $g(y) = \tanh y = \frac{e^y - e^{-y}}{e^y + e^{-y}}$

By utilizing simple algebra, equation (8) will be:

$$\sum_{i=1}^N \|x_i - W^T g(y)\|_2^2 \|x_i - W^T g(y)\|_2^{p-2} \quad (9)$$

$$= \sum_{i=1}^N \operatorname{tr} \{ (x_i - W^T g(y))^T * (x_i - W^T g(y)) \} d_i \quad (10)$$

$$= \sum_{i=1}^N \operatorname{tr} \{ x_i^T x_i - x_i W^T g(y) - g(y) W x_i + g(y)^T W W^T g(y) \} d_i \quad (11)$$

$$= \sum_{i=1}^N \operatorname{tr} \{ x_i^T x_i - x_i^T W^T g(y) \} d_i \quad (12)$$

where  $d_i = \|x_i - W^T g(y)\|_2^{p-2}$ .

Replacing Eq. (12) into the objective function (8), The objective function (8) turns into

$$\min_W \sum_{i=1}^N \operatorname{tr} \{ x_i^T x_i \} d_i - \sum_{i=1}^N \operatorname{tr} \{ W^T g(y) x_i^T \} d_i \quad (13)$$

Now our ultimate aim is to have a projection matrix  $W$  that can minimize the value of the objective function (13). 2 unknown variables  $W$  and  $d_i$  existing in the objective function (13). Hence, it will not accept a closed-form solution & it is hard to straightforwardly solve the solution of the objective function (13). Therefore, an algorithm could be elaborated for alternately updating  $W$  (while keeping  $d_i$  fixed) and  $d_i$  (while keeping  $W$  fixed). In more details, in the  $(t + 1)^{th}$  iteration, when  $d_i^{(t)}$  is known, accordingly  $W$  updated through the minimization the objective function (13). In cases like these, the first term in the function (13) turns into a constant. Therefore, Eq. (13) becomes

$$W^* = \operatorname{argmax} \operatorname{tr}(W^T g(Y) X^T D) \quad (14)$$

subject to  $W^T W = I_k$ .

where  $Y = XW$  and  $D$  is a diagonal matrix &  $d_i$  are its elements on diagonal. The column vectors in  $W$  of the objective function (14) which contains the eigenvectors of  $XDX^T$  matching the  $k$  largest eigenvalues. Afterwards, the element  $d_i$  is updated. Until the algorithm is converged, the prior iterative procedure will be repeated. Finally, we recapitulate the proposed method in Algorithm 2.

**Algorithm 2** Nonlinear  $\ell_{2,p}$ -norm PCA

Input:  $X = [x_1, x_2, x_3, \dots, x_N] \in \mathbb{R}^{m \times N}$ ,  $k, p$ , where  $X$  is centralized.

Initialize:  $W_1 \in \mathbb{R}^{m \times k}$  which satisfies the condition  $W^T W = I_k$ ,  $t = 1$ .

While not converge do

1. Calculate diagonal matrix  $D$  whose diagonal elements are  $d_i = \|x_i - W^T g(y)\|_2^{p-2}$ .
2. Solve  $W^* = \operatorname{argmax} \operatorname{tr}(W^T g(Y) X^T D)$ .  
The columns vectors of optimal projection matrix  $W_i$  that contain the first  $k$  eigenvectors of  $XDX^T$  matching to the  $k$  largest eigenvalues.
3.  $t \leftarrow t + 1$ ;

end while

Output:  $W_t \in \mathbb{R}^{m \times k}$

## 4 The Simulated Datasets

In this section, we present briefly the two datasets utilized during our experiments as well as the pre-processing step used to standardize these datasets.

### 4.1 KDDCup99 dataset

The KDDCup99 [30, 31] dataset was utilized in the KDD (Knowledge Discovery & Data Mining Tools Conference) Cup 99 Competition [32]. It is created & managed by DARPA Intrusion Detection Evaluation Program, since it was derived from the original DARPA dataset. It does contain several TCPdump raws, collected over 9 weeks.

The creation of the dataset was achieved in two phases: the first phase took over seven weeks and it was dedicated to create the training data. The full training data has almost 5 million connection records. The second phase took 2 weeks and it was dedicated to create the test data which represents two million connection records.

Almost 20% of the 2 datasets are normal connections (not attacks). Concerning attack connections, the 39 types of attacks are categorized into 4 categories: DOS, R2L, U2R, PROBE.

- Denial of Service (DoS) attacks, the attacker seeks to make some resources unavailable to handle simple and legitimate requests, consequently denying legitimate users access to a resource.
- Probe attacks, the attacker seeks to search out as much as possible network vulnerabilities and collect information via scanning a network.
- Remote-to-Local (R2L) attacks, the attacker seeks to gain access to the targeted system in an illegal way via transmitting packets to a machine over a network, then exploits machines vulnerability through compromising the security by the way of password guessing/breaking.

- User-to-Root (U2R) attacks, by utilizing Buffer overflow attack, the attacker seeks to gain root access to the system through just having a normal user account.

We can classify KDDCup99 features into three categories:

- Basic features: This group contains all the attributes that could be excerpted from a TCP/IP connection. The vast majority of these features usually engender an indirect delay in detection.
- Traffic features: The features belonging to this group are calculated in regard to a window interval.
- Content features: More than half of Probing and DoS attacks possess countless intrusion frequent sequential patterns, it is caused because these attacks carry out numerous connections to the host(s) in a lapse of time. On the other side, The U2R and R2L attacks are encapsulated in the payload of the packets, & usually contain uniquely a single connection, consequently, they do not have any intrusion frequent sequential patterns. To identify these types of attacks, some extra features are required to identify doubtful behavior in the packet payload. We call these features “content features”.

### 4.2 NSL-KDD dataset

NSL-KDD [33, 34] was proposed to solve some of KDDcup99 dataset drawbacks. The main ameliorations are:

1. An acceptable number of train sets (125973 samples) & test sets (22544 samples), that is rational and make it easier to make experiments with the entire data set.
2. The non-existence redundancy sample in the dataset & therefore it enables the classifiers to generate an un-skewed result.
3. The training set has different probability distribution compared to the test set.
4. Unknown attack types are existing in the test set, & they do not exist in the training set that makes it more reasonable.

It should be mentioned that the attack classes existing in the NSL-KDD database are also classified into 4 categories: DOS, R2L, U2R, PROBE and for every record there are 41 attributes developing different features of the flow & a label assigned to each either as an attack type or as normal.

### 4.3 The pre-processing step

Some classifiers generate a improved accuracy rate on normalized data sets, this is why the pre-processing step is more than crucial. This step was successfully carried out by replacing the discrete attributes values of the databases toward continuous values through exploring the same idea utilized in [36], the concept will be explained briefly as follow: for any discrete attribute  $i$  that accept  $k$  dissimilar values. The attribute  $i$  can be asserted via  $k$  coordinates include ones & zeros. For example, the attribute called protocol type, that takes 3 values, i.e. tcp, udp or icmp. According to the concept, the prior values turn into their equivalent coordinates (1,0,0), (0,1,0) or (0,0,1).



## 5 Experiments and Discussion

The current section is dedicated to present the experiments carried out to show the efficiency of our proposed algorithm. To examine our approach we used the popular datasets KDDcup99 & NSL-KDD, additionally we compare our method with PCA [3, 4],  $L_1$ -PCA [22],  $\ell_{2,p}$ -norm based PCA [29] to prove the effectiveness of our algorithm. The following measures are calculated: detection rate (DR), false positive rate (FPR) & F-measure presented underneath:

$$DR = \frac{TP}{TP + FN} * 100 \tag{15}$$

$$FPR = \frac{FP}{FP + TN} * 100 \tag{16}$$

$$FMeasure = \frac{2 * TP}{2 * TP + FP + FN} * 100 \tag{17}$$

where True positives (TP) are attacks successfully predicted. False negatives (FN) correspond intrusions classified as normal instances, false positive (FP) are normal instances badly classified, & true negatives (TN) are normal instances successfully predicted. We considerate that the most trustworthy and effective feature extraction algorithm is the one with the highest DR / F-measure & with the lowest FPR.

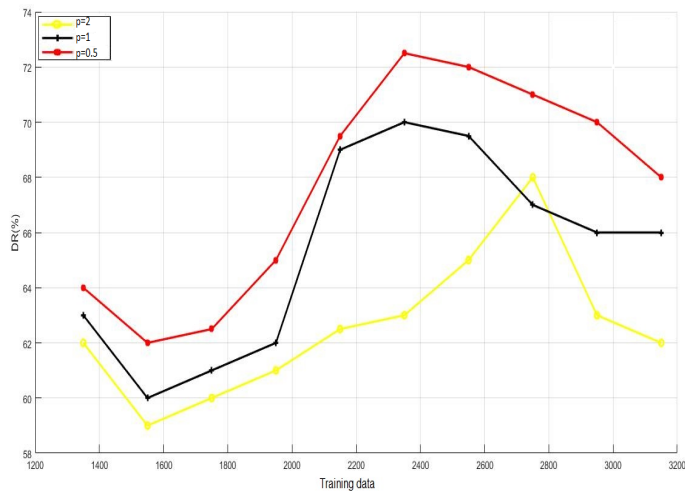


Figure 1: Evaluation of different values of the parameter p on KDDcup99 database.

In all our experiments, we utilized the nearest neighbour classifier since our goal is examining the efficacy of the feature extraction technique, and in order to obtain results that are more realistic we calculated the mean of twenty times. Hence, DR, F-measure and FPR took the average. We carried several experiments to test our approach, and each experiment has its own simulation settings.

Regarding the simulation settings of our first experiments, we choose to keep the test dataset intact with the following composition (100 normal data, 100 DOS data, 50 U2R data, 100 R2L data, and 100 PROBE), and vary the number of training samples. The main idea behind our first experiment is to evaluate all the techniques cited before under multiple training dimensionality.

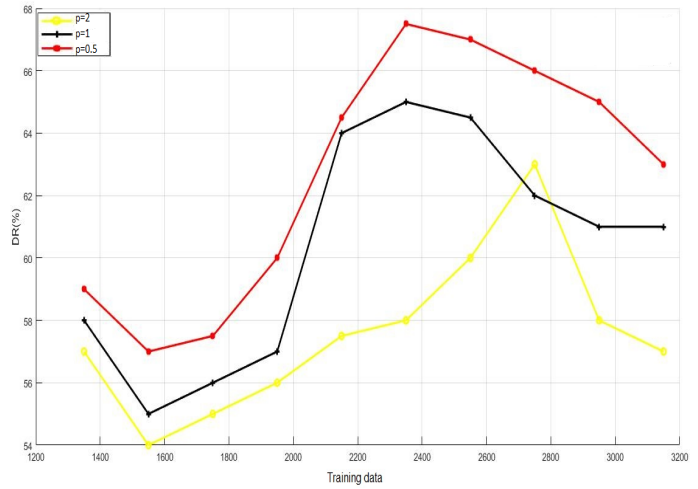


Figure 2: Evaluation of different values of the parameter p on NSL-KDD database.

The first two experiments were made to define the adequate initial parameter p as well as the nonlinear function that increases the efficiency of the nonlinear  $\ell_{2,p}$ -norm based PCA.

Figure 1 and Figure 2 plot the values of the detection rate versus the training data with different values of p for both datasets (KDD-Cup99 & NSL-KDD). As it is shown in Figure 1 and Figure 2, the detection rate is at its lowest values when p is 2. However, when p = 1 and p=0.5, it increases the detection rate for both datasets. The explanation for this is if we increase the value of p the impact of outliers will increase, thus the value of the reconstruction error will be huge and dominate the objective function 8. Consequently, we set p to 0.5 for the next experiments.

The second experiment, as explained before, aims to find the best nonlinear function that enhances the effectiveness of our proposed technique. The Figure 3 and Figure 4 depict the values of the detection rate using four different sigmoidal functions for both datasets (KDDCup99 & NSL-KDD). As it is seen in Figure 3 and Figure 4 the hyperbolic tangent is the function that produce the highest detection rate on KDDCup99 database as well as for the NSL-KDD database. Therefore, we will be using this function in the rest of our experiments.

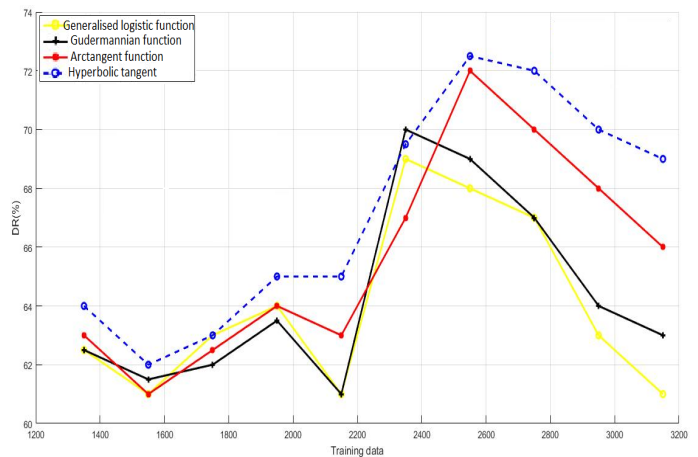


Figure 3: Evaluation of different nonlinear functions on KDDcup99 database.

Figures 5, 6 and 7 show the results we obtained when com-

paring our technique to the precedent linear PCA algorithms for KDDcup99 dataset. As stated in the Figure 5 and 7, we notice that Nonlinear  $\ell_{2,p}$ -norm based PCA surpasses all PCA algorithms once the training data exceed 2200. The simple reason that explain this phenomenon is that the more our training data is big the more outliers are visible. Since the other models except Nonlinear  $\ell_{2,p}$ -norm based PCA are linear, hence they will be more sensitive to outliers, which reduce their effectiveness.

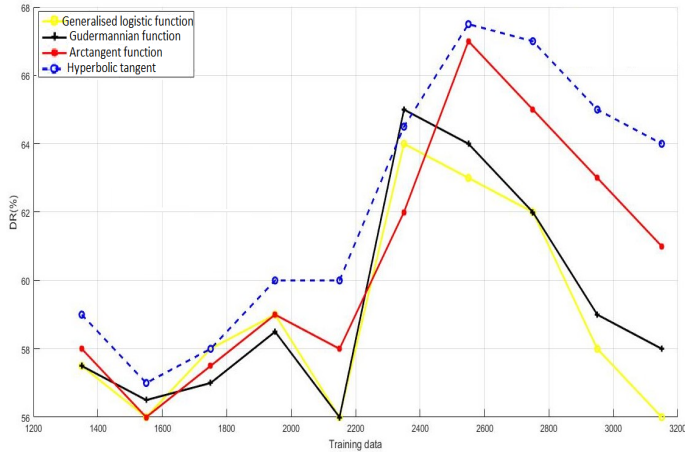


Figure 4: Evaluation of different nonlinear functions on NSL-KDD database.

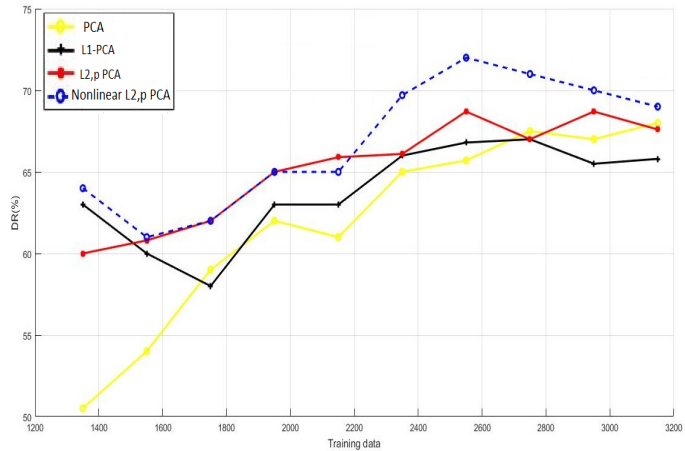


Figure 5: Training data vs. DR for KDDcup99 database.

Another explanation to that is the distribution nature of data, i.e if the structure of data is nonlinear or extra-complex, that can be badly resented in a linear space, the linear models of PCA will be less efficient.

Regarding FPR, Figure 6 exhibits that maximum value of the false positive rate of the suggested algorithm is around 3%. This proves that the approach has the ability to differentiate normal connections from attacks.

Concerning NSL-KDD dataset, Figures 8, 9 and 10 assert that the nonlinear  $\ell_{2,p}$ -norm based PCA overcomes the linear variants of PCA, and it enhances the detection rate by at least 6% over PCA and  $L_1$ -PCA, 3% over  $\ell_{2,p}$ -norm based PCA. Also, as we observe from the Figure 9 the proposed approach still produce the lower values for the false positive rate compared to linear PCA models.

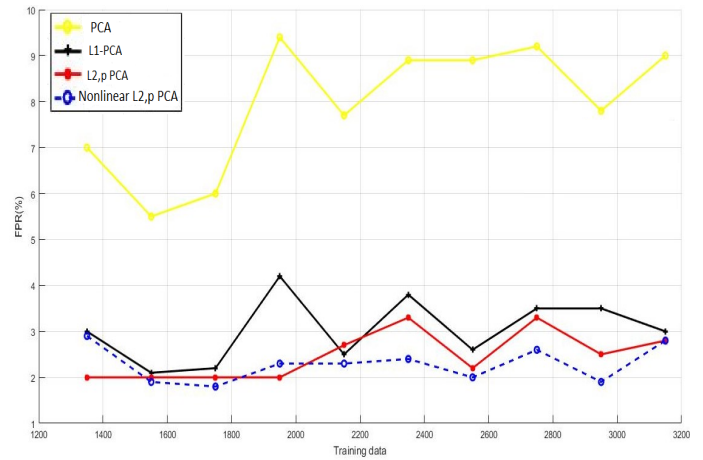


Figure 6: Training data vs. FPR for KDDcup99 database.

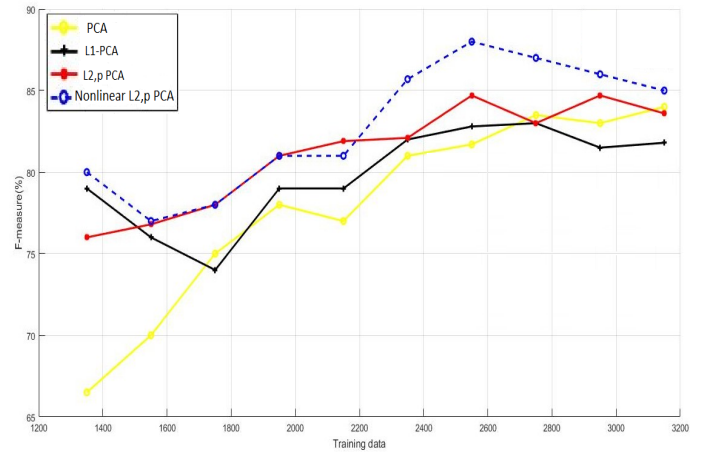


Figure 7: Training data vs. F-measure for KDDcup99 database.

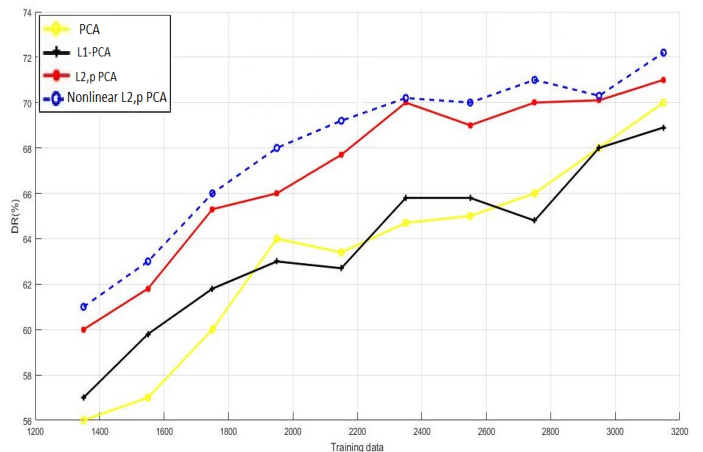


Figure 8: Training data vs. DR for NSL-KDD database.

In the second experiment, we examined the suggested approach when changing the number of principal component, we choose just 10 of the 41 principal components and increased their number during the simulation.

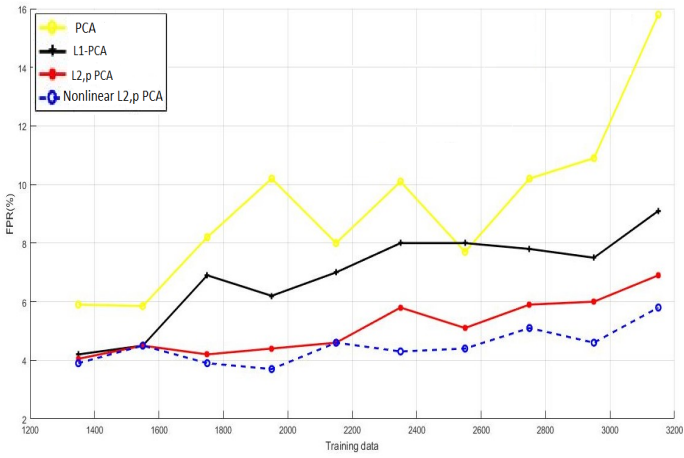


Figure 9: Training data vs. FPR for NSL-KDD database.

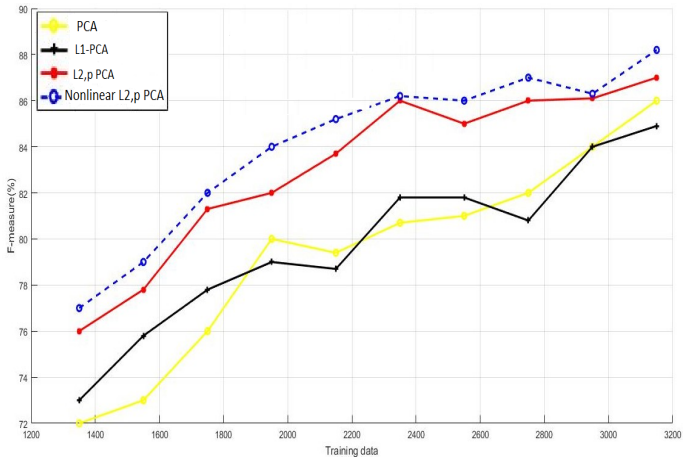


Figure 10: Training data vs. F-measure for NSL-KDD database.

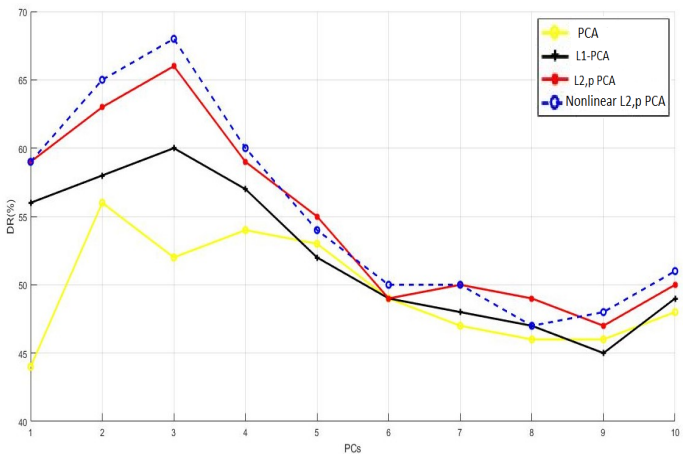


Figure 11: Principal Components vs. DR for KDDcup99 database.

data, 50 U2R data, 100 R2L data and 100 PROBE ) chosen in a random way for the two databases (KDDcup99 & NSL-KDD).

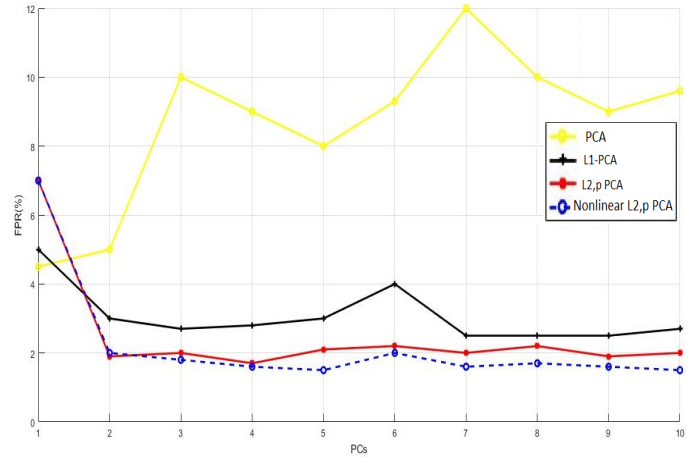


Figure 12: Principal Components vs. FPR for KDDcup99 database.

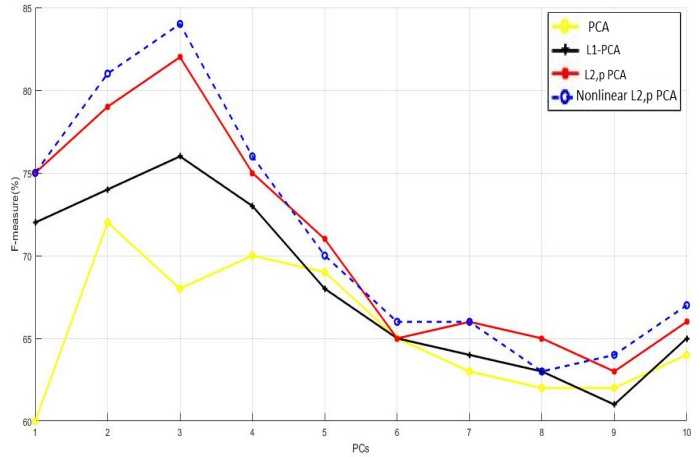


Figure 13: Principal Components vs. F-measure for KDDcup99 database.

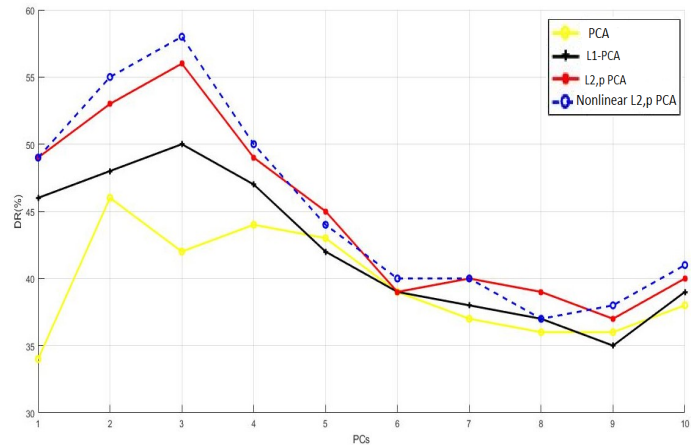


Figure 14: Principal Components vs. detection rate for NSL-KDD.

Note that the simulation settings were different from the first experiment. This time we utilized the following composition: for the training samples (1000 normal, 100 DOS, 50 U2R, 100 R2L and 100 PROBE) and for the test samples (100 normal data, 100 DOS

From Figure 11 and 13, we can see that nonlinear  $\ell_{2,p}$ -norm based PCA takes the lead over the linear PCA models and preserves

its superiority in producing high DR and F-measure values, it gives at least a 60 % for the first principal component and achieves around 67% as a maximum detection rate, and at least 75 % for the first principal component and gets around 84% as a maximum F-measure value. Additionally, we observe that the new approach outperform all the other linear PCA variants. Concerning the FPR, we can see from Figure 12 that nonlinear  $\ell_{2,p}$ -norm based PCA gives the lowest FPR unlike the aforementioned linear once.

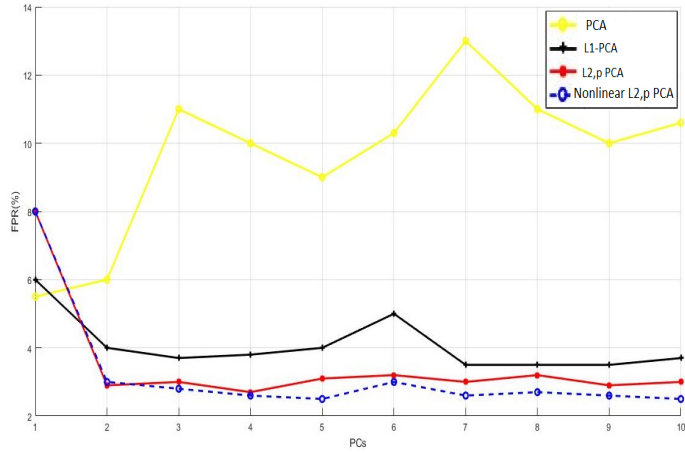


Figure 15: Principal Components vs. False positive rate for NSL-KDD.

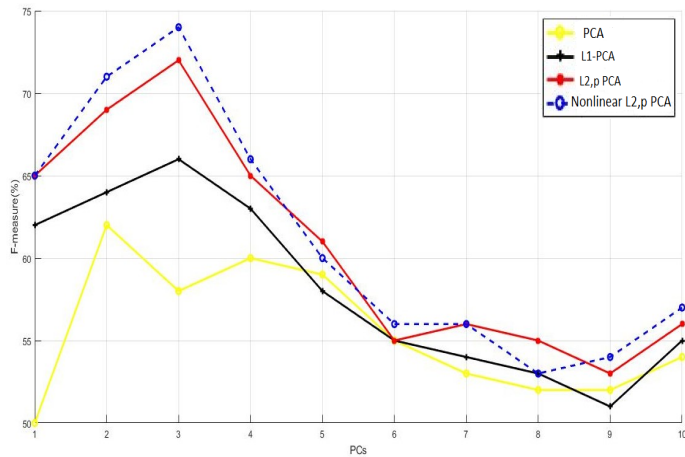


Figure 16: Principal Components vs. F-measure for NSL-KDD.

When we have performed identical experiment on NSL-KDD dataset, as it is clear from the Figures 14 and 16 that the new approach ensure improved rates of DR and F-measure over the linear PCA algorithm. For the false positive rate, as illustrated in the Figure 15, the proposed method has fewest false positive rate starting from the third principal component.

In the fourth simulation, we computed the amount of time consumed by each algorithm for both datasets. From figures 17 and 18, we observe that for both dataset the amount of time (CPU time) required is increasing proportionally as the principal components PCs number is increasing for all techniques. The only difference is that the suggested technique is a little more computationally speedy than the other algorithms which is expected.

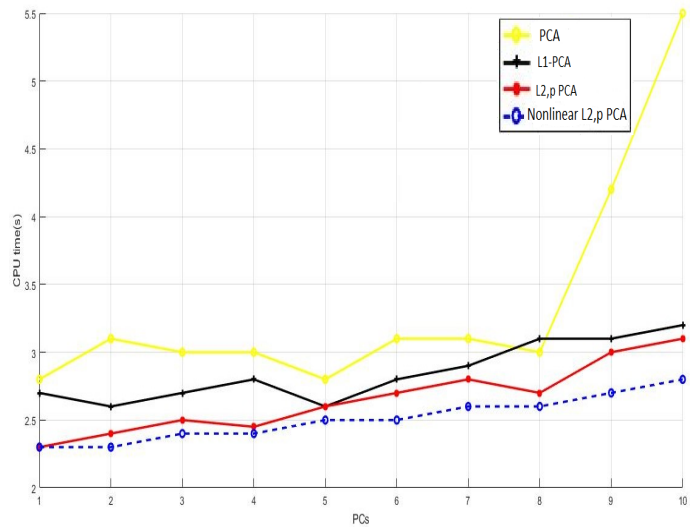


Figure 17: Principal Components vs. CPU time (s) for KDDCup99 database.

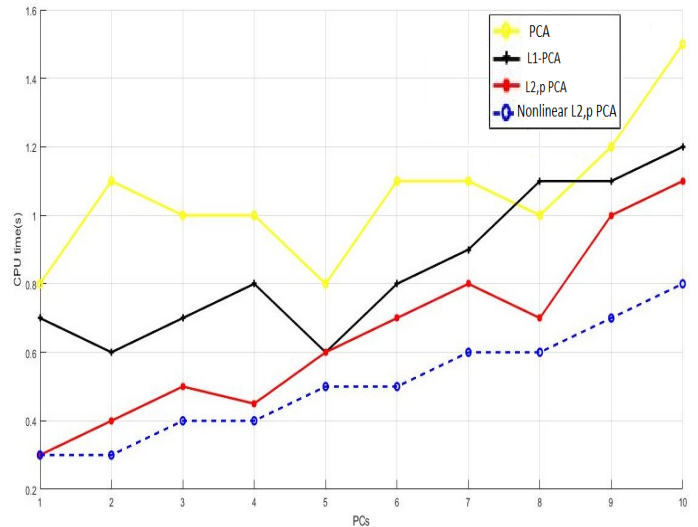


Figure 18: Principal Components vs. CPU time (s) for NSL-KDD database.

To get further insights about the effectiveness of the proposed approach, we carried an experiment where we calculated the detection rates for every single attack category for the aforementioned PCA variants as well as the proposed approach. We can observe clearly in Table 1 and Table 2 that the proposed technique outperform the others variants in identifying attacks for the KDDCup99 and NSL-KDD datasets.

Table 1: Attacks Detection Rate for PCA,  $L_1$ -PCA,  $\ell_{2,p}$ -norm PCA and Nonlinear  $\ell_{2,p}$ -norm PCA for KDDCup99 dataset.

	Method	DOS	U2R	R2L	Probing
DR(%)	PCA	68,7656	8,7329	4,7734	92,1342
	$L_1$ -PCA	72,3478	15,4635	4,1315	91,4325
	$\ell_{2,p}$ -norm PCA	74,9319	16,8951	4,1111	92,8325
	Nonlinear $\ell_{2,p}$ -norm PCA	76,5314	17,0132	4,6783	94,1311



Table 2: Attacks Detection Rate for PCA,  $L_1$ -PCA,  $\ell_{2,p}$ -norm PCA and Nonlinear  $\ell_{2,p}$ -norm PCA for NSL-KDD dataset.

	Method	DOS	U2R	R2L	Probing
DR(%)	PCA	67,6656	7,6319	4,6623	91,1142
	$L_1$ -PCA	71,3468	14,3525	4,1214	90,3215
	$\ell_{2,p}$ -norm PCA	73,8219	15,9715	4,0123	91,7523
	Nonlinear $\ell_{2,p}$ -norm PCA	74,3441	16,1023	4,7738	93,1125

## 6 Conclusion

In the current paper, we suggest a nonlinear variant of the  $\ell_{2,p}$ -norm based PCA, the suggested algorithm showed significant improvements compared to the original one. In addition, integrating nonlinear  $\ell_{2,p}$ -norm based PCA into our intrusion detection system (IDS) makes the prior more efficient and powerful against outliers. As we showed earlier, experiments on the popular datasets KDD-cup99 & NSL-KDD demonstrate that the nonlinear  $\ell_{2,p}$ -norm PCA outperforms and show its superiority over PCA,  $L_1$ -PCA and the original variant  $\ell_{2,p}$ -norm PCA. In the future works, we will attempt to test our IDS on recent datasets and and develop other variants of  $\ell_{2,p}$ -norm PCA.

**Conflict of Interest** The authors declare no conflict of interest.

**Acknowledgment** This work is supported by CNRST-MOROCCO under the excellence program, grant no. 15UIT2016.

## References

- [1] D. E. Denning, "An intrusion-detection model," *IEEE Transactions on software engineering*, (2), 222–232, 1987, doi:10.1109/TSE.1987.232894.
- [2] L. Portnoy, L. Eskin, S. Stolfo, "Intrusion detection with unlabeled data using clustering [CJ//Proc of ACM CSS Workshop on Data mining Applied to Security (DMSA-2001)]," 2001.
- [3] M. Ringnér, "What is principal component analysis?" *Nature biotechnology*, **26**(3), 303, 2008, doi:10.1038/nbt0308-303.
- [4] I. T. Jolliffe, J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **374**(2065), 20150202, 2016, doi:10.1098/rsta.2015.0202.
- [5] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern classification*, John Wiley & Sons, 2012.
- [6] D.-Q. Dai, P. C. Yuen, "Face recognition by regularized discriminant analysis," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **37**(4), 1080–1085, 2007, doi:10.1109/TSMCB.2007.895363.
- [7] H. Li, T. Jiang, K. Zhang, "Efficient and robust feature extraction by maximum margin criterion," in *Advances in neural information processing systems*, 97–104, 2004.
- [8] B. Schölkopf, A. Smola, K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural computation*, **10**(5), 1299–1319, 1998, doi:10.1162/089976698300017467.
- [9] Y. Koren, L. Carmel, "Robust linear dimensionality reduction," *IEEE transactions on visualization and computer graphics*, **10**(4), 459–470, 2004, doi:10.1109/TVCG.2004.17.
- [10] X. Wu, J. Zhou, "Fuzzy principal component analysis and its Kernel-based model," *Journal of Electronics (China)*, **24**(6), 772–775, 2007, doi:10.1007/s11767-006-0039-z.
- [11] S.-I. Xu, Q.-j. Zhang, "Gait recognition using fuzzy principal component analysis," in *2010 2nd International Conference on E-business and Information System Security*, 1–4, IEEE, 2010, doi:10.1109/EBISS.2010.5473671.
- [12] C. Sarbu, H. Pop, "Principal component analysis versus fuzzy principal component analysis: a case study: the quality of Danube water (1985–1996)," *Talanta*, **65**(5), 1215–1220, 2005, doi:10.1016/j.talanta.2004.08.047.
- [13] A. Hadri, K. Chougali, R. Touahni, "Intrusion detection system using PCA and Fuzzy PCA techniques," in *2016 International Conference on Advanced Communication Systems and Information Security (ACOSIS)*, 1–7, IEEE, 2016, doi:10.1109/ACOSIS.2016.7843930.
- [14] I. T. Jolliffe, N. T. Trendafilov, M. Uddin, "A modified principal component technique based on the LASSO," *Journal of computational and Graphical Statistics*, **12**(3), 531–547, 2003, doi:10.1198/1061860032148.
- [15] S. Yan, D. Xu, B. Zhang, H.-J. Zhang, Q. Yang, S. Lin, "Graph embedding and extensions: A general framework for dimensionality reduction," *IEEE transactions on pattern analysis and machine intelligence*, **29**(1), 40–51, 2006, doi:10.1109/TPAMI.2007.250598.
- [16] X. He, P. Niyogi, "Locality preserving projections," in *Advances in neural information processing systems*, 153–160, 2004.
- [17] X. He, D. Cai, S. Yan, H.-J. Zhang, "Neighborhood preserving embedding," in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, volume 2*, 1208–1213, IEEE, 2005, doi:10.1109/ICCV.2005.167.
- [18] M. Belkin, P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, **15**(6), 1373–1396, 2003, doi:10.1162/089976603321780317.
- [19] S. T. Roweis, L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *science*, **290**(5500), 2323–2326, 2000, doi:10.1126/science.290.5500.2323.
- [20] G. Liu, Z. Lin, Y. Yu, "Robust subspace segmentation by low-rank representation," in *ICML*, volume 1, 8, 2010.
- [21] N. Kwak, "Principal component analysis based on L1-norm maximization," *IEEE transactions on pattern analysis and machine intelligence*, **30**(9), 1672–1680, 2008, doi:10.1109/TPAMI.2008.114.
- [22] Q. Ke, T. Kanade, "Robust L/sub 1/norm factorization in the presence of outliers and missing data by alternative convex programming," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, 739–746, IEEE, 2005, doi:10.1109/CVPR.2005.309.
- [23] R. He, B.-G. Hu, W.-S. Zheng, X.-W. Kong, "Robust principal component analysis based on maximum correntropy criterion," *IEEE Transactions on Image Processing*, **20**(6), 1485–1494, 2011, doi:10.1109/TIP.2010.2103949.
- [24] Y. Wang, V. I. Morariu, L. S. Davis, "Unsupervised feature extraction inspired by latent low-rank representation," in *2015 IEEE Winter Conference on Applications of Computer Vision*, 542–549, IEEE, 2015, doi:10.1109/WACV.2015.78.
- [25] F. De La Torre, M. J. Black, "A framework for robust subspace learning," *International Journal of Computer Vision*, **54**(1-3), 117–142, 2003, doi:10.1023/A:1023709501986.
- [26] A. Y. Ng, "Feature selection, L 1 vs. L 2 regularization, and rotational invariance," in *Proceedings of the twenty-first international conference on Machine learning*, 78, ACM, 2004, doi:10.1145/1015330.1015435.
- [27] C. Ding, D. Zhou, X. He, H. Zha, "R 1-PCA: rotational invariant L 1-norm principal component analysis for robust subspace factorization," in *Proceedings of the 23rd international conference on Machine learning*, 281–288, ACM, 2006, doi:10.1145/1143844.1143880.
- [28] F. Nie, J. Yuan, H. Huang, "Optimal mean robust principal component analysis," in *International conference on machine learning*, 1062–1070, 2014.
- [29] Q. Wang, Q. Gao, X. Gao, F. Nie, "l2,p-Norm Based PCA for Image Recognition," *IEEE Transactions on Image Processing*, **27**(3), 1336–1346, 2017, doi:10.1109/TIP.2017.2777184.

- [30] K. Cup, "Data/The UCI KDD Archive, Information and Computer Science," University of California, Irvine, 1999.
- [31] M. Tavallae, E. Bagheri, W. Lu, A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, 1–6, IEEE, 2009, doi: 10.1109/CISDA.2009.5356528.
- [32] K. Cup, "Available on: <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>," 2007.
- [33] L. Dhanabal, S. Shantharajah, "A study on NSL-KDD dataset for intrusion detection system based on classification algorithms," *International Journal of Advanced Research in Computer and Communication Engineering*, **4**(6), 446–452, 2015, doi:10.17148/IJARCCCE.2015.4696.
- [34] X. Zong, Y. Sun, K. He, "Intrusion detection based on traffic research and application in Industrial Control System," 2018.
- [35] M. Kirby, L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces," *IEEE Transactions on Pattern analysis and Machine intelligence*, **12**(1), 103–108, 1990, doi:10.1109/34.41390.
- [36] Y. Bouzida, F. Cuppens, N. Cuppens-Bouahia, S. Gombault, "Efficient intrusion detection using principal component analysis," in 3<sup>ème</sup> Conférence sur la Sécurité et Architectures Réseaux (SAR), La Londe, France, 381–395, 2004.

## Buffering Supercapacitor Mechanism based on Bidirectional DC/DC Converter for Mini All-Terrain Vehicle Application

Syifaful Fuada<sup>\*1</sup>, Braham Lawas Lawu<sup>2</sup>, Bommegowda Kabbala Basavarajappa<sup>3</sup>

<sup>1</sup>Program Studi Sistem Telekomunikasi, Universitas Pendidikan Indonesia, Bandung, 40154, Indonesia

<sup>2</sup>PT. Perusahaan Listrik Negara (PLN) UPP Kitring, Muara Taweh, Barito Utara, Kalimantan Tengah, 73811, Indonesia

<sup>3</sup>Department of Electronics & Communication Engineering, N.M.A.M. Institute of Technology, Nitte, Udipi District, Karnataka, 574110, India

### ARTICLE INFO

Article history:

Received: 12 June, 2020

Accepted: 09 July, 2020

Online: 28 July, 2020

Keywords:

Non-isolated Bidirectional  
DC/DC Converter

Boost Mode

Supercapacitor

Mini All-terrain Vehicle (ATV)

### ABSTRACT

Fundamentally, the Bidirectional DC/DC converter consists of Buck and Boost modes, which working alternately. Hence, it has two output directions: Buck mode for decreasing the voltage and Boost mode for increasing the voltage at certain levels. In this work, we applied a non-isolated topology of the Bidirectional DC/DC converter for electric vehicle, that is mini all-terrain vehicle (ATV). We set a Buck mode to charge the Supercapacitor when the battery current and the Supercapacitor voltage are lower than considered level. Whereas the Boost mode was used to discharge as well as buffer the mini ATV when the battery current and the Supercapacitor voltage are higher than considered level. The discussion of Buck mode has been presented in previous work, so in this work, we focus on the Boost mode analysis only. This mode is set to increase the Supercapacitor's voltage. The Supercapacitor with  $25 V_{DC}/8$  Farad was used as the secondary main power inside the  $22.2 V_{DC}/5000$  mAh LiPo battery of the mini ATV motor. The mini ATV requires  $36 V_{DC}$  to work. Thus, it must be boosted first from  $22.2$  to  $36 V_{DC}$  using an external Boost converter. Moreover, it must be maintained at  $36 V_{DC}$ . Based on the requirement, we first design the bidirectional DC/DC converter involving the mathematical calculation and then simulate it into LTSpice®. The Printed-Circuit Board is then lay-outed and mounted. Later, we connected the designed system to mini ATV motor and tested the performance as well. According to the laboratory test, the Bidirectional DC/DC converter can increase (Boost) the voltage of the Supercapacitor from a certain level to  $36 V_{DC}$ . On the other hand, it can maintain  $36 V_{DC}$ . The central control in this system uses the STM32F4 Microcontroller, while the battery monitoring system employs the STMStudio.

## 1. Introduction

Battery Management System (BMS) is one of many ways or strategies to maintain battery in excellent performance and health conditions, which is widely implemented today, especially in the electrical vehicles, e.g., reported by [1]-[5]. Most of electric vehicles use DC source (rechargeable battery) as the main power. BMS contains several issues, e.g., 1) remaining run-time information; 2) battery-capacity monitoring; 3) charge control; and 4) charge-cycle counting [6]. BMS allows the battery of electric vehicles to have a longer life time, and it cannot be

damaged easily because the battery indicator is managed well and always monitored. For this reason, some research uses a Supercapacitor as a secondary power besides the use of the primary battery; it has very fast in charging-discharging times. The Supercapacitor has a high density so it can be utilized to prevent the primary battery from the sudden voltages or currents in a short time due to energy fluctuation. Prior works related to battery/Supercapacitor for energy storage system have been presented by [7]-[10].

This work will focus on the battery-capacity monitoring issue on the BMS. There are various BMS types; one of them is the Bidirectional DC/DC converter, which is a circuit formed by Buck

\*Corresponding Author: Syifaful Fuada, Program Studi Sistem Telekomunikasi, Universitas Pendidikan Indonesia, Bandung, [Syifafulfuada@upi.edu](mailto:Syifafulfuada@upi.edu)

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050430>

and Boost converters. This type has been reported in [11]-[14]. In this work, Buck-Boost mode is chosen due to easy operation; we can control the voltage/current output by Pulse Width Modulation (PWM). Moreover, Buck-Boost mode has a low-cost implementation and it has a high-efficiency. Based on the switching technique, DC/DC converter is divided into two parts, i.e., non-isolated and isolated topologies [15]. Non-isolated serves many advantages compared to isolated-topology one, such as simple circuit structure, and it contains little electronic components. Hence, the circuit dimension/size can be reduced. Furthermore, the copper losses in the Transformers component and heat waste caused by the power MOSFET' switching can be minimized.

In this work, a non-isolated topology is selected. The illustration of the converter power flow is illustrated in Figure 1(a) while Figure 1(b) shows the basic non-isolated topology. As we can see, the Bidirectional DC/DC converter have two ways (directions) of the current flow. If the  $V_A$  node is the input voltage (Buck mode), then the " $V_B$ " node acts as the output voltage, whereas if the " $V_B$ " node is input voltage (Boost mode), then the " $V_A$ " node acts as the input voltage.

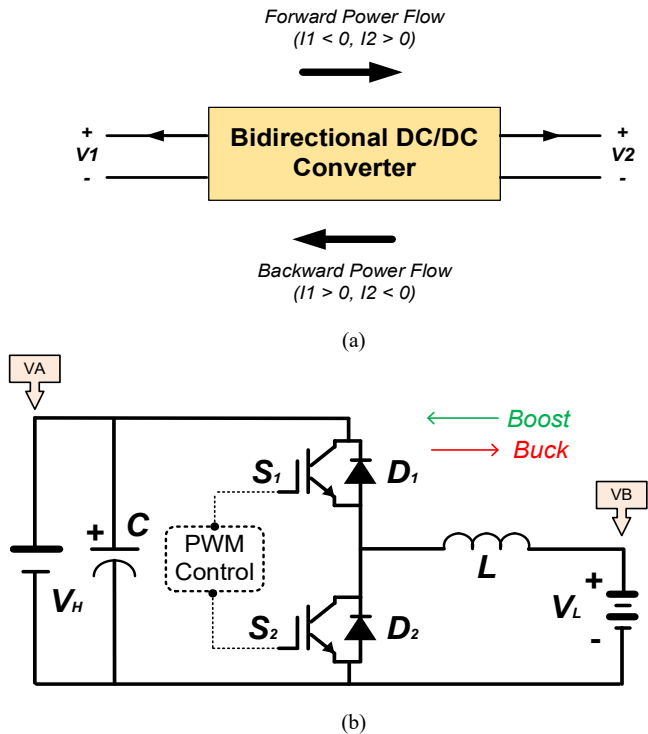


Figure 1: (a) Power flow illustration of the Bidirectional DC/DC converter, reproduced from [16] (b) Basic topology of Bidirectional DC/DC converter to charge and discharge exchange on Supercapacitor, reproduced from [16]

According to Figure 1(a), the basic circuit has 4 modes, which is [16]:

- Mode I: The DC/DC converter acts in buck mode when the high voltage ( $V_H$ ) is greater than the reference value  $V_{ref}$ . In this mode, the DC/DC converter controls the current to charge the Supercapacitor.
- Mode II: The DC/DC converter acts in boost mode when  $V_H$  falls below the reference value. In this mode, the Supercapacitor discharges.

- Mode III: When Supercapacitor is fully charged, the DC/DC converter shuts down to avoid damaging the Supercapacitor.
- Mode IV: When the Supercapacitor is fully discharged, the conditioner shutdown until the supply produces sufficient current to resume charging of the Supercapacitor.

Mode I and Mode III are used to configure the Bidirectional DC/DC converter in a Buck mode (charging mechanism). This mode has been used in previous work to decrease the voltage of the external Boost converter module from 36  $V_{DC}$  to 18  $V_{DC}$  and charge the Supercapacitor [17]. Whereas to discharge and buffer the voltage and current for an electrical All-Terrain Vehicle (ATV) motor, we will use Mode II and Mode IV. The difference between our work to [7-10] is the use of Microcontroller as the primary control. The STM32F4 Microcontroller was used due to serve a user interface feature, i.e., STMStudio. Therefore, we do not need to design a graphical user interface (GUI) from scratch to display the charge-discharge Supercapacitor mechanism. We just employ this software to save effort in designing and implementing the Bidirectional DC/DC converter. As stated by T. Adiono, et al., [18] and A.T. Agung, et al., [19] that an open-source software enables us to cut the required time during the design and implement a specified system.

The discussion of this paper is divided into four sections, which are: (1) Introduction which presents the research background and purpose; (2) Material and Methods which discusses the Bidirectional DC/DC converter design and its simulation employing LTspice® software; (3) Results and Analysis which shows the experimental data of Supercapacitor discharging mechanism in the real test; the last one is (4) Conclusion and References.

## 2. Materials and Methods

### 2.1. Boost Converter: Overview

The boost converter is one of the DC/DC Converter types functioned to increase certain voltage levels. Commonly, the Boost converter circuit has a topology, as illustrated in Figure 2 (a) and the working principle, as in Figure 2 (b). Based on Figure 2(a), we can see that Boost converter circuit contains several electronics components: a switch represented by  $S$ , a Diode represented by  $D$ , an Inductor represented by  $L$ , a Capacitor represented by  $C$ , a load represented by  $R$ , and  $V_s$  as a source voltage or input voltage.

Similar to the Buck converter circuit, the Boost converter circuit also has two modes: Continuous Conduction Mode (CCM) and Discontinuous Conduction Mode (DCM) with respect to the current flowing at the Inductor ( $i_L$ )

The difference between these modes lies in the value of the components inside and also the current passing through the Inductor. In CCM mode, the current flow through the inductor will never be zero. While in DCM mode, the current flow through the inductor will have a value when the current flow is zero. The relationship between the output and input voltage in the boost converter can be written as an Eq. (1),

$$V_{Out} = \frac{V_S}{1-D} \quad (1)$$



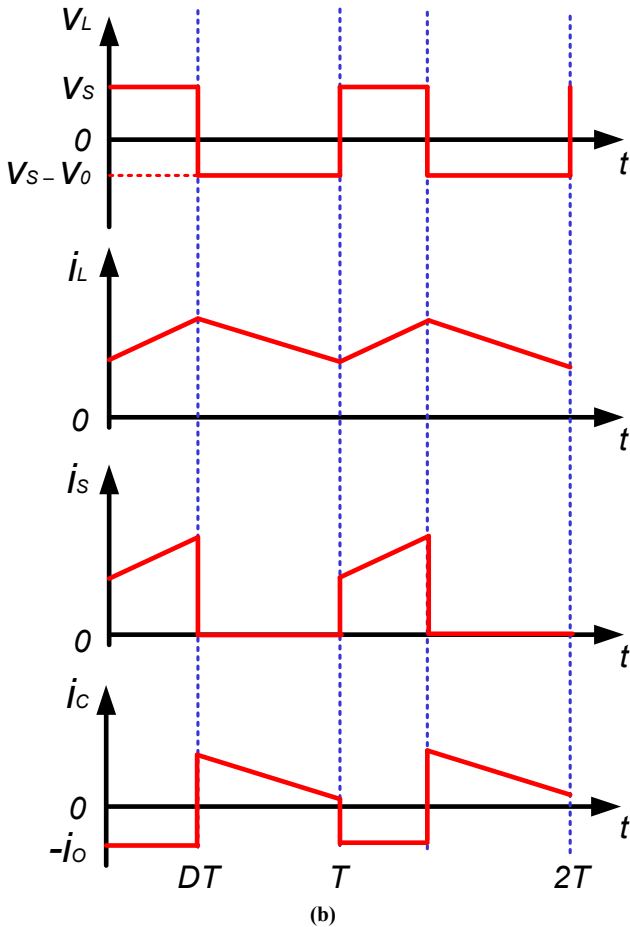
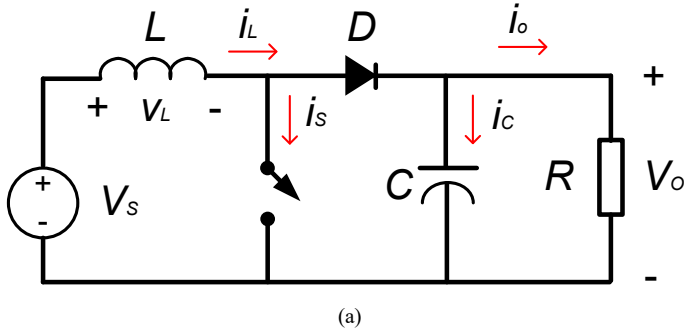


Figure 2: (a) Boost converter circuit configuration and (b) Boost converter waveform, obtained from [1]

Where  $V_o$  is output voltage,  $V_S$  is source voltage (input voltage), and  $D$  is the used Duty cycle in the Bidirectional DC/DC converter. In a Boost converter mode, the CCM and DCM limits are expressed by Eq (2),

$$L_b = \frac{(1-D)^2 DR}{2f} \quad (2)$$

Where  $f$  can be defined as  $F_{sw}$  and  $R$  is a dummy load. If  $L > L_b$ , then Boost converter will be on CCM mode. The use of capacitor as filters in this mode must be set to have a minimum value. Accordingly, it does not have a large DC ripple voltage ( $V_r$ ). To limit it, the Capacitor value used for Boost mode must be greater than Eq. (3),

$$C_{min} = \frac{DV_o}{V_r R f} \quad (3)$$

From Eq. (2) and Eq. (3) the Inductor and Capacitor values can be found, which is then applied in the circuit. Besides, we can also determine the operated frequency value in the circuit.

### 2.2. Bidirectional DC/DC Converter in Boost Mode: Design

We have a Supercapacitor designed by ourselves as the secondary power of the mini ATV motor rated in 25 V<sub>DC</sub> 8F. This Supercapacitor contains several commercially available Electrolytic Capacitor/Condensator (Elco) connected in series/parallel. We charged the Supercapacitor at least under 22 V<sub>DC</sub> to keep it healthy, precisely is 18 V<sub>DC</sub>. To buffer the DC motor of mini ATV, we need to increase the minimum voltage of Supercapacitor, that is 18 V<sub>DC</sub> to DC motor working voltage, that is 36 V<sub>DC</sub>.

Figure 3 shows the complete circuit of the Bidirectional DC/DC converter that has dual-mode in a single circuit: Boost and Buck.

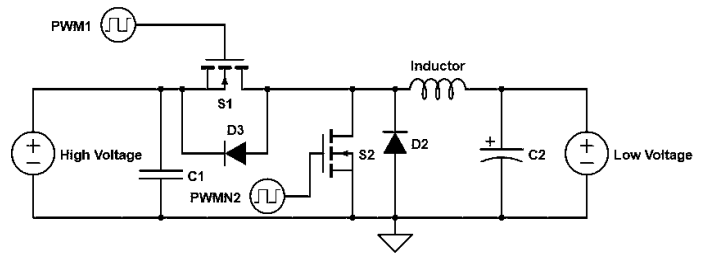


Figure 3: Basic structure of the Bidirectional DC/DC converter with Buck and Boost built-in

Afterward, we identify the desired system so that we can calculate the value of the electronic components needed in the circuit. The following are system specifications based on the specified identification.

- The bidirectional DC/DC converter consists of two P75NF75 power MOSFETs as the switching part, STM32F4 Discovery as the microcontroller unit, ACS712 as the current sensor, inductor with PQ type, and TLP250 as the gate driver.
- Input voltage ( $V_{in}$ ) = 18 V<sub>DC</sub>
- Desired output voltage ( $V_{out}$ ) = 36 V<sub>DC</sub>
- Desired discharging current in the Supercapacitor ( $I_{out}$ ) is about 4 A
- Switching frequency ( $F_{sw}$ ) used is 10 kHz based on the gate driver specification (IC TLP250)

As a reference calculation, we use the formula from [20]. From the above parameters, the Duty cycle value can be found by Eq. (1), which  $V_S$  is noted as  $V_{in}$

$$D = 1 - \frac{V_{in}}{V_{out}} = 1 - \frac{18}{36} = 0.5$$

where the determination of the minimum inductor value is referred to Eq. (2), so  $L_b$  would be

$$L_b = \frac{(1-D)^2 D * R}{2f} = 56.25 \mu H$$

While to determine the minimum value of Capacitor, Eq. (3) was used. Furthermore,  $C_{min}$  would be

$$C_{min} = \frac{DV_O}{V_r R_f} = 555.55 \mu F$$

Later, we calculate the values of inductor and capacitor needed for the circuit. The formula for the inductor is given by Eq. (4),

$$V_L = L \frac{di}{dt} \tag{4}$$

Where  $L$  is the value of the inductor,  $V_L$  is the voltage at the inductor,  $dt$  is the switching period,  $di$  is the ripple current.  $D$  is the selected Duty cycle. According to the specification, our experiment used  $F_{sw} = 10$  kHz and 50% of Duty cycle, then, the calculation for the inductor would be,

$$36 - 18 = L \frac{1.2}{0.5 * 10kHz}$$

$$L = 750 \mu H$$

To seek the Capacitor value, we used Eq. (5),

$$i_c = C \frac{dv}{dt} \tag{5}$$

where  $C$  is the used Capacitor value,  $i_c$  is the Capacitor current. By setting  $F_{sw} = 10$  kHz and  $D = 50\%$ , then the calculation for the capacitor would be,

$$C = \frac{1.2A * 0.5 * 10kHz}{0.18V} = 333.33 \mu F$$

But in practical design, we used 2.5 mH and 100  $\mu F$  of inductor and capacitor values, respectively.

We used TLP250 for gate driver on the Power MOSFET that has a maximum operating frequency at 25 kHz. As stated in the system requirement, we only used 10 kHz as the PWM switching frequency. The gate driver configuration is depicted as Figure 4,

The PWM output from the microcontroller will trigger the Gate of the MOSFETS. Our Bidirectional DC/DC converter circuit is needed a complementary PWM to work. To do this case, we configure the microcontroller in complementary PWM in which the STM32F4 microcontroller has a feature for it.

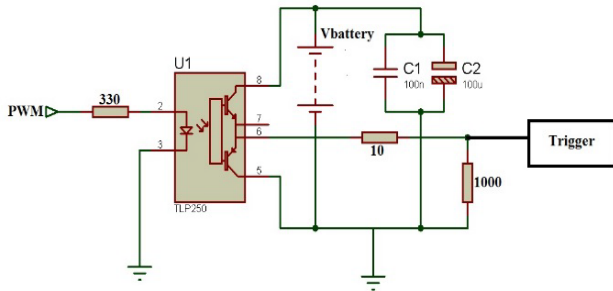


Figure 4: The Gate driver schematic of Bidirectional DC/DC converter in Boost mode

The Dead time also can be involved to prevent both of the two MOSFETS shorted at the same time. The dead time calculation can be written as expressed by [17]

$$t_{dead} = |(t_{D\_OFF\_MAX} - t_{D\_ON\_MIN}) + (t_{PDD\_MAX} - t_{PDD\_MIN})| * 1.2 \tag{6}$$

### 2.3. Bidirectional DC/DC Converter in Boost Mode: Simulation

We used LTSpice® software to simulate the Bidirectional DC/DC converter in boost mode as shown in Figure 5. The circuit in a simulation has two outputs, which are voltage and current outputs. We generated the PWM signal as an input signal to the MOSFET's gate by setting the voltage source as the PWM. In line with the system identification, Diode MUR1560 and MOSFET type P75NF75 were chosen. To obtain valid results in ideal conditions, the components model and its parameters inserted in LTSpice® that are precisely the same as the MUR1560 and P75NF75 datasheets. Other parameters set for the simulation is shown in Table 1

Table 1. Simulation Parameters

No	Parameters	Value
1	$V_{in}$	18 V <sub>DC</sub>
2	$C_{in}$	100 $\mu F$
3	$C_{out}$	100 $\mu F$
4	$R_{out}$	10 $\Omega$
5	$L$	2.5 mH
6	$F_{sw}$	10 kHz
7	Duty cycle	50% (0.5)
8	Supercapacitor	8 F

After the circuit has been constructed in LTSpice®, we convert it into a block as visualized in Figure 5, where X1 is our Bidirectional DC/DC converter, V1 is voltage input with 18 V<sub>DC</sub>, and R<sub>1</sub> is a dummy load with 10  $\Omega$ . Based on the real measurement using digital Multimeter, the DC motor of the mini ATV has total resistance about 10  $\Omega$ . Therefore, we set it as a load to represent DC motor in our simulation. The simulation result is shown in Figure 6, we can see that the output voltage is 35.026818 V<sub>DC</sub> when the Duty cycle is set 50%. The average of output current flows in the circuit is 3.2329 A and 3.3171 A of RMS current. The output voltage will swing in the initial state from 0V to the maximum of 40 V<sub>DC</sub> until it reached the steady state of  $\sim 36$  V<sub>DC</sub>, that is 35.026818 V<sub>DC</sub> while the output current will swing in the initial state from 0 A to the maximum of 4 A. In this Boost mode, the output voltage will buffer mini ATV. For data comparison, the Duty cycles in LTSpice® simulation are varied from 50% to 30%, 40% and 60% and we obtained 23.367980 V<sub>DC</sub>, 27.424739 V<sub>DC</sub>, and 38.787089 V<sub>DC</sub>, respectively.

### 3. Results and Analysis

After the simulation has been done perfectly, we tested the Supercapacitor performance whether it can increase the voltage from 18 to 36 V<sub>DC</sub> or not. As specified in system identification, the output voltage should stay at approximately 36 V<sub>DC</sub> even the input voltage from the Supercapacitor is decreasing. To meet the qualification, we set the experiment parameters as Table 2.

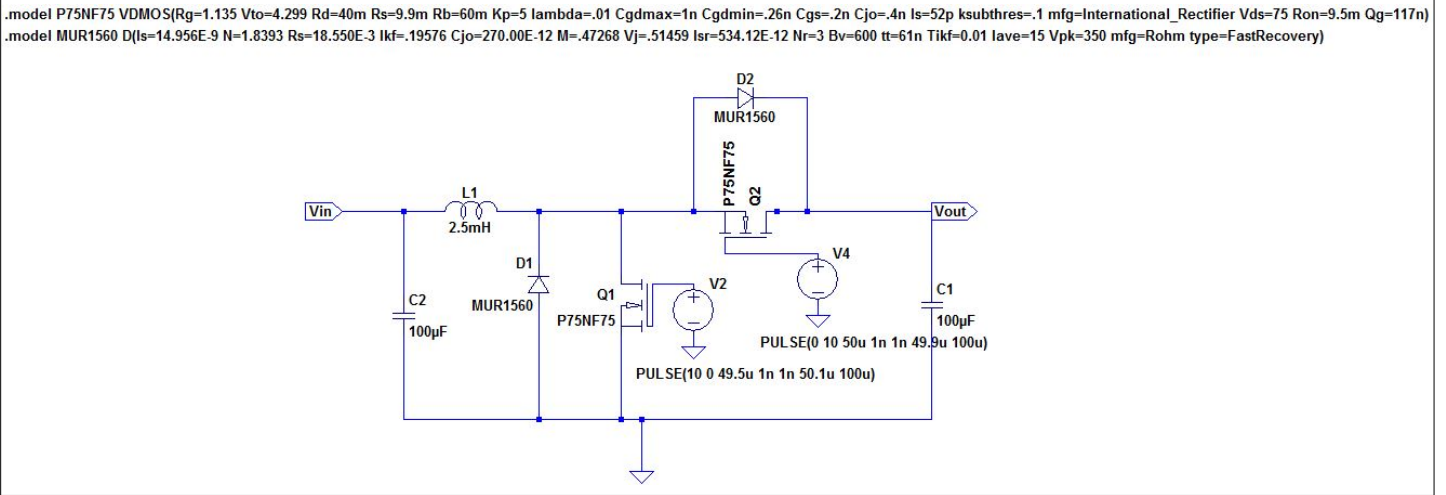


Figure 5: Bidirectional DC/DC converter circuit under the LTSpice® simulation, 18  $\rightarrow$  36 V<sub>DC</sub> for discharger and buffer operation, then 36 V<sub>DC</sub> (DC motor)  $\rightarrow$  18 V<sub>DC</sub> (Supercapacitor) for charger operation. This simulation will focus on discharger/buffer operation from 18 V<sub>DC</sub> (Supercapacitor) to 36 V<sub>DC</sub> (DC motor)

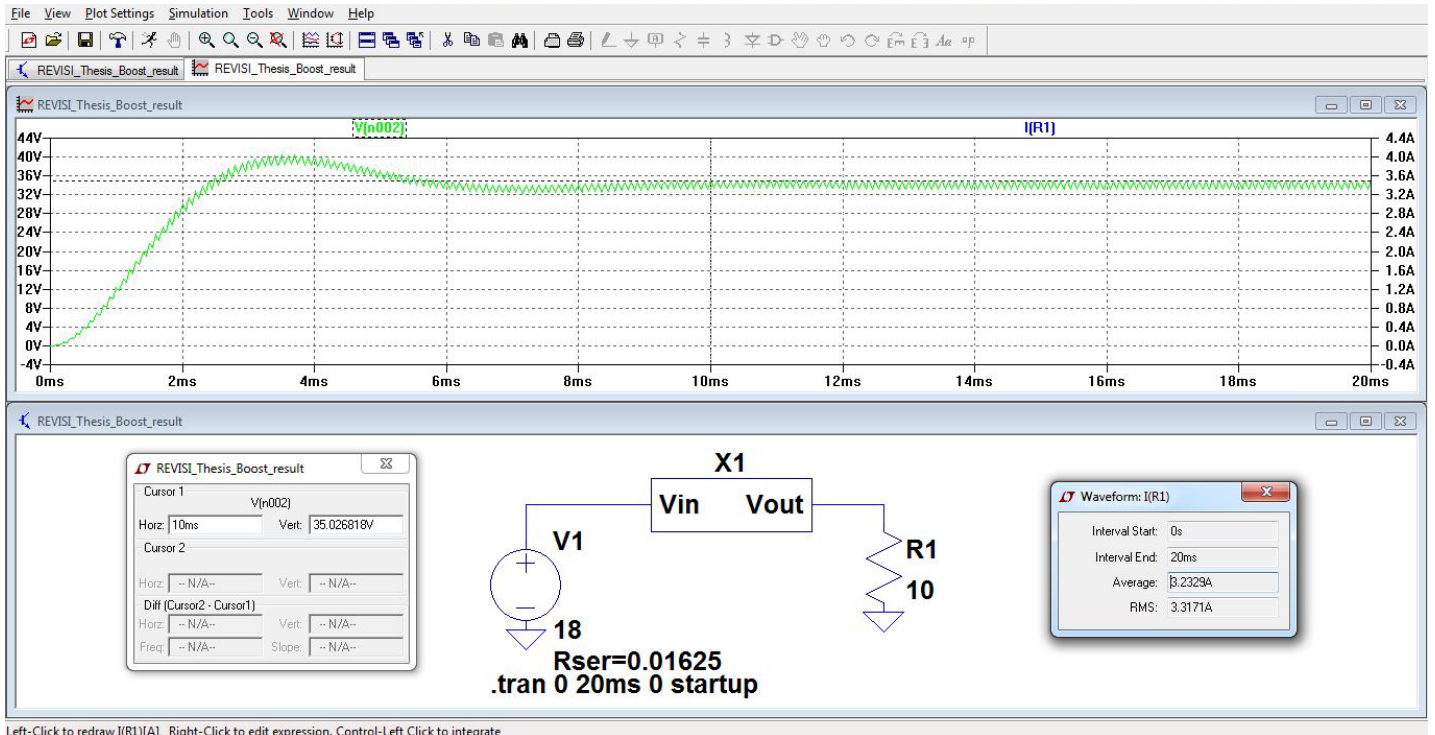


Figure 6: Simulation result of the Bidirectional DC/DC converter on Boost mode with Duty cycle = 50%

A Photograph of experimental setting to verify the circuit design and the printed-circuit board (PCB) of the designed bidirectional DC/DC converter are visualized in Figure 7(a) and Figure 7(b), respectively.

We set the voltage limit at the Supercapacitor to be 10 V<sub>DC</sub> minimum and 18 V<sub>DC</sub> maximum (fully-charged). Thus, when the DC/DC converter in a boost mode, it will be no longer in this mode if the Supercapacitor voltage reaches 10 V<sub>DC</sub> from 18 V<sub>DC</sub>.

The test scenario is the same as previous work [17]. It is shown in Figure 8 (reproduced from [17] with permission). An external Boost converter module was used to increase a 22.2 V<sub>DC</sub>/5000 mAh LiPo battery (series configuration of two 11,1 V<sub>DC</sub> LiPo

batteries). Afterward, this voltage output is connected to the mini ATV that requires 36 V<sub>DC</sub> of voltage to work.

Table 2. Experiment Parameters

No	Parameters	Value
1	$V_{in}$	18 V <sub>DC</sub> (Supercapacitor voltage when it is fully charged)
2	$L$	2.5 mH
3	$F_{sw}$	10 kHz
4	Supply at the Gate driver	single supply (11,1 V <sub>DC</sub> )

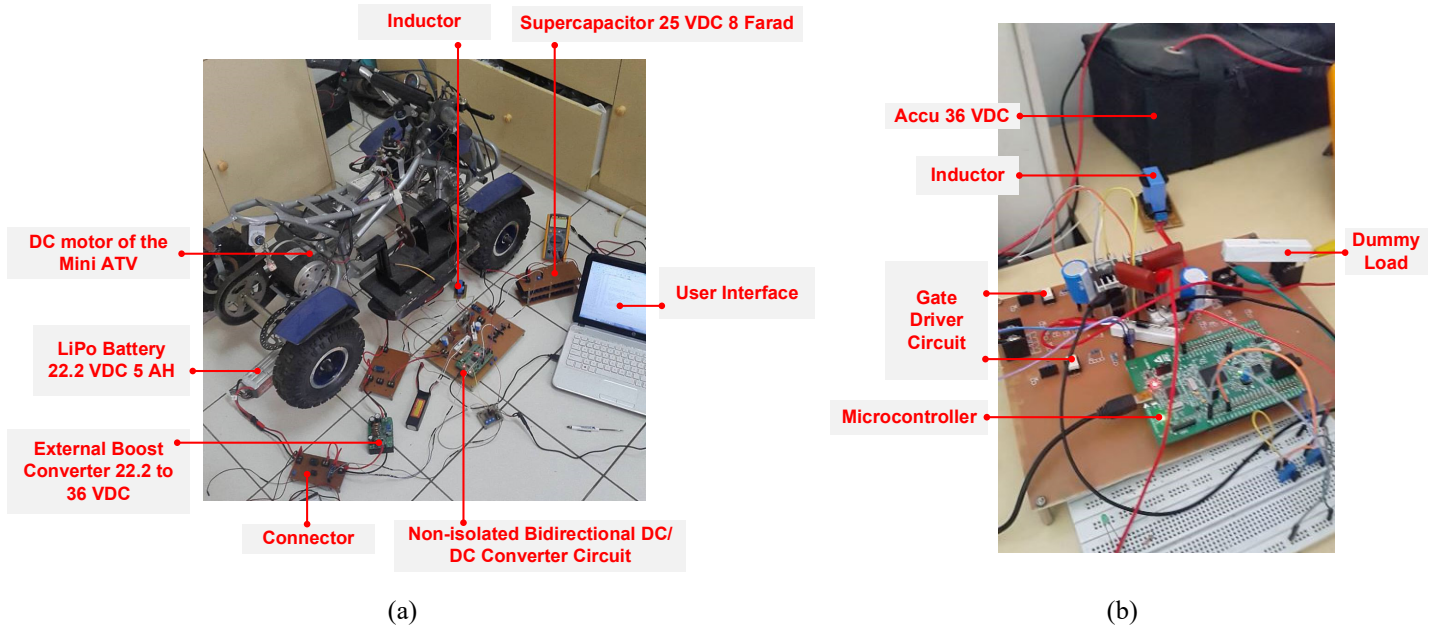


Figure 7: (a) Experimental setup containing mini ATV motor, Battery, Supercapacitor 25 V<sub>DC</sub>/8F, electronic circuit experiments, and Laptop as user interface; (b) fabricated the Bidirectional DC/DC converter circuit in single-sided PCB

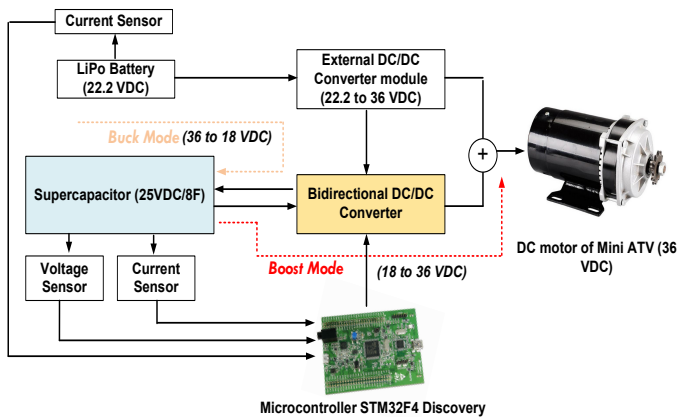


Figure 8: A Scenario for testing the performance of Bidirectional DC/DC converter in Boost mode

To monitor the Supercapacitor condition, we need a user interface. We employed the STMStudio to display the discharging voltage of the Supercapacitor. The experimental results are shown in Figure 9. The vertical axis represents Supercapacitor voltage while the horizontal axis represents Supercapacitor's discharging time. The pink line represents the Supercapacitor voltage while the orange line represents the output voltage of the Bidirectional DC/DC Converter in a boost mode. To easily understand the mentioned lines when this paper is printed in a grayscale format, we marked it as in Figure 9.

From the window, it can be seen that the output voltage stays at a certain level even the voltage at the Supercapacitor is decreasing from 17.5 V<sub>DC</sub> down to 10 V<sub>DC</sub> as pointed out by the GUI. However, the displayed values on the STMStudio is not accurate enough. Therefore, it can be used only to ensure that the output voltage is almost stable (information based on the visual observation). Digital multimeter was used to measure a real value; we got a stable voltage, i.e., ~36 V<sub>DC</sub> when 50% of the Duty Cycle

is set. This voltage level can be reached by controlling the PWM (Duty cycle) properly.

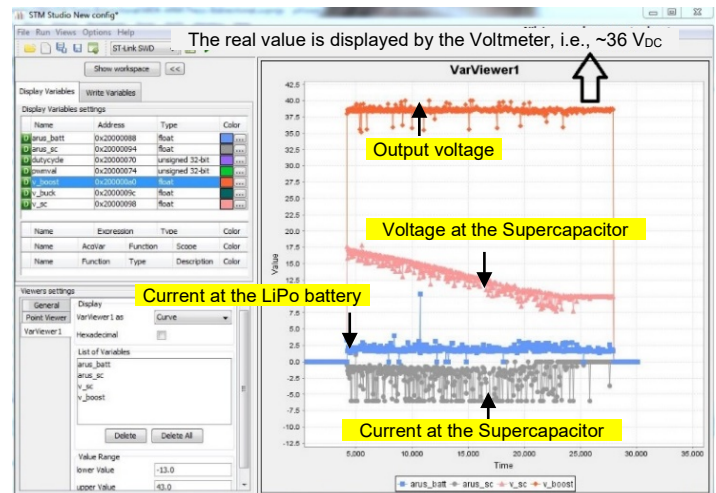


Figure 9: Plot of the Supercapacitor's voltage charging using STMStudio in Boost mode

Whereas to measure the Supercapacitors current accurately, we used the ACS712 current sensor reading in an Analog to Digital Converter (ADC). As a result, the current can be buffered to the mini ATV by the Supercapacitor varied from 0 to 6 A. In this experiment, our circuit can work well as expected: the input of the converter can be output, and vice versa, the output can be the input.

Afterward, the simulation LiPo results as presented in Section 2.3 are compared with the laboratory measurements. We adjust PWM variations according to the simulation settings, which are 30%, 40%, and 60%, and we obtained 25.1 V<sub>DC</sub>, 29.2 V<sub>DC</sub>, and 36.5 V<sub>DC</sub>, respectively. Table 3 is a comparison between simulation and real experiment of the Bidirectional DC/DC converter in a Boost mode.



To make it easier to see the difference level, a graph is served as shown in Figure 10, it can be seen that there is a difference between using simulations and implementing hardware. This can occur due to several factors like tolerance factor of the used components that make it non-ideal.

Table 3: Comparison between voltage output from a simulation and implementation (real condition) on the Bidirectional DC/DC converter circuit with Boost mode

No	Duty Cycle	Output voltage in Simulation	Output voltage in Implementation
1	30 %	23.367980 V <sub>DC</sub>	25.1 V <sub>DC</sub>
2	40 %	27.424739 V <sub>DC</sub>	29.2 V <sub>DC</sub>
3	50 %	35.026818 V <sub>DC</sub>	32.7 V <sub>DC</sub>
4	60 %	38.787089 V <sub>DC</sub>	36.5 V <sub>DC</sub>

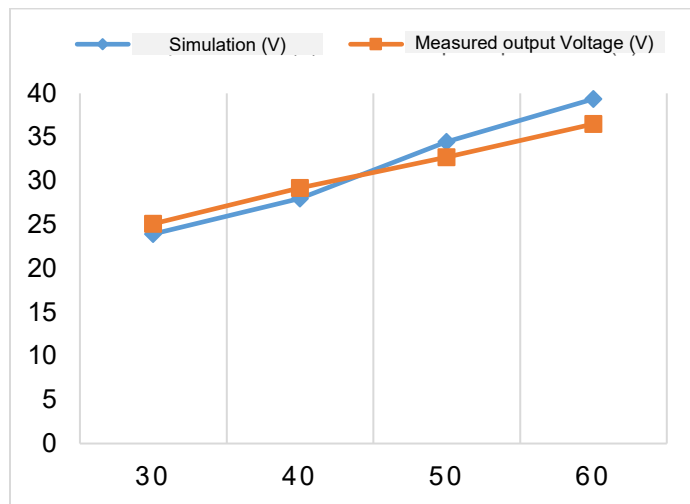


Figure 10: A Graph of simulation vs. measured output on the Bidirectional DC/DC converter circuit with Boost mode. Horizontal axis: Duty Cycle (%) and Vertical axis: Output voltage (V)

#### 4. Conclusion

Supercapacitor plays an essential role in the BMS of modern electric vehicles. By using Supercapacitor as a supplementary battery, the need for high current from the main battery can be buffered and maintained at the considered level. Thus, the main battery will have a long-time span. In this work, non-isolated bidirectional converter is chosen due to the use of fewer components than the isolated bidirectional converter. So, in terms of PCB size, it will be relatively smaller than the isolated Bidirectional DC/DC converter. The effect of the transformer's usage on isolated topology makes power losses due to overheating production. While for non-isolated, careful calculation needs to be done in determining the inductor as well as capacitor values. Thus, the circuit can operate properly, both for Boost or Buck modes. The use of the inductor on the non-isolated topology can reduce power losses.

In this work, we realize a BMS using a non-isolated bidirectional DC/DC converter controlled by the PWM. Our circuit is used as a current and voltage control system that operates in Buck and Boost. By changing the PWM Duty cycle in the Microcontroller

STM32F4, we can manually adjust the output voltage of the bidirectional DC/DC converter.

Based on the simulation result through LTSpice®, the circuit can work well as expected; the output voltage is ~36 V<sub>DC</sub> with 50% of Duty cycle. This voltage will be used to buffer/discharge the DC motor of mini ATV. According to the functional test, it can be seen that the supercapacitor can perform a discharging function. Hence, it can be used as a secondary power distributor to buffer voltage and current, which is then successfully implemented on the mini ATV. Discharging time of the Supercapacitor takes approximately 30 – 50 seconds, depending on the battery capacity with a current of up to 6A. This high current can be pulled from a 25 V<sub>DC</sub>/8F Supercapacitor to the mini ATV. The detail of this paper including the improved Buck mode using Resistor-Capacitor-Diode (RCD) Snubber circuit and Hybrid mode (Buck & Boost).

In future work, we will more emphasized the effects of different construction features upon the systems functioning.

#### Conflict of Interest

The authors declare no conflict of interest.

#### Acknowledgment

We would like to thanks to Dr. Arif Sasongko and Dr. Farkhad Ihsan Hariadi from School of Electrical Engineering and Informatics Insitut Teknologi Bandung (ITB), Indonesia, for their supervision.

#### References

- [1] I. Carlucho, R. de la Vega, M. Spina, and G. G. Acosta, 'A Modular Battery Management System for Electric Vehicles', in *2018 IEEE Biennial Congress of Argentina (ARGENCON)*, San Miguel de Tucumán, Argentina, 1–6, 2018. doi: 10.1109/ARGENCON.2018.8646227
- [2] A. P. Talie, W. A. Pribyl, and G. Hofer, 'Electric Vehicle Battery Management System Using Power Line Communication Technique', in *2018 14th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME)*, Prague, 225–228, 2018. doi: 10.1109/PRIME.2018.8430304.
- [3] L. K. Amifia, S. A. Widayat, A. I. Cahyadi, and O. Wahyunggoro, 'Fault detection design and simulation based on battery modelling', presented at the PROCEEDINGS OF THE 12TH INTERNATIONAL CONFERENCE ON SYNCHROTRON RADIATION INSTRUMENTATION – SRI2015, New York, NY USA, p. 090008, 2016. doi: 10.1063/1.4958526
- [4] D. A. Martinez, J. D. Poveda, and D. Montenegro, 'Li-Ion battery management system based in fuzzy logic for improving electric vehicle autonomy', in *2017 IEEE Workshop on Power Electronics and Power Quality Applications (PEPQA)*, Bogota, Colombia, 1–6, 2017. doi: 10.1109/PEPQA.2017.7981677
- [5] R. Xiong, Y. Zhang, J. Wang, H. He, S. Peng, and M. Pecht, 'Lithium-Ion Battery Health Prognosis Based on a Real Battery Management System Used in Electric Vehicles', *IEEE Trans. Veh. Technol.*, **68**(5), 4110–4121, May 2019. doi: 10.1109/TVT.2018.2864688
- [6] L. Buccolini, A. Ricci, C. Scavongelli, G. DeMaso-Gentile, S. Orcioni, and M. Conti, 'Battery Management System (BMS) simulation environment for electric vehicles', in *2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)*, Florence, Italy, 1–6, 2016. doi: 10.1109/EEEIC.2016.7555475
- [7] L. Kouchachvili, W. Yaïci, and E. Entchev, 'Hybrid battery/supercapacitor energy storage system for the electric vehicles', *J. Power Sources*, **374**, 237–248, 2018. doi: 10.1016/j.jpowsour.2017.11.040
- [8] A. Tahri et al., 'Management of fuel cell power and supercapacitor state-of-charge for electric vehicles', *Electr. Power Syst. Res.*, **160**, 89–98, Jul. 2018, doi: 10.1016/j.epr.2018.02.003.
- [9] B.-H. Nguyen, R. German, J. P. F. Trovao, and A. Bouscayrol, 'Real-Time Energy Management of Battery/Supercapacitor Electric Vehicles Based on



- an Adaptation of Pontryagin's Minimum Principle', *IEEE Trans. Veh. Technol.*, **68**(1), 203–212, Jan. 2019, doi: 10.1109/TVT.2018.2881057.
- [10] Z. Song, J. Hou, H. Hofmann, J. Li, and M. Ouyang, 'Sliding-mode and Lyapunov function-based control for battery/supercapacitor hybrid energy storage system used in electric vehicles', *Energy*, **122**, 601–612, 2017. doi: 10.1016/j.energy.2017.01.098.
- [11] R. H. Ashique and Z. Salam, 'A High-Gain, High-Efficiency Nonisolated Bidirectional DC–DC Converter With Sustained ZVS Operation', *IEEE Trans. Ind. Electron.*, **65**(10), 7829–7840, 2018. doi: 10.1109/TIE.2018.2802457.
- [12] H. Bahrami, S. Farhangi, H. Iman-Eini, and E. Adib, 'A New Interleaved Coupled-Inductor Nonisolated Soft-Switching Bidirectional DC–DC Converter With High Voltage Gain Ratio', *IEEE Trans. Ind. Electron.*, **65**(7), 5529–5538, 2018. doi: 10.1109/TIE.2017.2782221
- [13] Y. Zhang, X.-F. Cheng, C. Yin, and S. Cheng, 'A Soft-Switching Bidirectional DC–DC Converter for the Battery Super-Capacitor Hybrid Energy Storage System', *IEEE Trans. Ind. Electron.*, **65**(10), 7856–7865, 2018. doi: 10.1109/TIE.2018.2798608
- [14] M. Veerachary and O. K. Singh, 'Bidirectional non-isolated dc-dc converter for low-voltage DC-grid', in *2016 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, Trivandrum, India, 1–6, 2016. doi: 10.1109/PEDES.2016.7914408.
- [15] H. Ardi, A. Ajami, F. Kardan, and S. Nikpour, 'Analysis and Implementation of a Non-Isolated Bidirectional DC–DC Converter with High Voltage Gain', *IEEE Trans. Ind. Electron.*, **1**(1), 2016. doi: 10.1109/TIE.2016.2552139.
- [16] J. Zhang, 'Bidirectional DC-DC Power Converter Design Optimization, Modeling and Control', *Diss. Va. Polytech. Inst. State Univ.*, 3–5, 2008.
- [17] B. L. Lawu, S. Fuada, S. Ramadhan, A. F. Sabana, and A. Sasongko, 'Charging supercapacitor mechanism based-on bidirectional DC-DC converter for electric ATV motor application', in *2017 International Symposium on Electronics and Smart Devices (ISESD)*, Yogyakarta, 129–132, 2017. doi: 10.1109/ISESD.2017.8253318.
- [18] T. Adiono, S. F. Anindya, S. Fuada, K. Afifah, and I. G. Purwanda, 'Efficient Android Software Development Using MIT App Inventor 2 for Bluetooth-Based Smart Home', *Wirel. Pers. Commun.*, **105**(1), 233–256, 2019. doi: 10.1007/s11277-018-6110-x.
- [19] A. T. Agung, S. Fuada, and T. Adiono, 'IMPLEMENTASI ASSET MANAGEMENT DENGAN SNIPE-IT DI PUSAT MIKROELEKTRONIKA INSTITUT TEKNOLOGI BANDUNG', *Simetris J. Tek. Mesin Elektro Dan Ilmu Komput.*, **10**(1), 243–258, 2019. doi: 10.24176/simet.v10i1.2961.
- [20] M. H. Rashid, Ed., *Power Electronics Handbook*. Amsterdam, The Netherlands: Elsevier, 2007.

## Utilization of Data Mining to Predict Non-Performing Loan

Yosaphat Catur Widiyono<sup>\*1</sup>, Sani Muhamad Isa<sup>2</sup>

<sup>1</sup>Computer Science Department, Binus Graduate Program - Master of Computer Science, Bina Nusantara University, Jakarta 11480, Indonesia

<sup>2</sup>Computer Science Department, Binus Online Learning, Bina Nusantara University, Jakarta 11480, Indonesia

---

### ARTICLE INFO

Article history:

Received: 18 June, 2020

Accepted: 09 July, 2020

Online: 28 July, 2020

---

Keywords:

Data Mining

Non-Performing Loan

Machine Learning

---

---

### ABSTRACT

*In the banking industry, the existence of problem loans is inevitable. NPL (Non-Performing Loan) will certainly have an impact on the reduction in the capital of a bank. One good step in reducing the risk of credit default or the emergence of non-performing loans is to take proper care of debtors who begin to experience payment constraints. The main obstacle experienced in bank management, especially in the credit sector, is being unable to identify or detect potential debtors early due to a large amount of data and manual processing. In this study, the debtor payment history is presented as data to predict the existence of problem loans. History payment can be used to predict bad loans. The technic of data mining in this experiment is a new method. The results of research conducted using Naïve Bayes, Decision Tree, K-NN, Rule Induction, Logistic Regression, Random Forest, Generalized Linear Model, and Gradient Boosted Trees as a comparison then choose the method that has the highest accuracy to be implemented in making additional modules on the core banking system. Random Forest is the model that has the highest accuracy of 96.55%.*

---

## 1. Introduction

The banks are one of the economic drivers having an important role in participating in advancing the people's economy. Bank is a financial institution that has the authority to collect funds from the public and channel it back to the community in the form of working capital loans to improve the standard of living of the general public. In its operations, banking is overseen by the Financial Services Authority (OJK). The Financial Services Authority is a state institution formed under the Act that functions to organize a system of regulation and supervision that is integrated into all activities in the financial services sector both in the banking sector, capital market, and financial services sector. Credit debtors are bank customers who receive loan funds that have been agreed through a credit agreement made between the bank and the debtor [1]. NPL or Non-Performing Loan is one of the financial ratios that reflects credit risk. NPL is defined as a loan that has problems/difficulties in repayment or is often called bad credit [2]. Credit collectibility has 5 groups, namely current, special attention, substandard, doubtful, and loss. If it is associated with credit

ollectibility, then what is included in non-performing loans are credit that has a substandard, doubtful, and bad collectibility.

The challenge faced is how can the data mining process predict the emergence of problem loans? How are the results of trials conducted using several models for comparison?

The purpose of this study is to analyze the existing credit debtor payment data at XYZ Bank so that the payment patterns can be identified. This pattern can be used to predict non-performing loans in the next 3 months. The benefit of this research is that it helps companies to predict problem loans with active debtors so that they can be applied in credit management to make it more effective and efficient to use priority scale in determining actions to handle potential debtors.

## 2. Related Works

The research relates to the factors causing the occurrence of problem loans, states that from the 24 existing variables formed 8 factors that have a contribution and influence on problem loans. The factors that have the highest contribution weight are credit period (credit term), loan amount (ceiling), and loan interest rates to be aspects that need to be considered in granting credit. A study by building a multi-dimensional and multi-level credit risk

---

\*Corresponding Author: Yosaphat Catur Widiyono, Email: [yosaphat.widiyono@binus.ac.id](mailto:yosaphat.widiyono@binus.ac.id)

indicator system aims to find the most important credit risk characteristics that will cause serious default risks. With the existence of several algorithms that can be used in research related to the C4.5 algorithm it has been done [3], which is where they compare the performance of ID3 with C4.5. And the results are better using C4.5 [4]. The application of the C4.5 algorithm in evaluating scholarship granting to students has been conducted research [5]. In that study Wang et al. have the objective to analyze the relationship between student performance and the scholarships provided and by using the C4.5 decision tree algorithm in the scholarship evaluation system it is hoped that the scholarships will be efficient and fair and can be realized. In the study, testing was also carried out to compare the use of different algorithms, namely: C4.5, ID3, Fuzzy Mathematics and Set Pair Analysis. The comparison results of the 4 models C4.5 are the best. In [6], the author build and test classification models to predict student success in English exams. As in the study conducted in [7], they did an experiment using different values for the s parameter to execute C-C4.5. The results of this study indicate that the split criteria of the C-C4.5 algorithm is stronger against noise than the C4.5 criteria. Comparison with more models including Random Forest [8]. Shamsbur and Weill conducted a study by examining the impact of bank efficiency on credit costs [9]. In this study combining company-level data with bank-level data so that it can identify the level of efficiency of banks that lend money to each company. Estimation was then carried out using a large data sample from 240,000 companies from nine countries in Europe. Their main finding is that higher bank efficiency is associated with lower credit costs. They, therefore, support the view that the effectiveness of banks in minimizing costs is transferred to the borrowing company through lower credit costs. Ye et al. conduct research on P2P lending loans [10]. The experiment resulted in contributions in the form of methods that made it possible to improve the quality of credit in P2P loans provided to borrowers. Moradi et al. proposed a dynamic model for credit risk assessment that outperforms the models currently used. Our model has a dynamic engine that assesses the behavior of bad customers on a monthly basis and a fuzzy inference system (FIS) that includes the factors of credit risk, especially in economic crises[11]. Previous research has used data: Gender, Age, Amount of Credit, Monthly Income, Monthly Expenditures, Current Payments Per Month, Savings (Income Payment), Type of Collateral, Collateral Value, Loan Period, Type of Business Activity, Sources of Funds, Credit Status previous. Data mining is used to suggest a decision tree model for credit assessment as it can indicate whether the request of lenders can be classified as performing or non-performing loans risk. Using C 5.0 methodology, a new decision tree model is generated [12].

### 3. Methodology

Experiments carried out using the CRISP-DM methodology that offers a structured approach to data mining. This research will be carried out in a 5-step process, namely business understanding, data understanding, data preparation, modeling, and evaluation. Tools or software used to process data are MS Excel 2016 and Rapid Miner 9.6.000. Several classification algorithms are used in modeling and testing. The study was conducted based on the CRISP-DM framework as shown in Figure 1.

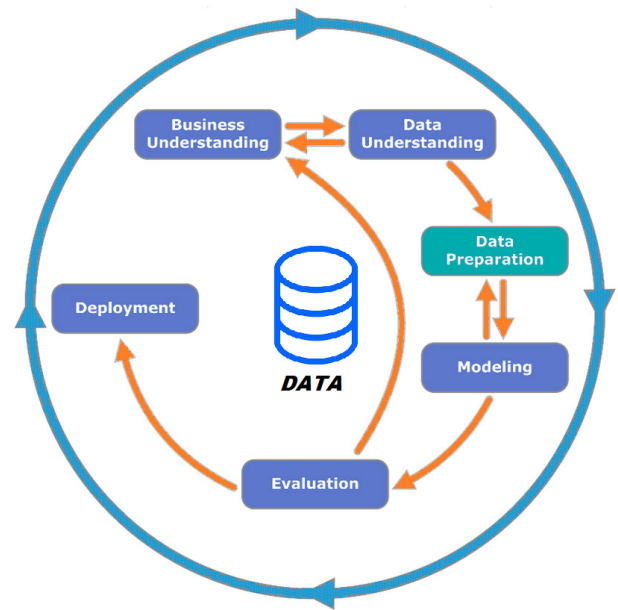


Figure 1: CRISP-DM Model

This study seeks to solve the challenge of how to predict the quality of debtors' credit based on payment history more precisely [13]. Some of the Algorithms we tested included the Decision Tree Algorithm which was chosen because it had many advantages. A trial was conducted to compare several methods in their ability to predict more accurately. The steps are taken in the collection of this data as shown in Figure 2.

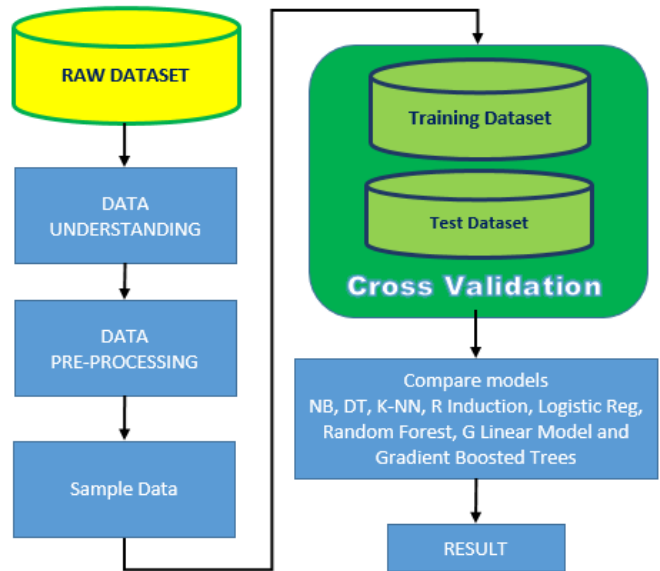


Figure 2: Experimental Design

### 4. Theory and Methods

Data mining is the activity of extracting to obtain important information that is implicit and previously unknown, from a data. The use of data mining for prediction has also been widely applied and continues to be studied. The algorithm used in this study, among others, the Decision Tree algorithm, is a series of algorithms for classification problems in a machine and data set. Next, the Naive Bayes algorithm is a classification algorithm based

on the Bayesian theorem in statistics. The Naive Bayes algorithm can be used to predict the probability of class membership. The Bayesian theorem is a fundamental statistical approach to pattern recognition. Naive Bayes is based on the simplification assumption that attribute values are conditionally independent if output values are given [14]. K-Nearest Neighbor enters the classification algorithm so this algorithm can be used to predict new classes from datasets that have classes. K-Nearest Neighbor (K-NN) algorithm is a method for classifying a set of data based on learning data that has been classified previously [15]. This algorithm is also one of the lazy learning techniques. K-NN searches the k group of objects in the training data closest to or similar to the object in new data or test data.

5. Proposed Method and Results

5.1. Business Understanding

XYZ Bank needs a way to do the Non-Performing Loan problem solving faster. Therefore, effective and early action is needed. With this effective action, the company can minimize losses that arise in the future due to Non-Performing Loans. Therefore, with debtor payment history data, machine learning will be made that can predict problem debtors in the future. By utilizing the results, the team that handles the problem of non-performing loans can take action earlier by using a priority.

5.2. Data Understanding

The data prepared is data on credit payment history every month from September 2017 to February 2020. Data collection is done by downloading arrears reports from XYZ Bank’s core banking system. Initial data processing using MS Excel software. The arrears data is then processed into credit history data for each credit account each period. Within 1 period is 9 months of payment history. Debtor payment history data structure as shown in Figure 3.

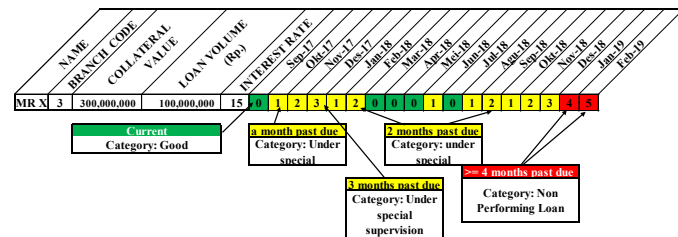


Figure 3: Debtor payment history data structure

The long historical data is then divided into periods per month with a stipulation that the length of the month is 15.

5.3. Data Preparation

The data that has been collected will be identified, selected with the help of MS Excel application software. For data classification, it is necessary to transform data on several data that will be used as attributes. Next is the evaluation of data requirements, attribute determination, data type, and class. The attributes that will be used in this study are:

- Branch Code (Representation of the region)
- Collateral Value (Classified as BIG and SMALL)
- Loan Volume (Classified as BIG and SMALL)

- Interest rate (Classified as BIG and SMALL)
- History of late payment for 12 months (M+11, ..., M+5, M+4, M+3, M+2, M+1, and M).

While the Label is M-3 (Prediction 3 months) with a value of {NOT and NPL}. NOT means it is not included in the NPL (Non-Performing Loan) category, which is a delay of <4 months. Whereas NPL means to be included in the category of Non Performing Loans, which is a delay of > 3 months. The process of collecting debtor delays data for each specified period can be seen in Figure 4.

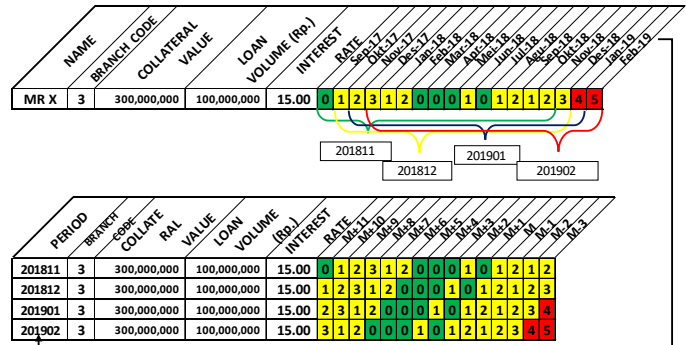


Figure 4: Data Collection With Conversion

After the data is collected, data conversion is required. Data conversion methods as shown in Figure 5.

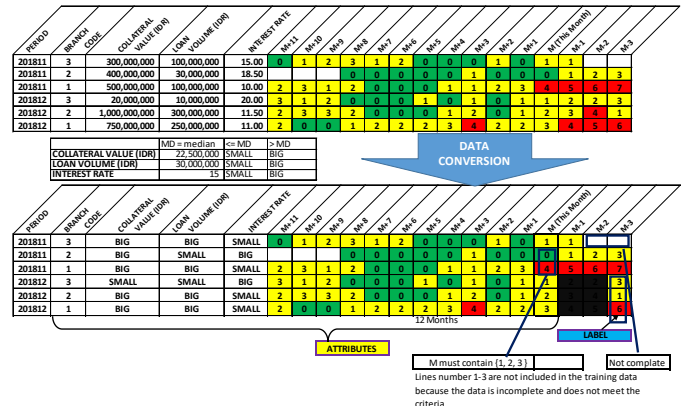


Figure 5: Data Conversion

From the data conversion, a total of 175,913 data were obtained, with the composition as shown in Table 1.

Table 1: Composition of Dataset

Class	Records	%
NOT	153,856	87%
NPL	22,057	13%

5.4. Modeling

By using the Rapidminer V 9.6 application, this study tested several models using the cross-validation method as shown in Figure 6. Because the composition of the data labeled NPL is far less than the data labeled NOT (not NPL), the sampling technique is done by taking all data labeled NPL and labeled NOT each as much as the amount of data labeled NPL. The implementation of the model in the Rapidminer application as shown in Figure 7.



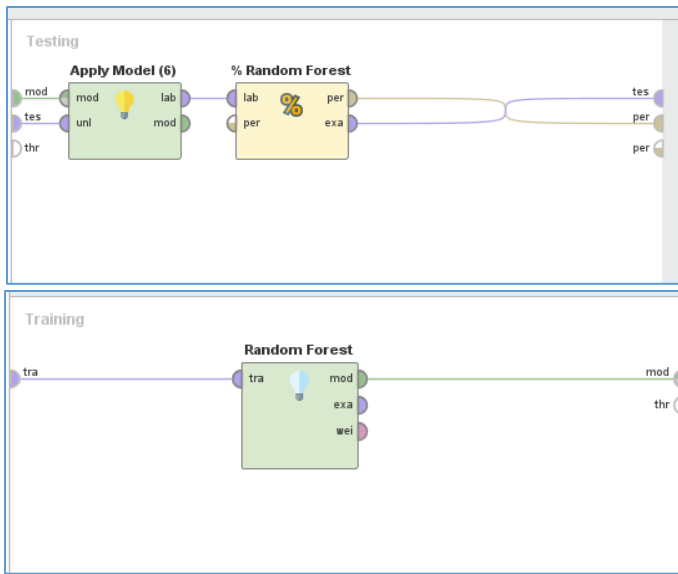


Figure 6: Cross-Validation Design

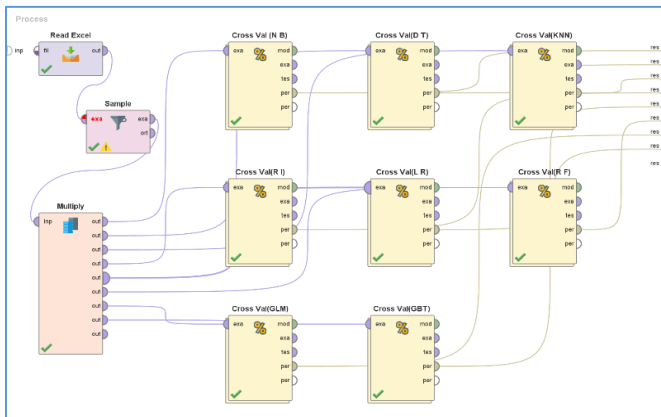


Figure 7: Model Implementation in the Rapidminer

Accuracy, AUC, F\_measure, Sensitivity, Specificity between each model can be compared [15]. The models that have been tested are Naïve Bayes (NB), Decision Tree (DT) [16], K-NN, Rule Induction, Logistic Regression, Random Forest, Generalized Linear Model dan Gradient Boosted Trees.

### 5.5. Evaluation

To see the accuracy of the model to each class, in this study using a confusion matrix [17]. By calculating the accuracy of some test data, the effectiveness of classification can be seen. From the several models tested, the results as shown in Figure 8.

In testing, the technique used in drawing data is balanced in each class {NOT and NPL}. From the comparison of test results, it was found that Random Forest was the highest accuracy. Accuracy is calculated from the total number of two correct predictions (TP + TN) divided by the total number of datasets (P + N). An excellent model has AUC near to the 1 which means it has a good measure of separability. A poor model has AUC near to the 0 which means it has the worst measure of separability. In fact, it means it is reciprocating the result. It is predicting 0s as 1s and 1s as 0s. And when AUC is 0.5, it means the model has no class separation capacity whatsoever. Random Forest - AUC: 0.996 +/- 0.000 (micro average: 0.996) (positive class: NPL) as shown in Figure 9.

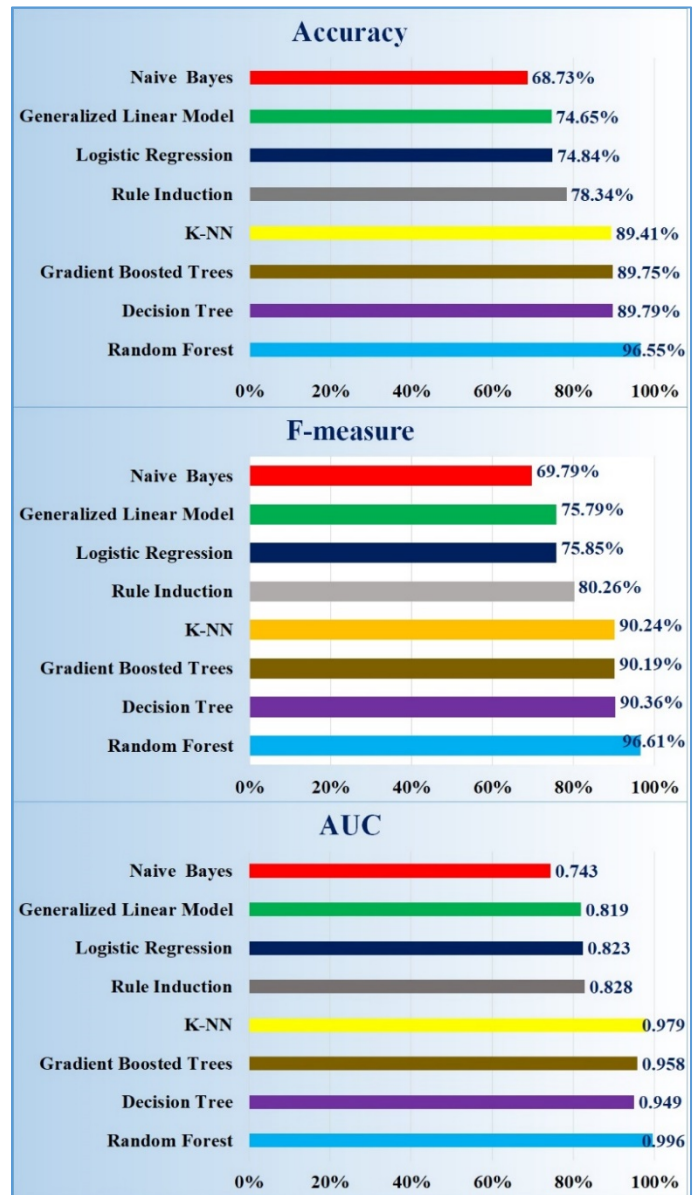


Figure 8: Graph for Accuracy, F-measure, and AUC on 8 models

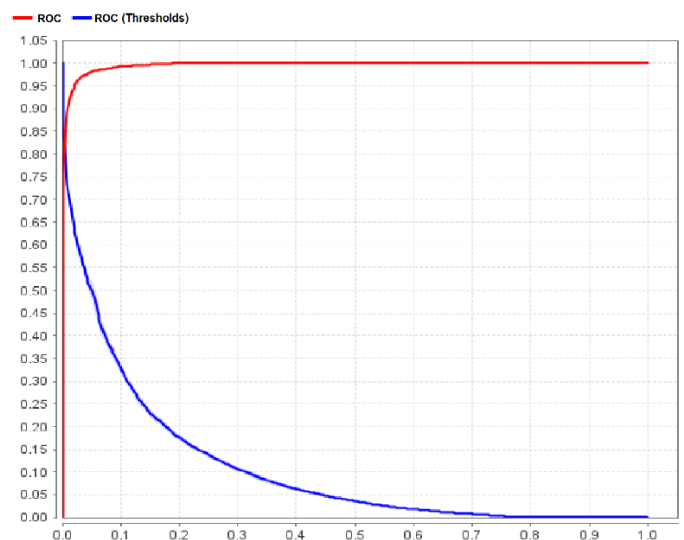


Figure 9: AUC - Random Forest



From the experiments conducted can show that the payment history can be used to predict the possibility of debtors becoming problematic in their payments. This is of course because the payment history is an orderly description of the debtor in fulfilling his obligations.

## 6. Conclusion

This research has discussed how historical data or loan repayment patterns can be studied through data mining and generate new knowledge to be able to predict future possibilities more accurately. In other studies it has been discussed that the dynamic behavior of the debtor during the last few years is more able to determine the condition of the debtor in the future which has an impact on the level of bad loans [11]. In this experiment, the data used are dynamic in the form of historical payments and payment patterns that can change from time to time, so that it can be better in terms of predicting uncertain conditions such as macroeconomic conditions, seasons, etc. Instead of using static data such as gender, marital status, employment, etc. less able to predict precisely [12]. The choice of attributes is very influential on the accuracy of the model made. One example is from this experiment found that the longer the payment history that we use, the more accurate the results obtained. In this experiment, it was found that factors such as branch codes also influence the determination of prediction results. The branch code represents the level of personal ability at the branch in handling this non-performing loan. Then the Collateral Value, Interest Rate and Loan Volume also have a correlation to the emergence of this NPL. Several models have been tested and compared to choose the best. Then it can be considered for implementation on the core banking system used by XYZ Bank. After predicting the NPL, we can use the data to be submitted to the relevant officers to follow up. Data mining in this study is still limited to credit loans with flat payment systems. From several models that have been tested in this experiment, it can be concluded that the Random Forest classification method is the most precise in predicting. It is expected that in subsequent studies it can predict types of credit models such as credit with seasonal payments, credit cards, and other models. Implementation of the model in the core banking system can be done by adding an NPL prediction module that processes data on the Database Server and provides reports and views to the user. The architecture of adding NPL prediction modules as shown in Figure 10.

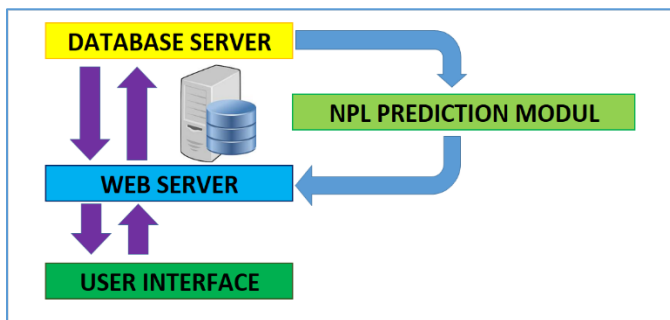


Figure 10: The architecture of adding NPL prediction modules

## References

- [1] M. B. Alexandri and T. I. Santoso, "Non Performing Loan: Impact of Internal and External Factor (Evidence in Indonesia)," *Int. J. Humanit. Soc. Sci. Invent.*, **4**(1), 87–91, 2015.
- [2] R. Agustiningrum, "Analisis Pengaruh Car, Npl, Dan Ldr Terhadap Profitabilitas Pada Perusahaan Perbankan," *E-Jurnal Manaj. Univ. Udayana*, **2**(8), 255030, 2013.
- [3] H. Elaidi, Z. Benabbou, and H. Abbar, "A comparative study of algorithms constructing decision trees: Id3 and c4.5," *ACM Int. Conf. Proceeding Ser.*, 3–7, 2018. doi: 10.1145/3230905.3230916.
- [4] J. S. Gil, A. J. P. Delima, and R. N. Vilchez, "Predicting students' dropout indicators in public school using data mining approaches," *Int. J. Adv. Trends Comput. Sci. Eng.*, **9**(1), 774–778, 2020. doi: 10.30534/ijatcse/2020/110912020.
- [5] X. Wang, C. Zhou, and X. Xu, "Application of C4.5 decision tree for scholarship evaluations," *Procedia Comput. Sci.*, **151**(2018) 179–184, 2019. doi: 10.1016/j.procs.2019.04.027.
- [6] W. Puarungroj, N. Boonsirisumpun, P. Pongpatrakant, and S. Phromkhot, "Application of data mining techniques for predicting student success in English exit exam," *ACM Int. Conf. Proceeding Ser.*, 1–6, 2018. doi: 10.1145/3164541.3164638.
- [7] C. J. Mantas, J. Abellán, and J. G. Castellano, "Analysis of Credal-C4.5 for classification in noisy domains," *Expert Syst. Appl.*, 2016. doi: 10.1016/j.eswa.2016.05.035.
- [8] H. Byeon, "A Prediction Model for Mild Cognitive Impairment Using Random Forests," *Int. J. Adv. Comput. Sci. Appl.*, **6**(12), 8–12, 2015. doi: 10.14569/ijacsa.2015.061202.
- [9] A. Shamshur and L. Weill, "Does bank efficiency influence the cost of credit?," *J. Bank. Financ.*, 2019. doi: 10.1016/j.jbankfin.2019.05.002.
- [10] X. Ye, L. an Dong, and D. Ma, "Loan evaluation in P2P lending based on Random Forest optimized by genetic algorithm with profit score," *Electron. Commer. Res. Appl.*, 2018. doi: 10.1016/j.elerap.2018.10.004.
- [11] S. Moradi and F. Mokhtab Rafiei, "A dynamic credit risk assessment model with data mining techniques: evidence from Iranian banks," *Financ. Innov.*, **5**(1), 2019. doi: 10.1186/s40854-019-0121-9.
- [12] I. G. N. N. Mandala, C. B. Nawangpalupi, and F. R. Praktikto, "Assessing Credit Risk: An Application of Data Mining in a Rural Bank," *Procedia Econ. Financ.*, **4**, Icsmed, 406–412, 2012. doi: 10.1016/s2212-5671(12)00355-3.
- [13] O. R. Devi, "International Journal of Advanced Trends in Computer Science and Engineering Available Online at <http://www.warse.org/ijatcse/static/pdf/file/ijatcse02422015.pdf>," **4**(2), 15–21, 2015.
- [14] N. Sun, B. Sun, J. (Denny) Lin, and M. Y. C. Wu, "Lossless Pruned Naive Bayes for Big Data Classifications," *Big Data Res.*, 2018. doi: 10.1016/j.bdr.2018.05.007.
- [15] A. Singh, M. N., and R. Lakshmiathan, "Impact of Different Data Types on Classifier Performance of Random Forest, Naïve Bayes, and K-Nearest Neighbors Algorithms," *Int. J. Adv. Comput. Sci. Appl.*, **8**(12), 1–10, 2017, doi: 10.14569/ijacsa.2017.081201.
- [16] M. Hoechstotter, A. Nazemi, and S. T. Rachev, "Recovery Rate Modelling of Non-performing Consumer Credit Using Data Mining Algorithms," **12**, 2012.
- [17] N. S. Buot, "Multiple intelligences and reading comprehension of senior high school students: A response evaluation through educational data mining technique," *Int. J. Adv. Trends Comput. Sci. Eng.*, **8**(6), 2871–2876, 2019, doi: 10.30534/ijatcse/2019/30862019.

## Diffusion of Technology for Language Challenges in the South African Healthcare Environment

Phathutshedzo Makovhololo, Tiko Iyamu\*

Department of Information Technology, Faculty of Informatics and Design, Cape Peninsula University of Technology, Cape Town 7700, South Africa

---

### ARTICLE INFO

#### Article history:

Received: 18 June, 2020

Accepted: 09 July, 2020

Online: 28 July, 2020

---

#### Keywords:

Healthcare

Diffusion of Innovation

Language translation

---

---

### ABSTRACT

Even though mobile technologies are increasingly used to enable communications between providers and receivers of healthcare services, there remain severity challenges owing to the use of various indigenous languages in many Africa countries. Some of the challenges get worst in a countries like South Africa where there are eleven official languages. As a result, interpreters and technology-enabled translation become inevitable. However, technology-enabled translation is even more complex because of the semantics that exists in the languages. The study was undertaken to identify the critical challenging factors that exists in the use of South African languages for healthcare services. The interpretivist approach was employed, and Rogers' diffusion of innovations (DOI) theory was applied in the analysis of the data, to examine and understand how indigenous languages affects healthcare services. From the analysis, six factors were found: confidentiality, translation, interpretation, synchronisation, spoken language, and integration, and categorised into three groups, namely ethics, services and systems.

---

## 1. Introduction

The activities and care of the health sector include diagnosis, medications, and treatment of disease and other forms of illness [1]. These activities rely on communication between actors (patients and healthcare practitioners) in the providing and receiving of healthcare services. As a result of significance of communication, information and communication technology (ICT) solutions, such as mobile technologies are explored as enabler. The ICT solutions are increasingly being used to facilitate communication of information across healthcare teams and groups, in the delivering of care and services [2]. Despite the increasing use of technology-enabled communication solutions, there remain many challenges. Some of the challenges arise from the use of various languages and dialect, which has different semantics. This challenge is worst in South Africa where there are eleven official languages.

Overcoming language barriers within the healthcare sector is considered one of the critical challenges threatening care services to patients [3]. At the time of this study, there were about 54 million people in South Africa. From which only 4.9 million people were English speakers, and the remainder of the population speak African languages such as IsiZulu, IsiXhosa, Sepedi, Sesotho, Tshivenda, and Xitsonga. Consequently, the healthcare

delivery system is hard-pressed to handle this diversity [4]. Qassim gives an example, that when capturing text using two different languages for the same thing, translation is required because different meanings can be created despite similarities in the intentions to say the same thing [5]. Another factor that influence the challenges of communication is cultural affiliation. In addressing some of these challenges, innovations can be employed and diffused. Rogers diffusion of innovation (DOI) has been applied in my studies where innovations or diffusion of technologies were investigated [6].

The DOI theory has been widely applied in studies including those of public health and medical care, in many areas such as diffusion, integration, innovation, and implementation [7]. In DOI, an *innovation* is referred to as “an idea, practice, or object that is perceived as new by an individual or another unit of adoption” [6: 11]. Diffusion is defined as “the process by which an innovation is communicated through certain channels over time among the members of a social system” [6]. One of the focuses of the DOI theory is to understand why “potential users make decisions to adopt or reject an innovation based on beliefs that they form about the innovation” [8]. Theories are concerned with the manner in which a new technological idea, artefact or technique is migrated from creation to use [9]. Rogers classifies the innovation-decision process into five categories: (1) knowledge; (2) persuasion; (3) decision; (4) implementation; and (5) confirmation [6].

---

\*Corresponding Author: Tiko Iyamu, Email: connectvilla@yahoo.com

Based on the discussion presented above, the objective of this study was to identify the factors that critically influence healthcare services in the use of South African languages. The remainder of this paper is structured into 5 main sections as follows: the literature review is followed by the research methodology. Thereafter, the analysis of the data, and the discussion of the findings are presented respectively. Finally, the conclusion is drawn.

## 2. Literature Review

The use of language enable and at the same time constraint communications and interactions in some instances. In South Africa, there are eleven official languages, which include English and indigenous languages. Of the languages, English is the most used as a medium of communication, which is a challenge for the majority who are neither literate nor fluent in the language [10]. This is a challenge that affects medical services in the country, which have more negative impact on patients' health conditions [11], [12]. For examples, some doctors' over-treat patients as result of limited English proficiency, resulting from lack of appropriate self-expression of health condition, and some patients are referred to get unwarranted additional tests and procedures [13], [14]. Language barrier is not only about proficiency, but complicated by semantics, which is often unique in translation. Translation of semantics in languages tend to be a more complex, awkward, detailed, and concentrated thought-process [15].

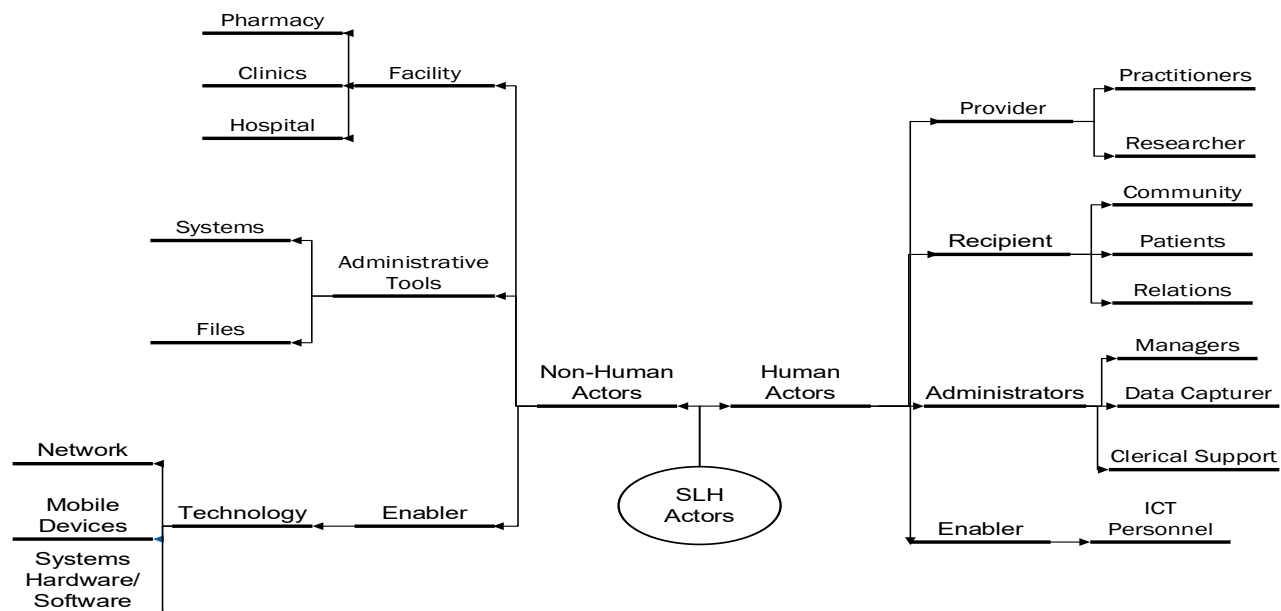
Based on challenges posed by language, ICT solutions such as mobile systems has been employed to accelerate access, and provide a wider range of health services in order to enhance the well-being of underprivileged people in South Africa [16]. Also, translation of semantics in some languages often change the original meaning that was associated with certain words [17], which has made it difficult for mobile systems to address. Healthcare service providers have increasingly employed mobile systems in delivering services [18].

Due to sensitivity of some health conditions, many patients prefer to be active participant in their own care by communicating directly with the practitioners, to influence decisions relating to their own care [19]. This includes accessing of information using language of their origins. This potentially improves communication between individuals and healthcare providers, ICT systems such as personal health records (PHR), which ultimately increases patients' participation in healthcare. Some ICT systems can be used to engage with patients toward improving communication [20]. The challenges posed by language affect integration of information systems, and access to health information.

Misunderstand between patients and health practitioners can result to fatality. An example is a case of 'Life Isidimeni' in South Africa where many patients lost their lives owing to distortion in communication, and other factors [21]. This once again demonstrates that the limits of language are the limits of a person's world, and the fact of the matter is that the 'real world' is to a large extent unconsciously built upon the language habits of the different communities of humans, speaking different languages. Different groups think differently to the extent that their languages differ from one another [22]. One of the challenges to be confronted by the healthcare industry is semantic interoperability. *Semantic interoperability* is the ability of a healthcare system to share information and have that information properly interpreted by the receiving system in the same sense as intended by the transmitting system. Semantic Web provides enabling technologies to achieving semantic interoperability [23].

## 3. Prepare Your Paper before Styling

Based on the objectives of the study, the interpretivist approach and qualitative methods were employed. Vosloo states that the interpretive approach enables access to reality through language and shared meanings [24]. The approach was employed in this study to seek explanation within the realm of individual consciousness and subjectivity, and within the socially constructed sphere of human relations [25]. This was done within the context of case study.



The case study approach allows focus on specific setting or environment, and multiple sources of evidence [26]. With the approach, it is also easier to extract cultural knowledge, identify actions and instruments that participants utilise in their everyday life in an organisation [27]. A health facility with the pseudo name 'Sweet Love Hospital' (SLH) was selected. The facility renders health services to all that lives in South Africa, irrespective of beliefs, cultures, and spoken language. Human and non-human actors in SLH were categorised into four phases, as shown in Figure 1: (1) provider, (2) recipient, (3) administrator, and (4) enabler. The facility is situated in Pretoria area of South Africa. The hospital focuses on patients who have been diagnosed of chronic conditions, such as tuberculosis, Human Immunodeficiency Virus (HIV), and cancer. In providing healthcare services to the patients, require regular and constant communication is vital. However, the patients are diverse in spoken languages, which sometimes affect the type of services that they receive.

The semi-structured interview was applied to gather data. This helps to gain an understanding about behaviours, beliefs, opinions, emotions, and relationships between individuals and groups [28]. The primary purpose of conducting interviews is to gain an understanding of the meaning of what the interviewees say. Participants in the study were from different unit of the hospital, which included medical practitioners (nurses and doctors), IT and administrative services. A total of eleven people were interviewed at the point of saturation. The interviewees were labelled 'SLH01 to SLH11' in order to preserve their identities.

The analysis of this data followed the interpretivist approach, which was guided by [6] Diffusion of Innovation (DOI). The use of DOI theory focuses on how decisions were made by both healthcare providers and recipients on languages preferences for health services. The DOI helps to explain how, over time, an idea or product gains momentum, and diffuses through a specific population or social system [29]. According to Striphas an innovation can be an idea, knowledge, a belief or social norm, a product or services, a technology or process, even a culture, as long as it is perceived as being new [30]. For Rogers, the innovation-decision process involves five steps: (1) knowledge, (2) persuasion, (3) decision, (4) implementation, and (5) confirmation [31]. These stages typically follow each other in a time-ordered manner, as applied in this study.

## 4. Data Analysis

### 4.1. Innovation decision process: Knowledge

The innovation-decision process starts with the knowledge, a stage where individuals learn about the existence of innovation, and seek information concerning the innovation. "What?" "How?" and "Why?" are critical questions at the knowledge stage. Also, during this stage, an individual attempts to determine "what the innovation is and how and why it works" [31]. According to Rogers, the questions form three types of knowledge: (1) awareness-knowledge, (2) how-to-knowledge, and (3) principles-knowledge.

Some health practitioners in SLH make use of IT solutions such as mobile system to access information, and engage in communicate with colleagues and patients. However, it becomes challenging for IT staff when some health practitioners (users) are not able to make use of the system appropriately. According to

SLH01 (2:66-67), "sometimes, it is a challenge when users are not knowledgeable about the application of technology". The emerging and changing technological development has led to a rethinking in the way technology can be effectively used in the delivering of healthcare. Technology-enabled systems allow healthcare providers to deliver better support at lower cost as well as contribute to quality health services. Therefore, healthcare practitioners must have a thorough knowledge of how different technologies in their environment work.

Technology certainly plays a part in the success or failure of these operations in SLH. Without a robust IT infrastructure, SLH would not be able to deliver the promised benefits of integration. This may not be immediately obvious to people in the healthcare industry, which were near the bottom of the ladder in terms of IT spending and uniform data standards. "Some employees were aware of the technology and any software application that can make their tasks including communication easier" (SLH01, 2:72-73). The SLH was challenges with language which was being used to communicate because not everyone in South Africa are able to speak all eleven official languages of the country. Furthermore, many people are not literate or knowledgeable enough to be able to assimilate communications understandingly. According to SLH03 (9:344-345), "we send messages to the patients in our own languages, and we don't get feedback, which is because some of the patients do not understand the language that we are trying to communicate with".

The SLH has many stakeholders, each with an agenda. Often these players have substantial resources and the power to influence public policy and opinion by helping or hindering an innovator. For example, hospitals and doctors sometimes end up blaming technology-driven product innovators for the healthcare system's high costs because they lack the knowledge of what has caused the problem. "They hardly acknowledge that they do not know how to use the application" (SLH02, 5:180). Also, many SLH clinics are situated in remote villages lacking easy access to hospitals and advanced medical facilities. However, the village inhabitants are now being educated through telemedicine and digitalised health information. This initiative is helping millions of citizens to improve their daily lives. This will address the knowledge and skill transfer of how to use mobile systems.

As medical technology evolves, the determination of how and when to adopt or invest in this technology increases in importance. Move too early, and the infrastructure needed to support the innovation may not yet be in place; wait too long, most of the SLH users are not using the technologies available them because they are not fully aware of them. While some are aware, they often lack the knowledge of language semantics used as part of the application. "We try to understand why some of the patients did not respond, and you find out that the problem was about language. Some of them did not understand what we have sent to them" (SLH01, 6:233-234).

### 4.2. Innovation decision process: Persuasion

The persuasion step occurs when an individual possesses an attitude, negative or positive, toward the innovation. [31] describes this as "the formation of a favourable or unfavourable attitude toward an innovation does not always lead directly or indirectly to an adoption or rejection". The individual shapes his attitude after learning about the innovation, so the persuasion stage follows the knowledge stage in the innovation-decision process. As in many African countries, most adults, aged 55 and above in SLH did not



find the use of technology friendly, which affect their use in accessing healthcare. The impact is not only from the use of technology but also the language that is used as part of these technologies. People, therefore, are more inclined to recognise ICT usefulness in healthcare, especially if they have had exposure to the systems.

The management of SLH resorted to awareness campaigns, through which they educated elderly patients on the use of technology in accessing health information. According to SLH03 (9:332-333), *“we always educate the users, and try to make them understand the benefits of using technology such as the pharmacy automation system for healthcare services”*. At SLH it was definitely challenging to persuade the user to accept systems before they could visibly see the functionality; some were persuaded while others waited for full implementation. For example, *“the pharmacists accept the use of the ‘mega four’ system. It was only then it occurs to many of the patients how important or beneficial it is to buy-in, and use the system”* (SLH03, 9:332-333). Persuasion becomes difficult because a number of patients are illiterate, escalating the difficulties which many healthcare providers endure. Furthermore, [31] states that while the knowledge stage is more cognitive (or knowing) centred, the persuasion stage is more affective (or feeling) centred. Thus, the individual is involved more sensitively with the innovation at the persuasion stage. This was revealed at SLH: those in the same unit, communicated better, and interaction got easier.

The degree of uncertainty about an innovation’s functioning, and the social reinforcement from others (colleagues and peers, for instance) affect individuals’ opinions and beliefs about the innovation, to an extent that IT personnel have to devise ways of making the users understand how this innovation works. According to SLH01 (4-5:160-162), *“As times goes by, as you deal with these challenges, you learn better ways of explaining things to the users at their level of knowledge”*. In most instances, communication about the innovation is silent which affect awareness and knowledgeability. As a result, many health practitioners as well as patients were not able to make use of some of the available systems

This was despite the fact that technologies were required, they were also barrier during interactions between healthcare workers and patients in providing or receiving healthcare service. If healthcare service providers were educated enough on the use of technologies, it had been easier to convince them to adopt the technology, so as to avoid the need for a third party. A third party as interpreter compromises privacy and confidentiality, which in turn affects the ethics code of conduct. Therefore, it is important to develop technologies that can translate and interpret for patients rather than the use of a person as a third party. One of the participants, SLH04 (10:406-407) expressed her view as follow: *“The communication team that always translate the messages into vernacular or indigenous languages were not always professional interpreters or translators”*. This demoralises some patients because information about their health was not treated with confidentiality.

#### 4.3. Innovation decision process: Decision

It is at the decision stage of the innovation-decision process an individual can choose to adopt or reject the innovation. While adoption refers to “full use of an innovation as the best course of action available,” rejection means “not to adopt an innovation” [31]. If an innovation has a partial trial basis, it is usually adopted

more quickly, as most individuals first want to try the innovation in their personal situation before arriving at an adoption decision. Healthcare providers in SLH were faced with a number of challenges, which included the increasing size of the aging population, shortage of healthcare workers, patients’ demands for increased access to health information, and participation in healthcare related decision-making, and rising healthcare costs. According to SLH03 (9:333-335), *“it was very challenging for the pharmacists to accept the ‘mega four’, when the system was initially deployed and introduced”*. The vicarious trial can speed up the innovation-decision process.

However, rejection is possible in every stage of the innovation-decision process. Most of the doctors were satisfied with adapting, as they want advance ways to lighten or hasten their jobs, but the nurses find it difficult as they were used to old ways of manual systems. Many people remain loyal with their old approach, but with time, they will become accustomed to new ways. In an attempts to increase productivity, some healthcare leaders and decision-makers tend to focus on short-term financial gains. However, the short-term gains derived from rash and quick decisions to solve a big problem can result in communication breakdowns that produce higher overall healthcare costs, such as inaccurate diagnoses that prompt unnecessary tests and services repeated multiple times. As per SLH11 (32:1269-1271), *“Across the sphere of South Africa, we work directly with healthcare practitioners and government representatives, to identify facilities, sub-districts, districts and provinces where we can deploy unique models in addressing individual and group needs”*.

Dealing with technologies can incur heavy financial challenges. There are times when decisions are made to rely on their own staff as interpreters forced by circumstances they found themselves in. They don’t have full-time interpreters on the premises, so if they need a professional interpreter, they must request this from public hospitals for assistance, which cause critical delays. However, these providers fail to consider both the consequences of not providing the services and the potential cost benefits of improving communication with their patients. Sometimes, they claim the challenge rests not necessarily with the language per say, but on the terminology used on the devices that are provided. According to SLH04 (12:454-455), *“I think the issue of semantics is, one of the challenges, which also depends on the kind of messages that you try to send to the patients”*. The SLH attempts to address the issue of language semantics translation, which requires decision and implementation of the decision that has been made.

#### 4.4. Innovation decision process: Implementation

At the implementation stage, an innovation is put into practice. However, an innovation brings the newness in which some degree of uncertainty is involved in diffusion. Uncertainty about the outcomes of an innovation can still be a problem at this stage. Thus, the implementer may need technical assistance from change agents and others to reduce the degree of uncertainty surrounding the consequences. At the time of this study, SLH has partnered with the Department of Health of the country, purposely to eradicate the uncertainty of users: *“from working with the Department of Health and our global donors, we have secured funding to rollout the e-Pharmacy system, which have helped in our health-related activities”* (SLH08, 27:1096-1097). Moreover, the innovation-decision process will end, since “the innovation loses its distinctive quality as the separate identity of the new idea disappears” [31].

The use of ICT solutions in the SLH healthcare sector has well-known advantages: promoting patient-centred healthcare, improving quality of care, and educating both health professionals and patients. However, implementation of ICT solutions remain difficult as it involves changes at a variety of levels: patients, healthcare providers, and healthcare organisations. In order for SLH’s IT Department to achieve their goal of implementing these technologies, they require the support of the users, who in this case were healthcare practitioners (physicians, nurses, councillors, researchers, fieldworkers, pharmacists, administrators, and executives). The experience of difficulty in communication or language in many instances, IT personnel proactively demonstrated interest by accommodating the users as they interact with them. The implementation of computer-aided detection of Tuberculosis at SLH, was an example that radiologists were scarce in South Africa, and to review each digital x-ray would have been time-consuming and costly. The facility, SLH searched for available technologies being used abroad towards addressing this challenge: *“We selected a Computer-aided detection for Tuberculosis (CAD4TB™) that was developed by Delft Imaging Systems. The system was designed to help (non) experts’ readers in the diagnosis of TB”* (SLH02, 28:1138-1139).

In SLH, the implementation of mobile health was a significant challenge in that it was attributed to the broad scope of services that mobile health need to address. This includes lack of resources such as infrastructure; and a lack of funds. In 2014, SLH identified the need to methodically document, evaluate, and share best practices, clinical principles, and innovations arising from the organisation’s programmes. The organisation established a dedicated operational research unit to meet these needs. Many of the unit’s solutions have resulted in significant positive changes to clinic functioning and consequently in the lives of patients. The health facility, SLH managed to implement systems such as *“an interface between TherapyEdge® and Trimed®, to eliminate the need for pharmacists’ to recapture prescriptions’ data”* (SLH07, 25:1020-1021). From the implementation, the transfer of data was managed via SHL07 messaging (an industry standard messaging system). This ensures up-to-date flow of patients’ demographic and prescription data between the systems, each system receiving and sending real-time information.

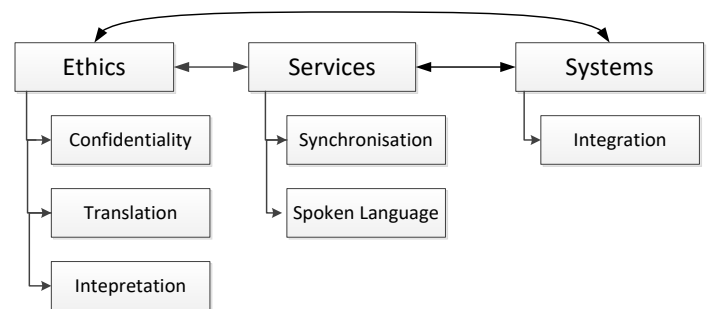
The innovation-decision has already been made, but at the confirmation stage an individual seeks support for his or her decision. According to [31], this decision can be reversed if the individual is “exposed to conflicting messages about the innovation” (189). However, the individual tends to stay away from these messages, and seeks supportive messages that confirm his or her decision. Thus, attitudes become more crucial at the confirmation stage. In this case, SLH accepted the innovation: *“The response was very positive. The e-Pharmacy dispensing unit made our lives easier because we no longer do somethings manual, such as standing in long queues and working long hours”* (SLH03: 338-340).

An efficient user-friendly customer relationship management (CRM) application was used to speeds up efficiency of wellness and HIV testing. The LoveHealth Mobile enables staff members to seamlessly synchronise data from a central database with mobile data capture devices. These systems enable testing process, while capturing accurate test information. One of the participants (SLH06, 22:929-931) explained: *“it can also save time and at the same time, it limits invasion of privacy, which makes some people feel free to express themselves without third party”*.

Finally, the widespread adoption of mobile technologies at SLH allowed for system connectivity and information exchange between practitioners, healthcare organisations, and ultimately between patients and community. However, the actual realisation of these benefits depended on how SLH implements and deploys the mobile technologies. As results were dependent on the support for adoption of the innovation and the attitude of the individual, later adoption or discontinuance happens during this stage. External suppliers with similar trucks were contracted and the digital x-rays routed into the same central Picture Archiving Communication System (PACS) for review and reporting. This innovation has empowered SLH, in partnership with LoveHealth, to x-ray 57 290 Department of Correctional staff and inmates. According to SLH07 (29:1157), *“after the success of the first truck, two more were built and rolled out”*.

## 5. Discussion of Findings

Based on the analysis of the data, six critical factors were identified for the translation of semantics in languages in providing and receiving healthcare services in South Africa. As shown in Figure 2, the factors – confidentiality, translation, interpretation, synchronisation, spoken language and integration – were categorised into three groups, namely ethics, services, and systems. The groupings were subjectively reached, based on the closeness of the factors in the provision and receipts of healthcare services. The three groupings are discussed below.



Factors that influences technology diffusion.

### 5.1. Ethics of health information

Healthcare practitioners are governed by code of conduct, which include keeping patients’ information private and confidential. This was critical to the hospital since it focuses on chronic conditions which include HIV/AIDS and tuberculosis. In the provision and receipt of service for these and other chronic conditions and the general healthcare, ethics are required to guide the activities of translation of semantics in languages. Chronic conditions are considered very sensitive, for a host of apparent reasons. Thus, confidentiality was high in the ethical code of conduct of the SLH. In addition, translation and interpretation were critical from the viewpoint of semantics, which must prioritise ethical behaviour. These factors – confidentiality, translation and interpretation – were challenging for both patients and healthcare practitioners owing to their inseparability nature.

Due to these challenges, compounded by lack of knowledge of the practitioners, some of the health practitioners began to form different groupings (networks), along spoken languages and area of specialisations. These networks were formed with the intention

of addressing practitioners and patients' communication barriers. Some practitioners tapped on actors in their networks as resource to play the role of interpreters of semantics that they themselves were unable to translate in some of the languages. Lack of knowledge about some spoken language led to making decisions, which became norm and culture of the hospital, irrespective of the fact that this culture or norm was unethical, in that the informal interpreters were not trained to do so. As a result, some of the interpreters did not have deep enough understanding of the languages to carry out semantics translation. Translation of semantics in languages requires a precise standard that can be followed for accuracy of meanings. For example, as expressed by one of the nurses, in Tshivenda, "Ndi tshitamba tshifhinga tshothe nowa yanga iya luma", which means "I suffer from period pain every time I menstruate". But when translated into English, it means "every time I take a bath my snake bites".

There were inaccuracies in translations of semantics in some of the languages. As a result, patients' levels of trust and confidence in the healthcare practitioners began to drop, and some patients involved their family members or friends to help with translation during consultations with health practitioners. This approach breaches practitioner-to-patient confidentiality, which compromises ethical behaviour. Also, there was no proper documentation on how translation was carried out at the SLH. This could have informed the creation of a repository, where reuse of translation can be enabled for accuracy and consistency purposes.

The fear of information leakage, sometimes caused some patients to lie about their actual health conditions, which often resulted to wrong diagnoses and medications. The code of conduct doesn't only apply to healthcare practitioners, but also to patients who need to be knowledgeable of ethical practices and behaviours on health matters. There is a clear need for education and awareness for practitioners concerning the issues of the code of conduct. This will minimise legal suits and prevent loss of lives, which result from misconducts.

#### *5.1.1. Confidentiality*

The SLH employs a dual confidentiality approach in accessing information, and interaction between healthcare practitioners and patients. However, the involvement of a translator, a third party, weakens this approach, and makes the goal of confidentiality less achievable. But all hopes are not lost as the use of a mobile system for translation of semantics in languages can reinforce security and privacy of information. The mobile system solution for semantics translation can instil trust and confidence in both practitioners and the patients. Thus, patients do not have to choose which practitioners they need to consult with because they were scared that privacy is at risk.

From two main perspectives, the mobile systems solution can be of critical importance for confidentiality in the translations of semantics, as follows: (1) the solution enables and support reuse, which can be efficient in the storage and retrieval of translated semantics in the languages that are prevalent for communication at SLH; and (2) the mobile system solution allows for validation of information at various levels, and grants various levels of access to information to both practitioners and patients. This improves manageability and confidentiality of patients' information. Also,

this approach enhances compliance to health ethics within SLH. Most importantly, the mobile system solution approach can improve health service delivery.

#### *5.1.2. Translation*

Translation of healthcare terminology in any language, especially South African languages, can be very difficult to virtually impossible. As revealed in the analysis, many meanings get lost in translation from English to some South African languages such as Tshivenda, Sepedi, Xitsonga, and IsiZulu. Even so, community members and patients were faced with the challenges of having to translate these words at critical times in their life. For example, if an Xitonga-speaking patient says to the health practitioner at SLH: "ndzi twa ndzi hlamba timbilu", the English translation means "I feel like my hearts are washing". However, what the patient was trying to say is that "I feel nauseous".

More often than not, patients and community members source health-related information from media such as radios, television, and the internet. Some of the information gets confusing instead of providing clarity to gain an understanding of health conditions. The confusion is often caused by challenges in attempts to directly translate certain words and phrases from English to the indigenous languages. In some cases, in attempts to address the confusion, networks of interested members are unconsciously formed. Despite the efforts, accurate meanings of some words were not realised because of semantics. Thus, the closest of meanings are accepted, sometimes based on assumptions, as language translation of semantics is sometimes based on assumptions and the little knowledge some members of the network have. This challenge is mostly with illiterate patients.

Those had advantage were mostly the literate patients or relations of illiterate patients who further carry out searches on relevant topics and hope for better translation through translation-triangulation: (1) accurate translation to a language such as Sepedi; (2) thereafter, translation from Sepedi to Xitsonga; and (3) double checking a word's meaning in English. However, the triangular approach does not necessarily accurately resolve the problem, for two main reasons: (1) it is not common to have interpreters who are deeply fluent in two or more South African languages; and (2) some semantics do not exist in or cannot be translated to other languages. Wrong or inaccurate translations can be dangerous in that it leads to wrong medical decisions that can seriously impact patient's health in the long run or fail to prevent an illness that could have been otherwise prevented.

#### *5.1.3. Interpretation*

Within healthcare ethics, interpretation of semantics in languages is vitally important, especially during consultation with patients. As a result, either the patients or practitioners get an interpreter if they realise that there is a communication challenge. Although this often solves the immediate problem, it is unethical for healthcare practitioners or patients to randomly choose an interpreter. Also, it is ethically wrong for healthcare practitioners to allow themselves to be influenced by a patient, to make use of their next of kin as an interpreter. This is mainly because it is highly unlikely that non-professional interpreters will adhere to



health ethics code of conduct. The non-professional interpreters do not sign the ethic code of conduct, also do not understand the details and implications of breaching the code of ethics within the sensitive health environment. In addition, there is no guarantee that non-professional interpreters will be accurate in their interpretation of semantics in some languages.

Owing to this challenge, some of the things patients had to say or tried to explain might be 'lost in translation' in that some semantics in some languages cannot be easily translated. Despite the risk of miscommunication, there is no standard method of translating semantics of languages. The depth of some of the languages do not help in addressing translation challenges. Also, in some instances, the interpreters were not appropriately trained to understand the semantics in the languages. Due to lack of knowledge and faulty understanding of the semantics in language, the words lose meanings. This may impact patients who are illiterate as they won't know the difference. If these patients were literate, they may not have even needed an interpreter in the first place. Therefore, in this respect we have two issues: translating what someone does not have acute knowledge of, and violation of the code of conduct.

## 5.2. Providing health services

The health services rendered by the hospital have to be of high quality irrespective of the patients' spoken languages, cultural or religious affiliations and background. Therefore, the intended quality should determine the innovations that the hospital employs. Providing healthcare services require collaboration among health practitioners, in their use of patient information, technologies and innovations. However, collaboration was not always synonymous to service at SLH. Some practitioners relied solely on their individual knowledge of an incident or wealth of experience. Through their stock of knowledge and wealth of experience, they made decisions to create an innovative approach. This impacts the services being rendered by the hospital as a whole. The danger was that some of those innovative approaches were not validated. The use of a mobile system can help address this challenge, in that the knowledge and experiences can be documented in a repository which can be accessed at any time and for various purposes by the practitioners.

The innovations for health services were embedded with terms (semantics) that were only understandable by the creator. This was another difficulty that both the creator of the innovation and the interested parties (other health practitioners) encountered. The challenge can be attributed to the main reason why the creators of the innovation had rather chose to make it exclusive. However, if they could communicate and enlighten other colleagues of the innovations, they had persuaded more interested persons in the adoption of the innovation. Duplication of services costs valuable hospital resources and time as people have to redo what has been done but in a different format. This might also have an impact on the use of technologies as each comes with its own terminologies. However, the hospital must be able to use the same terminology to avoid being lost in translation as some patients and health workers might not be able to understand the semantics in those systems of which they are unaware.

### 5.2.1. Synchronisation

Taking ethics into considerations, synchronisation is critical within the healthcare environment, primarily because it is experimental-based. Some healthcare practitioners, particularly doctors, bring into their practice what they think is the best innovation according to their knowledge without considering what is already available. For example, only one of the doctors at SLH made use of an open source system, to communicate with pregnant women who were on chronic medications. Some health practitioners sometimes do this without following any stringent process, which often affects an existing process. Such approach began to affect developmental processes, and knowledge sharing within the environment. If they had followed a process, they would have been able to persuade employees through relevant training on the use and application of the innovation. Then one particular system would have then be enhanced, if need be; however, with this process they end up with two systems which were not integrated, but running in parallel.

### 5.2.2. Spoken language

As revealed from the data analysis, spoken language was key to provision and receipts of healthcare by SLH and patients, respectively. At the time of this study, SLH did not have a system to identify the areas as per different spoken languages. Therefore, such identifications were done by guesswork. As a result, practitioners were allocated to areas for health services based on their availability, not based on whether they can communicate in the predominant language of the people in the area. The implication of this decision negatively impacted prescriptions, medications, and medical guidance provided to some patients. Some healthcare practitioners were not knowledgeable enough to make the responsible decision that they made when they sent practitioners out to the field. They sent people based on guesswork, availability of personnel, which were insufficient requirements for such sensitive environment.

### 5.3. ICT Systems for health services

At the SLH, there were various ICT systems. The use of systems did improve healthcare services to patients. However, the specialists were required to have a deep knowledge of the available systems before embarking on the implementation process. This was not always the case at SLH as some specialists lacked knowledge of the systems they intend to implement. As a result, groups of practitioners tried to find for themselves systems that worked best. Thus, some of the systems were implemented in isolation, without communicating or making their colleagues aware of the system's existence.

The implementation of numerous systems without proper valuation led to duplication of systems as well as services, as some practitioners made use of systems in isolation, while other practitioners employed their own systems. The systems then run in parallel without integration. Therefore, patient's information gets duplicated depending on the structure of the system. In some cases, patient's information was not captured appropriately and completely. In emergency cases, this caused frantic delay in trying to identify the patient by using parallel systems. It get worst if the patient's details such surname or address has changed in recent



times. The dual system was not the main challenge, but the lack of synchronisation, which could have regularly update the patients' details in the systems.

### 5.3.1. Integration

The hospital's systems were not integrated at the time of this study. From the inquiries to the general practitioners and pharmacy, parallel systems were operated. Unfortunately, these systems were not 'talking to each other'. As a result, some patients explored the gaps and often obtained more than one prescriptions. Such situations were not traceable because the systems were not linked. As was revealed from the data analysis, some patients' who live in informal settlement get medications from clinic in their area, and subsequently travel to another clinic to collect again the same medications because their systems cannot detect that the patient has already received this medication.

This practice by some patients to take advantage of the gap, lack of integration between the systems was influenced by two main reasons: (1) some of the patients sell the extra medicines to friends and interested persons; and (2) others consume more medicines than the prescribed doses. As for the latter, some of the patients held the erroneous belief that the more medicines they consume, the quicker they get healed. Another implication for the hospital (SLH) and its clinics was that they were often running short of medications. Shortage of medications led to fatalities as people were dying because of shortages in the supply of medications. These challenges were yet more devastations caused by inaccurate translation of languages or the semantics in languages.

The fact that the parallel systems were not in sync has led to poor performance of some of the staff members as they find themselves having to perform certain tasks repeatedly. The data capturing was hard-pressed with capturing of patients' information that came from the doctors and pharmacies. If the systems were integrated, several of these continuously repeated steps had been eliminated. As it stands, the process takes long, and patients queue for long hours. As revealed in the data analysed, some patients 'passed out' as a result of fatigue, which chronic patients are generally sensitive to. Mobile systems would be the solutions, as they can be used to integrate the parallel systems so that the information can be synchronised for improved efficiency and effectiveness of healthcare services.

## 6. Conclusion

This study identifies and presents some of the factors that influence the challenges some patients and health practitioners encounter in their attempt to communicate in the delivering and receiving services. As shown in Figure 2, the study puts the factors in perspectives as it demonstrates how they interrelate to influence actual communication between patients and health practitioners in the use of different languages. Both government and managers of healthcare services can make use of this study to trace the factors' influence in avoiding the delivering of poor services in the country.

The significant of this study can be viewed from three main perspectives: (1) health practitioners can now be more sensitive to patients' predicaments, based on the identified factors, which some

of them didn't know the extent of the challenges; (2) the factors are empirical evidence that can guide policymakers in formulating and promulgating policies for the enhancement of healthcare services; and (3) the factors can guide software developers in developing a system that translates semantics in South African languages, which can be used by both practitioners and patients. From academic front, the study adds to existing literature in the areas of IS, healthcare, and language. Currently, there are very rare literature in the area of IS and healthcare from language translation perspective, particularly in developing countries.

As work continues in improving healthcare service delivery, further studies can focus on measuring the impact of the factors identified in this study. This includes development of mobile systems for the translation of African languages.

## Acknowledgment

The authors wishes to acknowledge and thank the Cape Peninsula University of Technology (CPUT) for giving us an opportunity to explore and grow in the research world.

## References

- [1] K. S. Gupta, V. Rokade, Importance of quality in health care sector: A review. *Journal of Health Management*. 18(1):84-94, 2016. <https://doi.org/10.1177/0972063415625527>
- [2] O. P. Kogeda, N. Mpekoa, Model for A Mobile Phone Voting System for South Africa. In 2013 *Conference*. [https://www.researchgate.net/publication/256815434\\_Model\\_for\\_A\\_Mobile\\_Phone\\_Voting\\_System\\_for\\_South\\_Africa](https://www.researchgate.net/publication/256815434_Model_for_A_Mobile_Phone_Voting_System_for_South_Africa)
- [3] S. Naidoo. Transcultural and language barriers to patient care. *South African Dental Journal*. 69(9):425, 2014. <https://hdl.handle.net/10520/EJC160573>
- [4] L. R. Burns, ed, *The business of healthcare innovation*. Cambridge University Press, 2012.
- [5] A. H. Qassim, Translation, grammatically, viewed. University of Baghdad/College of Languages. 2014.
- [6] E. M. Rogers, "Diffusion of Innovations: modifications of a model for telecommunications." In *Die diffusion von innovationen in der telekommunikation*, 25-38. Springer, Berlin, Heidelberg, 1995. [https://doi.org/10.1007/978-3-642-79868-9\\_2](https://doi.org/10.1007/978-3-642-79868-9_2)
- [7] A. S. Lien, Y. D. Jiang, Integration of diffusion of innovation theory into diabetes care. *Journal of diabetes investigation*, 8(3), 259-260. <https://doi.org/10.1111/jdi.12568>
- [8] R. Agarwal, Individual acceptance of information technologies. *Framing the domains of IT management: Projecting the future through the past*, 85-104, 2000.
- [9] P. Nemutanzhela, T. Iyamu, "Theory of diffusion of innovation for analysis in information systems studies," *2015 Science and Information Conference (SAI)*, London, 2015, 603-608, doi: 10.1109/SAI.2015.7237205.
- [10] B. K. Daigle, Jr L.G. Kent, W. T. Daniell, W. H. Eason, S. N. Zellner, J. C. Liu, R. A. Koch, "Electronic message translations accompanied by indications of translation." U.S. Patent 8,027,438, issued September 27, 2011.
- [11] E. J. Pérez-Stable, A. Nápoles-Springer, J. M. Miramontes, The effects of ethnicity and language on medical outcomes of patients with hypertension or diabetes. *Medical care*, 35(12), 1212-1219, 1997. doi: 10.1097/00005650-199712000-00005.
- [12] T. M. Tocher, E. Larson, Quality of diabetes care for non-English-speaking patients. A comparative study. *Western Journal of Medicine*, 168(6), p.504, 1998. PMID: 1305066.
- [13] R. A. David, M. Rhee, The impact of language as a barrier to effective health care in an underserved urban Hispanic community. *Mount Sinai Journal of Medicine*, 1; 65:393-7, Oct 1998. PMID: 9844369.
- [14] P. K. Lee, C. N. Rosenberg, H. Tsao, A. J. Sober, Failure of Q-switched ruby laser to eradicate atypical-appearing solar lentigo: report of two cases. *Journal of the American Academy of Dermatology*, 38(2), 314-317, 1998. [https://doi.org/10.1016/S0190-9622\(98\)70572-9](https://doi.org/10.1016/S0190-9622(98)70572-9)
- [15] P. Newmark, *A textbook of translation*. Vol. 66. New York: Prentice hall, 1988.
- [16] N. L. Ruxwana, M. E. Herselman, D. P. Conradie, ICT applications as e-health solutions in rural healthcare in the Eastern Cape Province of South Africa. *Health information management journal*. 2010 Mar; 39(1):17-29. <https://doi.org/10.1177/183335831003900104>

- [17] W. A. Nord, Religion and American education: Rethinking a national dilemma. UNC Press Books, 2014.
- [18] C. Free, G. Phillips, L. Galli, L. Watson, L. Felix, P. Edwards, A. Haines, The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. *PLoS med*, 10(1), e1001362, 2013. <https://doi.org/10.1371/journal.pmed.1001362>
- [19] J. A. Effken, P. Abbott, Health IT-enabled care for underserved rural populations: The role of nursing. *Journal of the American Medical Informatics Association*. 2009 Jul 1; 16(4):439-45. <https://doi.org/10.1197/jamia.M2971>
- [20] G. M. Fix, T. P. Hogan, D. J. Amante, D. K. McInnes, K. M. Nazi, S. R. Simon, "Encouraging patient portal use in the patient-centered medical home: three stakeholder perspectives." *Journal of medical Internet research* 18, no. 11 (2016). doi:10.2196/jmir.6488
- [21] A. Dhai, The Life Esidimeni tragedy: Moral pathology and an ethical crisis. *South African Medical Journal*. 108(5):382-5, 2018. DOI: [10.7196/SAMJ.2018.v108i5.13232](https://doi.org/10.7196/SAMJ.2018.v108i5.13232)
- [22] L. Gleitman, A. Papafragou, New perspectives on language and thought. *The Oxford handbook of thinking and reasoning*. 2012; 2:543-68.
- [23] S. Iftikhar, A. M. Ishaq, H. F. Ahmad, K. Fatima, "Introducing semantics in DHTs for Grid services in a semantic registry." In 2010 6th International Conference on Emerging Technologies (ICET), 382-387. IEEE, 2010. DOI: [10.1109/ICET.2010.5638457](https://doi.org/10.1109/ICET.2010.5638457)
- [24] J. J. Vosloo, A sport management programme for educator training in accordance with the diverse needs of South African schools (Doctoral dissertation), 2014.
- [25] I. van Zyl, Disciplinary kingdoms: Navigating the politics of research philosophy in the information systems. *The Electronic Journal of Information Systems in Developing Countries*. 70(1), 1-7 Sep 2015. <https://doi.org/10.1002/j.1681-4835.2015.tb00501.x>
- [26] R. K. Yin, Validity and generalization in future case study evaluations. *Evaluation*. 2013 Jul; 19(3):321-32. <https://doi.org/10.1177/1356389013497081>
- [27] J. Sutton, Z. Austin, Qualitative research: Data collection, analysis, and management. *The Canadian journal of hospital pharmacy*. 68(3):226, May 2015. doi: 10.4212/cjhp.v68i3.1456
- [28] A. Fabijan, H. H. Olsson, J. Bosch, "Customer feedback and data collection techniques in software R&D: a literature review." In International Conference of Software Business, 139-153. Springer, Cham, 2015. [https://doi.org/10.1007/978-3-319-19593-3\\_12](https://doi.org/10.1007/978-3-319-19593-3_12)
- [29] P. Dutta, Communication of Innovations for Freedom. *International Journal of Research in Humanities, Arts and Literature*, 2(12), 15-32, 2014. [https://www.academia.edu/21193892/communication\\_of\\_innovations\\_for\\_freedom](https://www.academia.edu/21193892/communication_of_innovations_for_freedom)
- [30] T. G. Striphas, *Communication as...: Perspectives on Theory*. Sage; 2006.
- [31] E. M. Rogers, *Diffusion of innovations*. 5th ed. New York: Free Press, 2003.

## Exploring the Performance Characteristics of the Naïve Bayes Classifier in the Sentiment Analysis of an Airline's Social Media Data

Mba Obasi Odim\*, Adewale Opeoluwa Ogunde, Bosede Oyenike Oguntunde, Samuel Ayodele Phillips

Department of Computer Science, Redeemer's University, Ede 23210, Nigeria

### ARTICLE INFO

Article history:

Received: 20 April, 2020

Accepted: 08 July, 2020

Online: 28 July, 2020

Keywords:

Airline image branding

Naïve Bayes

Sentiment analysis

### ABSTRACT

Airline operators get much feedback from their customers which are vital for both operational and strategic planning. Social media has become one of the most popular platforms for obtaining such feedback. However, to analyze, categorize, and generate useful insight from the huge quantity of data on social media is not a trivial task. This study investigates the capability of the Naïve Bayes classifier for analyzing sentiments of airline image branding. It further examines the impact of data size on the accuracy of the classifier. We collected data about some online conversations relating to an incident where an airline's security operatives roughly handled a passenger as a case study. It was reported that the incident resulted in a loss of about \$1 billion of the company's corporate value. Data were extracted from twitter, preprocessed and analyzed using the Naïve Bayes Classifier. The findings showed a 62.53% negative and 37.47% positive sentiments about the incident with a classification accuracy of over 0.97. To assess the impact of training size on the accuracy of the classifier, the training sets were varied into different sizes. A direct linear relationship between the training size and the classifier's accuracy was observed. This implies that large training data sets have the potentials for increasing the classification accuracy of the classifier. However, it was also observed that a continuous increase in the classification size could lead to overfitting. Hence there is a need to develop mechanisms for determining optimum training size for finest accuracy of the classifier. The negative perceptions of customers could have a damaging effect on a brand and ultimately lead to a catastrophic loss in the organization.

### 1. Introduction

The evaluation of customers' perception of services offered by a firm is crucial to continuous patronage and subsequently, the growth of the business of the firm. How a customer feels about the service of a company over time is the sum of conscious events which is seen as a coordinated series of interactions between the customer and the brand of the service or product of the company. Thus, a customer's perception is entirely his/her personal view about the brand of service [1]. The major ways of capturing customers feedback of service or product include customer contact forms, customer feedback surveys (offline or online), social listening and interviews. Social listening is fast becoming the most prominent among the listed approaches.

Businesses can engage in social listening through social media platforms such as Twitter, Facebook, YouTube, Blog, and Email. Examples of brands that use Twitter to engage with their customers include Starbucks, PlayStation, Samsung Mobile US, Sony, Whole

Foods, MacDonald's, Amazon, JetBlue Airways amidst others. Social media for sentiments are probably the most trusted medium through which people express their view about various public issues compared to other media. This may be because most brands get talked about on social media by their customers, and such brands can obtain feedback from their customers without them necessarily knowing that such data are collected about them. Sentiment analysis is one of the ways of scrutinizing people's perception. In [2], it was described as an area of study concerned with examining the people's views, perception, assessments, judgements, attitudes and emotions towards the services, concerns, topical issues, just to mention a few. It was described in [3], as "contextual mining of text which recognizes and extracts subjective information in the source material to help an enterprise understand the societal perception of their brand, product or service through monitoring online conversations. It has become a major area of natural language processing (NLP). Its major task is the classification of texts-based conversations to obtain insight into the intention of the author of the text [4].

\*Corresponding Author: Mba Obasi Odim, Department of Computer Science, Redeemer's University, Ede, Nigeria +234-803-716-751, odimm@run.edu.ng

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050433>

New York Times [5] reported an incident that involved three security operatives and a sergeant who roughly handed and violently forced out a passenger of an Airlines Flight 3411 on April 9, 2017, from boarding the flight. It was further reported that the airline staff provided false information and consciously removed material facts from their reports. Consequently, the market value of the parent company of the airline lost about \$1 billion. Could the incident had contributed to the lost? What was people's perception of the incident? These and more are the questions we seek to answer in this work by analyzing peoples' comments on social media using artificial intelligence. Advances in artificial intelligence, particularly, Natural Language Processing, has considerably improved the ability of algorithms to analyze text. This study, therefore, explores the performance characteristics and the effect of data size on using a Naïve Bayes classification model to analyze people's sentiment on the social media data of a company.

## **2. Related Works**

There have been a number of research efforts ongoing on sentiments analysis in recent times through social media platforms. In this section, we reviewed selected literature. Hand-crafted and automatic models for removing factual or neutral comments that have no sentiments attached were reviewed in [6]. It was discovered that hand-crafted models perform well with strong sentiments but could not identify weak sentiments contents. Deep learning an automatic technique provides a meta-level feature illustration which generalizes new domains and languages. The multi-modal techniques merge audio and video social data with text using multiple kernels. a representation that generalizes well on new domains and languages. Multi-modal methods can combine the abundant audio and video forms of social data with text using multiple kernels. The high dimensionality of n-gram features and temporal nature of sentiments in long product reviews are identified as major challenges in sentiments mining.

In [4], Naive Bayes (NB) and Support Vector Machine (SVM) were implemented to classify sentiment analysis for movie reviews. In the study, SVM classifier outperformed the NB classifier in forecasting the sentiment of a review. Other supervised learning classification techniques such as stochastic gradient classifier, K-nearest neighbour, maximum entropy classifiers were suggested for future study. The classification of a Twitter sentiment review using the average semantic orientation of phrases containing adverbs and adjectives were conducted in [7]. Phrases with the good association are said to have positive semantic orientation while those with the bad association have a negative semantic orientation. The semantic orientation of a phrase is calculated as the difference of the mutual information between the given phrase and the word "excellent" and the mutual information between the given phrase and the word "poor". Experiments revealed that the proposed techniques outperformed the existing methods. The performance of Naive Bayes, Max Entropy, and Support Vector Machine (SVM), on twitter data streams were compared in [8]. SVM and naïve Bayes showed superior performance in terms of high accuracy and thus could be regarded as baseline learning methods. The review of mobile user' sentiments was analyzed in [9], statistical analysis was carried out on a large amount of real data, and four characteristics that distinguish phone reviews from that of PCs were identified,

namely: Short average length; Large span of length; Power-law distribution and significant difference in polarity of the mobile application reviews. The results further showed that for classification, the Bayesian algorithm outperformed the SVM algorithm. The extent to which sentiment analysis techniques could provide more useful new insights than those of traditional quality assessment methods was investigated in [10], using 4392 tweets dataset. Sentiment analysis techniques identified 23 attributes that could be used for comparison with other American Society For Quality (ASQ) scales. Results indicated that the rate at which passengers respond to the attributes of scale differed greatly in certain instances and that identifying these differences could provide insights for airport management to improve the airport service quality.

The performances of back propagation neural network (BPN), probabilistic neural network (PNN) and a homogeneous ensemble of PNN (HEN) were compared in [11], using varying levels of word granularity as features for feature level sentiment classification. Product review dataset collected from Amazon reviews websites was used to validate the algorithms. Performance analysis on the results of the ANN-based methods with two individual statistical methods was carried out using five quality parameters. The results showed that the homogeneous ensemble of the neural network method offers superior performance. In addition, probabilistic neural networks (PNNs) outperformed BPN during classification, and the combination of neural network-based sentiment classification methods with principal component analysis (PCA) as a feature reduction method offered superior performance in terms of training time. A hybrid approach for identifying product reviews offered by Amazon was proposed in [12]. The results showed that the hybrid approach outperformed the individual classifiers (Random Forest and Support Vector Machine) in this amazon dataset. Various sentiment analysis techniques and their performance at analyzing sentiments were presented in [13]. The study was aimed at developing a sentiment analysis technique that would categorize various reviews efficiently. The work described machine learning techniques like SVM, NB, maximum Entropy and other techniques that could improve the analysis process. They suggested the n-gram evaluation of semantic analysis instead of word by word analysis.

An evaluation of features on sentiment analysis carried out in [14] showed that that unigram was the best method to extract sentiment from the review. Unigram with stemming with stop-work gave 82.9% accuracy, and unigram with steaming without stop word gave 83% accuracy in positive class, while unigram with stemming and without stop word gave better accuracy of 83.1% in the negative class. The two classes offered improved performance with information gain. The authors suggested an ensemble feature selection technique to perform an additional experiment for further studies.

In [15], a sentiment analysis system was developed for a resource-poor language, namely Roman Urdu. The authors noted that most works on with sentiment analysis have been on the resource-rich languages of the world, while very few had been done on resource-poor languages. The work involved four different studies based on word-level, character level features, and feature union. The results showed that the error rate could be reduced by 12% from the baseline (80.07%), and the results of studies were statistically significant from the baseline.



The reviewed works show that sentiments analysis is an active research area in computing with diverse classification approaches that seek to provide an effective and efficient mechanism for analyzing sentiments. It was particularly noted that the popularity of machine learning techniques has been on the increase. However, little or no effort has been put in place to examine the impact of the data size on the classification performance of the various algorithms. We have, therefore, considered the relationship between data size and performance, using one of the highly-rated machine learning algorithms in the literature, the naïve Bayes, in the sentiment analysis of an airline's social media data.

### 3. Methodology

This section presents the sentiment classification process and methods applied in this study. A summary process is presented in Figure 1, and a description of methods follow. Tweets were extracted from Twitter using Python libraries, pyQuery, in the form of a JSON response, a process known as Web scraping, to a flat file. The data were loaded into the data frame for further processing. Figure 1 shows the sentiments classification process.

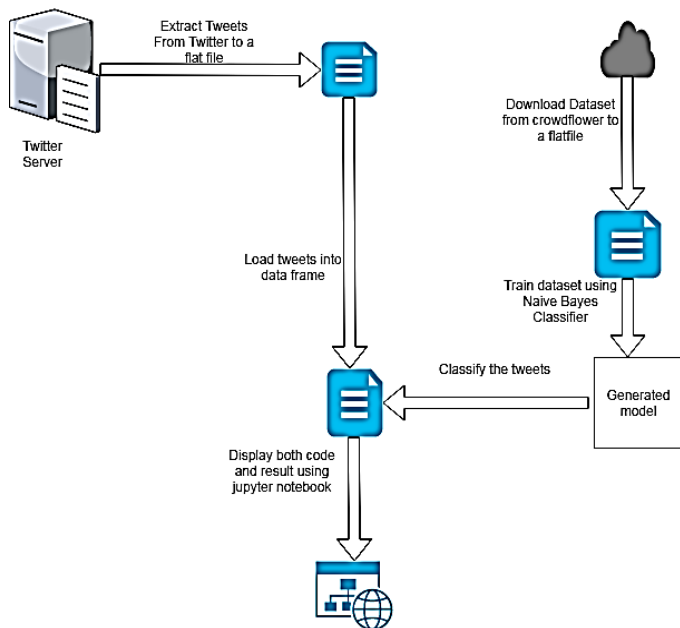


Figure 1: The Sentiments' Classification Process

Datasets for the training of the classifier were downloaded from crowdFlower.com, which contains a large amount of already labelled tweets texts in the areas as airline sentiments, hotel review, product reviews, and many other datasets, classified already into positive, negative or neutral. The extracted labelled airline tweets from crowdflower.com were used as a training set for the classifier and new downloaded tweets' datasets used for testing the classification accuracy.

#### 3.1. Data Collection and Description

Data collection was achieved using “forked” Tweet API, Jsoup library. The standard Tweet API allows retrieval of tweets for a maximum seven days lag time and could only permit for scraping 18,000 tweets per 15minute window”

(<https://towardsdatascience.com/how-to-scrape-tweets-from-twitter-59287e20f0f1>). However, the tweets used for this study went live between April and May 2017. Due to this restriction, a “forked program” available on GitHub (<https://github.com/Jefferson-Henrique/GetOldTweets-java>) was used instead to collect the data. The extracted data were saved in some files. Jsoup is a Java library. It works with data in HTML format. It scrapes and parses HTML from a URL, file or string using the WHATWG HTML5 specification. Forty thousand rows of tweets with the keyword “United Airlines” was scraped from the 12th May 2017 to 31st May 2017. This was done by querying Tweet with the command in the “forked” Tweet API: “java -jar got.jar QuerySearch=“United Airlines” since=2017-05-12 until=2017-05-31 maxtweets=40000”.

#### 3.2. Text Preparation and Cleaning

Stop word, that is words that do not influence the direction of the sentiments, were filtered out using the Natural Language Toolkit’s (NLTK) stop word corpus for the English language [7]. Stop words are words which are filtered out before or after processing of natural language data (text). Common stop words include ‘is’, ‘are’, ‘that’, ‘a’, ‘and’, ‘but’, ‘how’, ‘what’, ‘or’, ‘about’, ‘after’, ‘been’, ‘being’, ‘who’, ‘why’, ‘will’, ‘without’, ‘you’, ‘you’re’, ‘made’, ‘make’, ‘had’, ‘has’, ‘happen’, and so on.

##### 3.2.1. Filtering

Filtering was further carried out to eliminate the inclusion of metadata and some other irrelevant texts while classifying sentiments. The texts, such as Identity numbers, date, time, Irrelevant tags, Hyperlinks, #tags, punctuation, special characters, etc. were removed.

##### 3.2.2. Feature Extraction

Several methods could be used to classify texts as positive, negative or neutral. One of such is a subset of adjectives, which are manually classified as positives and a subset classified as negative are used as seed words [7]. Equation (1) describes such a method.

Suppose a seed set size is denoted respectively as  $ADJ_p$  and  $ADJ_n$  for positive and negative sets seed sizes. Let  $L(W_i, POS_j)$  represents a modified log-likelihood ratio for a  $W_i$  with part of speech  $POS_j$  as the ratio of its collocation frequency with  $ADJ_p$  and  $ADJ_n$  within a sentence. Then  $L(W_i, POS_j)$  can be computed as follows:

$$L(W_i, POS_j)^1 = \frac{\left( \frac{\text{Freq}(W_i, POS_j, ADJ_p) + \epsilon}{\text{Freq}(Wall, POS_j, ADJ_p)} \right)}{\frac{\text{Freq}(W_i, POS_j, ADJ_n) + \epsilon}{\text{Freq}(Wall, POS_j, ADJ_n)}} \quad (1)$$

where  $(Wall, POS_j, ADJ_p)$  denotes the collocation frequency of all words  $Wall$  of parts speech  $POS_j$  with adjective  $ADJ_p$  and  $\epsilon$  is smoothing constant.  $Wall, POS_j, ADJ_n$  is similarly defined for  $ADJ_n$ . Brill’s tagger was used in [7] to obtain part of speech information. In this study, the datasets for the training of the classifier were downloaded from CrowdFlower.com. The site

houses many datasets in areas as airline sentiments, hotel review, product reviews, and many other datasets. The dataset contains tweets already classified into positive, negative and neutral based on a cloud of words that shows that a new tweet should be classified as either positive or negative or neutral.

### 3.2.3. Training and Test Datasets

The training dataset contained tagged data from crowdflower.com, in which the tweets were already labelled as positive or negative tweets. df dataframe (which contains data loaded from airline\_sentiment\_main.csv) was split into training (80%) and test datasets (20%). df\_train contains the dataframe of the training datasets while df\_test is the dataset of the testing data. To begin training, the tweets with positive sentiments were copied into the dataframe df\_pos. The same was done for negative (df\_neg) and neutral sentiments (df\_neu). The dataframe df\_train was passed into the naïve Bayes (TweetNBClassifier) for the classification.

### 3.3. Naïve Bayes Classifier

The Bayes classification algorithm for text mining is based on the Bayesian rule, which defines the conditional probability that allows for twisting the condition in a convenient way. Given events X and Y, we compute that conditional probability, P(X|Y), of an event X on the condition that event Y had occurred as in equation (2) defined in [7] as:

$$P(X|Y) = \frac{P(X \cap Y)}{P(Y)} \quad (2)$$

where P(X|Y) denotes the Probability of X, given Y has occurred, X∩Y is the probability of the intersection of event X and Y, while P(Y) is the probability of event B occurring.

Applying the multiplicative law,

$$P(X \cap Y) = P(X|Y) \times P(Y) \quad (3)$$

Based on the multiplication rule, P(X∩Y)=P(X|Y)×P(Y), so P(Y∩X)=P(Y|X)×P(X) (4)

We applied Bayes' Law (Bayes' Rule or Bayes' Theorem) to help us understand the relationship between two events by computing the different conditional probabilities as follows: Since

P(X∩Y)=P(Y∩X), we substitute P(Y|X)×P(X) for P(X∩Y) in the conditional probability formula  $P(X|Y) = \frac{P(X \cap Y)}{P(Y)}$ . Thus

$$P(X|Y) = \frac{P(Y|X) \times P(X)}{P(Y)} \quad (5)$$

Hence, P(X|Y) can be computed more conveniently, applying the Bayes rule by equation (5) [13, 7].

The rule can be redefined with regard to the probability of a document occurring given it has been predetermined to be positive or negative since we have examples of positive and negative sentiments from our data set. Thus, we have

$$P(\text{Sentiment}|\text{Sentence}) = \frac{P(\text{Sentence}|\text{Sentiment}) \times P(\text{Sentiment})}{P(\text{Sentence})} \quad (6)$$

We computed P(Sentence|Sentiment) as a product of P(token|Sentiment) as in [7], and estimated P(token|Sentiment) as

$$P(\text{token}|\text{Sentiment}) = \frac{\text{Count}(\text{thistoken} \in \text{class}) + 1}{\text{Count}(\text{alltoken} \in \text{class}) + \text{Count}(\text{alltokens})} \quad (7)$$

The one and count of all tokens are referred to as add one or Laplace smoothing. The naïve Bayes Classifier chooses the most probable classification Vnb for a given set of attribute values a1, a2, ..., an [13, 15].

### 3.4. Measure of Accuracy

While scanning through the list of sentiments features, if the value of the feature (the predicted value) is equal to the actual value (target), the value 'correct' is appended to the list; otherwise, incorrect value is appended. Hence the classification accuracy was computed by equation 8.

$$\text{accuracy} = \frac{\text{correctprediction}}{\text{correctprediction} + \text{incorrectprediction}} \quad (8)$$

## 4. Results and Discussion

Some tweets with the number of both positive and negative sentiments are shown in Figure 2. There was a regular pattern of the sentiments. Negative sentiments were consistently higher in value than positive sentiments for each day. This suggests that the majority of the tweets oppose the action of the airline operatives on the passenger.

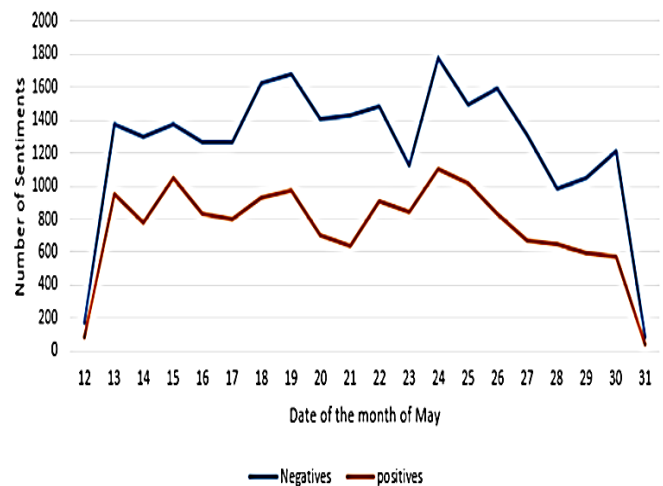


Figure 2: A screenshot segments of Tweets' Sentiments

Figure 3 shows a screenshot the code segment for loading the training data from the CSV file "Airline-Sentiment-2-main.csv" into the Dataframe called df to train Naïve Bayes classifier.

```
tweets_data_path='Airline-Sentiment-2-main.csv'
tweets = pd.read_csv(tweets_data_path, header=0)
df = tweets.copy()[['airline_sentiment', 'text']]
```

Figure 3: A snapshot of Code Segment for Loading Training Data

An extract of sample classification of ten tweets of the airline is depicted in Table 1. The tweets have been classified as either positive or negative sentiments.

Table 1: A sample output extract classification of ten tweets

Refined_Date	Refined_Tweets	Sentiments
2017-05-31	apologize lack compassion effective service presented today potential compensation urisite ad	Negative
2017-05-31	go back line near end fight announced room bags checked.	Negative
2017-05-31	clearly fit carry weighted lot less lbs contnum please get together	Negative
2017-05-31	also rude trying board early group contnum even though informed im disabled	Negative
2017-05-31	lets guess happen	Negative
2017-05-31	physically cannot lift remove carousel woman checking passengers informed flights attendance cant	negative
2017-05-31	yes appreciate flying experiences	Positive
2017-05-31	unitedairlines yet disrespecting customer refusing find take responsibilities delay caused	Negative
2017-05-31	please file format complaint via customer care well urisite ad	Negative
2017-05-31	way north long weather delays jetblue still got day offered food voucher without tweeting	negative

```
Mood = dataframeRefined['Sentiments'].value_counts()
Mood
negative    25013
positive    14986
Name: Sentiments, dtype: int64
```

Figure 4: A snapshot code for mood count

It was observed that the number of negatives, 25,013, outweighs the number of positives, 14,986. The count of moods was computed by counting the number of disparate values in the 'Sentiments' column of dataframeRefined.

A snapshot for the code of the mood distribution is shown in Figure 5, while the resultant graphical distribution of the mood follows in Figure 6.

```
def score(self, feature, target):
    compare = []
    for i in range(0, len(feature)):
        if feature[i] == target[i]:
            tmp = 'correct'
            compare.append(tmp)
        else:
            tmp = 'incorrect'
            compare.append(tmp)
    r = Counter(compare)
    accuracy = r['correct']/(r['correct']+r['incorrect'])
    return accuracy
```

Figure 5: A snapshot code for the mood distribution

There were more negative moods than positive. The first bar denotes nnegative Sentiments, while the second denotes positive moods. Again, the number of negative moods outweighed the positive. There were about 25,000 negative and 15,000 positive sentiments, respectively recorded.

```
#to draw the graph
index = [1,2]
plt.bar(index, Mood, color=['r', 'g'])
plt.xticks(index, ['Negative', 'Positive'])
plt.xlabel('Mood')
plt.ylabel('Mood Count')
plt.title('Mood Distribution')
```

Figure 6: Mood distribution of sentiments

Figure 7 shows a screenshot of the code segment that defined the method, called "score" for computing the classification accuracy using equation (8)..

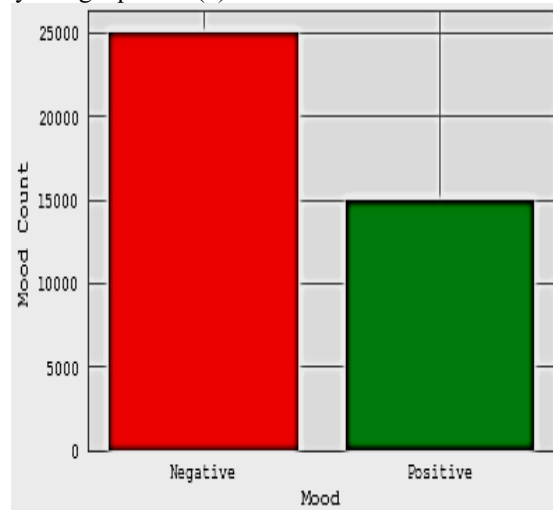


Figure 7: Code segment of the definition of the method "score".

To measure the accuracy of the classification on the test data sets, the following code segment shown in Figure 8 was used to call the method score for the computation of the classification accuracy. Accuracy of 0.9 was recorded based on 11,000 training datasets.

Figure 8 shows code segment for computation of classification accuracy calling the method in Figure 7.

```
tnb = TweetNBClassifier(df_train)
tnb = tnb.fit()
predict = tnb.predict(df_test)
score = tnb.score(predict,df_test.airline_sentiment.tolist())
print (score)
```

C:\Users\PHILLIPS\Anaconda2\lib\site-packages\ipykernel\_launcher.py:55: RuntimeWarning:   
in log

0.9

Figure 8 shows A code segment for computation of classification accuracy

The test result for an 11,000-training dataset produced an accuracy of 0.9. To measure the impact of training data size on the classification accuracy, various training dataset sizes were used experimented, and a subset of the result is depicted in Table 2.

Table 2: Classification of accuracy measure of various training datasets

Training Size	Accuracy (%)
3000	57
5000	68
7000	75
9000	80
11000	90
13000	97

Figure 9 showed a direct linear relationship between training size and classification accuracy. We also observed that the accuracy increases in direct proportion to the training size. It further revealed, however, that at a particular point, overfitting arose as we continued to increase the training size.

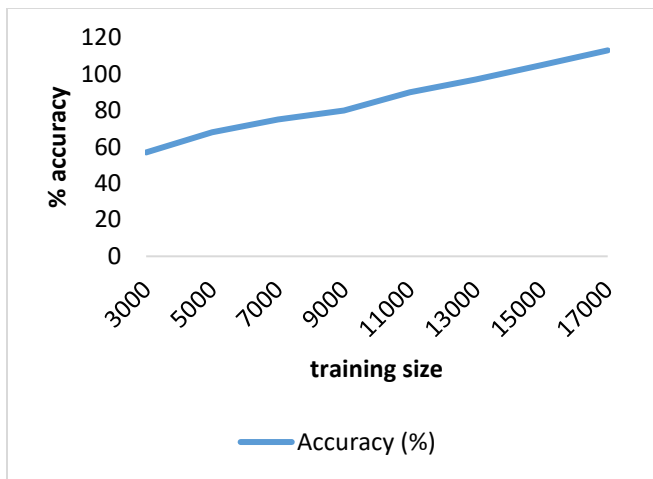


Figure 9: Relationship between training size and classification accuracy

This finding agrees with a number of machine learning techniques in the literature [16, 17]. Therefore, some experimentation with various data sizes may be required to determine optimal training size.

**5. Conclusion**

The study explored the Naïve Bayes classifier for analyzing airline image branding, a public relations crisis in which a customer was dragged off an airplane. The finding revealed 62.53%

negative sentiments and 37.47% positive perceptions of people about the image of the airline. The negative perceptions from customers feedback on the case study could lead to negative image branding and subsequently to a loss of patronage to the organization. Furthermore, experimenting the classification accuracy with a number of training sizes showed a direct linear relationship between training size and its accuracy. It was also observed that a continuous increase in the classification size could lead to overfitting. These findings are supported by the features of machine learning techniques as obtained in [16, 17]. Future works will consider the development of mechanisms for the optimal determination of the training size.

**Conflict of Interest**

The authors declare no conflict of interest.

**Acknowledgement**

The authors wish to acknowledge the Department of Computer Science, Redeemer’s University, Ede, Nigeria, for providing the laboratory facilities used for this study. No funding was received.

**References**

- [1] P. Greenberg, CRM at the Speed of Light, Fourth Edition: Social CRM 2.0 Strategies, Tools, and Techniques for Engaging Your Customers, 4th ed., New York City: McGraw-Hill Education, 2009.
- [2] B. Liu, Sentiment Analysis and Opinion Mining, Virginia: Morgan & Claypool Publishers, 2012.
- [3] S. Gupta, "Sentiment Analysis: Concept, Analysis and Applications," 2018.
- [4] A. Tripathy, A. Agrawal and S. K. Rath, "Classification of Sentimental Reviews Using Machine Learning Techniques," Procedia Computer Science, **57**, 821 – 829, 2015. <https://doi.org/10.1016/j.procs.2015.07.523>
- [5] M. Salam, "Security Officers Fired for United Airlines Dragging Episode," The New York Times, 17 October 2017.
- [6] I. Chaturvedi, E. Cambria, R. E. Welsch and F. Herrera, "Distinguishing between facts and opinions for sentiment analysis: Survey and Challenges," Information Fusion, **44**, 65-77, 2018. <https://doi.org/10.1016/j.inffus.2017.12.006>
- [7] A. Tamilselvi and M. ParveenTaj, "Sentiment Analysis of Microblogs using Opinion Mining Classification Algorithm," International Journal of Science and Research, **2**(10), 196 - 202, 2013.
- [8] V. A. Kharde and S. S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of," International Journal of Computer Applications, **139**(11), 5-15, 2016. <https://doi.org/10.5120/ijca2016908625>.
- [9] L. Zhanga, K. Huac, H. Wangd, G. Qiane and L. Zhanga, "Sentiment Analysis on Reviews of Mobile Users," Procedia Computer Science, **34**, 458 - 465, 2014. <https://doi.org/10.1016/j.procs.2014.07.013>
- [10] L. Martin-Domingoa, J. C. Martínb and G. Mandsberg, "Social media as a resource for sentiment analysis of Airport Service Quality (ASQ)," Journal of Air Transport Management, **78**, 106-115, 2019. <https://doi.org/10.1016/j.jairtraman.2019.01.004>
- [11] G. Vinodhini and R. Chandrasekaran, "A comparative performance evaluation of neural network-based approach for sentiment classification of online reviews," Journal of King Saud University – Computer and Information Sciences, **28**, 2–12, 2016. <https://doi.org/10.1016/j.jksuci.2014.03.024>
- [12] Y. AL Amrani, M. Lazaar and K. E. EL Kadiri, "Random Forest and Support Vector Machine based Hybrid Approach to Sentiment Analysis," Procedia Computer Science, p. 511–520, 2018. <https://doi.org/10.1016/j.procs.2018.01.150>
- [13] M. D. Devika, C. Sunitha and A. Ganesh, "Sentiment Analysis: A Comparative Study on Different Approaches," Procedia Computer Science, **87**, 2016. <https://doi.org/10.1016/j.procs.2016.05.124>



- [14] P. Shahana and B. Omman, "Evaluation of Features on Sentimental Analysis," *Procedia Computer Science*, **46**(2015), 1585 – 1592, 2015. <https://doi.org/10.1016/j.procs.2015.02.088>
- [15] K. Mehmood, D. Essam, K. Shafi and M. K. Malik, "Sentiment Analysis for a Resource Poor Language—Roman Urdu," *ACM Transactions on Asian and Low-Resource Language Information Processing*, **19**(1), 10.1-10.15, 2019. <https://doi.org/10.1145/3329709>
- [16] M. O. Odim and V. C. Osamor, "Required Bandwidth Capacity Estimation Scheme for Improved Internet Service Delivery: A Machine Learning Approach," *International Journal of Scientific & Technology Research*, **8**(8), 326 - 334, 2019.
- [17] M. O. Odim, J. A. Gbadeyan and J. S. Sadiku, "Modelling the Multi-Layer Artificial Neural Network for Internet Traffic Forecasting: The Model Selection Design Issues," in *ACM Computing Research and Innovations (CoRI 2016)*, Ibadan, 2016.

## Coastal Risk Modelling for Oil Spill in The Mediterranean Sea

Abdellatif Soussi<sup>\*1,2</sup>, Chiara Bersani<sup>1</sup>, Roberto Sacile<sup>3</sup>, Dounia Bouchta<sup>2</sup>, Ahmed El Amarti<sup>2</sup>, Hamid Seghioeur<sup>4</sup>, Driss Nachite<sup>5</sup>, Jaouad Al Miys<sup>6</sup>

<sup>1</sup>Italian Centre of Excellence on Logistics, Transport and Infrastructures (CIELI), University of Genova, Genova, 16126, Italy

<sup>2</sup>Laboratory of Materials and Interfacial Systems Faculty of Sciences, UAE Tétouan, 93000, Morocco

<sup>3</sup>Department on Informatics, Bioengineering, Robotics and System Engineering. University of Genova, Genova 16145, Italy

<sup>4</sup>Laboratory MOSIL, National School of Applied Sciences, UAE Tétouan, 93000, Morocco

<sup>5</sup>Department of Geology, Faculty of Sciences, UAE Tétouan, 93000, Morocco

<sup>6</sup>Horizon Tangier Terminal SA- Tangier Med Port, Tangier, 90053, Morocco

### ARTICLE INFO

Article history:

Received: 06 May, 2020

Accepted: 07 July, 2020

Online: 28 July, 2020

Keywords:

Maritime Risk

Oil spill model

### ABSTRACT

The accident probability estimation and the consequence analysis are based on statistical data about oil spill accident occurrence in the Mediterranean area, on the probability of different release sizes, and on the joint probability of wind speed and directions.

The risk model and its evaluation have been assessed for the Mediterranean littoral considering the time required by the oil slick to hit the coast in specific sensible target points assuming that an oil release accident potentially occurred in an accident sites located along the ship routes.

This approach has been applied on the area of the Strait of Gibraltar, which supports a significant volume of maritime traffic because it represents the navigational connection channel between the Atlantic Ocean and the Mediterranean Sea.

## 1. Introduction

The environmental risk analysis is increasingly used in the sector of maritime life, and many studies have been conducted in the framework of the maritime risk assessment with special attention to the hazardous material transportation [1,2]. Risk can be defined as a measure of an accident occurrence potentiality as well as its gravity [3]. In fact, risk analysis is widely acknowledged as a process for depicting risk systematically and scientifically [4,5]. The main objective of risk analysis is to prohibit the occurrence of accidents [1]. In order to define high-risk areas, it is necessary to quantify both the absolute level of risk and the relative significance of the different causes [6,7].

In recent decades, several methods and applications have been reported in the literature for maritime transport risk analysis. These approaches attracted growing interest both from international organizations that have suggested the use of specific risk analysis and management tools [8,9], and from researchers who focused on fundamental issues relevant to risk assessment. From the academic

viewpoint, the main topics referred to terminology, concepts and perspectives of risk analysis and management in the maritime context [10–25].

In [26], the authors proposed an environmental approach to assess the risk in a maritime area. In [27], the author presented a method to quantify the uncertainty related to traffic data in maritime risk assessment. The work in [6] defined the framework for risk analysis in a maritime area through a case study of RoPax vessels. In [28], the authors applied the so-called FMEA, failure mode and effects analysis method, for the risk of ship collision in a maritime area. The authors in [29] and [30] quantified the effect of risk reduction measures for the shipping in a waterway area. In [31], the researchers determined the relative risk of various coastal areas. Some studies determine the probability and consequences of a shipping accident [32, 33].

In the literature, specific works focused on the oil spill modelling. In [34], the authors developed a Lagrangian model to determine the trajectory of the oil spill on the Patos Lagoon in Brazil, considering factors influencing the region such as coastal ocean

\*Corresponding Author: Abdellatif Soussi, Email : [abdellatifsoussi90@gmail.com](mailto:abdellatifsoussi90@gmail.com)

currents, wind currents, tides and river flows. Another similar study proposed in [35] identified sensitive areas in a bay in the event of a hazardous substance spill accident. The work in [36] presented a model based on a Lagrangian approach to identify the areas most exposed to pollution risk in the Baltic Sea, Finland. In [37], a study has been carried out to protect the marine environment in the Bohai Bay region from spill accidents. This study simulated oil spill transport and fate in sea based on the particle approach.

Other studies examined oil spill modelling tools. The authors in [38] analyzed two tools for the deepwater region of the Campos Basin, Brazil: the General NOAA Operational Modeling Environment (GNOME) used by the Emergency Response Division of the Office of Response and Restoration (OR&R) and the Python Operational Ocean Forecasting Engine (OOFE).

Other approaches exist to estimate oil slick weathering. MEDSLIK predicts the oil spill on the surface of the Mediterranean Sea [39]. POSEIDON-OSM simulates evaporation, emulsification and sedimentation processes. It is used for the Greek Sea [40], the Baltic Sea [41], and the Aegean Sea [42,43]. ADIOS (Automated Data Inquiry for Oil Spills) is NOAA's oil weathering model. ADIOS models how different oil weathering processes occur in the marine environment. It has been used by many authors [44,45].

The oil spill is one of the most dangerous sources of pollution that threaten maritime safety because of its serious consequences to the ocean environment and the ecosystem as well as enormous economic losses and society impact [46–52].

A released petroleum product is subject to the effects of the environment which generates its dispersion in the marine environment and, simultaneously, it modifies its physical and chemical characteristics, the so-called "weathering" of the oil [53]. The behavior of oil drift at sea is the result of a set of interactions that occur between the spilled product and the external environment conditions [54]. When hydrocarbons are discharged at sea, they suffer a large number of transformation processes: drift and spreading, evaporation, dissolution, dispersion, emulsification, photo-oxidation, biodegradation, sedimentation, pouring, stranding and interaction with sea ice. While some processes are currently well-understood, such as spreading and evaporation, others remain poorly known (photooxidation and biodegradation) [54–57].

The risk of oil spill pollution in the Mediterranean Sea is high due to the significant traffic of oil and gas [58], where are listed more than 100 million gallons of crude oil spilled annually [59,60]. Statistically, 52 % of total oil spills in the Mediterranean come from shipping, compared to 48 % for other seas [61].

In case of an oil spill accident in the marine environment, it is mandatory to know the trajectory (movement and spreading) of the pollutant slicks under various weather conditions, so as to organize the oil recovery operations and to protect the areas exposed to the risk of pollution [62,63].

This paper provides an approach to rapid mapping for the analysis of the risk addressing the accidental of maritime transportation in a Strait of Gibraltar in the Mediterranean Sea. The purpose of the proposed paper is twofold. Firstly, it aims at defining a simple

methodology to classify the risk in marine and coastal areas due to maritime hazardous material transportation. Secondly, the proposed approach provides a useful tool that can support spill response teams and other operators in facilitating oil spill planning and preparedness.

The following section of the paper represents a review of the main shipping accidents which generated massive oil spills in the Mediterranean Sea. The third part introduces the proposed oil spill model, where a Lagrangian model has been proposed to identify risk areas that could be affected in the occurrence of a spill accident. The proposed model was applied to the real case occurred in the Mediterranean Sea toward the French coast in October 07, 2018, and the results have confirmed the reliability and relevance of the proposed model. In section 4, the application of the proposed model is described in the context of a potential maritime accident in the Strait of Gibraltar in the Mediterranean Sea. In section 5, the potential environmental risk was assessed on the basis of the time required by the oil slick to reach the coasts in the Strait, lastly, in section 6, conclusions.

## 2. State of art in the Mediterranean Sea

The Mediterranean Sea is a tragic theater of maritime accidents. In 1991, the Haven disaster in the coastal area of Genoa in Italy, which produced the release of 144,000 tons of hydrocarbons, has been ranked as the fourth most dangerous event among global shipping accidents [64,65]. In addition, the Mediterranean is threatened by accidents occurring outside its geographical area. For example, the maritime accident that occurred in the Atlantic Ocean as a result of a collision between the oil tanker «Seat Spirit», which was carrying heavy oil, and «Hesperus», which transported chemical products, caused a spill of 12,200 tons of oil. However, depending on weather conditions (wind speed and ocean currents), the contaminants were transferred by the Strait of Gibraltar to the Moroccan, Spanish and Algerian coasts [66-68].

Table 1: Number of oil spill accidents in the Mediterranean Sea between 1977 and 2019 (REMPEC, 2019).

Years	Number of accidents	Type of pollutant	
		Volatile Oil	Non-volatile Oil
Between 1977 and 1987	46	3	43
Between 1988 and 1997	73	5	68
Between 1998 and 2007	28	3	25
Between 2008 and 2019	124	82	42

According to recent statistics from the “Alerts and Accidents database” (REMPEC) [69], containing data on spills (quantity, type of spilled oil, location and on the ships involved), 268 tankers were involved in maritime accidents in the Mediterranean Sea, 93 accidents for “Volatile Oil” and 178 for “non-Volatile oil”, as shown in the table 1.

The aforementioned statistics shows that, the number of accidents involving accidental spills decreased between 1998 and 2007, with 46 and 28 accidents occurred respectively. This reduction can be

attributed to the implementation of international, regional and national legislations and, precisely, at the level of European countries. The European Union has put in place a series of strict measures (Erika I and Erika II)[70,71] to control ships entering in the European ports. Also, the new technologies application in the shipbuilding industry improves the quality and safety of this mean of transport [61]. Nevertheless, between the years 2008 and 2019, the number of accidents increased again especially in Greece maritime area where the 90% of Mediterranean oil spill accident ensued. Greece holds the record for oil spilled in the main accidents, with 378,027 tons, followed immediately by Italy with almost 364,823 tons and Spain with 333,492 tons [69].

Accidents occurring in the European sea area are more frequent than in other parts of the Mediterranean: on 268 accidents in the Mediterranean Sea, 232 involved European Mediterranean countries. This can be attributed to the increased trade and traffic and to the presence of the petrochemical industries on the coast: in Italy, as an example, there are 14 oil ports and 17 refineries [61]. The high maritime traffic density between Gibraltar and Sicily reflects the importance of the Western Mediterranean as a transit zone [72,73].

2.1. Statistics about oil spills from tank vessels

Oil spills are generally classified according to the estimated amount of released products. The different release sizes can be classified into three categories as follows:

spills smaller than 7 tons, spills with releases between 7 and 700 tons or spills greater than 700 tons. This information is available on several databases as REMPEC [69] or ITOPF [74], which contains data about 10,000 accidents: the most frequent events (84%) belong to the smallest category with releases inferior to 7 tons [74].

The table 2 shows that the number of large spills (> 700 tons) has decreased significantly over the last 20 years. The average number of large spills per year during the 2000s was less than one-third of the one observed during the 1980s.

Most accidents result from a combination of different causes and circumstances which contribute in different ways to the final event. These causes can be categorized as "operational" and "accidental" [75-77]. From the table 3, it may be noticed that:

- tanker spills in the Mediterranean Sea mostly come from accidental causes such as stranding, collisions and shipwrecks, which generally generate larger spills;
- tanker spills which result from habitual operations such as loading, unloading and bunkering normally occur at ports or in the oil terminals;
- the majority of these spills are small or medium-sized, approximately 53% of the accidents for quantities between 7 and 700 tons.

The Figure 1 summarizes the causes of spills in the Mediterranean Sea between 1977 and 2019.

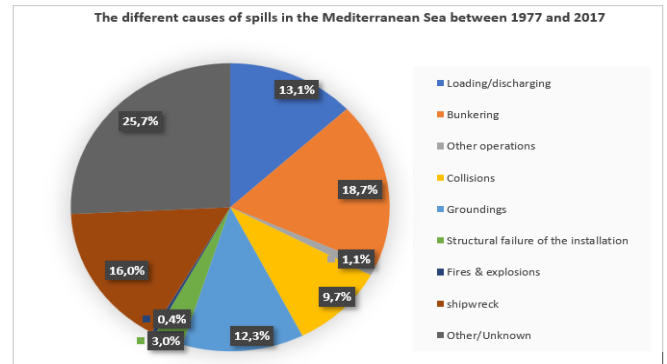


Figure 1: The different causes of spills in the Mediterranean Sea between 1977 and 2019. (REMPEC database).

3. A Lagrangian-based maritime and coastal risk model formulation

In the literature, the oil spill represents one of the main concerns in the context of risk analysis of maritime transportation due to the potential impact on marine ecosystems, to socio-economic activities and to the huge efforts in terms of recovery and clean-up operations [78-80].

Different risk definitions exist in the literature involving components such as probability, uncertainty, frequency of specific events, and/or related consequences. In [81], a review of methods and applications for maritime transportation risk analysis have been presented.

Table 2: Number of accidents and amount of spilled oil in the Mediterranean, between 1977 and 2019 (Source: REMPEC database)

Quantity spilled (tons)		Time horizon			
		Between 1977 and 1987	Between 1988 and 1997	Between 1998 and 2007	Between 2008 and 2019
<7	Number of spill accidents	20	36	17	51
	Total Quantity (t)	31	57	18	271
7<x<700	Number of spill accidents	18	31	9	63
	Total Quantity (T)	2595	5150	1835	9378
>700	Number of spill accidents	8	6	2	10
	Total Quantity (T)	283170	151700	3000	74700



In the proposed approach, the risk ( $R$ ) is associated with the expected value of the probability ( $P$ ) of an accident occurrence with a given spill size in a specific sea area and to the outcome arising as a consequence ( $C$ ) of the oil slick movement [47],[81]. The risk (*Risk*) is defined as a function of accident probability ( $P$ ) and consequences ( $C$ ) for the specific transportation hazard scenarios [82]:

$$Risk = f(Probability, Consequences, Scenario(s)) \quad (1)$$

Table 3: Number of oil spill accident according to spill size and operational/category causes in the Mediterranean, between 1977 and 2019 (REMPEC 2019)

Type of spill	Quantity spilt (Tons)			Total
	<7	7 < x < 700	>700	
<b>Operational</b>				
Loading / unloading	17	16	2	35
Leaking oil or gas	10	36	4	50
Other operations	2	1		3
<b>Accidental</b>				
Collision	8	13	5	26
Grounding	22	8	4	34
Structural failure of the installation	3	5		8
Fire or explosion	1	0	1	2
Shipwreck	22	17	4	43
Other	39	25	6	70

### 3.1. Accident probability analysis

The probability of maritime traffic accident occurrence is usually modelled by statistical approaches which are based on historical documentations about accident and non-accident rates, failures equipment, spill or release probabilities and container designs.

In the proposed model, the oil spill probability  $P_{i,s,h}$ , at the marine location  $i$ -th, for a specific spill size  $s$ -th, according to the weather scenario  $h$ -th generated by different meteorological conditions, is computed through the combination of three different components:

$$P_{i,s,h} = AR_i \times P_s^{spill\ size} \times P_{i,h}^{weather} \quad (2)$$

where:

-  $AR_i$  is the yearly oil spill accident rate for a specific water area which the location  $i$ -th belongs to. Assuming to be known a set of statistical data for a limited time horizon about accidents occurred in a predefined sea area,  $AR_i$  may be computed as:

$$AR_i = \frac{\# oil\ spill\ accident}{\# years * area} \left[ \frac{\# accident}{yr\ km^2} \right] \quad (3)$$

-  $P_s^{spill\ size}$ , is the probability of different release sizes in case of the accident occurrence. Three release sizes  $s$  were defined (small, medium, and large). These release sizes dictate the probability of the size of release in the probability analysis and in the consequence analysis.

$$P_s^{spill\ size} = \frac{\# accidents\ for\ the\ category\ of\ oil\ spill\ size\ s}{\# Total\ accidents} \quad (4)$$

-  $P_{i,h}^{weather}$ : the probability of weather stability, in the location  $i$ -th, for the weather scenario  $h$ -th, represents the probability of different combinations of atmospheric conditions for wind speed and wind directions based on frequency analysis. Statistical data are available in open-source database for different Mediterranean areas [83].

### 3.2. Consequence modelling

Currently, the consequence modelling is classified into two generations models according to their analysis in two or three dimensions (2D and 3D models). Those models have been developed to predict the evolution and behavior of hydrocarbons spilled on the surface and into the deepwater. The choice of dimension analysis directly influences the complexity of the model and the accuracy of the expected results. The 2D models [84–88] run quickly but they do not allow to obtain detailed information on the water column contamination [89] focusing only on the surface transport processes [90]. The 3D models [37], [39], [91] provide a description of the flow over the entire water column (surface, subsurface transport and fate processes) [90]. The latter models will give rise to more accurate results to simulate oil spills. Yet, more parameters have to be defined to get precise results. The decision to develop a 2D or 3D model strongly depends on the data that would be available to use as inputs. Upon the occurrence of a spill accident, the oil may stagnate as suspensions in the water column for a prolonged period due to the formation of emulsions. When emulsions processes are formed, the impact of the spill increases. As a consequence, the response and cleaning efforts become more complicated. Inclusion of the vertical movement of particles often makes the model very complex, as it will require detailed oceanographic information about the region for which the model is developed [92].

The 2D spreading models are mostly based on Lagrangian approaches [89]. The Lagrangian based models consider the oil slick as the movements of a set of small droplets subjected to wind, waves, and currents, and which can rise or sink due to buoyancy [93]. Several studies use the Lagrangian model to determine the areas that would be affected in the event of an oil spill [94–97].

In the proposed approach, a 2D Lagrangian based consequence model has been defined and used to compute maritime risk. The spreading, advection and diffusion processes which draw the oil spill trajectory and define consequently the impacted area of the spill accidents are described in the following paragraphs.

The vast majority of surface oil transport models use a random walk technique [46]. In this approach, the surface current field advects lagrangian elements representing the oil and disperses them through a random walk process used to represent horizontal dispersion.

#### 3.2.1. Spreading process

Spreading is one of the most relevant processes not only because it guarantees the prediction of the extent of oil slick area, but also, as it affects all other oil slick transformation processes.

Two physical phenomena lead to an oil slick movement on the surface water. First, the slick is subject to the spreading process under the influence of mechanical forces such as gravity, inertia, viscosity and interfacial tension and, on the other hand, to turbulent diffusion [53].

The oil slick extension in the wind direction is expected to increase with time proportionally to the wind speed, while the lateral elongation is always described by the gravity-spread equation proposed by [98].

$$A_0 = \pi \frac{k_2^4}{k_1^2} \left( \frac{V_0^5 g \Delta \rho}{\vartheta_m^2} \right) \quad (5)$$

where:

- $k_1$  and  $k_2$  are empirical coefficients ( $k_1 = 1.14$  and  $k_2 = 1.45$  [99]);
- $V_0$  is volume of oil spilled (m<sup>3</sup>);
- $\vartheta_m$  is the kinematic viscosity of water (m<sup>2</sup>/s);
- $g$  is gravitational acceleration (m/s<sup>2</sup>);
- $\Delta \rho$  is the relative density difference between the water and oil given by:

$$\Delta \rho = \frac{\rho_w - \rho_{oil}}{\rho_w} \quad (6)$$

where  $\rho_w$  is the density of water (g/cm<sup>3</sup>) and  $\rho_{oil}$  is the density of oil (g/cm<sup>3</sup>).

In the next gravity viscous spreading phase, the area  $A(t)$  of the oil slick during the time horizon may be computed using a correlation developed in [100], which assumes that oil slick spreading may have an elliptical shape on the water's surface with the major axis oriented in the direction of the wind.

The area covered by the oil slick (m<sup>2</sup>),  $A_t$ , at time t-th, from the time  $t_0$ , is described by:

$$A_t = \frac{1}{4} \pi Q_t R_t \quad (7)$$

The length of the minor ellipse axis (m),  $Q_t$ , is given by (8):

$$Q_t = 1.7(\Delta \rho V_0)^{1/3} t^{1/4} \quad (8)$$

where  $V_0$  is the volume of oil spill in barrels, the time t-th is the number of time units starting from the time  $t_0$ . The length of the major axis of the oil slick ellipse (m),  $R_t$ , is described by:

$$R_t = Q_t + 0.03 * (U_{wind})^{4/3} (t)^{3/4} \quad (9)$$

where  $U_{wind}$  is the wind speed in Knots, and t is time in minutes.

### 3.2.2. Advection and diffusion

The transport of an oil slick is generally induced by surface currents, wind, waves and turbulent diffusion [55]. Wind and currents are the two major processes composing the phenomenon of advection the slick. This surface current is largely wind generated, but in high tides regions tidal currents may dominate.

In this two-dimensional model of oil spill, the initial area of oil slick is divided into a large number of distinct Lagrangian particles in a XY plane reference at the water surface where (x<sub>t</sub>, y<sub>t</sub>) represent the position of a particle at a time step t.

It is assumed that these particles are connected to the surrounding body of water and therefore diffuse from a random process. The advection and diffusion properties of each particle can be calculated based on the flow fields at the water surface and the wind speed. Consequently, the speed, as well as the displacement of these particles, can be solved. Once their coordinates are determined at each time step, the shape and trace of the spill can be decided.

#### 3.2.2.1. Advection velocity

A large number of models use a constant parameter to associate the surface wind speed to the drift of the slick. This parameter is taken equal to about 3.5% [100–106] of the wind speed. The oil slick is also supposed to drift on water at 3.5% of the wind speed combined with 100% of the current speed [100–106].

In the proposed model, the advective velocity ( $U_a$ ) of the oil slick due to wind ( $U_{wind}$ ) and surface current ( $U_{current}$ ) effects is given [100]:

$$U_a = U_{current} + 0.035 U_{wind} \quad (10)$$

where  $U_a$  is the advective velocity of the oil slick (m/s),  $U_{current}$  is the surface current (m/s) and  $U_{wind}$  is wind speed (m/s).

#### 3.2.2.2. Horizontal turbulent diffusion

The Lagrangian approach predominantly represents the turbulent diffusion considering that the surface and the suspended particles of the slick are subjected to a random movement in addition to the regular movement due to the main current in the sea [37].

The translations, during a time step  $\Delta t$ , respectively  $\Delta X_{diff}$  and  $\Delta Y_{diff}$  due to the diffusion phase of the particles in the X and Y directions are based on [100]:

$$\Delta x_{diff} = [R]_0^1 \sqrt{12D_h \Delta t} \cos \theta \quad (11)$$

$$\Delta y_{diff} = [R]_0^1 \sqrt{12D_h \Delta t} \sin \theta \quad (12)$$

where  $[R]_0^1$  is a random number between 0 and 1 from a uniform distribution,  $D_h$ : horizontal diffusion coefficient (m<sup>2</sup>/s); and  $\theta$  is the directional angle  $\theta = 2\pi[R']_0^1$ . (where  $[R']_0^1$  is a random number between 0 and 1).

So, the displacement of the oil slick due to advection and horizontal diffusion is given as [106]:

$$X_{t+1} = x_t + u_{a,x} \Delta t + \Delta x_{diff} \quad (13)$$

$$Y_{t+1} = y_t + u_{a,y} \Delta t + \Delta y_{diff} \quad (14)$$

where:

- $x_t, y_t$  is the location of the particles at time step t-th;

- $u_{a,x}$  and  $u_{a,y}$  are the advective velocities in the X and Y directions respectively;
- $\Delta t$  is the time-step interval (s);
- $\Delta X_{diff}$ ,  $\Delta Y_{diff}$  are the displacements of the particles in the X and Y directions respectively.

### 3.3. Model treatment

The main required variables for the development of this model are related to water depth, spill rate, hydrocarbon density, current direction, and current speed. The determination of spill thickness and approximated area covered by the spill is based on the calculations of the estimated volumetric flow rate of oil from the source which is input for the simulation. The spill coverage is deduced through the use of a correlation developed by [100], which illustrated the elliptical spreading of oil on the water's surface with the major radius oriented in the direction of the wind.

This model finds the solution for the spreading and advection process functions numerically developed at discrete time. It also calculates the overall mass of oil for each time step based on the mass lost due to weathering process. Other oil properties, such as density, may be also recalculated after each time step. All of the aforementioned parameters are technically functions of time. Yet, in order to comply with the aims of numerical evaluation, the surface area and the overall mass of the oil volume are considered constant during a given time step and then updated in the next time step to reflect changes in mass and density.

The re-calculation of the center of mass of the oil slick generates the new local coordinate system for the center of a new set of concentric ellipses. The Fay method [98] is considered as the base of the spreading rate algorithm gravity-viscous spreading formula adjusted to cover wind effects. During the gravity viscous phase, the elliptical spreading of the oil slick is calculated. It should be mentioned that the calculation is only fulfilled when the terminal oil slick thickness is reached (as an example 0.1 mm [102]). The terminal oil slick thickness results in the removal of the spreading assumption. Then, the slick is allowed to spread under the influence of horizontal diffusion, surface winds and water current shears to represent complex, realistic surface movements.

The accuracy of the proposed model depends on specific input data such as product properties (e.g. density), wind and current speed and direction data, as well as the effect of other physical or chemical processes during the spill accident that may play a role in the correctness of the model results.

The collection of the actual data is a constraint for the implementation of the model, however, the reliability of the proposed approach for the consequences modelling has been positively verified in [107]. This latter paper analyzed to the oil spill accident occurred at the coast of Saint Tropez (France) at 2018, October the 7th. The accident occurred due to the collision between a Cypriot container ship and a Tunisian vessel generating an oil release of 600 cubic meters in the Mediterranean Sea. The proposed consequence model, from eq. (5) to (14), replicated the movements and trajectories of oil slick on the sea surface with good accuracy. Both in real and simulated cases, the oil slick reached the French coast within 9 days.

## 4. Case Study: Strait of Gibraltar

Due to its strategical position and its busy maritime traffic, represents the main maritime highway in West Mediterranean. The Strait occupies the space where the dense commercial traffic fueling between Europe and Asia intersects, revealing the historical links between Europe and Africa. These flows raise the Strait of Gibraltar to the ranks of the Straits of Pas de Calais or Malacca in terms of international maritime traffic (97,000 to 100,000 vessels per year) and, besides, the nature of these flows makes it an observatory area of global trade [108].

In this study, it is assumed that the proposed oil spill model to simulate the evolution of an oil slick and to compute the related risk assessment are based on maritime accidents or accidental ship collisions which may occur in four different potential accident test sites (ATs) in the body of water of the Gibraltar Strait. In detail, the four ATs have been identified in the intersections of the ship routes which covered the study area (Figure 2). The proposed risk assessment model aims at evaluating the impact of the simulated accident scenarios on 8 sensible Point of Interest (POIs), potentially exposed to oil spills from tankers, located on the African and European coasts, namely Tanger, Port Tanger Med, Dalia beach, Oued Mersa beach, Ceuta for Africa, and Tarifa, Algeciras and Gibraltar for Europe. Figure 2 represents the geographical position of the selected ATs and the locations of the sensible coastal Point of Interests (POIs). Table 4 shows the geographical coordinates of the ATs and the related distances from the specific POIs on the Moroccan and Spanish coasts.

The proposed model considers the movement of oil slick and its approaches to the coasts. The risk evaluation is based on the time required to the oil slick to hit the coastline on a given POIs in different accident scenarios.

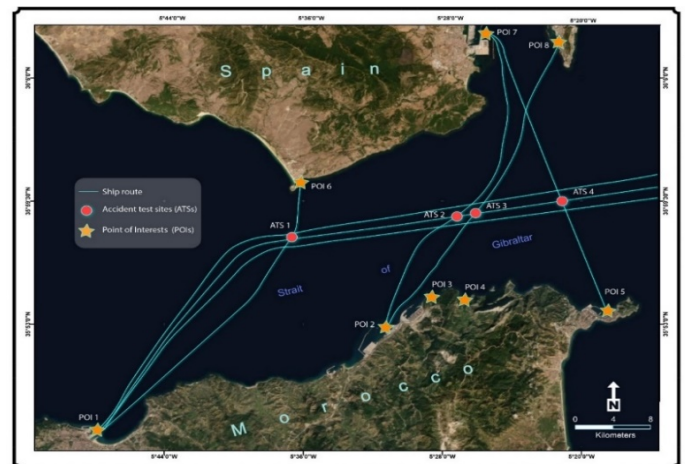


Figure 2: Locations of accident test sites (ATs) and point of interests (POIs) used to evaluate the environmental risk in the proposed simulations.

### 4.1. Model application

In the proposed study, the yearly accident rate for the case study area, computed by eq. 3, is based on REMPEC data [69] related to accidents occurred in 42 years, from 1977 to 2019, in Western Mediterranean Sea. Thus, for the four ATs  $i$ -th, considering the study area of about 653.245 km<sup>2</sup>, the accident rate is  $AR_i = 7.83 \times 10^{-06}$  accident/yr km<sup>2</sup>. The wind speed and wind directions

data have been collected by REMRO network database [83]. Those kinds of data related, as an example, to the ATS 1 appear in the table 5.

The consequence analysis has been carried out taking into consideration the following conditions:

- Three release sizes were defined: small (release <70 tons), medium (release between 70 and 7000 Ton) and large (release > 7000 Ton). Table 6 contains the  $P_s^{spill\ size}$  computed by eq. 4 according to accident data in the period 1977–2019.
- Twenty-four weather condition probabilities  $P_{i,h}^{weather}$ , associated to the h-th scenarios, in the four ATS i-th, have been used for the consequence modeling according to three wind speeds and eight wind directions combinations as follows:
  - $U_{wind} = 1.5\text{ m/s}, 14\text{ m/s}$  and  $24\text{ m/s}$ ;
  - 8 cardinal and intercardinal wind directions: north (N), northeast (NE), east (E), southeast (SE), south (S), southwest (SW), west (W), northwest (NW).

The probability of different atmospheric conditions, for  $U_{wind}=14\text{ m/s}$ , in the 4 different ATSs are listed in table 7. In the proposed simulations, the surface current in the four ATSs has been considered  $U_{current} = 0.6\text{ m/s}$  according to statistical data coming from [109, 110]. Concerning the horizontal diffusion coefficient,

the value  $D_h = 7\text{ m}^2/\text{sec}$  [88] has been adopted in the applications.

Table 8 indicates the oil spill probability  $P_{i,s,h}$  computed by eq.2 as the combination of the three different components in case of 700 Ton oil release category.

About 288 simulations have been carried out to evaluate the risk assessment on the sensible POIs on the African and European coasts. The simulations consist of 288 different scenarios related to three spill categories (70, 700 and 7000 tons), three different wind speeds (1.5, 14 and 24 m/s) in 8 wind directions in the four potential maritime accident locations (ATS1, ATS2, ATS3, ATS4).

For all simulations to observe the variations in the covered area, conditions such as wind velocity and current velocity are kept constant. The values have been input to generate spill location mappings over the period of time and are shown in the following figures.

The calculation of the axes of the oil slick spreading elliptical shape and related impacted area  $A_t$  computed by the model (5)-(14) for the three different release scenarios extended for a period of nine hours after the accident taken place (wind speed =14 m/s) is listed in table 9.

Table 4: The distances among the ATSs and the coastal POIs.

			Distance (km)			
			ATS 1 (35°57'06.5"N 5°37'04.2"W)	ATS 2 (35°58'11.7"N 5°28'00.8"W)	ATS 3 (35°58'22.5"N 5°26'30.2"W)	ATS 4 (35°59'03.9"N 5°20'44.8"W)
African coast	POI1	Tanger	24.56	36	38.5	46
	POI2	Port Tanger Med	12.42	9.19	10.46	17.22
	POI3	Dalia beach	13.84	7.25	8.13	14.78
	POI4	Oued Mersa beach	16.64	7.63	7.56	12.71
	POI5	Ceuta	28.86	16.78	15	11
European coast	POI6	Tarifa	6.34	12.87	14.77	23
	POI7	Algeciras	10	9.45	10.5	12.91
	POI8	Gibraltar	29.89	19	18	14

Table 5: Annual combined probability of wind direction and mean wind speed for ATS 1 (source: REMRO network database)

Direction	Average wind speed (m/s)									Total
	≤ 1.0	3.0	6.0	9.0	12.0	15.0	18.0	21.0	>21.0	
Calm	1.06E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.06E-02
N	0.00E+00	6.66E-03	1.16E-02	3.61E-03	2.80E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.21E-02
NNE	0.00E+00	6.43E-03	7.79E-03	1.15E-02	5.00E-05	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.57E-02
NE	0.00E+00	6.20E-03	9.48E-03	1.22E-03	1.90E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.71E-02
ENE	0.00E+00	5.54E-03	1.95E-02	1.64E-02	6.71E-03	1.55E-03	5.00E-05	0.00E+00	0.00E+00	4.98E-02
E	0.00E+00	6.34E-03	3.73E-02	7.82E-02	8.41E-02	4.58E-02	1.57E-02	3.47E-03	0.00E+00	2.71E-01
ESE	0.00E+00	4.51E-03	1.53E-02	1.25E-02	5.54E-03	1.60E-03	3.80E-04	0.00E+00	0.00E+00	3.98E-02



SE	0.00E+00	3.99E-03	7.42E-03	1.64E-03	2.30E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.33E-02
SSE	0.00E+00	3.99E-03	5.49E-03	1.69E-03	4.70E-04	5.00E-05	0.00E+00	0.00E+00	0.00E+00	1.17E-02
S	0.00E+00	4.32E-03	8.78E-03	4.46E-03	1.55E-03	5.00E-05	9.00E-05	5.00E-05	0.00E+00	1.93E-02
SSW	0.00E+00	5.77E-03	1.20E-02	9.86E-03	6.34E-03	1.83E-03	9.40E-04	9.00E-05	0.00E+00	3.68E-02
SW	0.00E+00	8.40E-03	2.65E-02	1.49E-02	8.82E-03	4.51E-03	6.60E-04	9.00E-05	0.00E+00	6.39E-02
WSW	0.00E+00	1.20E-02	4.16E-02	1.84E-02	6.05E-03	2.86E-03	4.20E-04	0.00E+00	0.00E+00	8.13E-02
W	0.00E+00	1.37E-02	5.89E-02	3.17E-02	8.49E-03	1.97E-03	5.00E-05	1.90E-04	0.00E+00	1.15E-01
WNW	0.00E+00	1.35E-02	6.25E-02	2.82E-02	8.07E-03	1.69E-03	5.20E-04	0.00E+00	0.00E+00	1.14E-01
NW	0.00E+00	1.03E-02	4.22E-02	1.92E-02	5.12E-03	8.90E-04	5.00E-05	0.00E+00	0.00E+00	7.77E-02
NNW	0.00E+00	8.96E-03	2.22E-02	8.03E-03	1.31E-03	9.00E-05	0.00E+00	0.00E+00	0.00E+00	4.03E-02
Total	1.06E-02	1.21E-01	3.89E-01	2.61E-01	1.43E-01	6.29E-02	1.89E-02	3.89E-03	0.00E+00	1.0

Table 6:  $P_s^{spill\ size}$  computed according to the accident data in the period 1977-2019.

Spill category	Release size (Tons)	Number of accident	$P_s^{spill\ size}$
s=1	70	261	6.92E-01
s=2	700	72	1.91E-01
s=3	7000	44	1.17E-01

Table 7: Probability of different combinations of atmospheric conditions for wind speed =14 m/s in 8 different directions in each ATS.

Wind speed m/s	Direction	Accident Tests Sites			
		ATS 1	ATS 2	ATS 3	ATS 4
14	N	5.00E-03	5.00E-03	5.00E-03	5.00E-03
	NE	5.00E-03	5.00E-03	5.00E-03	5.00E-03
	E	2.10E-01	3.20E-02	3.20E-02	1.50E-02
	SE	5.00E-03	5.00E-03	5.00E-03	5.00E-03
	S	5.00E-03	5.00E-03	5.00E-03	5.00E-03
	SW	2.00E-02	1.25E-02	1.25E-02	1.20E-02
	W	7.00E-03	5.00E-03	5.00E-03	9.00E-03
	NW	5.00E-03	5.00E-03	5.00E-03	1.50E-02

Table 8: The probability  $P_{L,s,h}$  for oil spill release category s=2, 700 Tons, according to the 24 scenarios dictated by wind speed and direction combined probability.

Wind speed m/s	Direction	$P_{ATS\ 1}$	$P_{ATS\ 2}$	$P_{ATS\ 3}$	$P_{ATS\ 4}$
1,5	N	1.71E-07	3.88E-07	3.88E-07	2.84E-07
	NE	2.69E-07	2.09E-07	2.09E-07	5.98E-08
	E	1.34E-08	7.47E-09	7.47E-09	7.47E-09
	SE	2.09E-07	4.33E-07	4.33E-07	3.58E-07
	S	2.69E-07	5.83E-07	5.83E-07	2.84E-07
	SW	1.71E-07	2.84E-07	2.84E-07	2.39E-07
	W	2.99E-08	7.47E-09	7.47E-09	7.47E-09
	NW	1.10E-07	1.34E-07	1.34E-07	1.34E-07
14	N	7.47E-09	7.47E-09	7.47E-09	7.47E-09
	NE	7.47E-09	7.47E-09	7.47E-09	7.47E-09
	E	3.14E-07	4.78E-08	4.78E-08	2.24E-08
	SE	7.47E-09	7.47E-09	7.47E-09	7.47E-09
	S	7.47E-09	7.47E-09	7.47E-09	7.47E-09
	SW	2.99E-08	1.86E-08	1.86E-08	1.79E-08
	W	1.04E-08	7.47E-09	7.47E-09	1.34E-08
	NW	7.47E-09	7.47E-09	7.47E-09	2.24E-08

24	N	7.47E-09	7.47E-09	7.47E-09	7.47E-09
	NE	7.47E-09	7.47E-09	7.47E-09	7.47E-09
	E	7.47E-09	7.47E-09	7.47E-09	7.47E-09
	SE	7.47E-09	7.47E-09	7.47E-09	7.47E-09
	S	7.47E-09	7.47E-09	7.47E-09	7.47E-09
	SW	7.47E-09	7.47E-09	7.47E-09	7.47E-09
	W	7.47E-09	7.47E-09	7.47E-09	7.47E-09
	NW	7.47E-09	7.47E-09	7.47E-09	2.24E-08

Table 9: The computation of the elliptical spreading area  $A_t$  for the oil slick during 9 hours after the incident occurred.

Wind speed (m/s)	Release size (Tons)	Time (hours)	$A_t$ (km <sup>2</sup> )	$Q_t$ (m)	$R_t$ (m)
14	70	1	0.41	625.72	2354.85
		2	0.65	744.11	3652.15
		3	0.85	823.49	4765.06
		4	1.04	884.90	5775.62
		5	1.22	935.67	6717.37
		6	1.40	979.30	7608.20
		7	1.56	1017.78	8459.13
		8	1.73	1052.33	9277.51
		9	1.89	1083.78	10068.61
	700	1	1.53	1346.36	3075.49
		2	2.23	1601.10	4509.14
		3	2.79	1771.91	5713.48
		4	3.28	1904.04	6794.76
		5	3.72	2013.28	7794.98
		6	4.13	2107.17	8736.07
		7	4.52	2189.96	9631.31
		8	4.88	2264.30	10489.48
		9	5.23	2331.97	11316.80
	7000	1	6.71	2900.03	4629.16
		2	9.56	3448.74	6356.78
		3	11.76	3816.66	7758.23
		4	13.64	4101.27	8991.99
		5	15.31	4336.56	10118.26
		6	16.82	4538.80	11167.70
		7	18.23	4717.13	12158.47
		8	19.54	4877.26	13102.44
		9	20.78	5023.01	14007.84

Table 10: Time (hours) required by the oil slick to reach the POIs on the Moroccan coastline for accidents occurred in one of the ATs according to different scenarios.

			Release quantity								
			70 Ton			700 Ton			7000 Ton		
Wind speed [m/s]			1.5	14	24	1.5	14	24	1.5	14	24
AFRICAN PART	Tanger	ATS 1	12	4.5	2.5	8	3.5	2.5	4.5	2.5	1.5
		ATS 2	17.5	5.5	4	12	5	3.5	7	4	3
		ATS 3	18.5	6	4	12.5	5.5	3.5	7.5	4.5	3
		ATS 4	21	7	4.5	15	6.5	4.5	8.5	5	3.5
	Port Tanger Med	ATS 1	6	2	1	3.5	1.5	<1	2	1	<1
		ATS 2	4	1	<1	2	1	<1	1	<1	<1
		ATS 3	4.5	1.5	<1	2.5	1	<1	1	<1	<1
		ATS 4	9	3	1.5	5.5	2.5	1.5	3	1.5	1
	Dalia beach	ATS 1	6.5	2	1	4	1.5	1	2	1	<1
		ATS 2	2	<1	<1	1	<1	<1	<1	<1	<1
		ATS 3	2.5	<1	<1	1.5	<1	<1	<1	<1	<1
		ATS 4	7	2	1.5	4.5	2	1	2	1.5	<1

	Oued Mersa beach	ATS 1	8.5	2.5	1.5	5	2.5	1.5	3	1.5	1
		ATS 2	2	<1	<1	1	<1	<1	<1	<1	<1
		ATS 3	2	<1	<1	1	<1	<1	<1	<1	<1
		ATS 4	6	2	1	3.5	1.5	<1	2	1	<1
	Ceuta	ATS 1	14.5	4.5	3	9.5	4	2.5	5.5	3.5	2.5
		ATS 2	8	3	1.5	5	2.5	1.5	3	1.5	1
		ATS 3	7.5	2.5	1.5	4.5	2	1.5	2.5	1.5	<1
		ATS 4	5	1.5	<1	3	1	<1	1.5	<1	<1

Table 11: Time (hours) required by the oil slick to reach the POIs on the European coastline for accidents occurred in one of the ATs according to different scenarios.

			Release quantity									
			70 T			700 T			7000T			
Wind speed [m/s]			1.5	14	24	1.5	14	24	1.5	14	24	
EUROPEAN PART	Time (hour)	Tarifa	ATS 1	1	<1	<1	1	<1	<1	<1	<1	<1
			ATS 2	6	2	1	3.5	1.5	<1	2	1	<1
			ATS 3	7	2.5	1.5	4.5	1.5	1	2	1.5	<1
			ATS 4	12	4	2.5	7.5	3.5	2.5	4.5	2.5	1.5
		Algeciras	ATS 1	4.5	1	<1	2	1	<1	1	<1	<1
			ATS 2	4	1	<1	2	1	<1	1	<1	<1
			ATS 3	4.5	1	<1	2.5	1	<1	1	<1	<1
			ATS 4	6	2	1	3.5	1.5	<1	2	1	<1
		Gibraltar	ATS 1	14.5	5	3	9.5	4.5	2.5	5.5	3.5	2.5
			ATS 2	10	3.5	1.5	6	2.5	1.5	3.5	2.5	1.5
			ATS 3	9.5	3.5	1.5	5.5	2.5	1.5	3.5	2	1.5
			ATS 4	7	2.5	1.5	4.5	2	1	2.5	1.5	<1

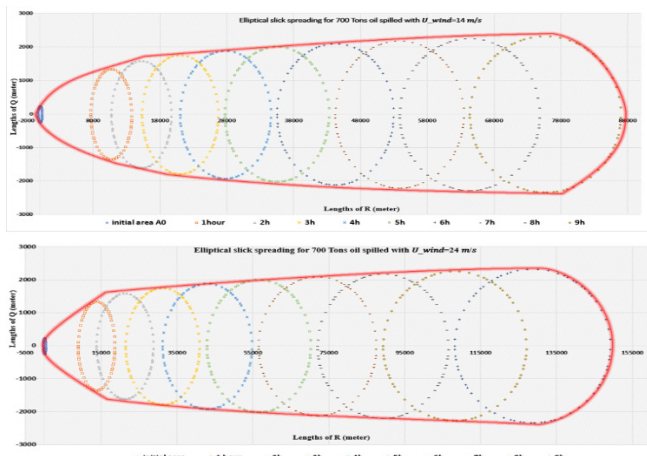


Figure 3: Elliptical slick spreading for 700 Tons oil spilled with  $U_{wind} = 14 \text{ m/s}$  and  $24 \text{ m/s}$ .

The plots show a series of elliptical shapes for the different scenarios that were interpreted according to the resulting vector of the wind-to-current vector interaction. Figure 3 and Figure 4 shows the applications of the consequence model to simulate the oil slick spreading and trajectory prediction in relation to different meteorological conditions. Specifically, Figure 3 shows the related graphical interpretation of the oil slick spreading for 700 Ton of spilled oil for different conditions of wind speed ( $U_{wind}=14\text{m/s}$  and  $24\text{m/s}$ ), which represents the development of

the oil spill area and movements according to the time in the wind direction since the origin. The main one is that the elliptical shape is moving significant distances from the origin. This is explained by having the factor applied to the wind and current vectors being too large.

Figure 4 represents the visualization, by a GIS, of the simulated scenario which concerns a maritime spill accident, occurred in the ATS4, with an estimated spill release of 700 tons, with  $U_{wind}=14 \text{ m/s}$  and major wind direction from Northeast-East (NE), which is estimated to impact the coast at POI1 within 6 hours of the accident release in ATS4.

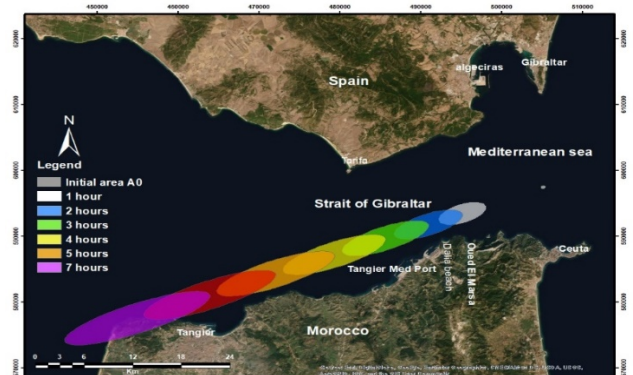


Figure 4: Estimated Spill location for 700 Tons from ATS 4 with wind speed  $14 \text{ m/s}$  (wind direction: Northeast-East).

## 5. Risk model application

In order to evaluate the environmental risk due to coastal pollution occurring in the event of a spill accident, it is necessary to identify risk's levels. In the proposed model, the potential environmental risk was assessed on the basis of the time required by the oil slick to reach the European and African coasts assuming that an oil spill accident may have occurred in one of the ATSSs.

The proposed Lagrangian-based maritime and coastal risk model can be used not merely in oil spill response and contingency planning but also in assessing risk impact. In the event of an oil spill, by a limited set of necessary meteorological data, the predictions of the slick movements and trajectories may be provided to the competent authorities. The main result of the proposed approach is the possibility to define a risk ranking for the coastal area based on the forecasted slick movements and to determine, under specified meteorological conditions, the time required by the oil slick to hit the littoral.

Tables 10 and 11 represent the estimated times (in hours) required by the oil spill to reach the POIs in the African and European coasts coming from the different ATSSs off the Strait of Gibraltar according to different scenarios.

The oil spill risk assessment is defined considering the cumulative probability, related to the overall simulated scenarios, that the oil slicks reach the coastal POIs in the successive hour time slots after the beginning of the oil spill accident releases.

Figures 5 and 6 represent, respectively, the coastal environmental risk for the selected African and European POIs. Among the selected POIs, Oued Mersa and Dalia represent the first locations to be affected in case of accidents in the Strait of Gibraltar in the first two hours. Besides, the region of Tarifa and Algeciras also appears to be subjected to a relevant risk on the European side.

According to the timing, after 4 hours, Port Tangier Med and the beaches of Oued Mersa and Dalia beach, on the African Mediterranean coast, have the main probability to be hit by an oil slick generated by maritime accidents in the study area. On the European side, the risk probability of oil beaching on the coasts, is growing, respectively, for Algeciras, Tarifa and Gibraltar.

Due to the interaction between the wind flow and the geography of the coast, the Gibraltar strait is very windy. It is exposed mainly to two types of winds [111,112]:

- East winds dominate in March, and from July to October, with a wind speed exceeding 8 m/s is 22% of the days;
- The west winds of Atlantic origin and important source of humidity and precipitation, dominate from December to April [111–113].

According to data provided by REMRO Network [83], the wind in the test points ATS1 and ATS4 has direction East-South-East and West direction (ESE-W) with a percentage of 53%, directed to European side, and direction West-North-West and East direction (WNW-E) in 47%. On the other hand, for the test points ATS2 and ATS3, the direction of the dominant winds, towards the Moroccan coast, is between WNW-E with a percentage of 61% (39% for ESE-W).

The importance of the proposed model appears to be most obvious due to this wind variability. Thus, it represents an added value in the real time applications in case of maritime accidents.

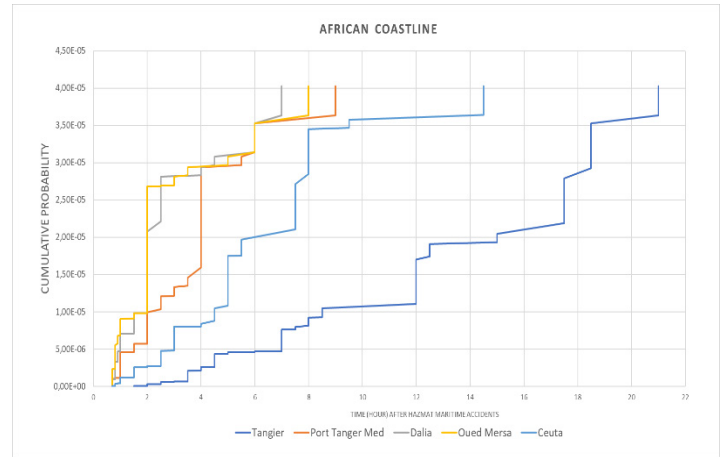


Figure 5: Risk evaluation on the African POIs

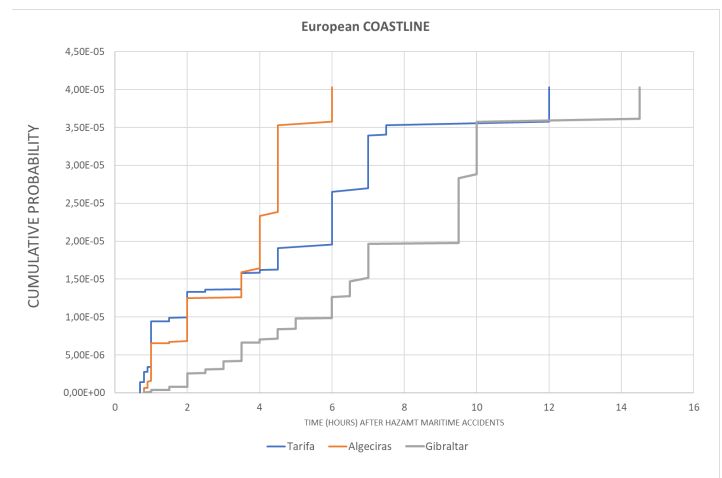


Figure 6: Risk evaluation on the European POIs.

## 6. Conclusion

The proposed risk model provides wide-ranging and fast information on the direction, spreading and magnitude of the oil spill, as well as it identifies the coastal areas which may be affected more or less quickly. This approach may be integrated with a GIS tool to generate detailed simulated maps on trajectories and oil slick spreading toward the relevant coasts.

The proposed approach presents a reliable methodology to classify the risk in marine and coastal areas due to the maritime transport of dangerous goods. The main contribution of our method is based on an integration of the probability of the maritime accident in each specific area and the related accident consequences, and, besides, it uses a simple 2D consequence model instead a more complicated 3D model. In the current literature, in fact, the mainly works focused on the movements of the oil slick and its impact without taking into account the accident probability.

The proposed model has been applied to 288 significant scenarios generated considering four potential accident sites in the West Mediterranean Sea. The accident probability analysis is function



of three components related to maritime oil spill accident frequency on the maritime routes included in the study area, probability of spill sizes, and a joint probability function of wind direction and speed. The consequences model related to the prediction of the oil slick trajectories and affected areas takes into consideration the spreading, the advection and the diffusion processes. To determine the oil spill path, data about the weather conditions and surface currents have been utilized.

Risk assessment for the coastal POIs is based on the cumulative probability to be impacted by an oil spill, over the time, starting from the initial accident event.

The main contribution of this paper is twofold. Firstly, the model application ranks the coastal locations according to higher hazmat risk to be strongly affected by the oil spill in each time intervals. Secondly, the proposed model, being connected to adequate ICT equipment to acquire in real time data on weather and sea currents, represents a useful tool to manage the immediate containment and recovery activities.

However, the validation test of the model in the framework of the Saint-Tropez (France) spill incident verified that the model algorithms provided an encouraging level of accuracy. Further developments have been identified to improve the accuracy and functionality of our model.

However, the proposed model may underestimate the oil slick impact area since it does not encompass all the possible physical processes in the water column. In the future development of the model, also dispersion, emulsification and dissolution phases may be taken into account.

Anyway, the main purpose of our work is to realize a simplified method useful for the public authorities which have quickly to intervene in case of accident.

The introduction of new parameters and variables in the model surely may provide the users with a more reliable prediction of the oil slick movements but also, they may complicate the applicability of the approach in case of accident, fast recovery after incidents, mitigation of consequences. In addition, accurate information on the environment such as spatially and temporally varying wind, water current and wave fields are essential for a reliable prediction of the transport and fate of the oil slick. This model currently has limited capability to incorporate operational predictions of wave impact. As a result, the model is unable to simulate the subsurface transport of dispersed oil droplets related to vertical dispersion due to wave action. The possibility to incorporate the underwater oil transport functionality will be explored in future improvements of the modeling system.

#### Conflict of Interest

The authors declare no conflict of interest.

#### Acknowledgment

The authors thank all the Port Authorities in Tangier Med (Morocco) especially the Oil Terminal HTT.SA and all its employees for their great support and valuable cooperation.

This work was supported by the project LOSE+ (Logistic and Safety of freight transport) in the framework of the Italy-France Interreg Maritime Program 2014-2020, European Project No. 276, Call n.3, Axis 2.

#### References

- [1] S. Kristiansen, "MaritimeTransportation. SafetyManagement and RiskAnalysis" Elsevier Butterworth-Heinemann, 2005.
- [2] J. F. Balmat, F. Lafont, R. Maifret, and N. Pessel, "A decision-making system to maritime risk assessment," *Ocean Eng.*, **38**(1) 171–176, 2011.
- [3] K. X. Li, J. Yin, H. S. Bang, Z. Yang, and J. Wang, "Bayesian network with quantitative input for maritime risk analysis," *Transp. A Transp. Sci.*, **10**(2) 89–118, 2014.
- [4] J. F. Balmat, F. Lafont, R. Maifret, and N. Pessel, "MARitime RiSk Assessment (MARISA), a fuzzy approach to define an individual ship risk factor," *Ocean Eng.*, **36**(15–16) 1278–1286, 2009.
- [5] J. R. W. Merrick, R. Van Dorp, and V. Dorp, "Speaking the Truth in Maritime Risk Assessment," *Risk Anal.*, **26**(1) 223–237, 2006.
- [6] J. Montewka, S. Ehlers, F. Goerlandt, T. Hinz, and K. Tabri, "A framework for risk assessment for maritime transportation systems — A case study for open sea collisions involving RoPax vessels," *Reliab. Eng. Syst. Saf.*, **124**, 142–157, 2014.
- [7] C. G. Soares and A. P. Teixeira, "Risk assessment in maritime transportation," *Reliab. Eng. Syst. Saf.*, **74**(3) 299–309, 2001.
- [8] F. Goerlandt and J. Montewka, "Maritime transportation risk analysis: Review and analysis in light of some foundational issues," *Reliab. Eng. Syst. Saf.*, **138**, 115–134, 2015.
- [9] IMO, "Degree of risk evaluation. SN.1/Circ.296.," 2010.
- [10] W. Guo, "Development of a statistical oil spill model for risk assessment," *Environ. Pollut.*, **230**, 945–953, 2017.
- [11] A. Al Shami, G. Harik, I. Alameddine, D. Bruschi, D. A. Garcia, and M. El-Fadel, "Risk assessment of oil spills along the Mediterranean coast: A sensitivity analysis of the choice of hazard quantification," *Sci. Total Environ.*, **574**, 234–245, 2017.
- [12] J. Fernández-Macho, "Risk assessment for marine spills along European coastlines," *Mar. Pollut. Bull.*, **113**(1–2), 200–210, 2016.
- [13] R. S. Kankara, S. Arockiaraj, and K. Prabhu, "Environmental sensitivity mapping and risk assessment for oil spill along the Chennai Coast in India," *Mar. Pollut. Bull.*, **106**(1–2), 95–103, 2016.
- [14] H. Landquist, L. Rosén, A. Lindhe, and I. M. Hassellöv, "VRAKA-A probabilistic risk assessment method for potentially polluting shipwrecks," *Front. Environ. Sci.*, **4**(JUL) 1–14, 2016.
- [15] P. F. Valdor, A. G. Gómez, V. Velarde, and A. Puente, "Can a GIS toolbox assess the environmental risk of oil spills? Implementation for oil facilities in harbors," *J. Environ. Manage.*, **170**, 105–115, 2016.
- [16] D. Depellegrin and P. Pereira, "Assessing oil spill sensitivity in unsheltered coastal environments: A case study for Lithuanian-Russian coasts, South-eastern Baltic Sea," *Mar. Pollut. Bull.*, **102**(1) 44–57, 2016.
- [17] T. M. Alves, E. Kokinou, G. Zodiatis, H. Radhakrishnan, and C. Panagiotakis, "Multidisciplinary oil spill modeling to protect coastal communities and the environment of the Eastern Mediterranean Sea," *Sci. Rep.*(May) 1–9, 2016.
- [18] R. Goldman, E. Biton, E. Brokovich, S. Kark, and N. Levin, "Oil spill contamination probability in the southeastern Levantine basin," *Mar. Pollut. Bull.*, **91**(1) 347–356, 2015.
- [19] X. Liu, R. Meng, Q. Xing, M. Lou, H. Chao, and L. Bing, "Assessing oil spill risk in the Chinese Bohai Sea: A case study for both ship and platform related oil spills," *Ocean Coast. Manag.*, **108**, 140–146, 2015.
- [20] S. Mokhtari et al., "Inferring spatial distribution of oil spill risks from proxies: Case study in the north of the Persian Gulf," *Ocean Coast. Manag.*, **116**, 504–511, 2015.
- [21] A. A. Sepp Neves et al., "Towards a common oil spill risk assessment framework - Adapting ISO 31000 and addressing uncertainties," *J. Environ. Manage.*, **159**, 158–168, 2015.
- [22] A. Jolma, A. Lehtikoinen, I. Helle, and R. Venesjärvi, "A software system for assessing the spatially distributed ecological risk posed by oil shipping," *Environ. Model. Softw.*, **61**, 1–11, 2014.
- [23] T. M. Alves, E. Kokinou, and G. Zodiatis, "A three-step model to assess shoreline and offshore susceptibility to oil spills: The South Aegean (Crete) as an analogue for confined marine basins," *Mar. Pollut. Bull.*, **86**(1–2) 443–457, 2014.
- [24] T. Aven and E. Zio, "Foundational Issues in Risk Assessment and Risk Management," *Risk Anal.*, **34**(7) 1164–1172, 2014.

- [25] N. Khakzad, F. Khan, and P. Amyotte, "Quantitative risk analysis of offshore drilling operations: A Bayesian approach," *Safety. Sciences.*, **57** 108–117, 2013.
- [26] A. Blokus-Roszkowska and L. Smolarek, "Collision risk estimation for motorways of the sea," *Reliab. Theory Appl.*, **2**(25), 58–68, 2012.
- [27] A. Talavera, R. Aguasca, and B. Galva, "Application of Dempster – Shafer theory for the quantification and propagation of the uncertainty caused by the use of AIS data," *Reliab. Eng. Syst. Saf.*, **111**, 95–105, 2013.
- [28] M. B. Zaman, E. Kobayashi, N. Wakabayashi, S. Khanfir, T. Pitana, and A. Maimun, "Fuzzy FMEA model for risk evaluation of ship collisions in the Malacca Strait: Based on AIS data," *J. Simul.*, **8**(1) 91–104, 2014.
- [29] J. R. Van Dorp, J. R. W. Merrick, J. R. Harrald, T. A. Mazzuchi, and M. Grabowski, "A Risk Management Procedure for the Washington State Ferries," *Risk Anal.*, **21**(1) 127–142, 2001.
- [30] Ö. S. Ulusçu, B. Özbaş, T. Ahtiok, and İ. Or, "Risk analysis of the vessel traffic in the strait of Istanbul," *Risk Anal.*, **29**(10) 1454–1472, 2009.
- [31] S. Hu, Q. Fang, H. Xia, and Y. Xi, "Formal safety assessment based on relative risks model in ship navigation," *Reliab. Eng. Syst. Saf.*, **92**(3) 369–377, 2007.
- [32] J. Akhtar and T. Bjørnskau, "Oil spill risk analysis of routing heavy ship traffic in Norwegian waters," *WMU J. Marit. Aff.*, **11**(2) 233–247, 2012.
- [33] J. Ylitlo, "Modelling Marine Accident Frequency," *Otaniemi, Finl. Helsinki Univ. Technol.* 2010. January, 2010.
- [34] Janeiro, J., Fernandes, E., Martins, Fernandes, "Wind and freshwater influence over hydrocarbon dispersal on Patos Lagoon, Brazil," *Marine Pollution Bulletin.*, **56**, 650–665, 2008.
- [35] H. Havens, M. E. Luther, and S. D. Meyers, "A coastal prediction system as an event response tool: Particle tracking simulation of an anhydrous ammonia spill in Tampa Bay," *Mar. Pollut. Bull.*, **58**(8) 1202–1209, 2009.
- [36] N. C. Delpeche-Ellmann and T. Soomere, "Investigating the Marine Protected Areas most at risk of current-driven pollution in the Gulf of Finland, the Baltic Sea, using a Lagrangian transport model," *Mar. Pollut. Bull.*, **67**(1–2), 121–129, 2013.
- [37] S. D. Wang, Y. M. Shen, Y. K. Guo, and J. Tang, "Three-dimensional numerical simulation for transport of oil spills in seas," *Ocean Eng.*, **35**(5–6), 503–510, 2008.
- [38] M. Marta-Almeida et al., "Efficient tools for marine operational forecast and oil," *Mar. Pollut. Bull.*, **71**, 139–151, 2013.
- [39] M. El-Fadel, R. Abdallah, and G. Rachid, "A modeling approach toward oil spill management along the Eastern Mediterranean," *J. Environ. Manage.*, **113**, 93–102, 2012.
- [40] P. Annika, T. George, P. George, N. Konstantinos, D. Costas, and C. Koutitas, "The Poseidon operational tool for the prediction of floating pollutant transport," *Mar. Pollut. Bull.*, **43**(7–12), 270–278, 2001.
- [41] A. Elizaryev et al., "Numerical simulation of oil spills based on the GNOME and ADIOS," *Int. J. Eng. Technol.*, **7**(2) 24–27, 2018.
- [42] L. Perivoliotis, G. Krokos, K. Nittis, and G. Korres, "The Aegean sea marine security decision support system," *Ocean Sci.*, **7**(5) 671–683, 2011.
- [43] K. Nittis, L. Perivoliotis, G. Korres, C. Tziavos, and I. Thanos, "Operational monitoring and forecasting for marine environmental applications in the Aegean Sea," *Environ. Model. Softw.*, **21**(2) 243–257, 2006.
- [44] A. C. Toz, B. Koseoglu, and C. Sakar, "Numerical modelling of oil spill in New York Bay," *Arch. Environ. Prot.*, **42**(4) 22–31, 2016.
- [45] J. Zhao, M. Temimi, M. Al Azhar, and H. Ghedira, "Satellite-based tracking of oil pollution in the Arabian Gulf and the Sea of Oman," *Can. J. Remote Sens.*, **41**(2) 113–125, 2015.
- [46] Y. Li, H. Chen, and X. Lv, "Impact of error in ocean dynamical background, on the transport of underwater spilled oil," *Ocean Model.*, **132**( August) 30–45, 2018.
- [47] P. Amir-Heidari and M. Raie, "Probabilistic risk assessment of oil spill from offshore oil wells in Persian Gulf," *Mar. Pollut. Bull.*, **136**( May) 291–299, 2018.
- [48] W. Guo et al., "A modified probabilistic oil spill model and its application to the Dalian New Port accident," *Ocean Eng.*, **121**, 291–300, 2016.
- [49] Witchaya Rongsayamanont et al., "Formulation of crude oil spill dispersants based on the HLD concept and using a lipopeptide biosurfactant," *J. Hazard. Mater.*, **334**, 168–177, 2017.
- [50] A. Al-Majed, A. Adebayo, and M. Hossain, "A sustainable approach to controlling oil spills," *J. Environ. Manage.*, 2012.
- [51] F. A. Vega, Emma F. Covelo, M. J. Reigosa, and M. Luisa Andrade, "Degradation of fuel oil in salt marsh soils affected by the Prestige oil spill," *J. Hazard. Mater.*, **166**(2–3), 1020–1029, 2009.
- [52] N. P. Ventikos and George Triantafyllou, "A high-level synthesis of oil spill response equipment and countermeasures," *J. Hazard. Mater.*, **107**(1–2) 51–58, 2004.
- [53] A. K. Mishra and G. S. Kumar, "Weathering of Oil Spill: Modeling and Analysis," *Aquat. Procedia*, **4**( Icwrcoc) 435–442, 2015.
- [54] National Research Council, *Oil in the sea III: inputs, fates, and effects*. 2003.
- [55] M. L. Spaulding, "State of the art review and future directions in oil spill modeling," *Mar. Pollut. Bull.*, **115**(1–2) 7–19, 2017.
- [56] Z. Liu, J. Liu, Q. Zhu, and W. Wu, "The weathering of oil after the Deepwater Horizon oil spill: Insights from the chemical composition of the oil from the sea surface, salt marshes and sediments," *Environ. Res. Lett.*, **7**, 3, 2012.
- [57] U. H. Yim et al., "Fingerprint and weathering characteristics of stranded oils after the Hebei Spirit oil spill," *J. Hazard. Mater.*, **197** 60–69, 2011.
- [58] G. Zodiatis et al., "The Mediterranean Decision Support System for Marine Safety dedicated to oil slicks predictions," *Deep. Res. Part II Top. Stud. Oceanogr.*, **133**, 4–20, 2016.
- [59] Hosein Bidgoli, Y. Mortaza, and A. A. Khodadadi, "A functionalized nano-structured cellulosic sorbent aerogel for oil spill cleanup\_ Synthesis and characterization," *J. Hazard. Mater.* **366**, 229–239, 2019.
- [60] F. Cláudio, D. F. Barros, L. Constantino, and G. Vasconcellos, "Removal of Petroleum Spill in Water by Chitin and Chitosan," *Electron. J. Chem.*, **6**(1), 2–6, 2014.
- [61] M. Albakjaji, "La pollution de la mer méditerranée par les hydrocarbures liée au trafic maritime," Thèse de doctorat. Paris Est. 2011.
- [62] Ş. N. E. Bozkurtoglu, "Modeling oil spill trajectory in Bosphorus for contingency planning," *Mar. Pollut. Bull.*, **123**(1–2), 57–72, 2017.
- [63] L. Karafyllidis, "A model for the prediction of oil slick movement and spreading using Cellular Automata," *Environ. Int.*, **23**(6), 839–850, 1997.
- [64] C. Bolognesi, E. Perrone, P. Roggeri, and A. Sciuotto, "Bioindicators in monitoring long term genotoxic impact of oil spill: Haven case study," *Mar. Environ. Res.*, **62**( SUPPL. 1), 287–291, 2006.
- [65] P. Guidetti, M. Modena, G. Mesa, and M. Vacchi, "Composition, Abundance and Stratification of Macrobenzothos in the Marine Area Impacted by Tar Aggregates Derived from the Haven Oil Spill (Ligurian Sea, Italy)," *Marin Pollution Bulletin*, **40**(12), 1161–1166, 2000.
- [66] R. Periañez, "Chemical and oil spill rapid response modelling in the Strait of Gibraltar-Alborán Sea," *Ecol. Modell.*, **207**(2–4), 210–222, 2007.
- [67] R. Periañez and A. Pascual-Granged, "Modelling surface radioactive, chemical and oil spills in the Strait of Gibraltar," *Comput. Geosci.*, **34**(2) 163–180, 2008.
- [68] M. LAOTBOZZI, "Répression et prévention de la pollution des navires de commerce en méditerranée," Thèse de doctorat. Paris Est. 2009.
- [69] REMPEC, "Historical Accidents in the Mediterranean Sea integrated in the Mediterranean Integrated Geographical Information System on marine Pollution Risk Assessment and Reponse (MEDGIS-MAR)," 2018. [Online]. Available: <http://medgismar.rempc.org/>. [Accessed: 20-Aug-2018].
- [70] F. J. M. Llácer, "Open registers: Past, present and future," *Mar. Policy*, **27**(6) 513–523, 2003.
- [71] H.P. Ventikos et al., "Spill accident modeling: a critical survey of the event-decision network in the context of IMO's formal safety assessment," *Journal of Hazardous Materials.*, **107**(1–2) 59–66, 2004.
- [72] A. Abdulla, *Maritime traffic effects on biodiversity in the Mediterranean Sea. Volume 1: review of impacts, priority areas and mitigation measures*. IUCN, 2008.
- [73] Ministère de l'Équipement, des Transports, du Logement, du Tourisme et de la Mer (France), "Étude du trafic maritime en Méditerranée occidentale," 2004. <https://www.viepublique.fr/sites/default/files/rapport/pdf/054000140.pdf>.
- [74] ITOPI, "Oil Tanker spill statistics 2013," 2013.
- [75] R. J. Bye and A. L. Aalberg, "Maritime navigation accidents and risk indicators: An exploratory statistical analysis using AIS data and accident reports," *Reliab. Eng. Syst. Saf.*, **176**( February) 174–186, 2018.
- [76] T. Baalisampang, R. Abbasi, V. Garaniya, F. Khan, and M. Dadashzadeh, "Review and analysis of fire and explosion accidents in maritime transportation," *Ocean Eng.*, **158**( September 2017) 350–366, 2018.
- [77] G. Psarros, R. Skjong, and M. S. Eide, "Under-reporting of maritime accidents," *Accid. Anal. Prev.*, **42**(2) 619–625, 2010.
- [78] J. R. Nelson and Tony H Grubestic, "Oil spill modeling: Risk, spatial vulnerability, and impact assessment," *progress Phys. Geogr.*, **42**, 112–127, 2018.
- [79] Xin Liu, Kai W. Wirtz, "The economy of oil spills: Direct and indirect costs as a function of spill size," *J. Hazard. Mater.*, **171**(1–3) 471–477, 2009.
- [80] P. Burgherr, "In-depth analysis of accidental oil spills from tankers in the context of global spill trends from all sources," *J. Hazard. Mater.*, **140**(1–2) 245–256, 2007.
- [81] F. Goerlandt and J. Montewka, "A framework for risk analysis of maritime

- transportation systems : A case study for oil spill from tankers in a ship – ship collision,” *Saf. Sci.*, **76**, 42–66, 2015.
- [82] CCPS, *Guidelines for Chemical Transportation Safety, Security, and Risk Management*. 2008.
- [83] REMRO Network, “Spanish Network of Measurements,” 2018. .
- [84] A. Hardeo, “Two-dimensional offshore oil-spill model for Eastern/Northern Trinidad and Tobago,” *J. Assoc. Prof. Eng. Trinidad Tobago*, **40**(2) 66–72, 2011.
- [85] P. Vethamony, K. Sudheesh, M. T. Babu, S. Jayakumar, and R. Manimurali, “Trajectory of an oil spill off Goa , eastern Arabian Sea : Field observations and simulations,” *Environ. Pollut.*, **148**, 438–444, 2007.
- [86] Wang et al, “Two-dimensional numerical simulation for transport and fate of oil spills in seas,” *Ocean Eng.*, **32**, 1556–1571, 2005.
- [87] C. J. Beegle-Krausel, “GENERAL NOAA OIL MODELING ENVIRONMENT ( GNOME ): A NEW SPILL TRAJECTORY MODEL,” *Int. Oil Spill Conf. Proc.*, **2001**(2) 865–871, 2001.
- [88] X. Chao et al., “Two- and three-dimensional oil spill model for coastal waters,” *Ocean Eng.*, **28**, 1557–1573, 2001.
- [89] Cedric Goery, “Modélisation du transport des nappes d’hydrocarbures en zones continentales et estuariennes,” Université Paris-Est, 2012.
- [90] Malcolm L.Spaulding, “State of the art review and future directions in oil spill modeling,” *Mar. Pollut. Bull.*, **115**( 1–2), 7–19, 2017.
- [91] P. Carracedo et al., “Improvement of pollutant drift forecast system applied to the Prestige oil spills in Galicia Coast (NW of Spain): Development of an operational system,” *Mar. Pollut. Bull.*, **53**( 5–7), 350–360, 2006.
- [92] A. Hardeo and D. P. Chakrabarti, “Vector-Based Oil Spill Model,” *Chem. Eng. Commun.*, **203**( 12), 1656–1665, 2016.
- [93] T. Nordam, R. Nepstad, E. Litzler, and J. Röhrs, “On the use of random walk schemes in oil spill modelling,” *Mar. Pollut. Bull.*, **146**( April) 631–638, 2019.
- [94] R. S. Kankara, S. Arockiaraj, and K. Prabhu, “Environmental sensitivity mapping and risk assessment for oil spill along the Chennai Coast in India,” *Mar. Pollut. Bull.*, **106**( 1–2) 95–103, 2016.
- [95] N. C. Delpeche, T. Soomere, N. C. D. Ellmann, and T. Soomere, “Using Lagrangian models to assist in maritime management of Coastal and Marine Protected Areas Using Lagrangian models to assist in maritime management of Coastal and Marine Protected Areas,” *J. Coast. Res.*, **65**, 36–41, 2013.
- [96] Q. Xu, X. Li, Y. Wei, Z. Tang, Y. Cheng, and W. G. Pichel, “Satellite observations and modeling of oil spill trajectories in the Bohai Sea,” *Mar. Pollut. Bull.*, **71**(1–2), 107–116, 2013.
- [97] Xu, Hong-lei, J. Chen, S. Wang, and Y. Liu, “Oil spill forecast model based on uncertainty analysis : A case study of Dalian Oil Spill,” *Ocean Eng.*, **54**, 206–212, 2012.
- [98] J. Fay, “Physical processes in the spread of oil on a water surface.,” *Proceeding Jt. Conf. Prev. Control oil spills*, 1971.
- [99] A. Berry, “Development of OILTRANS Model code,” 2011.
- [100] A. Berry, T. Dabrowski, and K. Lyons, “The oil spill model OILTRANS and its application to the Celtic Sea,” *Mar. Pollut. Bull.*, **64**( 11) 2489–2501, 2012.
- [101] R. Periañez and F. Caravaca, “A set of rapid-response models for pollutant dispersion assessments in southern Spain coastal waters,” *Mar. Pollut. Bull.*, **60**( 9) 1412–1422, 2010.
- [102] W. J. Guo and Y. X. Wang, “A numerical oil spill model based on a hybrid method,” *Mar. Pollut. Bull.*, **58**( 5) 726–734, 2009.
- [103] C. Ambjorn, “Seatrack web, forecasts of oil spills, a new version,” *Environ. Res. Eng. Manag.*, 60–66, 2007.
- [104] D. P. McCay-French, “Oil Spill Impact Modeling: Development & Validation,” *Environ. Toxicol. Chem.*, **23**( 10) 2441–2456, 2004.
- [105] ASCE, “State-of-the-art review of modeling transport and fate of oil spills. Committee on Modeling Oil Spills. Water Resources Engineering Division.,” *J. Hydraul. Eng.*, **122**( 11) 594–609, 1996.
- [106] A. H. Al-Rabeh, H. M. Cekirge, and N. Gunay, “A stochastic simulation model of oil spill fate and transport,” *Appl. Math. Model.*, **13**( 6) 322–329, 1989.
- [107] A. Soussi et al., “An oil spill trajectory model: Validation in the Mediterranean Sea,” *ISSE 2019 - 5th IEEE Int. Symp. Syst. Eng. Proc.*, 2019.
- [108] N. Mareï, “Le détroit de Gibraltar dans la mondialisation des transports maritimes,” *EchoGéo*, **19**, 0–15, 2012.
- [109] M. N. Tsimplis and H. L. Bryden, “Estimation of the transports through the Strait of Gibraltar,” *Deep. Res. Part I Oceanogr. Res. Pap.*, **47**( 12), 2219–2242, 2000.
- [110] H. Perkins, T. Kinder, and P. La Violette, “The Atlantic inflow in the Western Alboran Sea,” *J. Phys. Oceanogr.*, **20**, 242–263, 1990.
- [111] H. BENALI, “Evolution diachronique multitudes, sédimentologique, géomorphologique et occupation de l’espace du littoral marocain de Ksar Esghir et son arrière pays (Maroc),” Abdelmalek essaâdi, 2016.
- [112] A. El Gharbaoui, “La terre et l’homme dans la péninsule tingitane: étude sur l’homme et le milieu naturel dans le Rif Occidental,” Institut scientifique, 1981.
- [113] J. P. Thauvin, “Carte géotechnique de Tanger: Le climat à Tanger.,” *Notes du Serv. géologique du Maroc.*, **222**, 29–38, 1991.

## Optimization of the Procedures for Checking the Functionality of the Greek Railways: Data Mining and Machine Learning Approach to Predict Passenger Train Immobilization

Ilias Kalathas\*, Michail Papoutsidakis, Chistos Drosos

Department of Industrial Design and Production Engineering, University of West Attica, Athens, 15354, Greece

### ARTICLE INFO

Article history:

Received: 17 June, 2020

Accepted: 19 July, 2020

Online: 28 July, 2020

Keywords:

Machine learning

Railway

Train Immobilization

Data mining

Predictive Model

Malfunctions

Diagnosis

### ABSTRACT

Information is the driving force of businesses because it can ensure the ability of knowledge and prediction. The railway industry produces a huge amount of data, with the proper processing of them and the use of innovative technology, there is the possibility of beneficial information to be provided which constitute the deciding factor for the correct decision making. Safety is the railway comparative advantage that has to be reinforced by each business administration while making the optimum decisions. The main purpose of this paper is the investigation of the most important dysfunctions that arise in a train and can cause its immobilization at the main passenger rail, resulting in huge delays of conducting the routes setting the passengers at risk. Afterwards the total of malfunctions is assessed and the most important, potentially, malfunction is assessed, so as the executives of the Greek Railway company to plan and redefine the processes and the initial plan of the predictive maintenance. This paper demonstrates the effort of implementing innovative applications by making use of methods from the rapidly developed field of Data Mining to the Greek Railway Company that uses obsolete procedures for the control of the trains' functionality in order to investigate the data for the provision of specialized information which will be used as a tool for the faster, more accurate and precise decision making. This decision making approach is based on a specific algorithm's design in order to automatically detect faults and make periodic maintenance of trains easier. Holistic approach is performed in the management of real data from the Greek railway industry and a predictive model of Machine Learning is developed, for the optimization of the management's performance of the trains reinforcing the strategic target of the railway industry which is the transportation of citizens with safety and comfort.

### 1. Introduction

Businesses are more and more using sets of data in order to conduct their decisions. Developments in the field of Data Mining and the Machine Learning are expected to predominate in 2020 and to create, within the next decade, significant opportunities to all the companies [1]. The emerging technologies change the way that businesses collect and extract useful information from the data [2]. The Data Mining is an effective method for the analysis of a huge amount of collected data that extracts useful information. The Machine Learning is a field that was developed by the artificial intelligence and assists planning and development of algorithms and the eution of the performance related to empirical-operational data [3]. The railway industries are a field the performance of

which progressively depends on their ability to extract information from complex sets of data and take the optimum action in real time [4]. The Data mining in conjunction with the Machine Learning has the capacity to improve the operational progress raising the level of efficiency in decision making and the overall procedures [5].

The railway is a system of passengers' and merchandise transportation with wheeled vehicles (trains) that roll on rails. The railways as a mean of transport is defined by three components, the functioning utilization, the infrastructure, and the rolling stock. With the term rolling stock we refer to all kinds of vehicles pulled or driven on rails that perform railway transports [6]. The Railway Rolling Stock is a complicated system the smooth operation of which plays a leading role on the exploitation of the rail system and is subject to progressive lesions (wears, erosions, malfunctions

\*Corresponding Author: Ilias Kalathas, University of West Attica,  
Tel: +306974731434, i.kalathas@uniwa.gr



etc.) and therefore it is obliged to accept interventions of restoration – maintenance in order to ensure reliability, availability and security of its circulation [7]. The high quality of maintenance services of the rolling stock is an essential precondition for the smooth function of the whole railway system. The predictive models of machine learning are able to provide the management of railway industries with the appropriate information that will assist on making decisions for the restart of the process restoration - maintenance of rolling stock aiming to the increase of its reliability- effectiveness as well as the minimization of hazards and cost [8].

## 2. Problem Statement

For the smooth and orderly operation of the railway network necessary requirement is to avoid train immobilization on the main passenger track which can cause a traffic stop and consequently great delays in itineraries conduct and naturally question the most important goal, which is the passengers’ safety [9] [10].

In this research work it is presented a method in order to optimize the management of the Greek Railway Company STA.SY. S.A. The prediction, diagnosis and the dealing with malfunctions that can immobilize a train in the main passenger track is the purpose of the application. It is trying to discover – detect which malfunction or malfunctions combined are capable of immobilizing a train. The effort of this research focuses on the way the application of Data Mining for the construction of a predictive model of Machine Learning which will be used as a tool for the output of useful results to the management in order to ensure faster and right decisions. The model of machine learning will meaningfully assist in the infrastructure and the improvement of the maintenance of trains, in the effective detection of malfunctions and the simplification of the programming relevant to the train traffic.

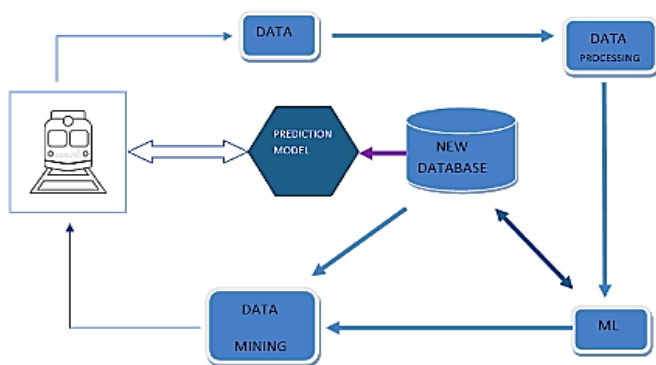


Figure 1: Predictive Malfunctions Solution

## 3. Theoretical background

Machine learning is a branch of artificial intelligence that covers the study or the construction of systems able to “learn” from a multitude of data. Machine learning comprises a field of information technology and belongs to the wider sector of artificial intelligence [11]. Algorithms of machine learning can learn from data and make predictions for new data that have not been dealt with.

### 3.2 Supervised learning

Supervised learning is a section of Machine learning and it is defined as the process of an algorithm constructing a function that depicts given inputs to known desired outputs, with ultimate aim to apply a model of prediction to new data that have not been seen before. It is used with problems of classification, prediction, and interpretation [12]. The creation of a model of prediction, which is necessary for the construction of a mechanism in order to make decisions or calculations, is achieved with supervised learning.

### 3.3 Data mining

The data mining is one of the stages of the process Knowledge Discovery in Databases – KDD. The assignments that are carried out at this stage incorporate algorithm applications for the construction of algorithmic models emphasizing on the discipline of Machine Learning [13]. Data mining is defined as the process of constructing a categorization model or prediction, with the use of algorithms aiming at the description and the prediction of data set for the production of exportable potentially useful knowledge.

### 3.4 Classification

Classification is one of the basic techniques of Mining Data, considered to be the most popular process because of the abundance of the effective applications. It is about a predictive technique aiming at the construction of a predictive model applying classification algorithms to present data and it is more easily translated than compared with the models which produce other techniques of data mining [14]. Classification is defined as the process of evaluating the most probable value a dependent variable will take, based on the already known data from the independent value of a known data set in order to predict a nominal discrete value. The classification algorithms are divided in five methods, Classification with Decision Trees, Classification with Neural Networks, Classification with Bayes Criteria, Classification with Support Vector Machines, and Classification with k – Nearest Neighbors [15].

### 3.5 Decision Trees.

The Decision Trees is one of the most widely known method of classification. It is one of the techniques of Machine Learning the construction of which is aiming at, the sequential split of a set of observations into subsets. It is easily apprehended by people because it applies simple representations for the classification of examples, and it is often used in the mining data sector [16]. The basic purpose of the Decision Trees is the prediction of the variable value that is trained by data created by other independent variables [17].

- Algorithm ID3 There are various algorithms for the data classification and the forecasting using the Decision Tree method, but the Iterative Dichotomies or ID3 has predominated along with its subsequent developments, the C4.5 and its commercial version C5.0 [18]. Its main characteristic is that it presupposes the existence of nominal fields only creating a tree with multiple courses (top down) that uses nodes with categorical variables which attract great sets of data at the beginning in order to be minimized during their courses.

- Algorithm C4.5 The algorithm C4.5 has been created by the ID3 which is its eution. Their main difference is the criterion of data separation that is called Gain Ratio and it is defined by the equation.

$$\text{GainRatio}(S,A)=\frac{\text{information}(S,A)}{\text{entropy}(s,a)} \quad (1)$$

The Gain Ration sets the rules, it regularizes the information gain as to the entropy. The Gain Ration improves the precision and minimizes the complexity of the trees [19]. An additional important difference between the two algorithms is that the ID3 produces a tree with multiple courses that uses nodes the categorical variables whereas the C4.5 can use a field of number signs setting border lines in order to split the data whether they are above the limit , under the limit or on the limit.

### 3.6 Weka

The WEKA is a software of Machine Learning of open source, its name comes from the initials Waikato Environment for Knowledge Analysis and it is used for Mining Data. It was developed at the University of Waikato in New Zealand and its development goes on from an international team of programmers [20]. It is publicly available according to the license terms GNU General Public License, which allows users its execution but also the making of free changes in the software and its wide acceptance [21]. WEKA provides graphical user interfaces which allow its application to users which don't provide knowledge of programming and it is capable of installing in all modern software platforms because it is written in Java programming language [22]. The software suit WEKA enables the preparation and visualization of the data and provides various techniques of knowledge extraction from data as the classification, clustering, association rules mining, and prediction [23]. The use of WEKA software is encountered in a great number of scientific papers and many books of mining Data, since the set of functions it offers in combination with the graphical reproduction of algorithms contributes to the analysis of data and the construction of predictive models.

## 4. Related work

In the present section of this work are briefly presented relevant researches and experiments regarding the techniques of machine learning, with the use or not of software WEKA aiming at data mining and prediction and they constitute the starting point for the development of a smart supporting tool of innovative processes. In the meanwhile, a comparison of machine learning models, which have been used in various transport sectors, takes place for the prevention and diagnosis of malfunctions in order to make the best decisions.

The article [24] refers to the detection and diagnosis of faults in parts of the rails of the railway network. It acknowledges that faults in railway tracks can cause catastrophic consequences to the passengers' safety. It presents a methodology for the in time detection of abnormalities in parts of the rails using sound data that have been collected by a rail vehicle model NS-AM Sehwa Company in Daejeon, South Korea, on the 1st of January 2016. Two different experiments take place (one for the total of data and one for the data characterized faulty) with the use of SVM algorithm at the WEKA software The results show that the system

allows the effective detection and precise diagnosis of malfunctions that exceed 94.1 %. The proposed method provides reliable means of investigation of the railways for the comprehension of the railway conditions.

This research [25] refers to the use of Apriori algorithm and to clustering algorithm of WEKA tool that are applied to a set of data for road accidents at the area Alghat that belongs to the town Riyadh Province, in Saudi Arabia. The aim of this research is to find new approaches and new rules of connection between the set of movement data using the WEKA tool in order to discover new factors that cause road accidents. The study acknowledges that the software of machine learning WEKA is a very useful tool for mining data, which allows the user to choose from various algorithms and to compare them, in order to accurately reach the demanding results.

In addition, in the traffic accidents analysis based on the algorithm C4.5 using the WEKA software of machine learning refers to the article [26]. The study acknowledges that the data mining contributes to the prevention of accidents and the safety management of traffic. The research was held in China, in a part of Wenli motorway which is located in the central part of Zhejiang province connecting Lishui town with Wenzhou and it was based on the traffic accidents data from 2006 to 2013. Using a Decision Tree and applying the algorithm C4.5 to WEKA the effect of various factors which affect an accident was tested. The model that was created can effectively handle large sets of data and achieve total precision of prediction 80% having as a result to be considered a method capable of traffic control.

The article [27] refers to the investigation of the use of algorithms of machine learning in shipping. It assesses the smart predictive methods in two stroke low speed marine engine aiming to the effective tracking and classification of the faults/malfunctions displayed since the in time tracking of the malfunction secures the nonstop rhythm of the ship but also less consumption of fuels. The potential that the tool of mining data WEKA offers is used for the application of the optimum predictive tracking and prediction system of faults/failures. The experimental research was based on the collection and processing of engine data MAN BW 7S60MC using the engines' simulator of the Faculty of Engineers and Merchant Academy of Aspropirgos / Greece. The methods presented showed that they provide reliable diagnostic tools setting the method "AdaBoost", as the right choice with rate 96.5 %.

The reference [28] applies four techniques of machine learning for the automatic prediction of traffic maintenance using functions and historical records. Its purpose is to create a list of prioritizing tasks in order to avoid its future decline. The case study was held at the Portuguese motorway network of 539 km total length managed by the Infrastructures de Portugal Company and four techniques of machine learning DT, KNN, SVM and ANN were used for the optimum selection in making decisions. The results show that with enforced data base from historical facts a satisfying model of prediction is provided for the avoidance of adverse future conditions in the motorway network.

The article [29] refers to the prediction of airline routes delay with the use of machine learning. It recognized factors that are able to cause delay in flights (e.g. visibility, temperature, wind strength,

age and type of aircraft) the techniques of machine learning were applied (decision tree, k-means, J48, random forest and Bayesians classification) for the prediction of delay and the size of the incident. The tests were held on American flights data and on a big Iranian airlines network. With the methodology suggested the estimations for the flights' delay reaches more than 70% precision in the U.S.A and Iran.

The doctoral thesis [30] aims to the investigation of the possibilities of a ship to reduce fuel consumption of pollutant emissions based on the data from machine records. It presents results in order to save energy using tools of machine learning in real conditions. The data comes from machines' records of the cruise ship M/S Birka Stockholm in the Baltic Sea during four journeys between the years 2014 and 2015. Using algorithms and the method of supervised learning with the regression technique and the results showed that there is satisfying prediction ability of fuel consumption. The main conclusion of the doctoral thesis is that the data from a ship's tasks can be used for the observation of maintenance and performances but also to provide a solid base for the construction of a model that will improve energetically and functionally the performance of a ship.

### 5. Data set and Pre-Processing

The data of the present Case Study came from the informative system BAAN – RSD that the STA.SY S.A. uses and especially the Sector of the Rolling Stock Line 1 (former ISAP). The initial set of data with no processing contained 1000 instances and 24 attributes. The useful – exploitable data were derived from data bases of the company using SQL questions and were extracted in Excel spreadsheet format, which contained fundamental data – information (480 instances and 8 attributes) for the trainsets condition in a 6month period from 1/6/2019 to 31/12/19

Table 1: Attributes Considered

a/s	Attributes	Explanations of attributes
1	Number of trainsets	It includes coding of all the trainsets and aims at its coordinate observation and recording.
2	Date's coding	It includes the date that the trainsets were received and started their operation.
3	Operational system	It includes the operating levels of the trainset that are divided based on the equipment used.
4	Subsystem	It includes the subsystem of the equipment of the trainset with the same features.
5	Indications	It includes the probable problem of the trainset in an operative system.
6	Failure	It includes the problem that sets the trainset out of order.
7	Kilometres	It includes the kilometres travelled of a trainset from one maintenance to the other.
8	Restoration of the failure	It includes the way that the problem was dealt with using specialized alternatives.

The addition / improvement or delete / extraction of inappropriate information from the set of data, with appropriate criteria is the necessary precondition for the creation of a satisfying data base, appropriate to the research [31]. The appropriate criteria to be met are:

- The antiquity of the trainsets (dates arranged in year and months of the delivery of the trainsets).
- Malfunctions that set in dispute the most significant target of the railways which is the passenger's safety.
- The gravity of the malfunctions affects the traffic of the trainsets.
- The periodicity of the malfunctions that the trainsets face.

The process continued with the discretion of the data from Excel spreadsheet applying special filters that are based on the appropriate criteria. The failures that arise and are capable of causing malfunction to the traffic of the trainsets are:

- Main switch overcurrent cut-off (code: SWITCH)
- Door malfunction (code: DOOR)
- Generator failure (code: GENERATOR)
- Braking during course (code: BRAKING)

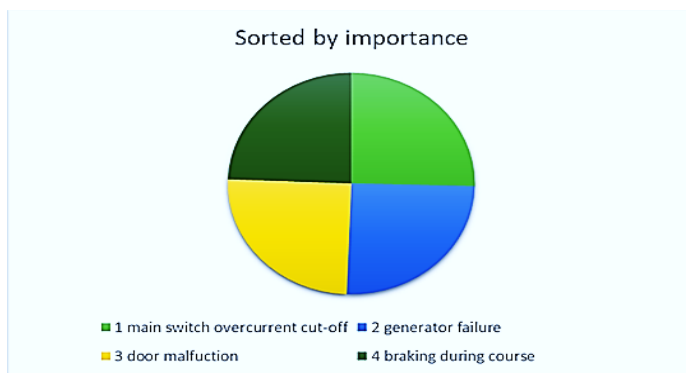


Figure 2: Results for malfunctions to the traffic of the trainsets

The braking during course (code: BRAKING) is divided in crucial (code: A) the cause of the failure is elicited from the malfunction of the engine of the vehicle and simple (code: B) the cause of the failure is elicited from the electronic display unit of the trainset.

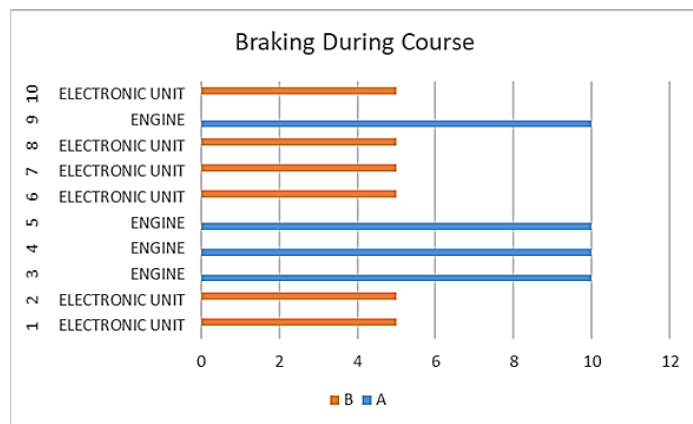


Figure 3: Results for malfunctions for braking during course

The number of repetitions, the importance of the failures that are able to cause malfunction to the traffic of the trainsets, during the 6month period of time, also defined the periodicity of the malfunctions.

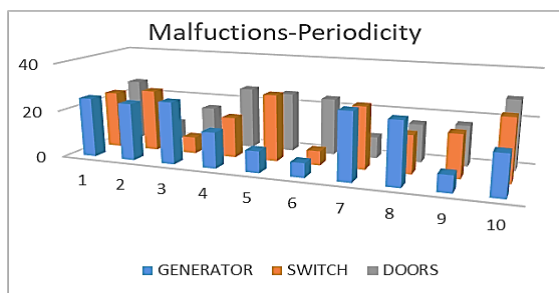


Figure 4: Results for Failure periodicity

- Frequent repetition: repetition from 20 to 30 times (code: FREQUENT)
- Moderate repetition: repetition from 10 to 19 times (code: MODERATE)
- Sparse repetition: repetition from 1 to 9 times (code: SPARSE)

Meanwhile, it is considered advisable to record that the Register of the trainsets referring to the date of delivery (antiquity) of the trainset.

Table 2: Receipt dates

Train register	Years	Months	Months	Months
8	1980	1 to 4	5 to 8	9 to 12
10	1992	1 to 4	5 to 8	9 to 12
11	2003	1 to 4	5 to 8	9 to 12

The result of discrimination in the division of the initial data in categories that can be easily apprehended and create a new base of data with malfunctions leading to immobilize a trainset on the main passengers' track.

TRAIN REGISTER	MALFUNCTIONS- PERIODICITY				RESULT
TRAIN INVETERACY	GENERATOR	SWITCH	DOORS	BRAKING	PROBLEM-STANDSTILL TRAIN
8	FREQUENT	FREQUENT	FREQUENT	B	YES
8	FREQUENT	FREQUENT	SPARSE	B	YES
10	FREQUENT	SPARSE	MODERATE	A	NO
10	MODERATE	MODERATE	FREQUENT	A	YES
8	SPARSE	FREQUENT	FREQUENT	A	YES
10	SPARSE	SPARSE	FREQUENT	B	NO
11	FREQUENT	FREQUENT	SPARSE	B	NO
11	FREQUENT	MODERATE	MODERATE	B	YES
11	SPARSE	MODERATE	MODERATE	A	YES
11	MODERATE	FREQUENT	FREQUENT	B	NO

Figure 5: New database

The criteria for the selection of the appropriate method and the optimum techniques of machine learning – data mining is:

- The correctness / reliability of the data
- The amount of data
- The suitability of data
- The dynamic performance of data
- The constant update
- The readjustment (with the entry of new data)
- The satisfactory response to the research
- The definition of the objective purpose

In the scientific – investigative research in order to find a predictive model that will pinpoint which malfunction or malfunctions combined (generator failure, main switch overcurrent cut off, door malfunction and braking during course) are able to immobilize a trainset, the Categorizing method is advised (data mining technique). Therefore, the next step is the selection and application of the appropriate Categorizing method. In the present study case, the method applied for the creation of a predictive model and malfunction diagnosis that meets the appropriate criteria is the method of decision tree using the algorithm C4.5. The techniques above are provided by the software of machine learning WEKA.

## 6. Data Processing

The operation of WEKA software requires specific file formats. The WEKA software converts and saves as an ARFF file the data base that contains all the data that will be used [32].

```

%relation ENGLISH

@attribute TRAIN INVETERACY real
@attribute GENERATOR {SPARSE, FREQUENT, MODERATE}

@attribute SWITCH {SPARSE, FREQUENT, MODERATE}

@attribute DOORS {SPARSE, FREQUENT, MODERATE}

@attribute BRAKING {B,A}

@attribute PROBLEM-STANDSTILL {NO, YES}

@data
8, FREQUENT, FREQUENT, FREQUENT, B, YES
8, FREQUENT, FREQUENT, SPARSE, B, YES
10, FREQUENT, SPARSE, MODERATE, A, NO
10, MODERATE, MODERATE, FREQUENT, A, YES
8, SPARSE, FREQUENT, FREQUENT, A, YES
10, SPARSE, SPARSE, FREQUENT, B, NO
11, FREQUENT, FREQUENT, SPARSE, B, NO
11, FREQUENT, MODERATE, MODERATE, B, YES
11, SPARSE, MODERATE, MODERATE, A, YES
11, MODERATE, FREQUENT, FREQUENT, B, NO
    
```

Figure 6: ARFF file

Afterwards the immediate presentation of data takes place in the form of diagrams

- The first table GENERATOR presents in red the immobilization of a trainset after “generator failure” malfunction, whereas in blue when the “generator failure” malfunction does not result in immobilizing the trainset. They are additionally presented, in order, the periodicity malfunctions “SPARSE”, “FREQUENT”, “MODERATE”.
- The second table SWITCH presents in red the immobilization of a trainset after “main switch overcurrent cut-off” malfunction does not result in immobilizing the trainset. They are additionally presented the choices of frequency malfunctions “SPARSE”, “FREQUENT”, and “MODERATE”.
- The third table DOORS presents in red the immobilization of a trainset after a “door malfunction”, whereas in blue when the “door malfunction” does not result in the immobilization of a train set. They are additionally presented, in order, the choices of frequency malfunctions “SPARSE”, “FREQUENT”, and “MODERATE”.



- The fourth table BRAKING presents overall immobilizations of trainsets, in red the immobilization of a trainset after the “braking during course” malfunction, and in blue when the “braking during course” malfunction does not result in immobilizing the trainset. It is additionally presented in red the significance of the malfunction “A” or “B”.
- The fifth table presents the results of the immobilization or not of a trainset, in blue the non-immobilization of a trainset, whereas in red the immobilization of a trainset.

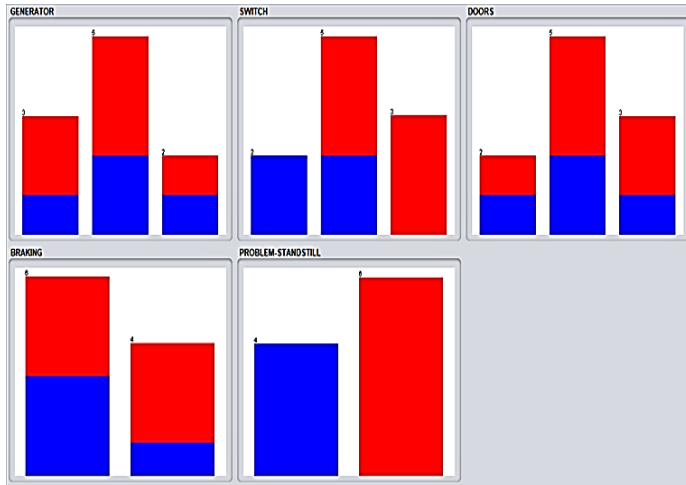


Figure 7: Data in the form of charts

The results after the completion of the algorithm are the following:

== Run information ==

```

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    ENGLISH-weka.filters.unsupervised.attribute.Remove-R1
Instances:   10
Attributes:  5
             GENERATOR
             SWITCH
             DOORS
             BRAKING
             PROBLEM-STANDSTILL
Test mode:   evaluate on training data
    
```

== Classifier model (full training set) ==

J48 pruned tree  
-----

```

SWITCH = SPARSE: NO (2.0)
SWITCH = FREQUENT: YES (5.0/2.0)
SWITCH = MODERATE: YES (3.0)
    
```

```

Number of Leaves :    3
Size of the tree :    4
    
```

Time taken to build model: 0 seconds

Figure 8: Run information of the algorithm

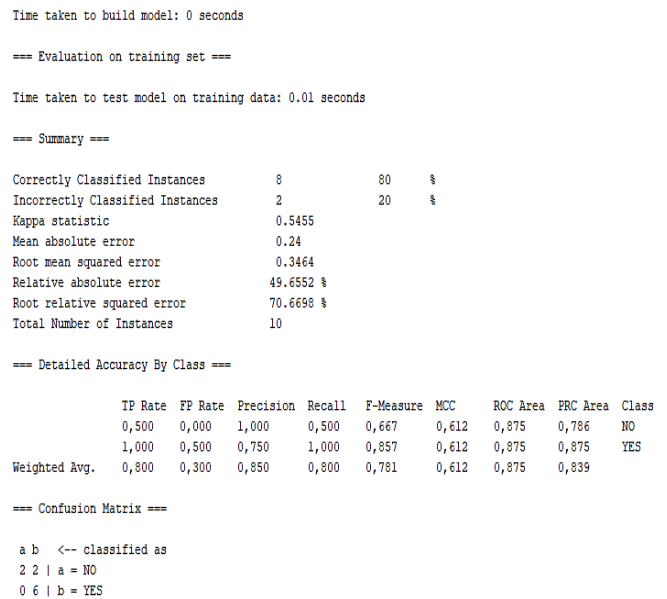


Figure 9: The result of the algorithm

The WEKA software creates – visualizes the requested decision tree

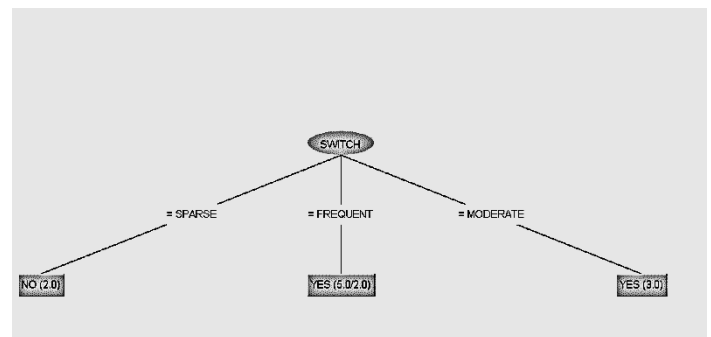


Figure 10: The output of C4.5 tree.

After the completion of the algorithm the conclusion that arises is that the “main switch” is the most important of malfunctions, which leads to the immobilization of a trainset. Setting the constant of antiquity to the algorithm an additional decision tree appears, whereas the root of the tree remains the same.

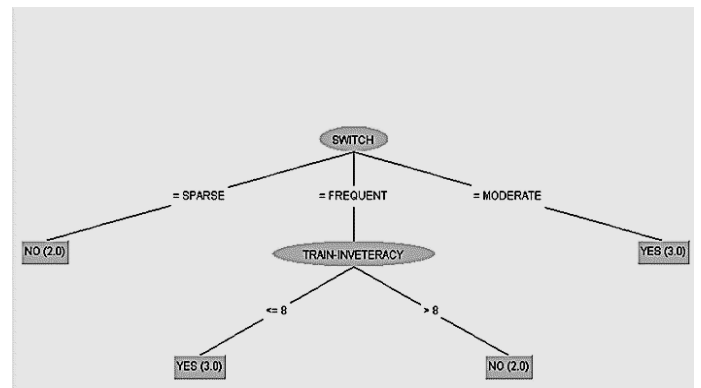


Figure 11: The output of C4.5 tree with variable ‘train inveteracy’

From the final decision tree, the outcome is that the most important problem for the immobilization of a trainset is the

malfunction of the “main switch” and the antiquity of the trainset follows.

If another occasional option of the WEKA software is used, like the random tree it seems that the decision tree is different, which means that the software cannot decide which malfunction or combination of malfunctions will immobilize the trainset.

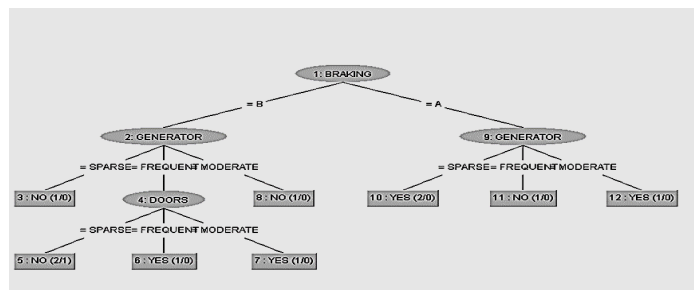


Figure 12: The output of random tree

### 7. Results and Discussion

For the precision of the characteristics (figure 12) used an assessment of data was held related to the periodic written reports of the maintenance technicians. The results presented totally 90 % success decreasing the danger of choice to minimum.

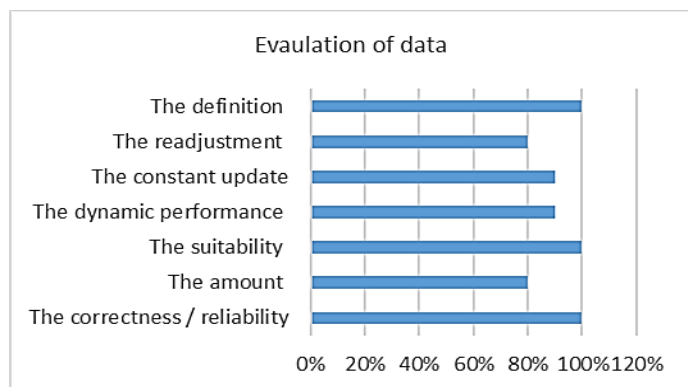


Figure 13: Success Rate Data

The algorithms assessment was held under specific criteria as the quantity, the adequacy, and the dynamic representation of data regarding the decision trees and the characteristic data of the clues.

The advantages of the C4.5 algorithm (figure 13) which are the speed and its independence from the classification model makes it the optimum choice. During the repeated procedure (5 times) the result appeared almost the same with an average of 80% success. [33]

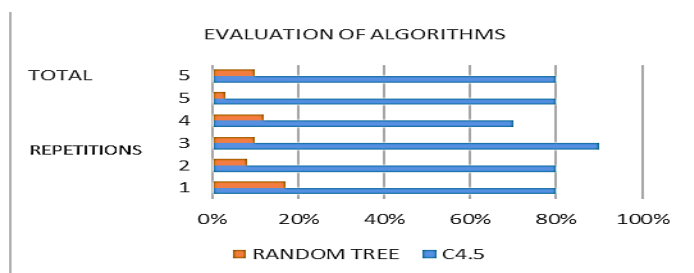


Figure 14: Success Rate Algorithms

During the pre-processing of data (figure 2) were presented the main malfunctions which affect the train movement and are capable of immobilizing the train at the main passenger rail. The malfunctions Main switch overcurrent cut-off (code: SWITCH), Door malfunction (code: DOOR), Generator failure (code: GENERATOR), Braking during course (code: BRAKING) appeared on every train in 25% each. The malfunction Braking during course (code: BRAKING) (figure 3) which is divided into crucial (code: A) and simple (code: B) appeared in a 6-month period as simple (code: B) 6 times and as crucial (code: A) 4 times. The periodicity of malfunctions (figure 4) in each train is not the same. Table 3 depicts the periodicity of each malfunction in every train.

Table 3: Periodicity-Malfunction

TRAIN	GENERATOR	SWITCH	DOOR
No 1	Repetition from 20 to 30 times	Repetition from 20 to 30 times	Repetition from 20 to 30 times
No 2	Repetition from 20 to 30 times	Repetition from 20 to 30 times	Repetition from 1 to 9 times
No 3	Repetition from 20 to 30 times	Repetition from 1 to 9 times	Repetition from 10 to 19 times
No 4	Repetition from 10 to 19 times	Repetition from 10 to 19 times	Repetition from 20 to 30 times
No 5	Repetition from 1 to 9 times	Repetition from 20 to 30 times	Repetition from 20 to 30 times
No 6	Repetition from 1 to 9 times	Repetition from 1 to 9 times	Repetition from 20 to 30 times
No 7	Repetition from 20 to 30 times	Repetition from 20 to 30 times	Repetition from 1 to 9 times
No 8	Repetition from 20 to 30 times	Repetition from 10 to 19 times	Repetition from 10 to 19 times
No 9	Repetition from 1 to 9 times	Repetition from 10 to 19 times	Repetition from 10 to 19 times
No 10	Repetition from 10 to 19 times	Repetition from 20 to 30 times	Repetition from 20 to 30 times

The results provide useful – necessary information, for the research conduction:

- The dysfunctions of the train can be pinpointed and immobilize the train are four: Main switch overcurrent cut-off (code: SWITCH), Door malfunction (code: DOOR), Generator failure (code: GENERATOR), Braking during course (code: BRAKING). The importance of each malfunction is the same with 25%.

- The malfunction Braking during course (code: BRAKING), is divided into crucial (code: A) and simple (code: B) and appears 10 times in total.
- The frequency of each malfunction Door malfunction, Generator failure, Main switch overcurrent cut-off in each train is different but equally important.

After the data analysis and processing with the use of WEKA software (figure 10 and figure 11) arises:

- From the four main and equivalent malfunctions, the most important malfunction, in potential, on which the executives of a company must focus is the Main switch overcurrent cut-off (code: SWITCH) and after that the trains' antiquity which is not referred as a dysfunction but it exists as data.
- The prediction which takes place for the most important malfunction, potentially, which can immobilize a train in the main passenger rail, offers to the executives of the company the opportunity to focus on the functional systems of the train related to the main switch, rescheduling the plan of predictive maintenance.

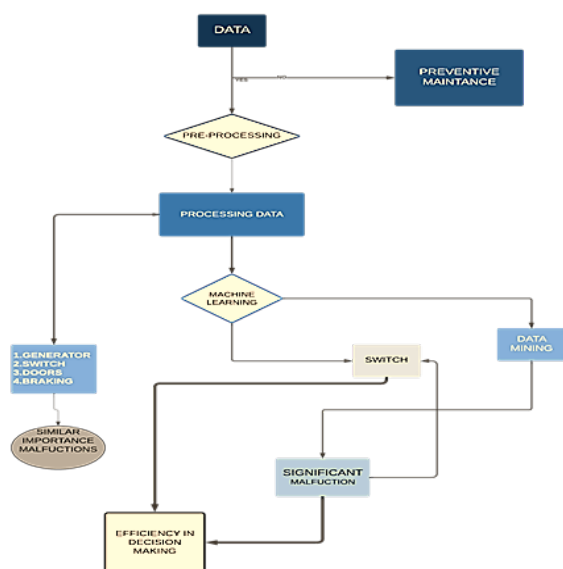


Figure 15:Flow chart diagram for Decision-making process

## 8. Conclusion

In the present paper the realization of the application of machine learning took place aiming to the prediction, diagnosis and dealing with malfunctions which immobilize a trainset on the main passengers' rail.

Detailed pre-processing of the data was carried out in the huge database provided by the Greek railway company STA.SY. S.A., resulting in the detection of malfunctions of equal importance (main device power failure, door failure, generator failure, braking during the course) capable of immobilizing the train on the main passenger rail.

Consecutively a new data base was created consisting of specialized - quality features, advantageous to the research like the trainsets' record, the significance, and the periodicity of malfunctions.

With the use of the Machine Learning software (WEKA) the most important factors which immobilize a trainset on the main passengers' rail were investigated and pinpointed. The results that arose from the method that was used are, that the malfunction "main switch overcurrent cut-off" and then the antiquity – year of the trainset, have the greatest significance on the immobilization of a trainset.

Additionally, the combination of simplicity and the clarity of the decision trees, with the use of strict criteria, for the accuracy, the quantity, the appropriacy, and the dynamic representation of the data, confirmed the accurate use of the C4.5 algorithm giving 80% result accuracy whereas the use of the random tree algorithm, with the same criteria, did not produce any quality results.

Finally, a new process was created which can successfully classify the malfunctions in order to make a precise and accurate prediction for the immobilization of the trains on the passengers' rail.

The suggested approach on the data analysis with the use of machine learning suggested to the STA.SY Company can develop a better method for the control of the trains' circulation.

The creation of a machine learning template for the prediction and mining data is capable of constituting the main tool for the improvement of the process of making decisions from the management, aim to the better programming and the effective management of the maintenance of trainsets, setting as dominant priority, the passengers' safety.

The new method of the malfunctions' classification with the use of innovative technologies, sets as main priority the passengers' safety even the maintenance procedures and provides the company's executives with new knowledge to take the right decisions planning new maintenance processes.

The Data Mining is located in the center of every "smart", adaptive system that we can think of, and can offer huge benefits to the Railway companies. Whatever produces data, constitutes a potential target of Data mining, revealing therefor its importance and usefulness as a technology of the Future. In the meantime, the Machine Learning as a new approach to the business performances of railways, constitutes an interesting idea questioning the traditional ways procedures that were reliable in the past but now they have started showing their limitations.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

All authors would like to thank the University of West Attica for the financial support provided to them to undertake this research project.

## References

- [1] M. Weske "Business Process Modelling Foundation. In: Business Process Management", Published by Springer, 2019
- [2] P. Aida-Maria, "Business Intelligence Methods for Sustainable Development of the Railways", Database Systems Journal 6(2), 48-55, 2015. <https://EconPapers.repec.org/RePEc:aes:dbjour:v:6:y:2015:i:23>

- [3] E. Kyrkos, 'Business intelligence and data mining' [eBook] Athens: Hellenic Academic Libraries Link. Chapter 4. 2015. <http://hdl.handle.net/11419/1231>
- [4] L. Dai 'A machine learning approach for optimization in railway planning, PhD Delft University of Technology, 2018
- [5] Petri, M., Pratelli, A., & Fusco, G. 'Data Mining and Big Freight Transport Database Analysis and Forecasting Capabilities'. Transactions on Maritimes Science, 2016. <https://doi.org/10.7225/toms.v05.n02.001>
- [6] J.C. Wagenaar, J.C., Kroon, L.G., Schmidt, M. 'Maintenance Appointments in Railway Rolling Stock Rescheduling'. Transportation Science, 51(4) 1138-1160., 2017. <https://doi.org/10.1287/trsc.2016.0701>
- [7] D. Ronanki, S. A. Singh and S. S. Williamson, 'Comprehensive Topological Overview of Rolling Stock Architectures and Recent Trends in Electric Railway Traction Systems,' in IEEE Transactions on Transportation Electrification, 3(3) 724-738, 2017. <https://doi.org/10.1109/TTE.2017.2765518>
- [8] E. Bosscha, 'Big Data in railway operation: using artificial neural networks to predict train delay propagation', University of Twente, PhD Thesis, 2016
- [9] I. Öztürk, G. Güner, E. Tümer, 'The Root Causes of a Train Accident: Lac-Mégantic Rail Disaster' Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018). Advances in Intelligent Systems and Computing, 823. Published by Springer, 2018. [https://doi.org/10.1007/978-3-319-96074-6\\_21](https://doi.org/10.1007/978-3-319-96074-6_21)
- [10] M. Heidarysafaa, K. Koesari, L. E. Barnes, Brown 'Analysis of Railway Accidents-Narratives using Deep Learning' IEEE International Conference on Machine Learning and Application., 2018. <https://doi.org/10.1109/ICMLA.2018.00235>
- [11] Judith Hurwitz, Daniel Kirsch 'Machine learning' IBM Limited Edition, Published by John Wiley & Sons, Inc., 2018.
- [12] G.R. Devi, Karpagam, V. Kumar, 'A survey of machine learning techniques, Int. J. of Computational Systems Engineering, 3(4), 203 – 212, 2017. <https://doi.org/10.1504/IJCSYSE.2017.089191>
- [13] R. Changala, D.R. Rao, T. Janardhana, P.K. Kumar, Kareemunnisa 'Knowledge Discovery Process: The Next Step for Knowledge Search' International Journal of Innovative Research in Computer and Communication Engineering 3(5), 1-6, 2015. <https://doi.org/10.15680/ijirccce.2015.0305127>
- [14] S.S. Bhaskaran 'An Investigation into the Knowledge Discovery and Data Mining (KDDM) process to generate course taking pattern characterized by contextual factors of students in Higher Education Institution (HEI) , PhD Thesis, Brunel University, London, 2017
- [15] Song, Yan-Yan, and Ying Lu. "Decision tree methods: applications for classification and prediction." Shanghai archives of psychiatry 27(2), 1-6, 2015. <https://doi.org/10.11919/j.issn.1002-0829.215044>
- [16] I.D. Mienyea, Y. Suna, Z. Wang 'Prediction performance of improved decision tree-based algorithms: a review' 2nd International Conference on Sustainable Materials Processing and Manufacturing, Published by Elsevier, 2019. <https://doi.org/10.1016/j.promfg.2019.06.011>
- [17] M. Batra, R. Agrawal, 'Comparative Analysis of Decision Tree Algorithms. In: Panigrahi B., Hoda M., Sharma V., Goel S. (Eds) Nature Inspired Computing. Advances in Intelligent Systems and Computing, 652. Springer, Singapore, 2018. [https://doi.org/10.1007/978-981-10-6747-1\\_4](https://doi.org/10.1007/978-981-10-6747-1_4)
- [18] S.B. Begenova, T.V. Avdeenko, 'Building of fuzzy decision trees using ID3 algorithm': Journal of Physics: International Conference Information Technologies in Business and Industry, 2018. <https://doi.org/10.1088/1742-6596/1015/2/022002>
- [19] A. Cherfi, K. Noura & A. Ferchichi 'Very Fast C4.5 Decision Tree Algorithm', Applied Artificial Intelligence, 32(2), 2018. <https://doi.org/10.1080/08839514.2018.1447479>
- [20] E.G. Kulkarni and R.B. Kulkarni, "Weka Powerful Tool in Data Mining. IJCA Proceedings on National Seminar on Recent Trends" in Data RTDM, 2, 10-15, 2016.
- [21] S. Akinola, O. Oyabugbe, 'Accuracies and Training Times of Data Mining Classification Algorithms: An Empirical Comparative Study'. Journal of Software Engineering and Applications, 8, 470-477, 2015. <https://doi.org/10.4236/jsea.2015.89045>
- [22] F.B. Márquez, 'Acquiring and Exploiting Lexical Knowledge for Twitter Sentiment Analysis', University of Waikato, PhD Thesis, 2017
- [23] F. Škegro, J. Zoroja, and V. Šimičević, "Credit Scoring Analysis: Case Study of Using Weka" (September 7, 2017). 2017 ENTRENOVA Conference Proceedings, Available at SSRN: <https://ssrn.com/abstract=3282504>
- [24] L. Jonguk, "Fault Detection and Diagnosis of Railway Point Machines by Sound Analysis." Sensors (Basel, Switzerland) 16(4), 549-555, 2016. <https://doi.org/10.3390/s16040549>
- [25] F.M.N. Ali & A.A.M. Hamed 'Usage Apriori and clustering algorithms in WEKA tools to mining dataset of traffic accidents', Journal of Information and Telecommunication, 2(3), 231-245, 2018. <https://doi.org/10.1080/24751839.2018.1448205>
- [26] J. Li, J. He, Z. Liu, H. Zhang, C. Zhang, "MATEC Web of Conferences 272, 01035 'Traffic accident analysis based on C4.5 algorithm in WEKA' School of Transportation", Southeast University, Nanjing 211189, Jiangsu, China, 2019. <https://doi.org/10.1051/mateconf/201927201035>
- [27] G. Tsaganos, D. Papachristos, N. Nikitakos, D. Dalaklis, A.I. Ölçer, "Fault Detection and Diagnosis of Two-Stroke Low-Speed Marine Engine with Machine Learning Algorithms", Conference: 3rd International Naval Architecture and Maritime SymposiumAt: Istanbul, Turkey, 2018. <https://www.researchgate.net/publication/324835430>
- [28] F. J. Morales, A. Reyes, N. Caceres, L. Romero, F. G. Benitez ' Automatic Prediction of Maintenance Intervention Types in Roads using Machine Learning and Historical Records' Transportation Research Record Journal of the Transportation Research Board, 2672(44), 2018. <https://doi.org/10.1177/0361198118790624>
- [29] H. Khaksar, A. Sheikholeslami, 'Airline delay prediction by machine learning algorithms' International Journal of Science and Technology, ume 26(5), 2019. <https://doi.org/10.24200/SCI.2017.20020>
- [30] A. Fredrik, 'Reducing ships' fuel consumption and emissions by learning from data' Linnaeus University, PhD Dissertation, 2018.
- [31] N. Barmounakis 'Investigating the decision-making process of drivers during overtaking by Powered Two Wheelers' National Technical University of Athens, Department of Transportation Planning & Engineering PhD Dissertation, 2017.
- [32] A. Kiranmai, J. Laxmi, "Data mining for classification of power quality problems using WEKA and the effect of attributes on classification accuracy", Prot Control Mod Power Syst 3, 29 2018. <https://doi.org/10.1186/s41601-018-0103-3>
- [33] J. Shen, O. Lederman, J. Cao, F. Berg, S. Tang, A. Pentland, "Gina: Group gender identification using privacy-sensitive audio data", IEEE International Conference on Data Mining (ICDM), 457-466 2018. <https://doi.org/10.1109/ICDM.2018.00061>



## Cluster Centroid-Based Energy Efficient Routing Protocol for WSN-Assisted IoT

Nalluri Prohess Raj Kumar\*, Josemin Bala Gnanadhas

Department of Electronics and Communication Engineering, Karunya Institute of Technology and Sciences, Coimbatore, 641114, India

### ARTICLE INFO

Article history:

Received: 31 May, 2020

Accepted: 19 July, 2020

Online: 28 July, 2020

Keywords:

Internet of Things

Wireless sensor networks

Energy management

Basestation

Clustering

Energy Centroid

### ABSTRACT

Wireless sensor network is highly resource constrained, where energy efficiency and network lifetime plays a major role for its sustenance. As the sensor nodes are battery operated and deployed in hostile environments, either recharging or replacement of batteries in sensor nodes is not possible after its deployment in inaccessible areas. In such condition, energy is the vital factor for the survival of sensor node in the sensing field. In order to increase the network lifetime and balance the energy consumption, robust routing protocols are required. The proposed network routing has three phases: 1. Network initiation phase to create a zone which enables the communication among local nodes 2. Zone co-ordinator selection phase algorithm to form zone cluster and its re-election procedure and 3. Zone head selection with its replacement phase based on energy centroid positional information and distance to the basestation to distribute load equally among zone co-ordinators, local sensor nodes. The data path between zone heads and basestation is distance centric and is optimized at one hop and dual hop levels to avoid data packet loss at zoneheads. Each zone is designed to own atmost  $\frac{1}{4}$  rth of deployed sensor node count through uniform random deployment. Simulations results when basestation is placed inside sensing field indicates that the proposed network algorithm outperforms when benchmarked against similar protocols like conventional LEACH, Traditional PEGASIS, existing PRRP, ES3 protocols in terms of performance metrics like Network energy consumption, Average energy consumed by sensor node, Packet delivery ratio, Packet loss percentage and Network throughput.

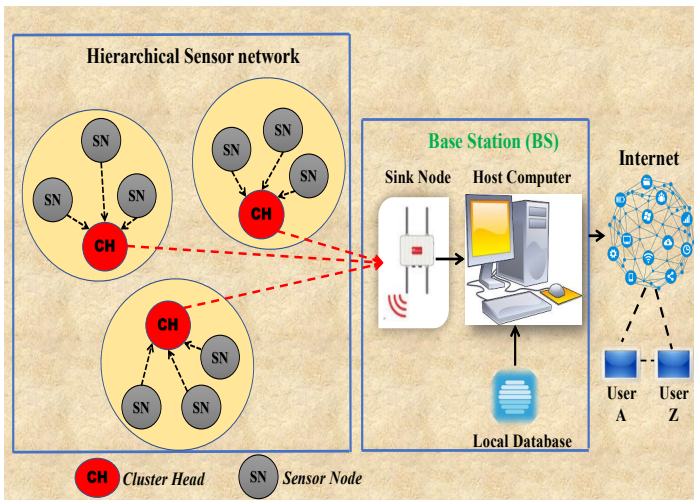
### 1. Introduction

Internet of Things (IoT) is a concept to interlink the mechanism of the object with the internet and communicate with them to get identified by others. IoT defines the world where a physical object can be connected and communicated in an intelligent way. IoT can build the network comprising physical objects and provide them the capability to gather and share their own information with others. Wireless sensor network is a subset of IoT as it has the technology that can be utilized within an IoT system to meet the user demands. So, a sensor network assisting IoT, in an application using wireless network may have an advantage of lesser cost, friendly deployment of sensor nodes and better scalability. But a main drawback is that, its energy resources are hard to replace as they operate in rugged environment. This makes the energy

management a key metric in wireless sensor network assisted IoT applications. A simple architecture of WSN- assisted IoT is shown in Figure 1.

Here a sensor node (SN) gathers information from environmental properties like temperature, pressure or moisture of air and convert it into digital form to send them wirelessly via *cluster* heads to the *base station* (BS). A basestation in contrast to sensor node possesses more computational power, large memory and connected to best source of energy unlike sensor nodes. In an IoT application, a basestation deployed can store, analyze and visualize the *sensed data* collected from cluster heads (CH). The basestation provides *graphical user interface* to interact with users directly or forward sensed data to a *remote server* via *Internet*. Then these *remote servers* will relay sensed data to the authorized users. Also, this sensed data can be saved as *web pages*, so that it can be accessed worldwide via *Internet*.

\* Corresponding Author: Nalluri Prohess Raj Kumar, Karunya Institute of Technology and Sciences, +91-8309329926 & proheessms2k17@gmail.com, nalluriprohess@karunya.edu.in



In order to simplify the network management, the *cluster/group/zone* concept is proposed by various researchers. CH nodes are the managers of grouped sensors nodes. CH has the responsibility to organize sensor nodes in its cluster, framing a routing table, collecting, aggregating and retransmitting the sensed data of the cluster. CH in sensor network will fast deplete their energy due to load imposed on them. To reduce the energy dissipation of CH nodes, the communication distance between CH and the basestation should be low or else information needs to be multi-hop forwarded from CH to basestation with the help of intermediate cluster heads.

Aiming at higher energy efficiency for the sensor network proposed, an extension of work originally presented in an IEEE conference [1] is modified as Cluster Centroid based Energy Efficient Routing (CEER) to balance energy consumption in cluster based WSN-assisted IoT. The main contributions of this research work are as follows:

- Producing a zone-clustering protocol which operates based on energy centroid position. The residual energy of sensor nodes and the communication distance to the basestation is taken into consideration here.
- To optimize the path between the zone head and the basestation, a dual hop or single hop communication routing algorithm is proposed. Zone heads count and distance metric are taken into consideration in a network area of  $A$  sq.m
- To optimize the path between the local sensor nodes in the zone and the basestation, zone co-ordinators are selected in the zone clusters. One among them acts as the zone head to forward sensed data to avoid local packet loss and long-distance communication among sensor nodes inside the zone.
- The user is provided with an option to opt for either 2 X 2 CEER (mostly suitable for dense environments) or 4 X 4 CEER (mostly suitable for sparse environments) to analyze the results based on their own preferred node deployment type. Provided the basestation should be placed inside the sensing field of the user defined area.

The research contributions reduced the average energy consumption of sensor nodes inside the zone clusters which eventually increased the lifetime of sensor nodes in the network when simulated under a simulation time  $t = 200$  Sec. Moreover,

the network lifespan is increased, as the sensor nodes balance the energy consumption among themselves (zone co-ordinator and zone head replacement strategy) to live for a longer period of time and contribute to the network throughput.

The rest of the paper is organized as follows: In section 2, the related work is discussed. The proposed CEER protocol and its phases along with its radio energy dissipation model is explained in section 3. The simulation results and discussions of the proposed protocol in comparison to the existing works is shown in section 4. Section 5 concludes the paper with further proceedings.

## 2. Related Work

Based on the network architecture, hierarchical routing protocols are classified into three types [2]: 1. Cluster-based; 2. Chain-based; 3. Tree-based protocols.

In *cluster-based* routing, network area is divided into clusters with the assistance of BS or sometimes through grid clustering and so on. A cluster head is selected for each cluster to transfer information collected from all the sensor nodes to the basestation either directly or with the assistance of other cluster heads or sensor nodes outside the cluster. In *chain-based* routing, the nodes are arranged in a chain like structure and only one node will act as the chain head to transmit the entire information of the chain to the basestation directly in the entire network area. In *tree-based* routing, all the sensed information by the leaf node is carried to the parent node (sink node) and at last, the information is sent to the root node i.e., BS.

### 2.1. Conventional Clustering approach

The first and foremost hierarchical routing protocol designed for wireless sensor networks is the LEACH protocol. In the open system interconnections (OSI) reference model, cluster routing protocols works on the network layer which is connected to both the data link layer and the transport layer. The primary function of the network layer is to permit different networks to be interconnected. It also translates logical network addresses into physical address. So, a network layer forms the major part in forwarding and routing the information across the network.

The LEACH [3-6] protocol is explained in detail in section 4.1.1. Its main characteristic is the local cluster generation, dynamic CH node rotation with data fusion inside the cluster. By generating a random number which is between 0 or 1, all sensor nodes are given provision to be the CH and the threshold set for selection of node is calculated as in Eq. (1)

$$T(n) = \begin{cases} \frac{P}{1 - P[r \bmod (\frac{1}{P})]} & n \in N \\ 0 & \text{Otherwise} \end{cases} \quad (1)$$

The generated random number and  $T(n)$  are compared for each sensor node. If the random number generated by a sensor node is less than the calculated threshold value  $T(n)$ , then that particular sensor node is selected as a cluster head for the current round. In Eq. (1)  $P$  is the desired percentage of sensor node inside the local cluster to be selected as cluster head,  $r$  is the round number,  $n$  is

node count and  $T(n)$  is the threshold. Number of cluster head nodes are not constant in LEACH due to its random selection of cluster head nature. The drawback of LEACH [7] is that, the different cluster count noted in each round will make number of sensor nodes in each cluster different. Uneven sensor node count in clusters dissipate uneven energy in each round. During CH selection, residual energy of sensor node is not taken into consideration in LEACH protocol.

### 2.2. Energy based clustering approach

An Energy Based Clustering – Self organizing map (EBC-S) to bring the effectiveness in cluster-based routing algorithms is proposed in [8] for topological clustering and incorporating a topological energy-based clustering technique to achieve extended lifetime and network coverage. The assumption made on BS is that it has no constraints on energy resources. A hierarchical and distance-based clustering technique is proposed in [9] which utilizes a new rank-order distance measure for agglomerative hierarchical clustering. Here, authors generate a rank order list by sorting all other sensor nodes in near neighbors by absolute distance. The distance based on rank order of two sensor nodes is computed using their rank order. The algorithm designed by them by grouping sensor nodes into small number of sub-clusters is similar to CEER networking like the formation of zone clusters inside the zone in CEER. The word *centroid* for wireless sensor networks is used by authors in [10-12] for  $k$ -means clustering algorithm. This algorithm is mainly based on Euclidean distance of nodes and CH selection based on residual energy of nodes. In this method, sink node/BS collects information about the *identifier*, *position* and *residual energy* of all nodes and store that information.

The steps in  $k$ -means clustering algorithm for WSNs are as follows:

- To form  $k$  clusters of sensor nodes,  $k$  centroids are to be selected initially at different locations inside the network area  $A$  sq.m.
- Euclidean distance from each sensor node to the selected centroids is calculated and information of a sensor node is saved in its nearest centroid. Thus,  $k$  initial clusters are formed in network area.
- After initial round, the position of centroid in each cluster is calculated again to check for position change from the previous one. If it is so, again calculation of Euclidean distance step repeats, else clusters are finalized.

### 2.3. Mobile sink node clustering approach

Data generated by an individual sensor may not appear significant, but overall data generated by sensor nodes of network area in dense environments is big. So, utilizing sink node's mobility is the concept proposed by authors in [13] to enable data gathering. The proposed technique by authors may reduce energy consumption of sensor nodes, but create additional challenges like

determining sink node's trajectory and cluster formation prior to data aggregation. The mobile sink data collection process, cluster head selection problem and *mobile sink* path optimization is discussed in [14]. The mobile sink path optimization is formulated as shortest path finding problem. So, artificial bee colony algorithm is used to find optimal solution and the shortest path of mobile sink to improve data collection efficiency in network area.

### 2.4. Grid based Clustering approach

*Grid clustering* for energy optimization of sensor nodes is proposed in [15,16], where size of grid is directly related to the transmission coverage range of the node. So, as the grid size is reduced, the transmission range for sensor node also decrease and thus conserves the energy. But it may lead to more control overhead and consumes extra bandwidth. Moreover, if the grid size is too small, CH has to dissipate more energy to transmit its information to basestation. The similar simulation is done in proposed CEER protocol in the form of 2 X 2 CEER and 4 X 4 CEER protocols. But depending on concentration of network nodes in network area the concerned protocol is implemented in proposed research.

### 2.5. Centroid (Midpoint) based routing approach

An energy efficient clustering protocol based on K-means algorithm named Energy Efficient Clustering Protocol (EECPK-means) has been proposed by authors in [17] for WSNs where midpoint algorithm is used to improve initial centroid selection procedure. It considers residual energy and Euclidean distance as the parameter metrics for CH selection. A clustering algorithm using spatial correlation is proposed in [18], which groups sensor nodes with similar readings into one cluster and reporting the same, as the reading of the entire group. One node is selected as CH using centroid method here. The sensor node which has minimum distance to cluster centroid point is chosen to be the cluster head of the similar reading sensor group. Similarly, an energy efficient clustering protocol to prevent unbalanced clusters based on firefly and midpoint algorithms is proposed in [19] which uses residual energy and Euclidean distance as the performance metrics. It produced balanced clusters to balance CHs load and increase network lifetime. Similarly, the term *gateway node* is used in [20] to reduce the data load on cluster heads and forward the data to BS. But their assumption that every gateway node should be in the range of its neighbor gateway node in their proposed topology proves high algorithm complexity and limits user defined random sensor node deployment process. The term *energy centroid* for sensor networks is used in [21] where each cluster is designed to own 25% of sensor nodes using *distance centroid* algorithm. Here CH selection is based on *energy centroid* while the communication between CHs and the basestation is distance centric. An *energy centroid* algorithm for WSN-Assisted IoT is proposed by authors in [22] with the EECRP (Energy Efficient Centroid based Routing Protocol) algorithm when the BS is placed inside the network. This algorithm produced better results than conventional LEACH routing algorithm but no proper routing phase is described for the

communication between the CHs and the BS. Moreover, the EECRP algorithm finds the CHs, and transfer the information of cluster to the BS directly in single hop. But the proposed finds the help of other ZHs to relay the information of ZH which is far from BS and thus has the provision to opt for dual hop communication to the BS. Most of the existing centroid routing approaches has following drawbacks when choosing the initial centroids randomly for CHs.

- An empty cluster.
- Residual nodes
- User desired clusters as input to algorithm.
- Unbalanced workload on the CHs
- Non-selection of optimal CHs count.
- Unsuitable for Multi-hop routing.

### 3. Cluster Centroid based Energy Efficient Routing (CEER) protocol

In this section, deployed sensor node network either densely or sparsely distributed in uniform random fashion is utilized to obtain optimal path from source node to basestation/sink using cluster centroid-based routing. Subsequently the radio energy dissipation model is introduced. Moreover, few assumptions are presented for better understanding of our proposed work.

#### 3.1 Assumptions

Wireless sensor network (WSN) in CEER consists of Sink node/Basestation along with other sensor nodes. The communication established between sensor nodes and basestation is multihop based. The routing structure is hierarchy. The sensor network is grouped to zone clusters (Interzonal communication) and each cluster has four zonal quadrants (Intrazonal communication). The *cluster / zonehead* collects information from four *zone co-ordinators* which is collected from each of four *zone-quadrants* or *zone-clusters*. The cluster/zone head communicates with the basestation either in single hop (directly) or through multihop (*relaying zone heads*) communication. In the proposed work the following assumptions are made.

- Distribution of sensor nodes in network is uniform random.
- There is only one sink node located inside the sensing field. (Placed at centre of sensing field).
- Both sink and sensor nodes position is static/fixed i.e., the position of sensor nodes will not change once the sensor network arrangement is done.
- All nodes can adjust their transmission power according to distance of transmission by using free space model or multipath fading model.
- All sensor nodes deployed are isomorphic i.e., equipped with equal energy levels when deployed and has the same processing and communication capabilities.
- Sensor nodes are aware of their locations through some localization techniques i.e., the locational co-ordinates are already fed into them.
- Sensor nodes are capable of transceiving information.
- Energy cost for zone formation is on BS and no control packets for sensor nodes.

- BS has complete knowledge about the energy level and locational co-ordinates of sensor nodes in network area of  $A$  sq.m.

Once the sensor network formation is done, position of sensor node is not changed. The location co-ordinates of sensor nodes are fed into the node already during network deployment. We assume that every sensor node knows the basestation co-ordinates after initial broadcasting as well as their residual energy at any time. The shape of sensing area in which sensor nodes are distributed is in rectangle. Network region is constructed using cartesian co-ordinate system (xlocation, ylocation). The notations and its explanation are provided in Table 1.

Table 1: Notations in CEER protocol and its explanation

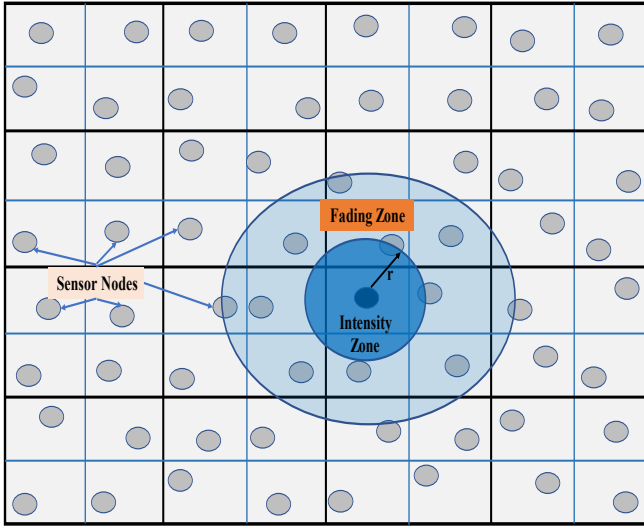
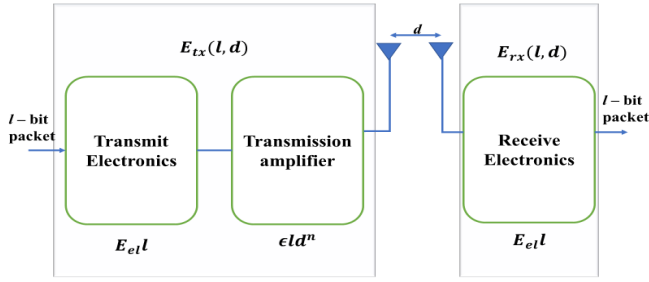
Notation	Explanation
$l$	data packet
$d$	distance
$A$	Area of sensing field in Sq.m
$d_{th}$	distance threshold
$E_{tx}$	Energy of transmitter
$E_{rx}$	Energy of receiver
$E_{el}$	Energy dissipated at tx or rx
$\epsilon_{fsm}$	Free space model ( $d^2$ power loss)
$\epsilon_{mpf}$	Multi-path Fading ( $d^4$ power loss)
<b>BS</b>	Base Station
<b>ZH</b>	Zone Head
<b>ZC</b>	Zone Co-ordinator
$E_{simR}$	Energy consumed in a Simulation Round
$E_{ZH}$	Energy consumed by Zone Head
$E_{SN}$	Energy consumed by a Sensor Node
$K_i$	$i$ Zone Count
$E_{ZH\ to\ BS}$	Energy consumed in transferring data from ZH to BS
$E_{Reception}$	Energy consumed by ZH while receiving data from ZC
$E_{Aggregation}$	Energy consumed by ZH in transferring $l$ bit data packet to BS
$d_{ZH\ to\ BS}$	distance between ZH and BS
$dist_{MAX}$	Maximum broadcast distance of BS in Sensor Network of Area $A$ sq.m
$L_{x,y}$	Cartesian locational co-ordinates of Sensor Node
$d_{S\ to\ D}$	distance from source node to destination node
$NS$	Total network size in Sq.m
$Z_S$	Zone Size in Sq.m
$Z_q$	Zone quadrant/Zone cluster size in Sq.m
$(\bar{X}_C, \bar{Y}_C)$	Mathematical centroid co-ordinates for sensor field
$(\bar{X}_{ec}, \bar{Y}_{ec})$	Energy centroid co-ordinates of CEER network sensor field

#### 3.2 Radio Energy Dissipation Model of CEER Network

The clustering in CEER is dependent on the energy model implemented below in figure 2.

From the radio energy model shown in figure 2, the required energy to transmit  $l$  bit data to distance  $d$  can be formulated as:





$$E_{tx}(l, d) = \begin{cases} l \cdot E_{el} + l \cdot \epsilon_{fsm} \cdot d^2 & \text{if } d < d_{th} \\ l \cdot E_{el} + l \cdot \epsilon_{mpf} \cdot d^4 & \text{if } d \geq d_{th} \end{cases} \quad (2)$$

The energy used in receiving  $l$  bit packet at the receiver is formulated as Eq. (3)

$$E_{rx}(l, d) = l \cdot E_{el} \quad (3)$$

where  $E_{el}$  is per bit dissipated energy at the transmitter or receiver,  $\epsilon_{fsm}$  and  $\epsilon_{mpf}$  reflect “free-space model” ( $d^2$  power loss) and “Multi-path fading model” ( $d^4$  power loss) conditions. As shown in Figure 3, the data communication process of sensor nodes to the intensity zones uses “free-space model” and “Multi-path fading model” to its fading zones.  $d_{th}$  is the distance threshold set for both models and  $d$  represents the distance between source sensor to destination sensor. From Eq. (2) distance threshold may be expressed as follows in Eq. (4)

$$d_{th} = \sqrt{\frac{\epsilon_{mpf}}{\epsilon_{fsm}}} \quad (4)$$

Now consider  $A$  sq.m sensor network area with  $N$  sensors deployed in uniform random fashion and divided into  $K$  zones or clusters. In hierarchical approach the total utilized energy within a simulation round  $E_{SimR}$  is calculated as follows Eq. (5)

$$E_{SimR} = \sum_{i=1}^K E_{ZH_i} + \sum_{j=1}^K E_{SN_j} \quad (5)$$

where  $E_{ZH_i}$  is the utilized energy by Zone Head (ZH) when it receives information from Zone Co-Ordinators and relaying information on behalf of other Zone Heads to base station along with its own zonal information to basestation.  $E_{SN_j}$  is the energy consumed by sensor nodes in zone quadrants of a zone including zone co-ordinators.

Energy consumed by zoneheads in each zone/cluster is defined as Eq. (6)

$$E_{ZH_i} = E_{(ZH \text{ to } BS)_i} + E_{Reception_i} + E_{Aggregation_i} \quad (6)$$

where  $E_{(ZH \text{ to } BS)_i}$  is the Energy utilized when the ZH of zone  $i$  transfer information to basestation,  $E_{Reception_i}$  is the utilized energy while receiving information from the zone co-ordinators within the zone cluster  $i$ ,  $E_{Aggregation_i}$  is the energy consumed to process  $l$  bit packet to the basestation by ZH. Now  $E_{(ZH \text{ to } BS)_i}$  is calculated as Eq. (7)

$$E_{(ZH \text{ to } BS)_i} = \begin{cases} l \cdot E_{el} + l \cdot \epsilon_{fsm} \cdot d_{to \text{ BS}_i}^2 & \text{if } d < d_{th} \\ l \cdot E_{el} + l \cdot \epsilon_{mpf} \cdot d_{to \text{ BS}_i}^4 & \text{if } d \geq d_{th} \end{cases} \quad (7)$$

$$E_{Reception_i} = |N_i| \cdot l \cdot E_{el} \quad (8)$$

$$E_{Aggregation_i} = |N_i| \cdot l \cdot Z_{pc} \quad (9)$$

where  $|N_i|$  in Eq. (8) is the node count in zone  $i$  and  $Z_{pc}$  in Eq. (9) is the zone path cost of a bit reporting the basestation.

From the assumptions made above, some zone heads operate in free space mode and others in amplification mode to reach basestation with the zonal information. Here zone heads using amplification mode to transmit their information uses relaying zone heads to reach basestation. Let total ZH count in both modes be  $m$ . Then  $\sum_{i=1}^K E_{(ZH \text{ to } BS)_i}$  is formulated as Eq. (10)

$$\sum_{i=1}^K E_{(ZH \text{ to } BS)_i} = l \left( K \cdot E_{el} + \epsilon_{fsm} \cdot \sum_{i=1}^{K-m} d_{to \text{ BS}_i}^2 + \epsilon_{mpf} \cdot \sum_{i=1}^m d_{to \text{ BS}_i}^4 \right) \quad (10)$$

So, total consumed energy by all ZH’s in sensing area  $A \times A$  m is given by Eq. (11)

$$\sum_{i=1}^K E_{ZH_i} = l \left( (K + N) \cdot E_{el} + N \cdot Z_{pc} + \epsilon_{fsm} \cdot \sum_{i=1}^{K-m} d_{to \text{ BS}_i}^2 + \epsilon_{mpf} \cdot \sum_{i=1}^m d_{to \text{ BS}_i}^4 \right) \quad (11)$$

Now the Energy consumed by all sensor nodes ( $N-K$ ) other than  $K$  zone heads is given by  $\sum_{j=1}^K E_{SN_j}$  in Eq. (13)

Energy consumed by single sensor node inside a zone is given by Eq. (12)

$$E_{(SN)_j} = \begin{cases} l \cdot E_{el} + l \cdot \epsilon_{fsm} \cdot d_{to \text{ BS}_j}^2 & \text{if } d < d_{th} \\ l \cdot E_{el} + l \cdot \epsilon_{mpf} \cdot d_{to \text{ BS}_j}^4 & \text{if } d \geq d_{th} \end{cases} \quad (12)$$

Similar to Eq. (11) let  $n$  be the node count operating in both free-space and amplification modes. So,

$$\sum_{j=1}^{N-K} E_{SN_j} = l \left( (N - K) E_{el} + \epsilon_{fsm} \cdot \sum_{j=1}^{N-K-n} d_{to ZH_j}^2 + \epsilon_{mpf} \cdot \sum_{j=1}^n d_{to ZH_j}^4 \right) \quad (13)$$

From Eq. (11) and (13), the total energy consumed in a simulation round of  $t$  sec in proposed hierarchical approach routing is concluded in Eq. (14)

$$E_{simR} = l \left( 2 \cdot N \cdot E_{el} + N \cdot Z_{pc} + \epsilon_{fsm} \left( \sum_{i=1}^{K-m} d_{to BS_i}^2 + \sum_{j=1}^{N-K-n} d_{to ZH_j}^2 \right) + \epsilon_{mpf} \left( \sum_{i=1}^m d_{to BS_i}^4 + \sum_{j=1}^n d_{to ZH_j}^4 \right) \right) \quad (14)$$

### 3.3 Zone or cluster creation and Node deployment

To implement CEER protocol zone formation plays a major role, as it reduces overhead and energy consumption in data transmission within the zone/cluster. By further dividing the zone into 4 quadrants each the energy consumption of zoneheads/cluster heads is reduced. Zone co-ordinators are created within a quadrant for collecting the entire information of quadrant and sending it to zone head. From Figure 4 one can see the formation of a zone, out of  $K$  zones with four equal quadrants ( $Q_n$ , where  $n = 1,2,3,4$ ) displayed inside the sensor network. So, inside a 2 X 2 CEER network i.e., a sensor network formed with 2 zones in a row and 2 columns in a sensor area of  $A \times A$  m, will give better results when CEER is implemented with a greater number of sensors mostly suitable for dense environment. If a CEER has to be implemented in a sensor network with lesser number of nodes then zone formation of 4 X 4 or 8 X 8 finds more suitable for application in sparse environments. Let the intersection point of all the quadrants within a zone be ZQI (Zone Quadrant Intersection) located at co-ordinates  $(x, y)$ . Now zone quadrant formation algorithm is coded as Algorithm 1

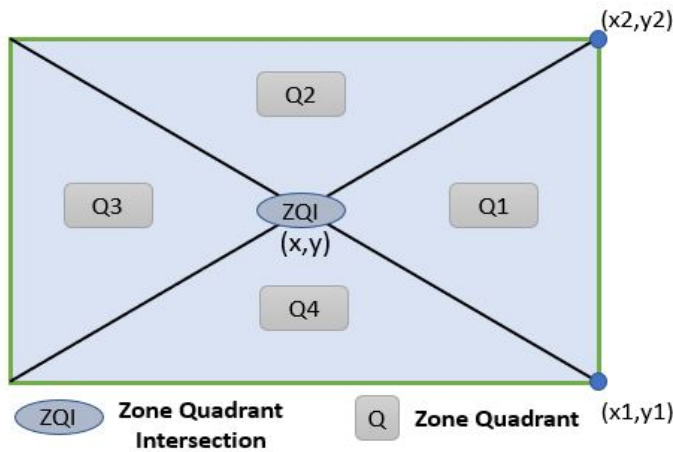


Figure 4: Zone formation in CEER Network

#### Algorithm 1: Zone Quadrant Formation

**Step 1:** theta1 = compute angle  $(x, y, x_1, y_1)$ ;

theta2 = compute angle  $(x, y, x_2, y_1)$ ;

**Step 2:** compute angle  $(x_1, y_1, x_2, y_2)$ ;  $x = x_1 - x_2$ ;  $y = y_1 - y_2$ ;

**Step 3:** angle = atan2(y, x);

**Step 4:** angle = angle\*180/pi;

**Step 5:** while (angle < 0) angle += 360;

**Step 6:** while (angle > 360) angle -= 360;

**Step 7:** return angle;

Assuming that the sensor nodes are static throughout the simulation round, they are distributed uniform randomly within the zones. In the sensor network, the locational co-ordinates of sensor nodes are obtained using the following Eqs. (15), (16)

$$X\text{-coordinate } X_c = \text{sensingField}_{\text{Length}} * \text{random} \quad (15)$$

$$Y\text{-coordinate } Y_c = \text{sensingField}_{\text{breadth}} * \text{random} \quad (16)$$

Where  $\text{sensingField}_{\text{Length}}$  and  $\text{sensingField}_{\text{breadth}}$  are the network size locational parameters and  $\text{random}$  denotes positional deployment of node in random fashion inside sensing field and is the number between 0 and 1. Since the simulation is done through Network simulator 2 (NS2 version 2.32), nodes are deployed using  $\text{scen}(N)$  command where  $N$  is the Node count in Network area  $A$  sq.m

The distribution of energy in entire network is done equally among all sensor nodes inside sensing area. Each sensor computes the distance ( $d_{S \text{ to } D}$ ) between its own location  $L_{x,y}$  to destination sensor or the Base station  $B_S$ . The distance ( $d_{S \text{ to } D}$ ) is estimated using the Pythagorean Theorem. So, finding the distance between sensors by using the coordinating points of source sensor  $(X_S, Y_S)$ , and the destination sensor  $(X_D, Y_D)$  as shown in Figure 5 and then computing the distance between them is as follows in Eq. (17)

$$(d_{S \text{ to } D}) = \sqrt{|X_S - X_D|^2 + |Y_S - Y_D|^2} \quad (17)$$

In  $R \times C$  CEER network formed with  $R$  row zones and  $C$  column zones inside  $A$  sq.m network field, let total network size be  $NS = A^2 m$  as represented in fig.5. Consider  $Z_S$  as one zone size in sq.m which is split into further equal quadrants of each size  $Z_q$  sq.m. Now  $NS$  is formulated from Eq. (18), (19)

$$Z_S = (4 * Z_q) \text{ sq.m} \quad (18)$$

$$NS = (R * C * Z_S) \text{ sq.m} \quad (19)$$

Suppose if the 2 X 2 CEER network is implemented in 100 sq.m network field.  $Z_S$  is  $4 * Z_q = 25$  sq.m. Then total network size  $2 * 2 * 25 = 100$  sq.m. from Eq. (19)

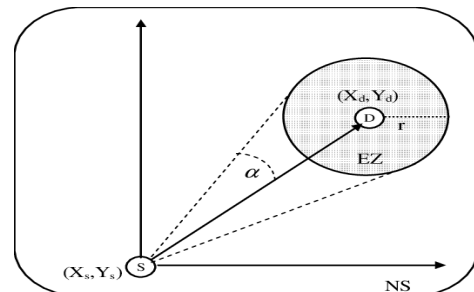


Figure 5: Expected Zone (EZ) formed in radius "r" for the destination sensor w.r.t source sensor locational co-ordinates at an angle  $\alpha$

When the CEER network of network size 2 X 2 is considered i.e., with 2 rows and 2 columns totally comprising of 4 zones. Each zone  $Z_s$  will occupy an equal area of 25 sq.m and so total network size NS would be  $25 * 2$  (Rows of zones formed) \*  $2$  (Columns of zones formed) = 100 Sq.m

### 3.4 Zone Clustering Algorithm Scheme

The proposed CEER algorithm find the most appropriate zone head (ZH) node for the zone based on energy centroid. The algorithm has 3 phases: 1. Network initiation phase. 2. Zone co-ordinator selection phase and 3. Zone head selection and its replacement phase

#### 3.4.1 Network initiation phase

The aim of this phase is to create mutual message exchange path between sensor nodes and basestation. These messages will have nodes positional, energy level information along with the average energy of deployed network. Also, information of the zone co-ordinators (ZC's) is chosen by the basestation in its first round and the longest transmission distance too. Firstly, sensor nodes will send their *position* message to the *basestation* (BS). The packet format of *position* message is shown in Figure 6.

<i>msg type</i>	<i>Sensor ID</i>	<i>X-coordinate</i>	<i>Y-coordinate</i>	<i>Energy Level</i>
-----------------	------------------	---------------------	---------------------	---------------------

Figure 6: Position message packet format

The *msg type* header shows that the packet has a sensor node locational information. The *Sensor ID* is given to identify the sensor, sending its location update to base station. The position of sensor node in the sensing field is given in cartesian co-ordinate (X, Y). Here *X-coordinate* and *Y-coordinate* of sensor is sent. *Energy level* in the packet format shows the residual energy of sensor node at particular time *t*.

When this location message packet is received from all sensor nodes in the sensing field within the time limit set by timer, basestation will start to estimate the distance from itself to all the sensor nodes which have updated their location message. Consequently, the zone formation will be done with its quadrants, inside the sensing field area. Now the basestation starts identifying the *zone clusters* based on the estimated distances calculated previously. Subsequently, it updates the *node table* with the position information of node and its energy level. Then, the basestation broadcast the *ACK* (Acknowledgement) message specifically to the sensor nodes in one zone. The *ACK* packet format is shown in Figure 7

<i>msg type</i>	<i>dist<sub>MAX</sub></i>	<i>ZC ID</i>	<i>Average Energy</i>
-----------------	---------------------------	--------------	-----------------------

Figure 7: Pack format of ACK message

*msg type* here is used to send the intended information to particular zone cluster nodes from *basestation / sink*. *dist<sub>MAX</sub>* field provides the information on maximum broadcast range to each node in particular *zone cluster* and is calculated by basestation and set it as a *communication threshold* for CEER network. Note that, *dist<sub>MAX</sub>* is calculated by taking the *Average Energy* from Eq. (2) *ZC ID* is needed here, so that the basestation can send the above-mentioned information to the particular *zone co-ordinator*. The *Average*

*Energy* indicates the CEER network average energy intended for each node. The *ACK* message receiving node will update its routing table according to *ACK* packet content. At the end of this phase, the mutually exchanged information is updated in routing tables of both basestation and sensor nodes memory. Also, the routing table information is updated in real time, as the CEER network enhances to further phases.

#### 3.4.2 Zone co-ordinators selection phase

The main aim of this phase is to select *zone co-ordinator* for every *zone cluster*. After receiving the *position* message and delivering *ACK* message, the basestation selects the node which has the *Energy Level* greater than the *Average Energy*. In the initial round, ZC selection is mostly random as the sensor nodes are equipped with equal energy levels at the deployment stage. In the consecutive rounds as the basestation updates the broadcast information, it is saved by all sensor nodes. After receiving the updated broadcasting information from basestation which has the information on node that can act as the ZC, the sensor node will check its own ID whether it is the ZC. So, if the ID matches then that node claims its election as the ZC for that transmission round. It also acts as a transceiver by receiving the information of all sensor nodes in a *zone cluster* and transmitting it to the ZH from then. Zone Head is elected later based on energy centroid in routing phase as shown in Figure 8. If suppose a sensor node finds its own ID different from BS broadcasted information, it saves the energy by activating its receiving antenna and closing the transmitter antenna to wait for further updates from the basestation through ZC.

#### 3.4.3 Zone head selection and its replacement phase

After selecting the zone co-ordinators, the ZCs' broadcasts the schedule message with their IDs and the positional information to neighbour ZC nodes. Based on ZC ID in *ACK* message and schedule message received by neighbour nodes to determine whether they belong to that zone quadrant/cluster, zone clustering phase is completed. Now the sensor nodes inside each zone cluster will send their positional and energy level information to their ZC. Now zone coordinators which fall under zone  $K_i$  (where  $i=1, 2, \dots, (R*C)$  as shown in Figure 8) has to find the energy centroid location of the zone  $K_i$ . The ZC node which is nearest to that energy centroid will act as ZH for that zone. So, ZH will be one among the four zone coordinators of a selected zone.

By electing the ZH in the above-mentioned way, the CEER network can balance the energy consumption of networking sensor nodes inside the sensing field area. Now the CEER network satisfies the four aspects of our protocol framework. Firstly, in default simulation round, the temporary ZCs' are chosen by basestation itself, thereby gaining an overall picture of CEER network. When the simulation is processing, official ZC is selected in the zone cluster, which shows that CEER network is self-organizing. Also, one among the selected ZCs' of a zone closer to the energy centroid can act as zone head (ZH), which improves coverage of CEER network. The position of energy centroid calculation is described in next section which is done purely based on positional and residual energy information of sensor nodes inside CEER network. So, for a sensor node to become a ZH it should reach the ZC stage first and its energy level must prove worthy to become ZH after considering the closest distance parameter.

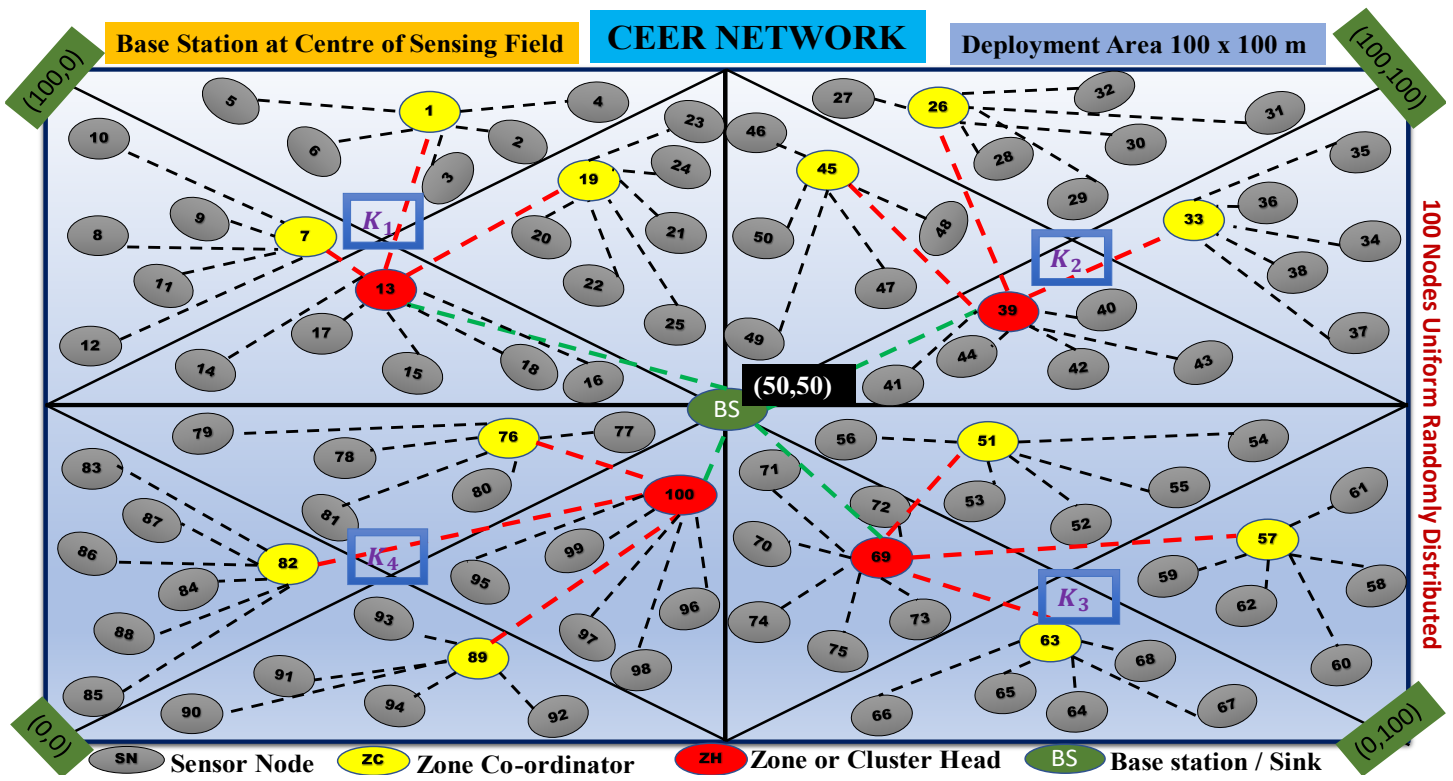


Figure 8: Zone formation and routing process in the CEER network when the basestation (BS) is placed at centre co-ordinates of Network area  $A$  sq.m

### 3.5 Zone Energy Centroid Selection Algorithm

In a concept of mathematics, centroid is the point of concurrency. Indirectly, it is the intersection point of all the medians in a geometric figure. Median is the value separating the higher half from the lower half of a data sample i.e., a population or a probability distribution. Now consider the ZC as a median of all the sensor nodes in a zone cluster. These zone coordinators from each zone quadrant are the set of medians. Now all these zone coordinators fulfil their duty to find the centroid position and which would be an imaginary point of mass concentration. Then the ZC near to that point will be selected as ZH for one simulation round. Here the concept is introduced as *Zone Energy Centroid* slightly named different from traditional mathematical centroid for few reasons. Firstly, median in terms of sensor nodes doesn't make sense. Secondly, finding the intersection point of the medians in terms of sensor operation to find zone head is meaningless. Finally, residual energy of sensor nodes is the only metric which changes in the sensor network operation. So, energy centroid in CEER network operation is to display the distribution of remaining energy of entire network.

Here, to calculate the Energy Centroid, the residual energy level and location of zone co-ordinators  $i$  is taken into consideration. In the zone cluster, if the weight of ZC is known then position of centroid is calculated using Eq. (20) and (21). In the field of mathematics, calculation for centroid in cartesian co-ordinate system  $(\bar{X}_C, \bar{Y}_C)$  is calculated as

$$\bar{X}_C = \frac{dM_y}{M} = \frac{\iint_{A_z} x \cdot D \, dw}{\iint_{A_z} D \, dw} \quad (20)$$

$$\bar{Y}_C = \frac{dM_x}{M} = \frac{\iint_{A_z} y \cdot D \, dw}{\iint_{A_z} D \, dw} \quad (21)$$

where  $\bar{X}_C$  and  $\bar{Y}_C$  are the results of centroid in mathematical prospective for CEER network. Here  $A_z$  is the zone geometrical area.  $dw$  is the centroid weight differential.  $D$  is the density of nodes weight.  $dM_y$  and  $dM_x$  are the static moments of  $x$  and  $y$  axis respectively.

In CEER network, when the weight of ZC is unknown i.e., as the term "weight" makes no sense in ZH formation, ZC locational and energy level information is gathered to obtain position of *Energy Centroid*. When its position is obtained, ZC nearer to that *Energy Centroid* can take the responsibility of particular zone by acting as a ZH. Once ZH is selected for that particular simulation round, its locational information is updated in the routing tables of ZH itself, the basestation and the other ZH's in the network area of  $A$  sq.m by basestation. The position of centroid is given as  $(\bar{X}_{ec}, \bar{Y}_{ec})$  in zone  $K_i$  as obtained from Eq. (22) and (23)

$$\bar{X}_{ec} = \frac{\sum_{p=1}^4 \frac{RE_{ZC_p}}{E_0} \cdot X}{N_j(p)} \quad (22)$$

where  $RE_{ZC_p}$  is the Residual Energy of Zone Co-Ordinator  $p$  in zone  $K_i$ .  $X$  and  $Y$  are the location co-ordinates of Zone Co-Ordinator  $p$ .  $N_j(p)$  is the  $j$  node count in the zone cluster with zone coordinator  $p$ .

When the Zone Heads of  $K$  zones are ready to send packets to basestation, they estimate their own distance  $d_{ZH \text{ to } BS}$  to basestation placed at center of network area with  $dist_{MAX}$  sent by the basestation in *ACK* packet. If  $d_{ZH \text{ to } BS} < dist_{MAX}$  then



ZH's will start sending the packets to base station, else ZH's will find their nearest neighbor ZH from pool of  $(ZH_i)$  where  $i = 1, 2, \dots (K - 1)$  in terms of distance, to transmit its information in maximum dual hop to the basestation. Then the condition to be checked is whether  $d_{ZH_i \text{ to } BS} < dist_{MAX}$  or not. If fulfilled then proceed to transmit the information to the nearest ZH and it will send that information to basestation on its behalf, else store the packets in ZH buffer queue and wait for the next round. This process of storing packets in ZH and waiting for consecutive round will be the least case, and if it has to happen there might be packet loss in rare scenario. So, a trading exists in CEER network between energy consumption and packet loss to avoid the long-distance communication of ZH nodes which may lead to more packet loss. So, the energy consumed in a simulation round of  $t$  sec is calculated through Eq. (14) after updating the Zone path cost  $Z_{pc}$ .

---

**Algorithm 2:** CEER Network Routing Process

---

- Step 1:** Sensor Nodes  $SN_j \xrightarrow{\text{position PACKET}}$  Basestation  $BS$
- Step 2:** Calculate Average Energy and  $dist_{MAX}$
- Step 3:**  $BS \xrightarrow{\text{ACK PACKET}}$   $SN_j$
- Step 4:** *if* (Zone Head ZH prevail)
- Step 5:** *then Begin* Zone Clustering  $\xrightarrow{\text{Result}}$  Zone co-ordinator ZC
- Step 6:** ZCs'  $\xrightarrow{\text{Schedule Message}}$  neighbour ZC nodes,
- Step 7:** Recalculate Energy Centroid  $\xrightarrow{\text{Result}}$  updated ZH ID
- Step 8:** timer is set,
- Step 9:** *and if* (Node count  $N \geq$  Zone Cluster count)
- Step 10:** *and if* (pre-set timer time not reached *then*
- Step 11:**  $SN_j \xrightarrow{\text{sensed data}}$  ZC ( $p$ ), where  $p = 1, 2, 3, 4$  in a zone
- Step 12:** ZC ( $p$ )  $\xrightarrow{\text{Energy Centroid information}}$  forward data to ZH  $K_i$
- Step 13:** *if* ( $d_{ZH \text{ to } BS} < dist_{MAX}$ )
- Step 14:**  $K_i$  (where  $i = 1, 2, \dots, (R * C)$ )  $\xrightarrow{\text{forward Data}}$  BS directly
- Step 15:** *elseif* ( $d_{ZH_i \text{ to } BS} < dist_{MAX}$ );
- Step 16:** *then*  $\xrightarrow{\text{forward Data}}$  nearest  $ZH_i$  where  $i = (1, 2, \dots, (K - 1))$
- Step 17:** return
- Step 18:** *end if*
- Step 19:** return
- Step 20:** *end if*
- Step 21:** return
- Step 22:** *end if*

**Step 23:** Recalculate Average Energy and  $dist_{MAX}$

**Step 24:**  $BS \xrightarrow{\text{ACK PACKET}}$   $SN_j$

**Step 25:** Select Zone Head ZH nodes

**Step 26:** return

**Step 27:** *end if*

**Step 28:** return

**Step 29:** *end if*

---

As shown in Eq. (14), the energy consumed by CEER network in the initial stage is for the basic information exchange i.e., for sending *position* message and *ACK* message. But when the network enters the stable stage of transferring the sensed information by sensors, the energy consumed by the above messages can be ignored. As it would be very less when the routing proceeds further rounds. We know that hierarchical sensor networks are good at achieving better results in local zone clusters (in our case). So, local sensor nodes i.e., nodes inside a *zone cluster* can only communicate with their ZCs'. Also, the average transmission distance of the local sensor is purely dependent on the location of ZC inside a *zone cluster*. In most of the conventional-clustering protocols like LEACH (Low Energy Adaptive Clustering Hierarchy), the location of sensor node is not considered while choosing the cluster head (CH). As a result, when the CH is chosen it may be located at edges of the cluster and it makes the sensor node inside the cluster to consume most of its energy for communicating with the cluster head. Contrarily, the proposed CEER chooses the ZC and ZH in an entirely different way. Also, the ZC of the next round is chosen by ZC of current round based on metrics encoded in the script. The ZH is chosen as the centre of energy network to optimize energy consumption parameter.

**4 Simulation Results and Discussions**

The simulation tool used for simulating CEER network is NS-2 version 2.32. Network Simulator (*ns*) is a discrete event simulator targeted at networking research. Based on few sources in NS-2, few assumptions are made for the CEER simulation. The simulation is based on fact that there may be no loss of sensed data from environment by sensor nodes, since the transmission medium used is reliable by default in NS2. Since the omnidirectional antenna is used in simulation, the radio coverage will be in circular direction. The configured node uses free space model /multipath fading model as its propagation *type*.

Here the physical layer is used as wireless channel transmission medium, where the node transmits and receive information. Interface Queue type can be either Drop Tail /Priority Queue based on Routing. If the channel stays busy/channel unavailable i.e., the basestation is far reachable for ZH when  $d_{ZH \text{ to } BS} > dist_{MAX}$  and even neighbour Zoneheads  $d_{ZH_i \text{ to } BS} > dist_{MAX}$  in our case, the packet will be saved to the queue. Here the number of packets that can be hold in queue is set to 50 packets. Here the topology loaded is of Flat Grid Type. Two sensor nodes can mutually interchange their data with each other if they lie within their sensing range. But as the topology changes there may be a change

in the nodes position so that the node may move out of its neighbours sensing range. Then (GOD) General Operation Director) will monitor Sensor Nodes information in simulation of Wireless Networks. But GOD needs to be re-created for every instance of simulation. GOD creation requires node count to be predefined, so that memory could be reserved for those nodes to capture information. Simulation area of CEER network is 100 x 100 m. The node configuration used for NS-2 simulation of CEER network is shown in Figure 9.

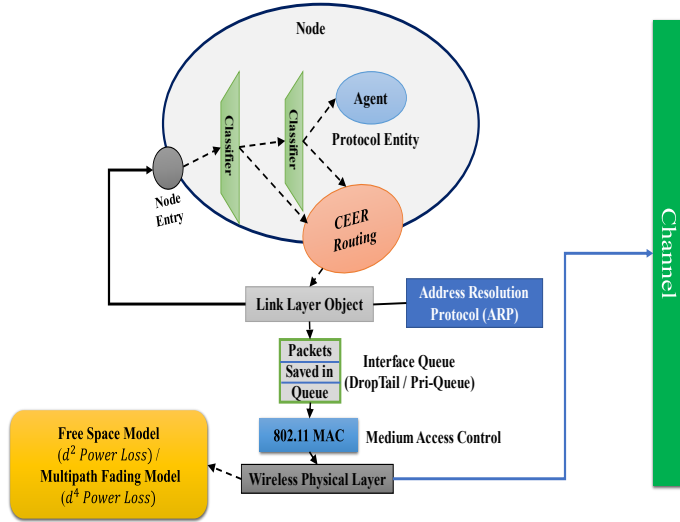


Figure 9: Node configuration of CEER network in NS2 simulator

The “free space model” assumes the ideal propagation condition with only one LOS (Line-of-sight) path between transmitter and receiver. It basically represents the communication range as a circle around the transmitter. The receiver receives all packets, if its location is within the circle. This model is best suitable for short distance communication. Otherwise, packets are lost. So, according to free space model the received signal power in free space at distance  $d$  from the transmitter is given as Eq. (24) [23]

$$\epsilon_{fsm} = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L} \quad (24)$$

where  $P_t$  is transmitted signal power.  $G_t$  and  $G_r$  are the transmitter and receiver antenna gain respectively.  $L$  ( $L \geq 1$ ) is the system loss and  $\lambda$  is the wavelength. It is common to select  $G_t = G_r = 1$  and  $L = 1$  in ns simulations.

In “multipath fading model” both the LOS i.e., direct path and the obstacle reflection path are considered. This model gives more accurate prediction for long distance communication than the free space model. In this model the received power at distance  $d$  is given as Eq. (25) [24]

$$\epsilon_{mpf} = \frac{P_t G_r G_t h_t^2 h_r^2}{d^4 L} \quad (25)$$

where  $L = 1$  in case of multipath fading model model.

$P_t$  is the signal transmitted power,  $G_r$  is the receiver antenna gain,  $G_t$  is the transmitter antenna gain,  $h_t$  and  $h_r$  are the heights of transmit and receive antennas respectively.

Since multipath fading model doesn't provide good result for short distance communication due to oscillation caused by constructive

and destructive combination of two rays, free space model best suits the situation when  $d$  is small. From Eq. (25) as the distance  $d$  increases the sensor node experience faster power loss. So, to utilize the benefits of both models, we calculate the distance threshold  $d_{th}$  from Eq. (4). When  $d < d_{th}$  Eq. (24) is used and when  $d \geq d_{th}$  Eq. (25) is used. So, distance threshold is calculated as Eq. (26)

#### 4.1 Simulated protocols for comparison with CEER network

The simulation results of proposed CEER protocol is compared with conventional LEACH (Low-Energy Adaptive Clustering Hierarchy), Traditional PEGASIS (Power Efficient Gathering in Sensor Information Systems), Existing PRRP (Position Responsive Routing Protocol) and Pre-proposed ES3 (Energy Efficient Sink Selection Scheme). The results of CEER protocol are taken from two scenarios when R X C is 2 X 2 and 4 X 4 when the BS is placed at center of the sensing field. The evaluated performance metrics in the simulation are Total Energy Consumption of network, Average Energy Consumed by a sensor node, Packet Delivery Ratio (PDR), Dropping ratio of packet and the Network Throughput.

##### 4.1.1 LEACH protocol (Low Energy Adaptive Clustering Hierarchy) [3-6]

**Aim:** To Improve the lifespan of Wireless Sensor Networks by reducing Energy Consumption in the Network.

- LEACH protocol is a TDMA based MAC Protocol.
- It is a clustering-based proactive routing Protocol.

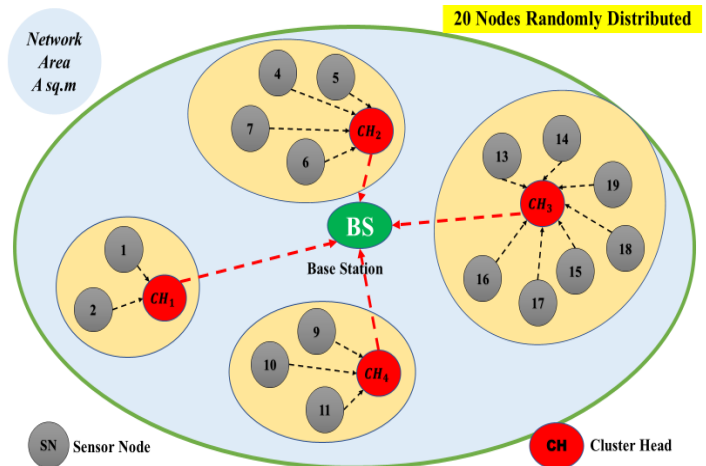


Figure 10: Routing in LEACH Protocol

#### Protocol Description:

- LEACH is self-organizing and adaptive clustering Protocol.
- In Figure 10 sensor nodes (SN) are organized into local clusters of their choice, where one of them will be a cluster head.
- This randomized approach is used to reduce the energy consumption of one particular sensor node in relaying the information collected from the sensor nodes to the sink.

- The responsibility of cluster head is not only to collect the information from its cluster, but to aggregate the data to be sent to base Station (BS) to reduce the quantity of messages, resulting in less consumption of energy which increases the network lifetime.
- LEACH uses randomized rotation of cluster heads to distribute the energy load among the sensor nodes deployed in the sensing field.

**Advantages:**

- Cluster Head (CH) aggregates the entire information which reduces the traffic load of the network.
- Since this protocol uses routing with single hop between sensor nodes to cluster heads it saves energy.
- Thus, increasing the lifetime of sensor network.
- Also, geographical location of networking node used to create a cluster is not required.
- LEACH uses no control information from BS and does not require any global knowledge of network.

**Drawbacks:**

- The protocol does not provide the information on the cluster head count in the network to the Basestation
- Also, if cluster head dies due to any reason the cluster become useless as there is no possibility of that particular cluster info reach the base station.
- Division of clusters is random and results in uneven distribution of clusters.
- E.g. Some clusters have more nodes and some clusters have lesser nodes.
- Some clusters heads are placed at the center of the cluster while some are in the border of cluster.
- This results in an increase in energy consumption and reduces the performance of the network.

4.1.2 PEGASIS protocol (Power Efficient Gathering in Sensor Information Systems)

**Aim:** To gather the information or data received from sensor nodes and transmit the same to the close neighbors in distance and iterating the leader nodes for transmitting received data to the basestation.

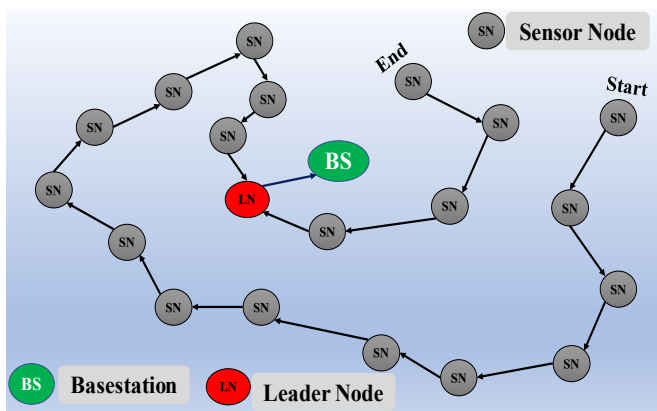


Figure 11: Routing in PEGASIS Protocol

**Protocol Description:**

- PEGASIS [25,26] is a near optimal chain-based routing protocol as shown in Figure 11. which is an improvement over LEACH.
- It facilitates distribution of energy load evenly among sensor nodes.
- The deployment of sensor nodes in sensing field is random and those sensor nodes arrange themselves into a chain using greedy algorithm.
- Also, the computation of this chain is done by basestation and broadcasts it to all sensor nodes.
- In PEGASIS, every sensor node receives and transmit atleast one packet in every round and be a leader in atleast once in  $n$  rounds and  $n$  is also the node count.

**Advantages: (Over LEACH)**

- Secondly, the leader node can receive atleast two messages from its neighbor nodes which is not the same in the case of LEACH protocol.
- For E.g. In the network of 100 nodes, cluster head (CH) in LEACH receives 20 messages at a time if the 20 nodes transfer their information at a time to the cluster head inside a cluster.
- For every round of communication one node i.e., leader node takes the entire responsibility to transmit the message to the basestation.
- In PEGASIS, data fusion helps in reducing the quantity of data being transmitted between sensor nodes and the basestation.
- Data fusion combines one or more data packets from different sensor measurements to produce a single packet for better inference.
- PEGASIS protocol is used mostly to characterize and monitor the quality of environment.

**Drawbacks:**

- The distance between the basestation and the leader node is not predefined and so the transmission distance between the leader node and basestation can be far.
- Energy level of the leader node is not considered for its selection.
- Since only one leader node can exist for one instance of communication, it may be the bottleneck for the network in causing delay.
- Redundant transmission (same data or information received more than once) exists as only one node takes the responsibility of transmitting entire chain network information to the basestation.

4.1.3 PRRP protocol (Position Responsive Routing Protocol)

**Aim:** PRRP protocol [27,28] is mainly designed to reduce the energy consumed by each node by minimizing the time the sensor

node is in idle listening state and by reducing the average communication distance over the network.

**Protocol Description:**

Sink/Basestation builds the tree-based sensor network by broadcasting the control message containing six fields *sender, position, type, level, parent and energy*. The BS creates network tree and each node receives tree information. The protocol finds logically, how many nodes lies in a particular tier  $T_n$  as shown in Figure 12.

Energy threshold  $E_{th}$  in PRRP is calculated using Eq. (27).

$$E_{th} = lE_{el} + lE_{amp}d^2 + 8lE_{el} \tag{27}$$

Where  $lE_{el}$  is the dissipated energy in processing  $l$  bits,  $E_{amp}$  is the energy of the amplifier electronics used in simulation. Energy consumption of transmitter  $E_{tx}$  and receiver  $E_{rx}$  are given by Eq. (28) and Eq. (29) respectively.

$$E_{tx} = lE_{el} + lE_{amp}d^2 \tag{28}$$

$$E_{rx} = lE_{el} \tag{29}$$

Where  $lE_{el}$  is the dissipated energy in processing  $l$  bits,  $E_{amp}$  is the Energy of the amplifier electronics used in simulation.

The lifespan of cluster head (CH) in PRRP is calculated in terms of sampling interval number  $N_s$ , as in Eq. (30)

$$N_s = \frac{E_{in} - E_{th}}{E_t + lE_r + (T_s - l + 1)T_t P_s} \tag{30}$$

Where  $E_{in}$  denote initial energy of sensor field.  $E_t$  is energy consumed in transmitting one data bit.  $T_s$  and  $T_t$  are sampling time and transit time respectively and  $P_s$  is sensor node power.

The sensor nodes are deployed in a random fashion across sensing field as shown in Figure 11 and are assumed that they are aware of its location using GPS or by any other location means. The gateways for the base station are the nodes within tier  $T_0$ . Each tier is having a vertical cross-sectional area on either side of Base station. Each vertical cross-section is provided with *tier-head* which acts as a sink for all the nodes in that vertical cross-sectional tier. *Tier-heads* are selected based on node energy level, its distance from the base station.

**Advantage:**

- It uses TDMA (Time division multiple access technique) to transfer the data from *tier heads* to the Basestation.
- No delay constraints in data transmission.

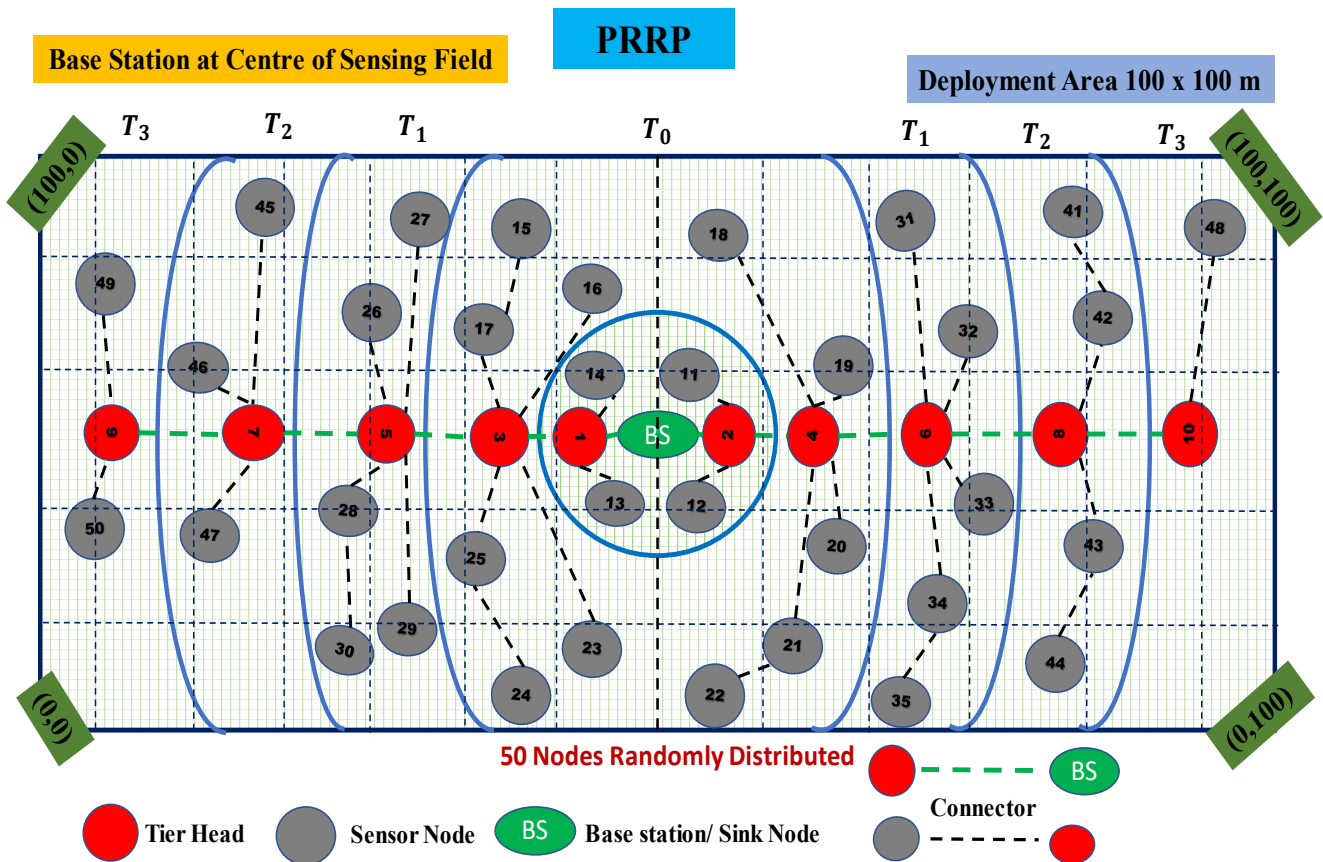


Figure 12: Tier formation and routing process in PRRP (Position Responsive Routing Protocol) when the basestation is placed at the center of the sensing field



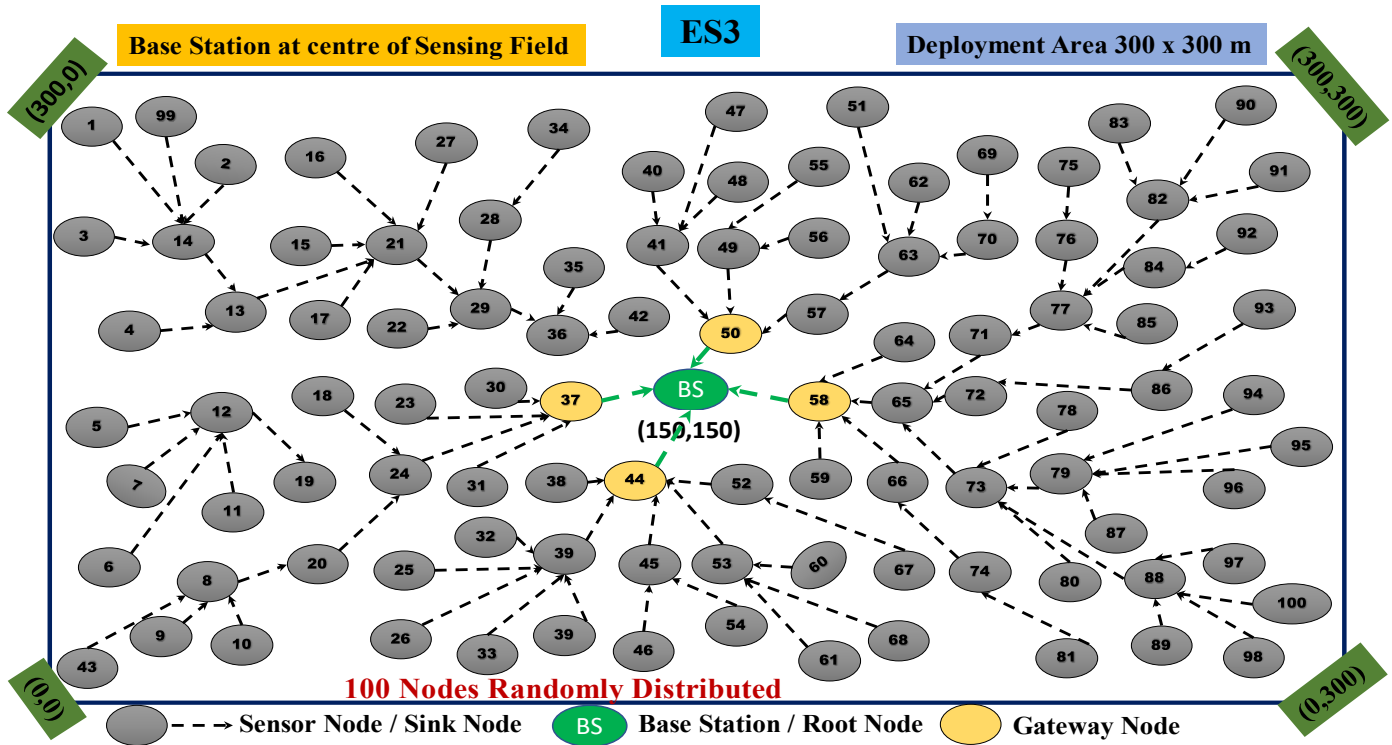


Figure 13: ES3 (Energy Efficient Sink Selection Scheme) protocol routing process when the BS is placed inside the sensing field

**Drawback:**

- It uses only *free-space model* ( $d^2$  power loss) for propagation of sensor information.
- The energy consumption of tier heads located in tier  $T_3$  which is far from BS will deplete its energy faster than normal and so it leads to reselection of tier heads frequently in the tiers far from BS.

4.1.4 ES3 protocol (Efficient Sink Selection Scheme)

**Aim:** To increase network performance and reduce average energy consumed by the node in large-scale sensor networks.

**Protocol Description:**

ES3 [29,30] uses *tree-based* routing. Here the nodes in ES3 network are classified into three types: Base station / Root Node, sink node/sensor node and Gateway node. Root node is the first node existed in the network with its hop zero. This node search for child nodes by sending broadcast packets. Here every sensor node can act as a sink node for relaying information of cluster of sensor nodes. Gateway nodes are utilized to handle the communication load of BS. The routing process in ES3 when BS is placed at center of sensing field is shown in Figure 13.

The energy model used in ES3 is same as in CEER network as explained from Eq. (2) and Eq. (3). The weight of sensor node when acting as a sink, differs from its calculation of CEER network as shown in Eq. (31) [31]

$$W_{S_n} = \frac{\rho}{dist_m} + \frac{\vartheta}{C_{m+1}} + \sigma R_m + \frac{\tau}{HoP_{m+1}} \quad (31)$$

Here,  $W_{S_n}$  is the weight of sink node  $S_n$ .  $C_m$  is the current count of child nodes connected to sink node  $S_n$ .  $R_m$  is sink node  $S_n$  left over energy.  $dist_m$  is the distance between current node and Sink node  $S_n$ . We define the root node hop to be 0 and gate way nodes hop by default is set to 1.  $HoP_m$  is the hop of sink node  $S_n$ .  $\rho$ ,  $\vartheta$ ,  $\sigma$ , and  $\tau$  are standardized parameters of four variables  $dist_m$ ,  $C_m$ ,  $R_m$  and  $HoP_m$  and are set at values 15, 6,  $1/20$ , 11 respectively during the simulation.

**Advantages:**

- Suitable for Large-scale sensor networks
- Provides balance between various network performance metrics like Energy consumption, communication distance, hop count and child node count.

Table 2: Comparison between the PEGASIS, PRRP, ES3 and CEER protocols

	PEGASIS	PRRP	ES3	CEER
<b>Scalability</b> (Increasing number of nodes after network establishment)	very low	Low	Medium	High
<b>Transmission delays</b>	Very high, since it uses a simple control token passing approach	Medium, as only the tier heads of the extreme tiers have to use the intermediate tier heads to relay the information to BS	Quite high, as no proper clustering structing is maintained during the data transmission phase	Quite low, as the local quadrant information is routed directly to zone coordinators and one among them will take the responsibility to relay the information to basestation
<b>Distribution of sensor Nodes in sensing field</b>	Randomly Distributed	Uniform Random distribution	Randomly Distributed	Uniform Random Distribution
<b>Control Message Overhead</b>	Medium	Low as the network area is divided into tiers and overhead is added to data when the information is relayed through neighbour tier heads	Medium	Very low as the overhead is added to the data only when an entire zonal information has to be relayed through neighbour zonal heads in single or dual hop to reach the basestation
<b>Uniform Distribution of Energy</b>	High, when implemented will lesser node count as it has cain network architecture	Medium, due to its tier-based network architecture as the extreme right or extreme left tier heads faces few problems	Quite High, as it approach is tree based hierarchical approach	Quite High, as the network area is uniformly divided into equal zones and further into equal quadrants
<b>Energy Efficiency</b>	Low	Low as the sensor nodes in extreme tier heads has to dissipate more energy.	Very high, as it balances distance, hop, number of child nodes and residual energy	Very high, as it uses centroid clustering approach to find the best zonal head of a particular zone
<b>Inter cluster/ quadrant structure</b>	1-Hop	1-Hop	Multihop	1-Hop
<b>Algorithm Complexity</b>	High	Medium	High	Medium

Table 3: Simulation Parameter Table of ES3, PRRP and CEER networking protocols

Parameter	Value		
Protocol	PRRP	ES3	CEER
<b>Network area (m)</b>	100 × 100 m	100 × 100 m	100 × 100 m
<b>Number of nodes</b>	100	100	100
<b>Sensor deployment</b>	Uniform Random	Random	Uniform Random
<b>Location of the sink</b>	Centre of the Sensing Field	Centre of the Sensing Field	Centre of the Sensing Field
<b>Channel Access Mechanism</b>	MAC IEEE 802.11	MAC IEEE 802.11	MAC IEEE 802.11
<b>Propagation Model</b>	Free space Model	Free space/Multipath fading	Free space/Multipath fading
<b>Total Simulation Time (sec)</b>	200 Sec	200 sec	200 Sec
<b>Initial Energy of each Node (J)</b>	2 Joule	1.5 Joule	2 Joule
<b>Packet Size (l)</b>	64 Bytes	64 Bytes	64 Bytes
<b>Transmitting/Reception Energy</b>	400nJ/byte	0.4μJ/byte	50nJ/bit
<b><math>\epsilon_{fsm}</math> (<math>d^2</math> power loss) (Amplifier Energy) <math>E_{amp}</math></b>	10pJ/bit/m <sup>2</sup>	0.01nJ/bit/m <sup>2</sup>	10pJ/bit/m <sup>2</sup>
<b><math>\epsilon_{mpf}</math> (<math>d^4</math> power loss)</b>	NA	1.3nJ/bit/m <sup>2</sup>	1.3nJ/bit/m <sup>2</sup>
<b>Energy Threshold (initial) <math>E_{th}</math></b>	1.5J	0.75J	1J

<b><math>E_{th}</math> (For one simulation Round)</b>	0.4 mJ	0.4 mJ	0.4 mJ
<b>Bandwidth</b>	0.2 Mbps	0.2 Mbps	0.2 Mbps
<b>Beam Forming Energy (nJ/bit)</b>	5	5	5
<b>Capture Threshold</b>	10	10	10
<b>Carrier Sense Threshold</b>	3.652e-8 i.e., 250 m	22.427e-8 i.e., 50 m	3.652e-8 i.e., 250 m
<b>Receiver Threshold</b>	3.652e-8 i.e., 250 m	22.427e-8 i.e., 50 m	3.652e-8 i.e., 250 m
<b>Interface Queue Length for Node</b>	50 Packets	2 Packets	50 Packets
<b>Receiving Power of Node</b>	0.1 W	0.395 W	0.1 W
<b>Transmitting Power of Node</b>	0.2 W	0.66 W	0.2 W
<b>Route Timeout</b>	10 sec	10 sec	10 sec
<b>Active Route Timeout</b>	10 sec	10 sec	10 sec
<b>Reverse Route Life</b>	6 sec	6 sec	6 sec
<b>Broadcast ID save</b>	Up to 6 Times	Up to 6 Times	Up to 6 Times
<b>Route Request Retries</b>	3 Times	3 Times	3 Times
<b>Maximum Route Request Timeout</b>	10 Sec	10 Sec	10 Sec
<b>Node Traversal Time</b>	0.03 Sec	0.03 Sec	0.03 Sec
<b>Local Repair Wait Time</b>	0.15 Sec	0.15 Sec	0.15 Sec
<b>Network Diameter</b>	30 Hops	10 Hops	30 Hops
<b>Route Reply Wait Time</b>	1 Sec	1 Sec	1 Sec
<b>Address Resolution Protocol Delay</b>	0.01 Sec	0.01 Sec	0.01 Sec
<b>Hello Interval</b>	5 Sec	5 Sec	5 Sec
<b>Allowed Hello Loss</b>	3 Sec	3 Sec	3 Sec
<b>Bad Link Lifetime</b>	3 Sec	3 Sec	3 Sec
<b>Max Hello Interval</b>	(1.25 * Hello Interval)	(1.25 * Hello Interval)	(1.25 * Hello Interval)
<b>Min Hello Interval</b>	(0.75 * Hello Interval)	(0.75 * Hello Interval)	(0.75 * Hello Interval)
<b>RTQ Maximum Length</b>	64 Packets	64 Packets	64 Packets
<b>RTQ Timeout</b>	30 Sec	30 Sec	30 Sec
<b>Max Errors Allowed</b>	100	100	100

#### Drawback:

- Delay in data transmission due to more control overheads and normalized overheads added to sensed data transfer process.

#### 4.2 Simulation performance metrics with results

The simulation parameters and their values used in NS-2 simulation of protocols PRRP, ES3 and CEER are displayed in Table 3 and their comparison results with conventional LEACH and traditional PEGASIS are shown in graphs under respective sections below.

##### 4.2.1 Comparison in terms of Total Energy Consumption of Sensor Network

Augmenting the energy balance in sensor nodes minimizes the energy dissipation during network communications. Total energy consumption of sensor network is the energy consumed to perform transmission, reception and data aggregation.

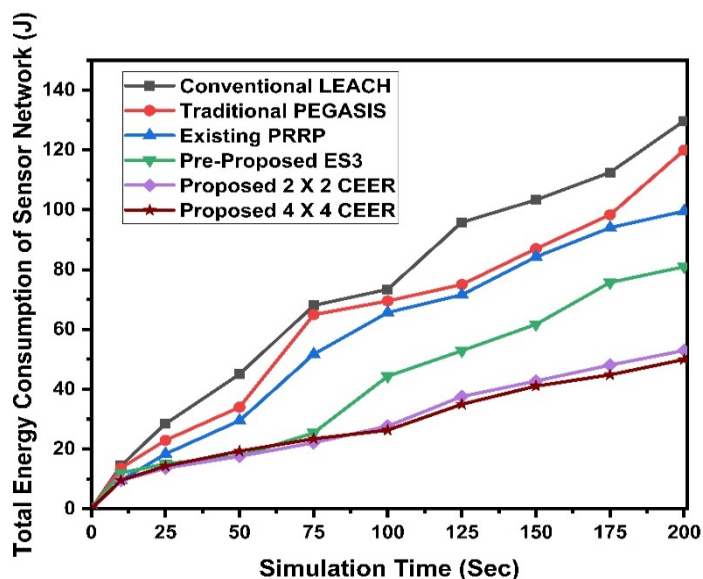


Figure 14: Total energy consumed by proposed and pre-existing sensor networks

in a simulation round of  $t = 200$  Sec when BS is placed at centre of sensing area with co-ordinates (50,50)

When the protocols LEACH, PEGASIS, PRRP, ES3 and CEER are simulated under the same platform of NS-2 version 2.32 for a simulation round of  $t = 200$  Sec, CEER network proved to consume less energy than the other network as shown in Figure 14. The energy saving table of CEER networks when compared to others is shown in Table 4.

Table 4: Outperformance by CEER network in terms of Energy at the end of simulation round

	Comparison with ES3	Comparison with PRRP	Comparison with PEGASIS	Comparison with LEACH
2 X 2 CEER	Saves 28 J	Saves 46.627 J	Saves 66.9947 J	Saves 76.80 J
4 X 4 CEER	Saves 31 J	Saves 49.632 J	Saves 70 J	Saves 79.80 J

#### 4.2.2 Comparison in terms of Average Energy Consumption of Sensor Node in Sensing Field [32,33]

The performance metric “Average Energy consumption of sensor node” is vital in CEER as it determines its network lifetime. In any other protocol this metric has its importance as it shows how efficient is the proposed or implemented algorithm to compete with the existing works. In our scenario, when this metric result is compared with existing works as shown in Figure 15, CEER network saves energy consumed by a sensor node as its algorithm finds energy effective sink for every sensor node deployed in the network area and transmit the data through it to reach the basestation. Indirectly, the energy load is distributed equally and effectively among all the sensor nodes in the network.

From Figure 15, CEER network shows that a sensor node deployed inside its network can have longer lifespan i.e., around 62% than conventional LEACH, 56% than traditional PEGASIS, 38% than existing PRRP and 22% than our pre-proposed ES3.

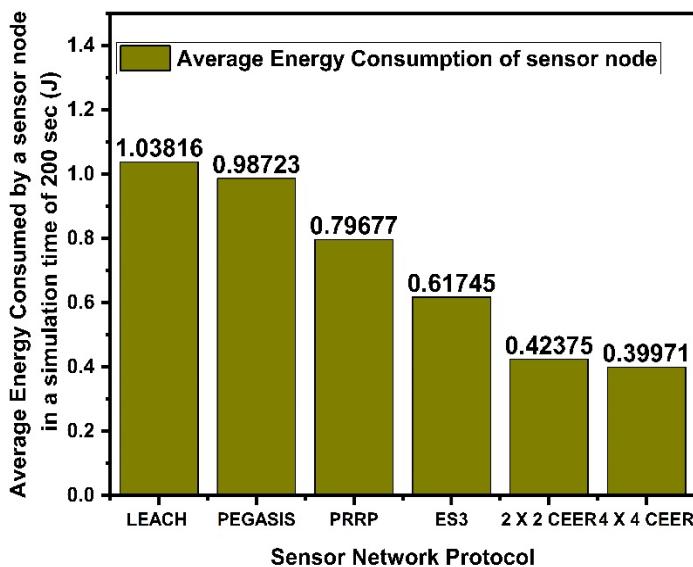


Figure 15: Average Energy Consumed by a sensor node in simulated network protocols at the end of simulation time of  $t = 200$  sec when the basestation is placed at co-ordinates (50,50) in the network area.

#### 4.2.3 Comparison in terms of packet delivery ratio (PDR) of Sensor Network

Packet Delivery Ratio (PDR) is defined as the percentage ratio of packets received by the basestation and the generated packets by the source sensors. In *ns* terms it is defined as the ratio between the client of packages originated by the “application layer” CBR (Constant Bit Rate) sources and the package count received by CBR sink (BS) at the final destination.

$$\text{Packet Delivery Ratio (PDR)} = \frac{(\sum \text{Total packets received by BS at time } t) * 100}{\sum \text{Total packets sensed by sensor nodes in network at time } t} \quad (32)$$

PDR in the CEER network is high as this hierarchical protocol is constructed with stable routing operational procedure. As shown in Figure 16, at the beginning of the network initialization and initial data transfer phase, the PDR of CEER network might be low, but as the network enters the stable operation phase as mentioned in section 3, the PDR is quite good and more as compared to other simulated works under simulation time  $t$  and BS at the centre of the sensing field.

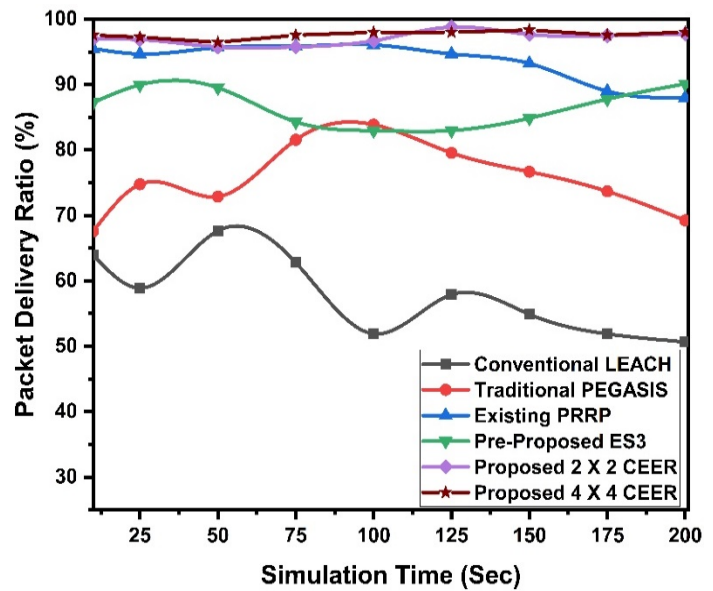


Figure 16: Packet delivery ratio (PDR) of simulated sensor networks in a simulation time  $t = 200$  Sec, when the BS is placed at centre of sensing field at location co-ordinates (50,50)

#### 4.2.4 Comparison in terms of packet loss percentage in Sensing Field

Packet loss rate is calculated in wireless sensor networks to check the reliability of a communication network path. This packet loss percentage is equal to the unreceived packets by the BS to the total packets sensed by sensor nodes in the network. Packet loss occurs due to failure of sensor node that carries data across the sensing area of the network.

$$\text{Packet loss (P.L)} = \left( \text{Total packets sensed by sensor nodes in network in simulation time } t \right) - \left( \text{Total Packets received by BS in simulation time } t \right) \quad (33)$$



$$\text{Packet Loss Percentage} = \frac{(P.L * 100)}{\text{Packets sensed by sensor nodes in network in simulation time } t} \quad (34)$$

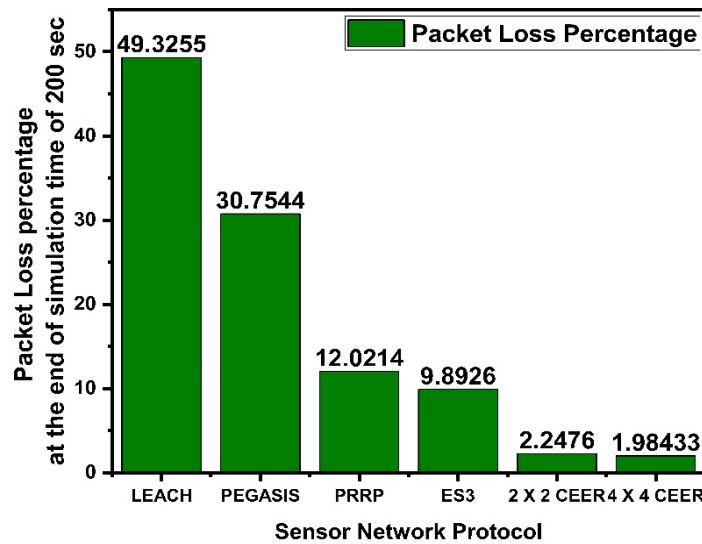


Figure 17: Packet Loss Percentage of sensor networks simulated in time  $t = 200$  sec

Results from Figure 17 shows acceptable packet loss for CEER networks. Reliable protocols react to packet loss automatically and so as the CEER network as its nodes save the excess packets received in one transmission round in the queue memory to avoid further packet loss. Moreover, when compared to conventional LEACH and traditional PEGASIS protocols, CEER network packet loss is more than 47%, 28% lesser respectively. Similarly, when compared to existing PRRP protocol, it is around 10% lesser.

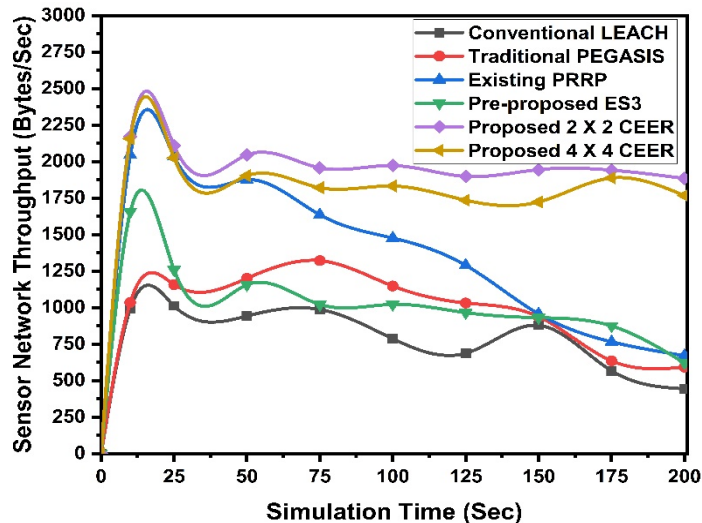


Figure 18: Sensor network throughput measured in bytes/sec of the protocols simulated in a simulation time of  $t = 200$  sec when the BS is placed at locational co-ordinates (50,50)

#### 4.2.5 Comparison in terms of Sensor Network Throughput

Throughput is the rate at which information is sent through the network. Network throughput in CEER networks is an amount of data transferred successfully from source sensors to base station at time  $t$  sec and its result graph is shown in Figure 18. High ratio of unsuccessful packet delivery leads to lower throughput and reduces the network lifespan. Throughput capacity is usually measured in *bits/sec*.

$$\text{Throughput (bits/sec)} = \frac{(\text{Packet count successfully received at BS} * \text{packet size})}{\text{Total time } t \text{ spent in delivering those packets from source sensor to base station}} \quad (35)$$

Packet loss, latency and jitter are all related to reduce the sensor network throughput. All these three metrics are proved to be lesser in CEER network than the compared protocols when simulated based on the assumptions made in section 3.

## 5 Conclusions and Further proceedings

In this research work, we propose a clustering centroid-based energy efficient routing (CEER) protocol to solve the problem of forming zones and its clusters to transmit information energy efficiently to the basestation or sink node. An optimized path is achieved through CEER networks in the zones formed, by selecting energy efficient zone co-ordinators and zone heads. Energy centroid routing is proposed in this paper by considering the residual energy of sensor nodes and communication distance to the sink/basestation. The data path between zone heads and basestation is optimized at both one-hop level and two-hop levels by relaying the information to basestation via intermediate zone heads to avoid packet loss at the zone heads and sometimes directly to basestation itself if the transmission distance between the zone head and the basestation is less. From the simulation results, when the basestation is placed at center of the sensing field, CEER networks can transmit an ample information by consuming less amount of energy. Its simulation results proved that the average energy consumption of a sensor node inside CEER network consumes energy i.e., atleast 60% and 55% less than the conventional LEACH and traditional PEGASIS protocols respectively and atleast 35% than the existing PRRP protocol.

In future work, we have planned to optimize the path between the Zone Heads (ZH) and the Basestation (BS) by introducing multi-hop communication between them unlike the current dual hop and to implement it in IoTN scenario when the BS is placed outside the sensing field at the co-ordinates (50,125). In case if the basestation is still far reachable that the ZH has to dissipate most of its energy for data transfer, we planned to optimize the path using ANT routing strategy to avoid the packet loss.

## Conflict of Interest

The authors declare no conflict of interest.

## References

- [1] N. P. R. Kumar and G. J. Bala, "An Energy Efficient Quadrant Based Position Responsive Routing Protocol," 2019 2nd International Conference on Signal Processing and Communication (ICSPC), Coimbatore, India, 5-10, 2019. <https://doi.org/10.1109/ICSPC46172.2019.8976781>.
- [2] H. Ali et al., "Clustering methods for cluster-based routing protocols in wireless sensor networks: Comparative study", Comparative study. International Journal of Applied Engineering Research, 12(21), 11350-11360, 2017. <http://www.ripublication.com>
- [3] W. Xinhua and W. Sheng, "Performance Comparison of LEACH and LEACH-C Protocols by NS2," 2010 Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science, Hong Kong, 254-258, 2010. <https://doi.org/10.1109/DCABES.2010.58>.
- [4] N. Wang and H. Zhu, "An Energy Efficient Algorithm Based on LEACH Protocol," 2012 International Conference on Computer Science and Electronics Engineering, Hangzhou, 339-342, 2012. <https://doi.org/10.1109/ICCSEE.2012.150>.

- [5] L. Li and C. Liu, "An Improved Algorithm of LEACH Routing Protocol in Wireless Sensor Networks," 2014 8th International Conference on Future Generation Communication and Networking, Haikou, 45-48, 2014. <https://doi.org/10.1109/FGCN.2014.18>.
- [6] S. Gupta and N. Marriwala, "Improved distance energy based LEACH protocol for cluster head election in wireless sensor networks," 2017 4th International Conference on Signal Processing, Computing and Control (ISPC), Solan, 91-96, 2017. <https://doi.org/10.1109/ISPC.2017.8269656>.
- [7] R. Dasondhi, M. Singh, D. Kulhare, "An Algorithm for Balanced Cost Cluster-Heads Selection for Wireless Sensor Network" International Journal of Engineering Research & Technology, **01**(10), 2012. IJERTV1IIS10526
- [8] N. Enami, R. A. Moghadam and K. D. Ahmadi, "A new neural network based energy efficient clustering protocol for Wireless Sensor Networks," 5th International Conference on Computer Sciences and Convergence Information Technology, Seoul, 40-45, 2010. <https://doi.org/10.1109/ICCIT.2010.5711026>.
- [9] S. Mostafavi, V. Hakami, "A new rank-order clustering algorithm for prolonging the lifetime of wireless sensor networks" Int J Commun Syst., 33(e4313), 2020, <https://doi.org/10.1002/dac.4313>
- [10] P. Sasikumar and S. Khara, "K-Means Clustering in Wireless Sensor Networks," 2012 Fourth International Conference on Computational Intelligence and Communication Networks, Mathura, 140-144, 2012. <https://doi.org/10.1109/CICN.2012.136>.
- [11] W. Fakhret, S. E. Khediri, A. Dallali and A. Kachouri, "New K-means algorithm for clustering in wireless sensor networks," 2017 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC), Gafsa, 67-71, 2017. <https://doi.org/10.1109/IINTEC.2017.8325915>.
- [12] Jorio A. Sanaa El Fkihi, Brahim Elbhiri, and Driss Aboutajdine "An Energy-Efficient Clustering Routing Algorithm Based on Geographic Position and Residual Energy for Wireless Sensor Network" Journal of Computer Networks and Communications, 2015, <https://doi.org/11.10.1155/2015/170138>.
- [13] D. Takaishi, H. Nishiyama, N. Kato and R. Miura, "Toward Energy Efficient Big Data Gathering in Densely Distributed Sensor Networks" IEEE Transactions on Emerging Topics in Computing., 2(3), 388-397, 2014. <https://doi.org/10.1109/TETC.2014.2318177>.
- [14] R. Vijayashree and C. Suresh Ghana Dhas., "Energy efficient data collection with multiple mobile sink using artificial bee colony algorithm in large-scale WSN", *Automatika*, **60**(5), 555-563, 2019, <https://doi.org/10.1080/00051144.2019.1666548>
- [15] G. Kumar and J. Singh, "Energy efficient clustering scheme based on grid optimization using genetic algorithm for wireless sensor networks," 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, 1-5, 2013. <https://doi.org/10.1109/ICCCNT.2013.6726634>.
- [16] N. Thakur and R. K. Chauhan, "Conservation of energy by using grid clustering in wireless sensor networks," 2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), Wagnaghat, 591-596, 2016. <https://doi.org/10.1109/PDGC.2016.7913192>.
- [17] A. Ray and D. De, "Energy efficient clustering protocol based on K-means (EECPK-means)-midpoint algorithm for enhanced network lifetime in wireless sensor network," IET Wireless Sensor Systems, **6**(6), 181-191, 2016. <https://doi.org/10.1049/iet-wss.2015.0087>.
- [18] A. A. Shaikh and D. J. Pete, "Spatial Correlation and Centroid Based Clustering in Wireless Sensor Network," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 1-5, 2018. <https://doi.org/10.1109/ICCUBEA.2018.8697416>.
- [19] R. Daniel and K. N. Rao, "EEC-FM: Energy Efficient Clustering based on Firefly and Midpoint Algorithms in Wireless Sensor Network," KSII Transactions on Internet and Information Systems, **12**(8), 3683-3703, 2018. <https://doi.org/10.3837/tiis.2018.08.008>
- [20] Q. Kashif et al., "Optimized Cluster-Based Dynamic Energy-Aware Routing Protocol for Wireless Sensor Networks in Agriculture Precision", *Journal of Sensors*. 1-19, 2020. <https://doi.org/10.1155/2020/9040395>.
- [21] S. Loganathan, J. Arumugam, "Energy centroid clustering algorithm to enhance the network lifetime of wireless sensor networks", *Multimed Syst Sign Process* 2019. <https://doi.org/10.1007/s11045-019-00687-y>
- [22] J. Shen, A. Wang, C. Wang, P. C. K. Hung and C. Lai, "An Efficient Centroid-Based Routing Protocol for Energy Management in WSN-Assisted IoT," *IEEE Access*, **5**, 18469-18479, 2017. <https://doi.org/10.1109/ACCESS.2017.2749606>.
- [23] H. T. Friis, A note on a simple transmission formula. *Proc. IRE*, **34**, 1946.
- [24] T. S. Rappaport, *Wireless communications, principles and practice*. Prentice Hall, 1996.
- [25] R. Hetal, V. Sangeeta, A. Mohammad, "Comparative Study of PEGASIS Protocols in Wireless Sensor Network. *IOSR Journal of Computer Engineering*," **16**, 25-30, 2014. <https://doi.org/10.9790/0661-16512530>
- [26] R. K. Yadav and A. Singh, "Comparative study of PEGASIS based protocols in wireless sensor networks," 1st India International Conference on Information Processing (IICIP), Delhi, 1-5, 2016. <https://doi.org/10.1109/IICIP.2016.7975320>.
- [27] N. Zaman, A.B. Abdullah, "Energy Optimization through Position Responsive Routing Protocol (PRRP) in Wireless Sensor Network," *International Journal of Information and Electronics Engineering*, **2**(5), 748-751, 2012. <https://doi.org/10.7763/IJIEE.2012.V2.199>
- [28] N. Zaman and A. B. Abdullah, "Position Responsive Routing Protocol (PRRP)," 13th International Conference on Advanced Communication Technology (ICACT2011), Seoul, 644-648, 2011.
- [29] T. Qiu, X. Liu, L. Feng, Y. Zhou and K. Zheng, "An Efficient Tree-Based Self-Organizing Protocol for Internet of Things," *IEEE Access*, **4**, 3535-3546, 2016. <https://doi.org/10.1109/ACCESS.2016.2578298>.
- [30] Nalluri, Raj Kumar & Bala, G. "An Efficient Energy Saving Sink Selection Scheme with the Best Base Station Placement Strategy Using Tree Based Self Organizing Protocol for IoT", *Wireless Personal Communications* **109**(2), 869-895, 2019. <https://doi.org/10.1007/s11277-019-06595-5>.
- [31] H. Zhang, P. Chen and S. Gong, "Weighted spanning tree clustering routing algorithm based on LEACH," 2010 2nd International Conference on Future Computer and Communication, Wuha, 223-227, 2010. <https://doi.org/10.1109/ICFCC.2010.5497381>.
- [32] S. Krit & L. Elmaimouni, "Energy consumption in wireless sensor network: simulation and comparative study of flat and hierarchical routing protocols", *IADIS International Journal on Computer Science and Information Systems*. **12**, 109-125, 2017.
- [33] H. Oudani, J. Laassiri, S. Krit and L. El Maimouni, "Comparative study and simulation of flat and hierarchical routing protocols for wireless sensor network," 2016 International Conference on Engineering & MIS (ICEMIS), Agadir, 1-9, 2016. <https://doi.org/10.1109/ICEMIS.2016.7745357>.

## Deep Learning Model for A Driver Assistance System to Increase Visibility on A Foggy Road

Samir Allach\*, Mohamed Ben Ahmed, Anouar Abdelhakim Boudhir

*LIST Laboratory, FSTT, UAE University, Tangier 90000, Morocco*

---

### ARTICLE INFO

*Article history:*

*Received: 22 May, 2020*

*Accepted: 13 July, 2020*

*Online: 28 July, 2020*

---

*Keywords:*

*ADAS*

*Deep learning*

*Image dehazing*

*Foggy road*

*Object detection*

---

### ABSTRACT

*For many years, a lot of researches have been made to develop Advanced Driver Assistance Systems (ADAS) that are based on integrated systems. The main objective is to help drivers. Hence, keeping them safe under different driving conditions. Visibility for drivers remains the biggest problem faced on the road in an atmosphere of fog. In this paper, we examine a system that can be employed to substantially enhance visibility through using deep neural networks. Researches done recently- which are based on deep learning for eliminating image fog- have made clear that an end-to-end proposed system is such an effective model. However, it becomes a must to extend the idea to end-to-end real-time video dehazing. In this paper, we introduce a model of image dehazing. It is based on Convolutional Neural Networks (CNN) as a basis for developing the video dehazing model. As in addition, we concatenate our model with the faster RCNN for detecting objects on the road in real time. The experimental results on our image datasets shows the performance of our model with regard to Peak Signal to Noise Ratio (PSNR=19.823) and Structural Similarity (SSIM =0.8501). On the dataset of the synthesized videos, our model achieved a performance of PSNR = 21.4032 and SSIM = 0.9354. Moreover, with the concatenation of our dehazing model with Faster R-CNN (regions with convolutional neural networks), our proposed system displays desirable visual quality and a remarkable progress of the object detection achievement on blurred images with mean Average Precision (mAP) equal to 0.933 during the day and 0.804 at night.*

---

### 1. Introduction

Advanced Driver Assistance Systems (ADAS) are mainly designed to offer vehicle drivers help, thereby minimizing a potential threat to their safety. The majority of these systems are based on image processing algorithms, such as those allowing detecting of nearby vehicles and pedestrians and the recognition of signs. Because of this, quite a lot of systems are fixed into vehicles. Though such systems are widely employed to draw drivers' attention once a potential threat appears for vehicle drivers, they perform less effectively under certain adverse weather conditions where the vision weakened. This happens most notably in the presence of fog.

Eliminating or reducing the fog of an image captured by an ADAS system seems to be difficult and somehow ill-posed phenomenon. It is worth mentioning here that significant developments that took place in deep learning mainstream paved the way to get considerable results in improving vision degraded

by fog [1]. Therefore, the elimination of fog requires the estimation of the depth map. What is more, previous assumptions are necessary to estimate the depth map for systems using single images as input.

Very lately, a lot of algorithms have been suggested for detecting objects and eliminating fog [1, 2]. It is worth mentioning here that traditional algorithms need two crucial elements: gathered facts on the environment and developed learning as well. In addition, most object detection and fog elimination algorithms are not suited for real-time uses on account of they consist of considerable arithmetic time.

In this way, authors in this area of study have introduced deep learning approaches for degraded vision so that images can be restored and reconstructed [1]-[3]. Yet, these methods cannot directly be applied for removing the fog from the image. The authors F. Hussain, and J. Jeong, came up with an approach, using Deep Neural Network (DNN). They also assumed that an anonymous complex function could model mathematically the fog in an image [4]. Li Chongyi et al. proposed a cascading CNN

---

\*Corresponding Author: Samir Allach, LIST-FSTT Bp 416 Tangier-Morocco, Tel: +212615571400 & Email: allach.samir@gmail.com

model composed of three components: a component of the shared invisible layers that extract the common features, the work of the global estimate of the atmospheric light and that of the subnet of the average transmission [5]. The researchers Li Boyi and others introduced a model for dehaze of image constructed with a convolutional neurons network (CNN), named “all-in-one dehazing network” and “An All-in-One Network for Dehazing and beyond” (AOD-Net) [1, 6]. It is designed on the basis of a reformulated atmospheric diffusion model. The researchers B.Cai et al. found an end-to-end formable model an excellent alternative for calculating the average transmission, called DehazeNet [2], it takes a hazy image at the input and leaves its transmission matrix, and then image dehazed is recovered by the atmospheric diffusion model. W. Ren et al. used a multi-scale CNN (MSCNN) that created and improved a coarse scale transmission matrix [7].

With regard to the video dehazing, most approaches depend crucially on the phase of improving temporal inconsistencies. The authors Kim et al. suggested inserting a temporal coherence in the cost function, with a clock filter to accelerate the processing [8]. The authors Cai et al. conceived a spatio-temporal optimization for de-hazing video in real time [2].

In recent years, an interest has begun to grow in video modeling using CNN for a huge number of tasks. And here are three examples: super-resolution (SR) [9], blur [10] and classification [11, 12] and style transfer [13]. In [9], the author studied various structure configurations for SR video. In [11, 12], the authors made similar attempts by exploring different connectivity options for video classification. The researchers Liu et al. introduced such adjustable system. In this context, they placed a spatial alignment network between the images [14]. Others introduced an end-to-end CNN to learn how to accumulate information on several hazy images / video [10]. With reference to video style transfer, Chen et al. integrated short- and long-term coherence. This is not to mention they indicated the superiority of multi-image approaches over single-image approaches [13].

In this paper, we present a video dehazing system based on an image-dehazing model. We use convolutional neural networks (CNN). Then, we concatenate our model with the faster RCNN to detect objects on the foggy road in real time. Moreover, we deploy our system to validate the quantitative and visual results obtained.

## 2. Network Design

Our proposed work is made of two important steps. Firstly, we worked hard to develop an end-to-end CNN model [15] that explicitly learns the mapping relationships between the raw images and their associated transmission maps to recover fog-free images. And then we integrated the Faster R-CNN proposed model [16] for objects detection in video. The algorithm presents the general steps of our model for dehazing and detecting objects in a video:

### 2.1. Atmospheric diffusion model

The model of atmospheric diffusion is the usual description of the hazy image production process, proposed by McCartney [17] and developed later by Nayar and Narasimhan [18]. The model is written as follows [1,2,19]:

$$I(x) = J(x) t(x) + A(1 - t(x)) \quad (1)$$

---

### Algorithm: Video dehazing and object detection

---

**Result:** video dehazing and object detection

**Let's**  $i$  : counter;

$n$  : number of images after splitting video to images;  
 $V_{in}$  : input video ,  $M$ : Resulting of dehazed image();  
 $Dehz()$ : Dehaze Function;  
 $Detect()$ : Object Detection Function;  
 $Save()$ : Saving resulting image;  
 $Concat()$ : Image concatenation;  
 $V_{split}()$ : Splitting video to images;  
 $V_{out}()$ : Video generating function;  
 $Image1()$ : Output image of  $V_{split}()$ ;  
 $Image2()$ : Output image of  $Detect()$ ;

```

begin
  Insert  $V_{in}$  ;
  Splitting Video to n images;
   $V_{split}(V_{in})$ ;
  Dehazing images and detecting objects using Faster R-CNN;
   $i \leftarrow 0$  ;
  while ( $i \leq n$ )
     $M \leftarrow Dehz(Image1(i))$  ;
     $Image2(i) \leftarrow Detect(M)$  ;
     $Save(Image2(i))$  ;
    if  $i > 0$  then
       $Concat(Image2(i), Image2(i - 1))$  ;
    end
     $i \leftarrow i + 1$ ;
  end
   $V_{out}(image2(i))$ ;
end

```

---

$$t(x) = e^{-\beta d(x)} \quad (2)$$

where  $I(x)$  is an observed hazy image,  $J(x)$  is the image that recovered after the estimation of  $A$  and  $t(x)$ ,  $A$  is the global atmospheric light and  $t(x)$  is the transmission map.

In equation (2),  $\beta$  is the atmosphere diffusion coefficient and  $d(x)$  is used to refer to the distance between the image and the camera.

Equation (2) indicates that when  $d(x)$  moves to infinity,  $t(x)$  comes near to zero. With equation (1), we have:

$$A = I(x), d(x) \rightarrow inf \quad (3)$$

In reality,  $d(x)$  cannot be infinite, but can be a long distance, which gives a very weak transmission  $t_0$ . Then, the global atmospheric light ( $A$ ) is estimated by the following formula

$$A = \max_{y \in \{x | t(x) \leq t_0\}} I(y) \quad (4)$$

### 2.2. Construction of the proposed CNN model

The proposed model for the image frame consists of an estimation module. This module uses convolutional layers to



estimate transmission map  $t(x)$  (see Figure 1), followed by a clean image generation module that consists of a multiplication layer of several element by element and the addition layers for generating the recovery image (see Figure 2).

The estimation module is the essential component, responsible for estimating the depth and level of relative disorder. As Figure 1 shows, we use five convolutional layers by merging filters of various sizes (see Table 1).

After comparing the results obtained by different architectures of CNN (see table 1), we concluded that the CNN model that consists of five convolutional layers with the number of  $3 \times 3$  filters is the most efficient, with a PSNR of 19.8231 and a SSIM 0.8501.

The authors B. Cai et al. used parallel convolutions with filter sizes varied [2]. W. Ren et al. concatenated the distinctive features of the large-scale network with an intermediate layer of the fine-scale network [7]. Influenced by them, we concatenated in our model the layer "pool1" entities with the layers "conv1" and "conv2". The same thing for "pool2" with those of "conv2" and "conv3"; and "pool3" with those of "conv1", "conv2", "conv3" and "conv4". This model captures the characteristics of images at different scales and at the intermediate connections, compensating for the loss of information during convolutions. Notably, in our proposed model each convolutional layer uses only three filters. Consequently, our model is lighter and performs well in terms of PSNR and SSIM values in comparison with other existing deep methods.

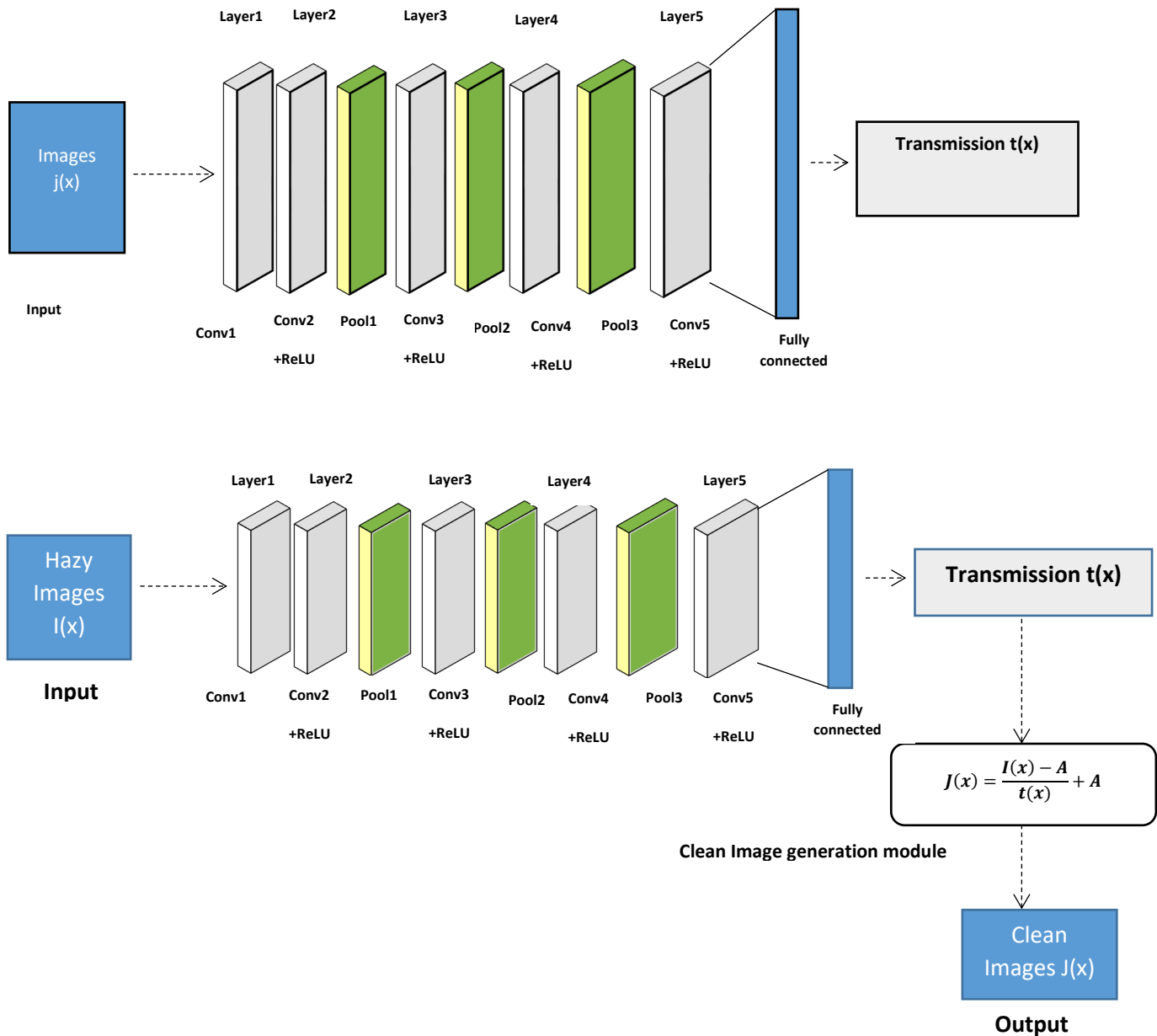


Table 1: Average of PSNR and SSIM results for different architectures of CNN model

Number of convolution layers	Number of filters	PSNR	SSIM
2	3 x 3	-	-
	5 x 5	14.0321	0.6801
	7 x 7	14.1321	0.7032
3	3 x 3	15.6706	0.7791
	5 x 5	16.7653	0.7356
	7 x 7	16.0362	0.7452
4	3 x 3	16.8521	0.7937
	5 x 5	17.7653	0.7745
	7 x 7	18.2381	0.8012
5	3 x 3	<b>19.8231</b>	<b>0.8501</b>
	5 x 5	19.6364	0.8125
	7 x 7	19.5364	0.8029
6	3 x 3	18.5381	0.8101
	5 x 5	18.2381	0.7998
	7 x 7	18.3381	0.8111

Table 2: Comparison of detection methods on the 2007 and 2012 PASCAL VOC dataset

Method	mAP	FPS	Batch size
Fast R-CNN [21]	0.70	0.5	1
YOLO [20]	0.634	45	1
SSD300 [21]	0.743	46	8
Faster R-CNN VGG-16 [21]	0.732	7	1
Faster R-CNN-ResNet [20]	0.764	5	1

Table 2 shows comparison measurements of mAP and the FPS (frames per second) speed with the batch size, using the dataset PASCAL VOC between Fast R-CNN, YOLO, SSD300, Faster R-CNN VGG-16 and R- Faster CNN-ResNetet. Our chosen detection method (Faster R-CNN-ResNetet) surpasses all methods in terms of mAP. Though a set of certain methods may work at higher speeds, they have lower accuracies (mAP). The Faster R-CNN-ResNetet remains the best method in real time to reach more than 76.4% of mAP.

Outside the dehazing part, the temporal coherence should be taken into consideration for object detection so that we can reach satisfactory results. With our proposed model, we can hereby-advancing Faster R- model CNN ideally suited for video [22]. The first two convolutional layers of Faster R-CNN model to an image, was divided into three equivalent sections to insert the earlier, present and upcoming images, in this order. The last three images are concatenated in the next of the second convolutional layer and go through the resting layers for predicting the bounding boxes of the objects for the present image.

### 3. Results and discussion

#### 3.1. Dehazing an image

First, we created foggy images synthesized from equations (1) and (2), using the ground truth images with the NYU2 inner depth database.

We put different atmospheric lights (A), selecting them consistently each channel between [0.5, 1.0] and selecting  $\beta \in \{0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6\}$ .

The input data is RGB images with a resolution of 640 x 480 and the output was a depth map with a resolution of 320 x 240.

In our model, about 12 epoch are enough for it to converge, and it works well enough after these times. During the learning phase, the weights of our network are initialized in a random way. We used the ReLU function as a more efficient neuron in our specific context. We opted for the function of Mean Square Error loss (MSE), and that it stimulates the PSNR (Peak Signal to Noise Ratio) and the SSIM (Structural Similarity) as well as the visual quality.

#### 2.3. Model of Object detection

Faster R-CNN [16] is proposed to detect objects on images accurately. With ResNet, Faster R-CNN achieved a mean average Precision (mAP) of 76.4% on the PASCAL VOC dataset 2007 and 2012 (see Table 2). By combining region and classifier propositions in a large network, it becomes possible to automatically learn good representations of the features for the task.

In this paper, we examine the detection and recognition of objects in the presence of fog with a view to improve high-level vision tasks to combine with the model for dehazing of video. We opted for the Faster R-CNN model as the basic algorithm for detecting robust objects in real time (see Table 2), approved on synthetic and natural blurred images.

We modified the Faster R-CNN Caffe source code and used ResNet [20] as the Faster-RCNN convolution backend. We came to the conclusion that using ResNet provides a substantial improvement over other architectures. The time reweighting of the chain gives further improvements.

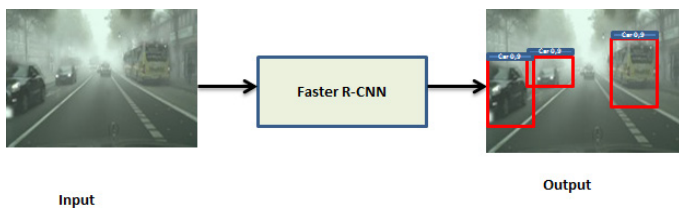


Figure 3: Object detection with Faster R-CNN model

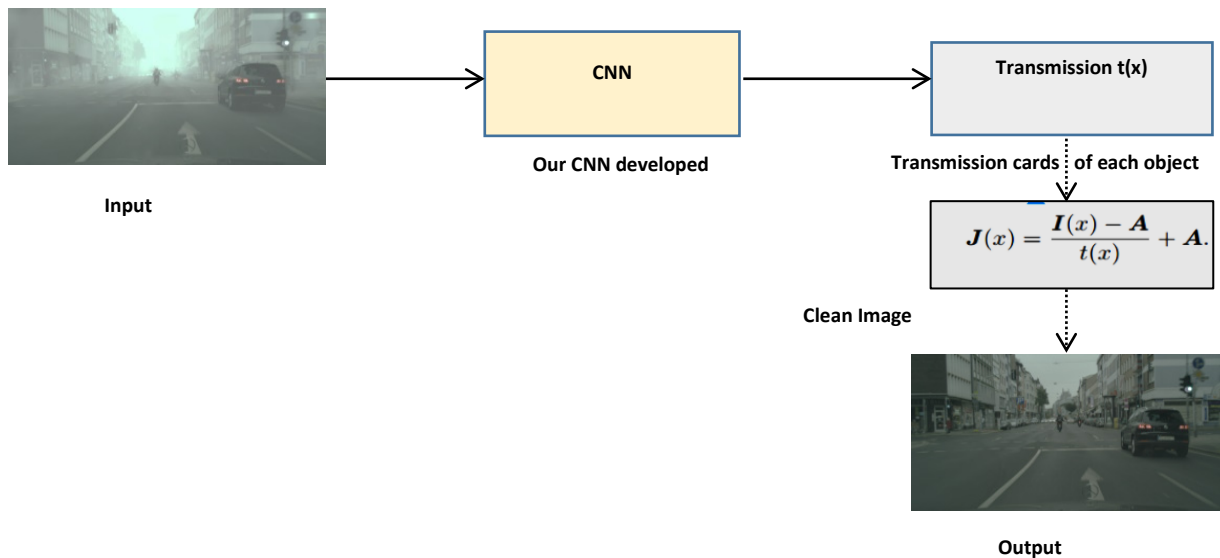


Figure 4: Our model of image dehazing

Table 3: Average PSNR and SSIM results on Test1

	ATM	BCCR	NLD	FVR	DCP	DehazeNet	CAP	MSCNN	AOD-Net	<b>OUR MODEL</b>
PSNR	14.1475	15.7606	16.7653	16.0362	18.5385	18.9613	19.6364	19.1116	19.6954	<b>19.8231</b>
SSIM	0.7141	0.7711	0.7356	0.7452	0.8337	0.7753	0.8374	0.8295	0.8478	<b>0.8501</b>

We compared our basic model proposed with several dehazing methods, in particular: Automatic atmospheric light Recovery (ATM) [24], Regularization of the context constrained to limits (BCCR) [25], No-local Image Dehazing (NLD) [26], Fast visibility restoration (FVR) [27], Dark channel priority (DCP) [28], MSCNN [7], DehazeNet [2], Color Attenuation Prior (CAP) [29] and AOD-Net [1].

Our synthesized hazy images come along with ground truth images, paving the way for us assess PSNR and SSIM and check out if the results stay accurate.

As our model is optimized from start to finish in case of MSE loss, one should not be amazed to see its PSNR performance superior to other methods. Further, even if SSIM is not directly called the optimization criteria, our model obtains SSIM advantages superior to the other models compared.

It is well known that SSIM more accurately reflects human perception, as SSIM measures apart from pixel level errors. We become faithful to our model with such a consistent SSIM improvements we achieved.

Table 3 shows the promising performance of our model compared to the others, in terms of PSNR and SSIM.

Our method has an advantage greater than 0.2 dB in the PSNR and 0.03 in the SSIM when we compare it to approach AOD-Net.

### 3.2. Video sequences dehazing

First of all, we brought into existence a dataset of foggy synthetic video from an equation(1), employing 20 chosen videos in the TUM RGB-D dataset [30] that consists of different video sequences. Depth information is refined by means of the filling model of Silber man et al. [31]. After that, we broke our dataset into a learning set made up of 12 videos with 120,000 images and a set of non-overlapping tests called Test2, consisting of 8 short video sequences containing a total of 450 images. Finally, we collected a set of natural hazy video sequences to validate and evaluate the performance of our model.

During the training of our model, we adopted the loss of the mean squared error (MSE), which is aligned with the SSIM, the PSNR and the visual quality.

Table 4: PSNR average and SSIM results on Test2.

	STMRF	EVD-Net	OUR MODEL
PSNR	18.9956	20.9908	21.4032
SSIM	0.8707	0.9087	0.9354

Concerning video-based methods, we compared our model with EVD-Net [19] and STMRF [32] methods (not based on CNN). Then, we got a higher advantage over the two previous approaches. Furthermore, we observed that our advanced system performs quite remarkably showing a difference of 0.5 dB in the PSNR and 0.03 in the SSIM.

### 3.3. Object detection model

We reformed our adaptive Faster-RCNN model on a set of training data provided by Foggy Cityscapes [33] and on a set of synthesized personal data. Foggy Cityscapes is a foggy synthetic data set. It causes fog to operate on real scenes. The images are rendered utilizing the Cityscapes images and depth maps [8]. It contains 2,975 images in the learning set and 500 images in the validation set. In this experiment, we reported our findings on categories such as person, car, truck, bus, motorcycle and bicycle. We opted for the average score of the channel separately for each class based on validation performance.

On the validation set, we achieve a mAP of 0.777 (see Table 5 for a full breakdown between classes).

Table 5: Average Precision (AP) and Mean Average Precision (mAP) between original Faster R-CNN and adaptive Faster R-CNN with ResNet.

Metrics	Original Faster R-CNN	Adaptive Faster R-CNN with ResNet
Person (AP)	0.731	0.779
Car (AP)	0.809	0.812
Truck (AP)	0.750	0.803
Bus (AP)	0.692	0.792
Motorbike (AP)	0.663	0.723
Bike (AP)	0.621	0.756
mAP	0.711	0.777

### 3.4. The proposed model Architecture for object detection in hazy video

The merging of our model for video dehazing and the adapted Faster R-CNN model (Figure 5) has given birth to our general

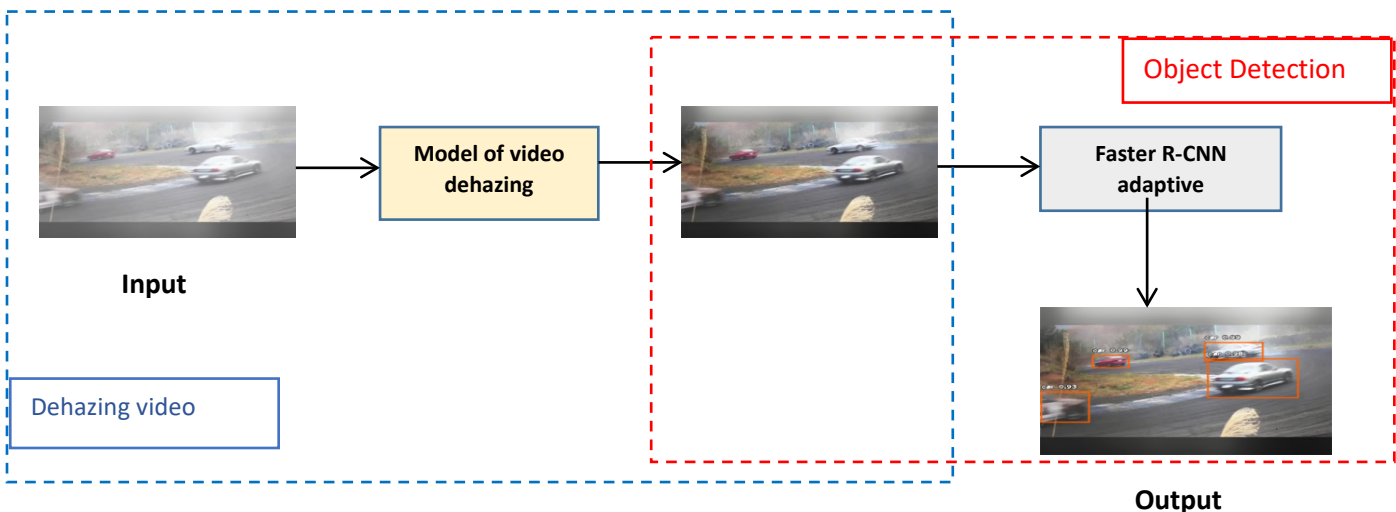


Figure 5: The proposed general model

model, which naturally displays an interesting tree structure that is locally linked and is subjected to joint and more crucial optimization.

On the validation set of our proposed model, we achieved a mAP of 0.929 (see Table 6).

Table 6: Average Precision (AP) and mean Average Precision (mAP) results with adaptive Faster R-CNN before and after the dehazing operation.

Metrics	Adaptive Faster R-CNN with ResNet	
	Before the dehazing	After the dehazing
Person (AP)	0.779	0.902
Car (AP)	0.812	0.911
Truck (AP)	0.803	0.909
Bus (AP)	0.792	0.970
Motorbike (AP)	0.723	0.981
Bike (AP)	0.756	0.901
mAP	0.777	0.929

## 4. Deployment of our system

As an embedded platform, we used a Raspberry Pi 3 Model B+, for our experiments we made predictions on new video sequences using the raspberry Pi with camera designated v2.1, based on the Sony IMX219 CMOS type sensor with a resolution of 8 MP (3280 × 2464 pixels) and we downloaded our pre-trained model on our pi. We needed Raspbian Stretch 9, because TensorFlow 1.9 officially supports Raspberry Pi if you use Raspbian 9.



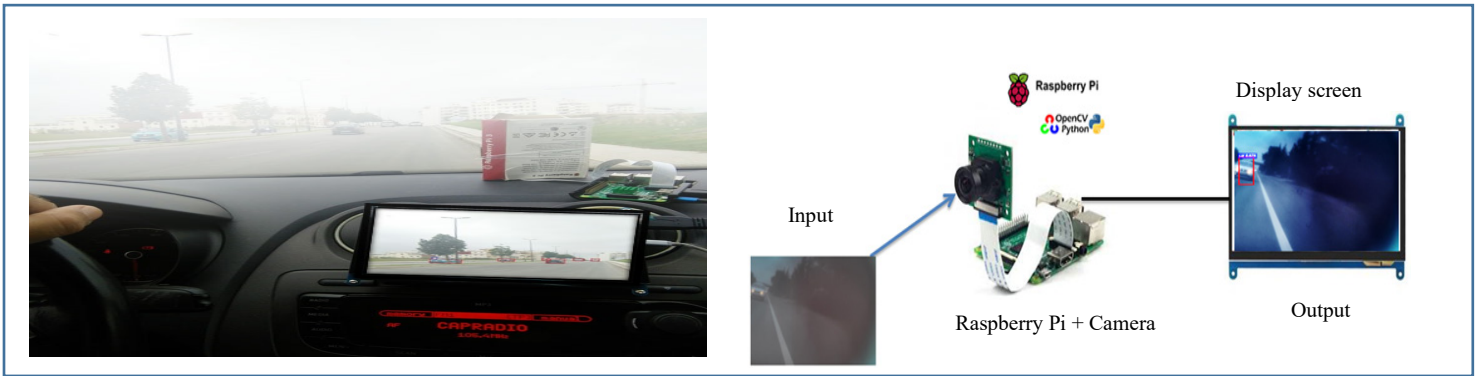


Figure 6: structure of system deployment

Table 7: Results of mAP for 5 states: Heavy haze (day), Medium haze (day), Light haze (day), Medium haze (night), Light haze (night), before and after dehazing.

	During the day			At night	
	Heavy haze	Medium haze	light haze	Medium haze	light haze
Faster R-CNN (adaptive)	mAP = 0.501	mAP = 0.663	mAP = 0.815	mAP = 0.531	mAP = 0.678
Dehazing + Faster R-CNN (adaptive)	mAP = 0.773	mAP = 0.891	mAP = 0.933	mAP = 0.723	mAP = 0.804

Table 8: the visual results of our model on images taken by the raspberry camera

	During the day			At night	
	Heavy haze	Medium haze	Light haze	Medium haze	Light haze
Hazy image					
Objet detection in hazy images					
Objet detection in Image dehazing					

We performed tests on images taken during the day and at night using the Faster R-CNN model adaptive at first, and then the proposed general model (image dehazing + Faster R-CNN adaptive). Table 7 and 8 show the results obtained.

#### 4.1. The Quantitative results

For quantitative results, we calculated the mean Average Precision (mAP) across all images, using both the adaptive Faster R-CNN model and the dehazing model concatenated with the adaptive Faster R-CNN. (See Table 7). During the day, the heavy haze degrades mAP to around 0.27. Adding the dehazing model, mAP is improved by 0.12 for the detection of objects under a light veil conditions, 0.23 in medium fog and 0.27 in thick fog. During the night, with the dehazing model the mAP improves by 0.28 for detecting objects in light haze conditions and 0.21 in the middle fog (see table 7).

#### 4.2. The results visualized

For the visualized results, Table 8 makes a visual comparison of the results when the object is detected on hazy images. And here we are illustrating five cases:

- **During the day:**

- Detection of objects in the Heavy Haze image with adaptive Faster R-CNN before and after dehazing.

- Detection of objects in the Medium haze image with adaptive Faster R-CNN before and after dehazing.

- Detection of objects in the Light haze image with adaptive Faster R-CNN before and after dehazing.

- **At night:**

- Detection of objects in the Medium haze image with adaptive Faster R-CNN before and after dehazing.

- Detection of objects in the Light haze image with adaptive Faster R-CNN before and after dehazing.

Experiments we conducted lead us to this conclusion: as soon as the haze becomes heavier at night, detecting objects becomes less dependable. Most of all, under whichever fog conditions- light or medium or high- our highly advanced system can regularly ameliorate the process of detecting objects.

## 5. Conclusion

In this paper, we present an intelligent system with a view to enhance visibility quality in atmospheric fog, allowing dehaze and object detection. The proposed system is based on the deep learning model using a light CNN model of five convolutional layers works well in terms of PSNR and SSIM compared to other existing deep methods.

We modified the Faster R-CNN model, using ResNet as its convolution backend. We concluded that using ResNet offers a substantial improvement over other architectures. Further, we concatenated the CNN video dehazing model with the faster RCNN as a robust model compared to the others. Our objective was detecting objects on the road in real time with significant performance in terms of mAP.

Based on the qualitative and visual results, our system demonstrated both efficiency and superiority over other existing systems in terms of PSNR, SSIM, mAP and visual quality.

It should be noted that, though the proposed system was much more applicable in enhancing visibility for road drivers when there is fog, there were some limitations. First of all, we did not test our proposed system in case of a rainy foggy environment. Secondly, our system is not applicable in case of very dense fog. Lastly, it is not also applicable when detecting so many objects. Consequently, in our future work we shall definitely spare no effort to enhance the performance of our system by an automated Chabot. This Chabot interprets the captured image and announces audio warnings in the event of critical situations, a project that mainly aims at developing a system in the situation where the presence of fog and rain come at the same time.

## References

- [1] B.Li, X.Peng, Z.Wang, J.Xu, D. Feng, "Aod-net: All-in-one dehazing network." In Proceedings of the IEEE International Conference on Computer Vision, 4770-4778, 2017. DOI: 10.1109/ICCV.2017.511
- [2] B.Cai, X.Xu, K.Jia, C.Qing, D. Tao, "Dehazenet: An end-to-end system for single image haze removal." IEEE Transactions on Image Processing, 25(11), 5187-5198, 2016. DOI: 10.1109/TIP.2016.2598681
- [3] C.Dong, C. C.Loy, K.He, X.Tang, "Image super-resolution using deep convolutional networks." IEEE transactions on pattern analysis and machine intelligence, 38(2), 295-307, 2015. DOI: 10.1109/TPAMI.2015.2439281
- [4] F.Hussain, J. Jeong, "Visibility enhancement of scene images degraded by foggy weather conditions with deep neural networks." Journal of Sensors, 2016. <https://doi.org/10.1155/2016/3894832>.
- [5] C.Li, J.Guo, F.Porikli, H.Fu, Y.Pang, "A cascaded convolutional neural network for single image dehazing." IEEE Access, 6, 24877-24887, 2018. DOI: 10.1109/ACCESS.2018.2818882
- [6] B.Li, X.Peng, Z.Wang, J.Xu, D.Feng, "An all-in-one network for dehazing and beyond." arXiv preprint, 2017. <https://arxiv.org/abs/1707.06543>
- [7] Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., & Yang, M. H. "Single image dehazing via multi-scale convolutional neural networks." In 2016Springer/Cham, European conference on computer vision, 154-169, 2016. [https://doi.org/10.1007/978-3-319-46475-6\\_10](https://doi.org/10.1007/978-3-319-46475-6_10)
- [8] J. H.Kim, W. D.Jang, J. Y.Sim, C. S. Kim, "Optimized contrast enhancement for real-time image and video dehazing." Journal of Visual Communication and Image Representation, 24(3), 410-425, 2013. <http://dx.doi.org/10.1016/j.jvcir.2013.02.004>
- [9] A.Kappeler, S.Yoo, Q.Dai, A. K. Katsaggelos, "Video super-resolution with convolutional neural networks." IEEE Transactions on Computational Imaging, 2(2), 109-122, 2016. DOI: 10.1109/TCI.2016.2532323
- [10] S.Su, M.Delbracio, J.Wang, G.Sapiro, W.Heidrich, O.Wang, "Deep video deblurring for hand-held cameras." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1279-1288, 2017. DOI: 10.1109/CVPR.2017.33
- [11] A.Karpathy, G.Toderici, S.Shetty, T.Leung, R.Sukthankar, L. Fei-Fei, "Large-scale video classification with convolutional neural networks." In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 1725-1732, 2014. <https://doi.org/10.1109/CVPR.2014.223>
- [12] H.Shen, S.Han, M.Philipose, A.Krishnamurthy, "Fast video classification via adaptive cascading of deep models." In Proceedings of the IEEE conference on computer vision and pattern recognition, 3646-3654, 2017. DOI: 10.1109/CVPR.2017.236
- [13] D.Chen, J.Liao, L.Yuan, N.Yu, G.Hua, "Coherent online video style transfer." In Proceedings of the IEEE International Conference on Computer Vision, 1105-1114, 2017. DOI: 10.1109/ICCV.2017.126
- [14] D.Liu, Z.Wang, Y.Fan, X.Liu, Z.Wang, S.Chang, T. Huang, "Robust video super-resolution with learned temporal dynamics." In Proceedings of the IEEE International Conference on Computer Vision, 2507-2515, 2017. DOI: 10.1109/ICCV.2017.274

- [15] S.Albawi, T. A.Mohammed, S.Al-Zawi, "Understanding of a convolutional neural network." In 2017IEEE International Conference on Engineering and Technology (ICET), 1-6, 2017. DOI: 10.1109/ICEngTechnol.2017.8308186
- [16] Y.Xu, G.Yu, Y.Wang, X.Wu, Y. Ma, "Car detection from low-altitude UAV imagery with the faster R-CNN." Journal of Advanced Transportation, 2017. <https://doi.org/10.1155/2017/2823617>
- [17] E. J. McCartney, "Optics of the atmosphere: scattering by molecules and particles." New York, John Wiley and Sons, Inc., 421, 1976. <https://doi.org/10.1063/1.3037551>
- [18] S. G.Narasimhan, S. K. Nayar, "Chromatic framework for vision in bad weather." In 2000 Proceedings IEEE Conference on Computer Vision and Pattern Recognition, Vol. 1, 598-605, 2000. DOI: 10.1109/CVPR.2000.855874
- [19] B.Li, X.Peng, Z. Wang, J.Xu, D. Feng, "End-to-end united video dehazing and detection." In 2018 AAAI Thirty-Second Conference on Artificial Intelligence, 2018. <https://arxiv.org/abs/1709.03919>
- [20] K.He, X.Zhang, S.Ren, J. Sun, "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778, 2016. DOI: 10.1109/CVPR.2016.90
- [21] W.Liu, D.Anguelov, D.Erhan, C.Szegedy, S.Reed, C. Y.Fu, A. C. Berg, "Ssd: Single shot multibox detector." In 2016 Springer/Cham, European conference on computer vision, 21-37, 2016. [https://doi.org/10.1007/978-3-319-46448-0\\_2](https://doi.org/10.1007/978-3-319-46448-0_2)
- [22] S.Ren, K.He, R.Girshick, J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks." In Advances in neural information processing systems, 91-99, 2015. DOI: 10.1109/TPAMI.2016.2577031
- [23] Z.Li, P.Tan, R. T.Tan, D.Zou, S.Zhiying Zhou, L. F. Cheong, "Simultaneous video defogging and stereo reconstruction." In Proceedings of the IEEE conference on computer vision and pattern recognition, 4988-4997, 2015. DOI: 10.1109/CVPR.2015.7299133
- [24] M.Sulami, I.Glatzer, R.Fattal, M.Werman, "Automatic recovery of the atmospheric light in hazy images." In 2014 IEEE International Conference on Computational Photography (ICCP), 1-11, 2014. DOI: 10.1109/ICCPHOT.2014.6831817
- [25] G.Meng, Y.Wang, J.Duan, S.Xiang, C.Pan, "Efficient image dehazing with boundary constraint and contextual regularization." In Proceedings of the IEEE international conference on computer, 617-624, 2013. DOI:10.1109/ICCV.2013.82
- [26] D.Berman, T.Treibitz, S.Avidan, "Air-light estimation using haze-lines." In 2017 IEEE International Conference on Computational Photography (ICCP), 1-9, 2017. DOI: 10.1109/ICCPHOT.2017.7951489
- [27] J. P. Tarel, N. Hautiere, "Fast visibility restoration from a single color or gray level image." In 2009 IEEE 12th International Conference on Computer Vision, 2201-2208, 2009. DOI: 10.1109/ICCV.2009.5459251
- [28] K.He, J.Sun, X. Tang, "Single image haze removal using dark channel prior." IEEE transactions on pattern analysis and machine intelligence, 33(12), 2341-2353, 2010. DOI: 10.1109/TPAMI.2010.168
- [29] Q.Zhu, J.Mai, L.Shao, "A fast single image haze removal algorithm using color attenuation prior." IEEE transactions on image processing, 24(11), 3522-3533, 2015. DOI: 10.1109/TIP.2015.2446191
- [30] J.Sturm, N.Engelhard, F.Endres, W.Burgard, D.Cremers, "A benchmark for the evaluation of RGB-D SLAM systems." In 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, 573-580, 2012. DOI: 10.1109/IROS.2012.6385773
- [31] F. Mal, S. Karaman "Sparse-to-dense: Depth prediction from sparse depth samples and a single image." In 2018 IEEE International Conference on Robotics and Automation (ICRA), 1-8, 2018. DOI: 10.1109/ICRA.2018.8460184
- [32] B.Cai, X.Xu, D. Tao, "Real-time video dehazing based on spatio-temporal mrf." In 2016 Springer/Cham, Pacific Rim Conference on Multimedia, 315-325,2016. DOI: 10.1007/978-3-319-48896-7\_31
- [33] C.Sakaridis, D.Dai, L. Van Gool, "Semantic foggy scene understanding with synthetic data." International Journal of Computer Vision, 126(9), 973-992, 2018. <https://doi.org/10.1007/s11263-018-1072-8>

## Logistics Solutions in the Preparation Phase for the Appearance of Disasters

Erika Barojas-Payán<sup>\*1</sup>, Diana Sánchez-Partida<sup>2</sup>, Miguel-Josué Heredia-Roldan<sup>3</sup>, Victorino Juárez-Rivera<sup>1</sup>, Jesús Medina-Cervantes<sup>1</sup>

<sup>1</sup>Faculty of Engineering, Veracruz University, Xalapa 94466, Mexico

<sup>2</sup>Department of Logistics and Supply Chain Management, Universidad Popular Autónoma del Estado de Puebla, Puebla 72410, Mexico

<sup>3</sup>Gustavo A. Madero Technological Institute, Technological National of Mexico, Mexico 07470, Mexico

### ARTICLE INFO

Article history:

Received: 14 April, 2020

Accepted: 03 June, 2020

Online: 28 July, 2020

Keywords:

Warehouse location

Inventory levels

Routing Problem

### ABSTRACT

Natural disasters have caused not only economic but also human losses. These events bring with them, among other things, deficiencies in the supply of food, clothing, health and cleaning products, to name a few. This situation makes it imperative to locate facilities that can supply the needs, as mentioned earlier, to the victims in the shortest possible time. This document presents the evaluation of a logistic model of the literature whose foundations are: a) the classic  $p$ -median problem for the location of a pre-positioning warehouse; b) an extension of the  $(q-R)$  model for calculating inventories of different products, according to different types of demand, and c) the problem of multiple vehicles routing with the capacity to establish delivery routes from warehouses to the affected municipalities. This model is evaluated with 90 municipalities belonging to Veracruz, Mex. The results show that the location of the warehouse falls in the municipality of Fortín de las Flores, the inventory levels for five demands and four product kits, and product delivery routes, which are a total of 12, with which a favorable cost minimization is obtained.

## 1. Introduction

This document is an extension of an original work presented at the “2019 IEEE International Conference on Engineering Veracruz (ICEV)” [1], which derives from a global problem, such as the impact of *natural phenomena of hydrometeorological* type on the world. The Office of the United Nations Organization (UN) for Humanitarian Affairs (2020), mentions that from the year 2000 to date, 152 million Latin Americans and Caribbean people have been affected by 1205 disasters, including floods, hurricanes and storms, earthquakes, avalanches, droughts, fires, extreme temperatures and volcanic events [2]. It should be mentioned that natural phenomena do not affect society in the same way; the so-called hydrometeorological phenomena such as heavy rains and floods have a higher frequency compared to other types of natural phenomena and tend to affect the same region in relatively frequent periods. In the last 20 years, these were responsible for causing 90% of economic losses worldwide.

In Mexico, the effects of hydrometeorological phenomena appear depending on the distribution of the population throughout

its territory. In a broad sense, northern Mexico is characterized by low levels of precipitation, frequent droughts, and recurring heat waves. In contrast, the south of the country has the highest records of recurrent precipitation and flooding, in addition to being the region most affected by the effects of hurricanes and other tropical storms. During 2016 these were responsible for causing 86% of the damage in Mexico [3, 4].

The National Civil Protection System (SINAPROC, for its acronym in Spanish), was created to respond to this and other types of natural phenomena. The said system, together with various federal and state agencies, organizations, and entities, creates the necessary instruments and mechanisms in order to carry out coordinated actions in the area of civil protection [5].

Within SINAPROC, there is the National Fund for Natural Disasters (FONDEN, for its acronym in Spanish), which is an inter-institutional mechanism that aims to execute actions, authorize and apply resources to mitigate the effects of a disturbing natural phenomenon, in regarding of the SINAPROC. For access to that fund, the Federal Government's Ministry of the Interior may issue: a) an *emergency declaration*, i.e., an act by which the

\*Corresponding Author: Erika Barojas-Payán, Ixtaczoquitlán, Ver., Mex., Email: [eribarojas@gmail.com](mailto:eribarojas@gmail.com)



Ministry recognizes that one or more territorial demarcations or municipalities are in imminence, high probability of presence of an abnormal situation generated by a disturbing agent. Therefore, immediate assistance is required to the population whose safety and integrity are at risk, or *b) a disaster declaration*, i.e., the act by which the Secretary recognizes the presence of a natural agent. Severe disturbance in particular municipalities, whose damage exceeds the local financial and operational capacity for their care, in order to access resources from the financial instrument for natural disaster care [6, 7, 8].

Based on the previous information, it should be noted that the southern states and the Gulf of Mexico coasts such as Veracruz, Guerrero, Tabasco, Chiapas, and Oaxaca concentrate the expenses to support the reconstruction of damaged or destroyed infrastructure by FONDEN due to the effects of hydrometeorological events. While, in the northern region of Mexico, the states of Nuevo León and Baja California Sur were the ones that required the largest amounts [9].

Likewise, Table 1 contains a synthesis of the declarations issued during the period covered from 2000 to January 2020, by FONDEN, for the states of the southern zone and the coast of the Gulf of Mexico. It can be seen: *a) the number of municipalities that the state houses; b) the number of disaster declarations; and c) the number of declarations of emergency.* In that Table, it can be identified that the most significant number of declarations are issued on the state of Veracruz.

Table 1: Declarations from 2000 - 2020

State	# of municipalities	disaster declaration	emergency declaration
Chiapas	118	59	75
Guerrero	81	34	36
Oaxaca	570	61	95
Tabasco	17	14	31
Veracruz	212	114	205

Own elaboration with information from [9, 10].

Based on that information, the present document evaluates the model presented in [1], whose bases fall on the problems of *a) location of facilities, b) inventory levels, and c) multiple vehicle routing with capacity.* All this, to generate a feasible solution that allows the establishment of a pre-positioning warehouse, with product inventory levels that allow supplying people affected by a hydrometeorological phenomenon, as well as their delivery routes. In order to provide a decision-making tool to institutions in the face of disaster.

Unlike the conference document, this research expands the geographical space, carrying out the zoning of the state of Veracruz, and the application of the model to one of them. This central area is made up of 90 municipalities, of which 66 of them were affected by some hydrometeorological phenomenon during 2016, as well as the expansion of products to be supplied, introducing the health one.

The document is made up of the following sections from this: *a) Literary Review*, contains studies carried out by researchers immersed in the area of emergency logistics; *b) Methodology*, shows step by step how the research was carried out until the results were obtained; *c) Conclusions*, to which the authors arrived,

after an analysis of results and *d) Bibliographic sources*, for support and foundation of the research.

## 2. Literate Review

Relief operations often focus on response rather than preparedness due to the unpredictability of natural disasters. It leads to reactive rather than proactive systems. The effectiveness of the response may be determined by the structure of the supply chain [11]. These supplies can be said to be preparatory elements if some (or many) of them are available when needed, within strategic establishments [12]. Since the main cause and magnitude of the event of most disaster scenarios in Mexico is unknown, it is very complex to make some kind of forecast, but it is possible to prepare for it.

In recent years, many researchers from various countries have addressed the problem of the location of facilities in various fields and areas, an example of this is the health area, in which optimal location of hospitals, clinics, warehouses of medicines, and natural disaster centers that allow them to provide their services to the highest number of people in the shortest time. Decision making regarding the optimal or feasible location of a facility involves both significant investment and operating costs from the start of the project, which is why this is a crucial issue to be studied [13].

Research has been carried out over time regarding the issue of humanitarian logistics, and these documents reflect the application of different methodologies, methods, and tools in any of the stages of the disaster: (1) prevention; (2) mitigation; (3) preparation; (4) alert; (5) response; (6) rehabilitation, or (7) reconstruction [14].

These inquiries take as a basis for study one or several elements of logistics, such as the location of facilities, which have been captured in [15], where researchers examine the problem through a survey focused on four different models: 1. Deterministic; 2. Stochastic; 3. Dynamic, and 4. Robust.

In the same way in [16], they present two models, one deterministic, which considers the dependence of the damages about the distance between the disaster response facilities and the location of the affected population, and another stochastic, which extends the first concerning the intensity of the damage. A cost advantage of the second model over the first is demonstrated. Likewise, the researchers of [17], raise the impact of the mean time between disasters (MTBD) in the strategy of pre-positioning of medical supply inventory. The authors carry out a compensation model implemented on a platform based on a spreadsheet. In this area, we can mention the authors in [18], who propose a relocation of sites for the delivery of support in Mexico, through the problem of the p-median modeled in the LINGO software, thereby achieving the minimization of the current path at 25%. Meanwhile, in [19], the authors develop a stochastic programming model for shelter location in disaster preparedness, introducing a Genetic Algorithm (GA) for larger instances.

In addition to the above, some authors carry out studies focused on the location of facilities, adding *the capacity term*, in order to supply the victims assigned to it, such as the case of [20] whose study focuses on the location of warehouses for the distribution of humanitarian aid and its capacity, through a mixed-integer linear programming model.

Another element of logistics through which different investigations have been carried out under different optimization models, heuristics and hybrids, is the problem of vehicle routing, among which, it is worth mentioning [21], where multiple vehicle routing is presented with stochastic demand, in which researchers develop a nonlinear mathematical model with a probability constraint approach. This model is linearized with separable programming methods. While researchers in [22] develop a greedy iterative heuristic for vehicle routing with stochastic post-disaster times. Likewise, the researchers in [23] propose a model that determines the policy of resupply of products in disaster containers and generates the optimal route for vehicles to carry out the activity, maximizing the total benefits from sales revenue and costs of transport in a planning horizon. It is worth mentioning the investigation of [24], It is developed in a municipality in the state of Veracruz, Mexico, within which the heuristics of the closest neighbor (NN) generate product delivery trajectories to localities in the emergency phase of the disaster.

Traditional approaches had analyzed production, inventory, and transport operations separately rather than considering the entire operation as a whole [25, 26]. For example, when deciding inventory levels and product shipments in a given logistics network, inventory levels were generally determined first, and then product shipments were determined. However, as industries have become global and more competitive, simultaneous analysis of complex interactions between production and distribution operations has become essential to achieve efficiency [27].

Other integrated network models include [28], which studies an integrated production and transportation planning model with multiple retailers, products, and time frames. He used a two-phase heuristic model to solve the resulting mixed-integer program. Heuristics first gets an initial solution for distribution and production decisions.

However, the literature is scarce in terms of integrative approaches that consider a comprehensive plan that covers both aspects: plant location, inventory levels, and product supply. Only recent works take into account these two dimensions of the problem [29, 30].

In this scale, three logistical elements are introduced, such as location, inventories, and routing, such as the case [31], in which researchers address a plan to transfer of people affected by a flood to safe areas, they formulate a linear and mixed programming model, which allows defining the location and number of shelters and distribution centers with a pre-positioned inventory of aid packages, in addition to the evacuation routes for the victims.

### 3. Methodology

The research considers two of the regions immersed in the subject state of Veracruz that is used as a case study, called the *Capital Region* and the *High Mountains Region*, which are located in the central area of the entity. This region is composed of 90 municipalities. Moreover, 66 of its municipalities have been emergency declarations by the National Fund for Natural Disasters during 2016. Figure 1, shows the methodology to be followed, begins with the collection, selection, and analysis of the information, continues with the programming of the logistics model that allows the location of a host municipality for the installation of the pre-positioned warehouse, the calculation of the

levels of each product for said warehouse and the set of routes for the delivery of products, to end with the analysis of the results obtained.

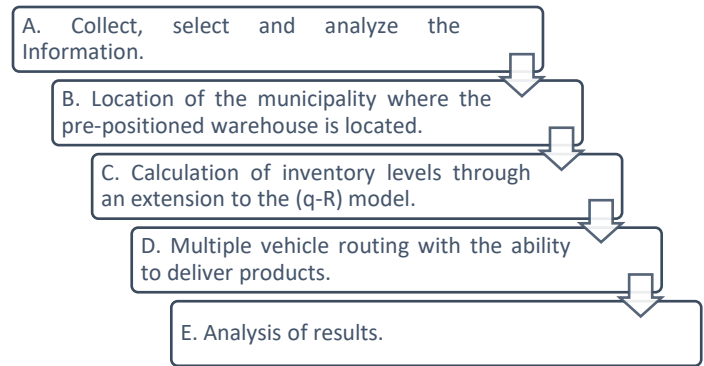


Figure 1: Applied methodology

#### 3.1. Collect, select, and analyze the information.

An analysis of the information necessary to feed the model from [1] and its evaluation is carried out, which is described in the following paragraphs.

- *Matrix of distances between municipalities with the emergency declaration and municipalities without the emergency declaration*, using the Google Earth Pro® application, the distances are obtained that will allow the creation of the base matrix to locate a municipality without an emergency declaration in 2016 for the installation of the pre-positioned warehouse and the routes to the municipalities affected by the natural phenomenon in the same year, for the delivery of products.
- *Positional weight*, is the value made up of the degree of satisfaction that a municipality meets according to its attributes of *a) highway infrastructure, highway type; b) services, water, network, telephone, drainage, electricity, c) not being a reason for issuing an emergency declaration, that is, not having been impacted by the hydrometeorological phenomenon, and d) constructed buildings*. This variable, together with the distance, will allow finding the municipality where the pre-positioning warehouse is located. The data was extracted from the pages of [32, 33, 34].
- *Demand*, five types of demand: *a) early childhood; b) childhood; c) adolescence; d) adulthood, and e) old age*; taking as a reference to the World Health Organization, ruling out the peculiar stages such as pregnancy. The calculation of victims is based on the emergency declarations issued and the number of people by type of demand that lives in the area [35].
- *Products*, the items to be supplied, will be calculated in the form of kits: *a) food; b) equipment (clothing, hygiene, and protection); c) health, and d) water*. The first two, calculated according to the type of demand and the last two for total demand.
- *Total Inventory cost includes a) ordering cost, b) purchasing cost, c) cost per shortage, and d) holding cost*, [1, 36, 37, 38].

#### 3.2. Location of the municipality where the pre-positioned warehouse is located

For the location of the host municipality of the pre-positioned warehouse, the p-median problem is taken as a base, adding the

positional weight variable, with which not only minimum routes are sought, but also adequate conditions for the management of the warehouse. Below is the mathematical formulation:

$$\min \sum_{i=1}^n \sum_{j=1}^m w_i d_{ij} x_{ij} \tag{1}$$

**subject to**

$$\sum_{j=1}^m x_{ij} = 1, \quad \forall i, \tag{2}$$

$$x_{ij} \leq y_j, \quad \forall ij, \tag{3}$$

$$\sum_{j=1}^m y_j = p, \tag{4}$$

$$x_{ij}, y_j \in \{0, 1\} \tag{5}$$

Where: *equation (1)* objective function, seeks to minimize the sum of the distances  $d_{ij}$  traveled for each location of the demand,  $w_i$  being the variable that qualifies the attributes of each location with the possibility of hosting the facility; *equation (2)* confirms that each client is assigned to a single facility  $n$ ; *equation (3)* confirms that each customer is supplied by a single facility; *equation (4)* assign  $p$  facilities, and *equation (5)* indicates that the variables are binary [18].

### 3.3. Inventory level calculation through an extension to the (q-R) model

The following mathematical formulation shows how the calculation of inventory levels for the different types of demand is carried out, which is an extension of the (q-R) model.

$$\min \sum_{i=1}^I \sum_{j=1}^J \sum_{p=1}^P \left[ \frac{co_i^p d_j^p}{Q_{ij}^p} + Cs_i \left( \frac{Q_{ij}^p}{2} + z_{cls} * S_{D_j}^p \right) + C_i^p D_j^p + \frac{D_j^p}{Q_{ij}^p} (Cf_i^p + S_{D_j}^p * E_{z_i}^p) \right] \tag{6}$$

**subject to**

$$S_{D_j}^p = S_{D_j}^p \sqrt{L_i} \quad i = 1, 2, \dots, I \quad j = 1, 2, \dots, J \tag{7}$$

$$R_{D_j}^p = D_j^p * L_i + z_{CSL} * S_{D_j}^p \quad i = 1, 2, \dots, I \quad j = 1, 2, \dots, J \quad p = 1, 2, \dots, P \tag{8}$$

$$Q_{ij}^p = \sqrt{2Co_i \frac{\sum_{j=1}^J y_j^p D_j^p}{Cs_i^p}} \quad i = 1, 2, \dots, I \quad j = 1, 2, \dots, J \quad p = 1, 2, \dots, P \tag{9}$$

$$E_{z_i}^p = z [F_s^p(z) - 1] + f_s^p(z) \quad i = 1, 2, \dots, I \quad p = 1, 2, \dots, P \tag{10}$$

$$FR_i^p = 1 - \frac{S_{D_j}^p E_{z_i}^p}{Q_{ij}^p} \quad i = 1, 2, \dots, I \quad p = 1, 2, \dots, P \tag{11}$$

$$Q_{ij}^p \in Z^+ \tag{12}$$

The equation (6), seeks to minimize costs and is made up of:

- a) ordering cost,  $\frac{co_i^p d_j^p}{Q_{ij}^p}$ ; b) holding cost  $Cs_i^p \left( \frac{Q_{ij}^p}{2} + z_{cls} * S_{D_j}^p \right)$ ; c) purchasing cost,  $C_i^p D_j^p$ , and d) Stockout cost  $\frac{D_j^p}{Q_{ij}^p} (Cf_i^p + S_{D_j}^p * E_{z_i}^p)$ ; the *equation (7)*, performs the calculation of the adjusted standard deviation for each demand ( $S_{D_j}^p$ ). It is mentioned that the demand has a normal distribution of its data; the *equation (8)*, performs the calculation to obtain the point to which the inventory of products ( $p$ ) of each of the demands of the affected municipalities ( $D_j$ ) must reach, to issue a purchase order for these; *equation (9)*, determines the number of products ( $p$ ) to order for the pre-positioning warehouse ( $i$ ) that will allow supplying the different demands of the affected municipalities ( $j$ ); *equation (10)*,

evaluates the probability of shortages to victims of products ( $p$ ); *equation (11)* measures the level of product delivery service ( $p$ ) of the pre-positioned warehouse located in the municipality ( $i$ ) for municipalities ( $j$ ), and finally, *equation (12)*, indicates that ( $Q_{ij}^p$ ) is the integer variable.

### 3.4. Multiple vehicle routing with the ability to deliver products

Through the problem of vehicle routing with capacity, the necessary routes are obtained that will allow supplying products to people affected by the natural phenomenon of hydrometeorological type, and the mathematical formulation starts from:

The two-index vehicle flow formulation type linear programming model uses  $O(n^2)$  binary variables of  $x_{ij}$  to indicate with a value of 1 if the arc  $(i, j) \in A$  belongs to the optimal solution and takes a value of 0 otherwise [39, 40].

$$\min \sum_{i \in V} \sum_{j \in V} c_{ij} x_{ij} \tag{13}$$

**subject to**

$$\sum_{i \in V} x_{ij} = 1 \quad \forall j \in V \setminus \{0\} \tag{14}$$

$$\sum_{j \in V} x_{ij} = 1 \quad \forall i \in V \setminus \{0\} \tag{15}$$

$$\sum_{i \in V} x_{i0} = K, \tag{16}$$

$$\sum_{j \in V} x_{0j} = K, \tag{17}$$

$$\sum_{i \notin S} \sum_{j \in S} x_{ij} \geq r(S) \quad \forall S \subset V \setminus \{0\}, S \neq \emptyset \tag{18}$$

$$x_{ij} \in \{0, 1\} \quad \forall i, j \in V \tag{19}$$

Where:  $V = \{v_0, v_1, \dots, v_n\}$  is the set of vertices of the graph, where  $v_0$  corresponds to the warehouse;  $C$  is the distance or cost matrix  $c_{ij}$  between customers  $v_i$  and  $v_j$ ;  $K$  is the number of capacity  $C$  vehicles needed to load all the demand;  $S$  is a set of vertices of the graph where  $S \subset V \setminus \{0\}$ , and  $r(S)$  is the minimum number of vehicles necessary to serve all customers in  $S$ .

In this way, *equations (14-15)*, ensure that a single arch enters and a single arch leaves each vertex, clients are visited only once; while *equations (16-17)* ensure that the number of vehicles leaving the warehouse (vertex) return to it; while *equation (18)* satisfies connectivity and non-violation of vehicle capacity.

### 3.5. Analysis of the results

- Impacted and non-impacted municipalities.

Table 2 shows the zoning of the state of Veracruz that is established to carry out this document, as can be seen, part of its regionalization. Table 2 shows the number of municipalities that each area houses and the number of municipalities impacted by a natural phenomenon of hydrometeorological type during 2016, the percentage of emergency declarations issued by FONDEN of a total is shown in the last column of 147 [41]. The central zone is taken as a reference for evaluating the model, derived from the fact that it is the one with the most significant number of municipalities and more than 50% of declarations issued.

- Vehicle capacity.



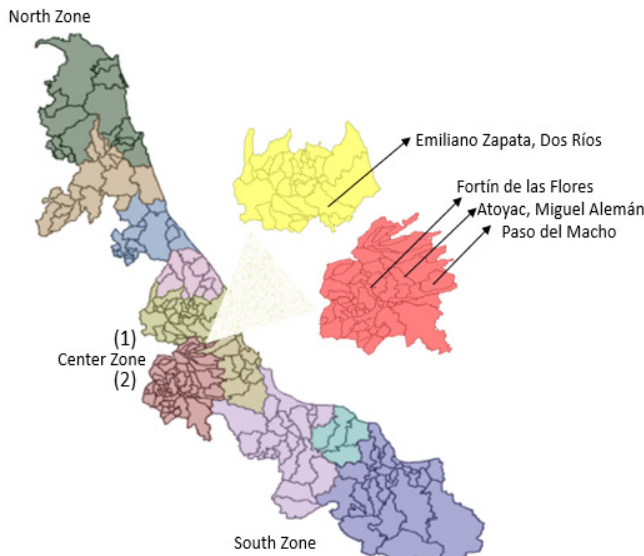
The same vehicular capacity is set as in [1], 4913 kits.

Table 2: Zoning of the state of Veracruz

Zone	Regions	# municipalities	Impacted municipalities	% declarations
North	Huasteca Alta, Huasteca Baja, Nautla, Totonaca	59	26	21.77%
Center	Capital, Altas Montañas	90	66	51.70%
South	Los Tuxtlas, Olmeca, Papaloapan, Sotavento	63	31	26.53%
Total		212	123	100%

- Location of the host municipality of the pre-positioned warehouse.

For the location of the host municipality, four different scenarios are proposed for the location of host municipalities of the pre-positioned warehouse. The first scenario designates the municipality of Fortín de las Flores as the headquarters of the aforementioned warehouse, which is obtained from the 24 municipalities without issuing an emergency declaration and the 66 municipalities impacted in 2016. The second scenario, part of rule out the municipality obtained in the first scenario (Fortín de las Flores) to give rise to the municipality of Emiliano Zapata as headquarters, the discard is from the assumption that the municipality first obtained, is impacted at a particular moment by a natural phenomenon of the type hydrometeorological. The third scenario is carried out in the event that the two previous municipalities cannot house the pre-positioned warehouse, Paso del Macho is obtained as the host municipality. A fourth scenario is captured, suppressing the variable of positional weight, that is, without taking into account the attributes of each municipality except that of affectation by a natural hydrometeorological phenomenon, thus the municipality of Atoyac, Miguel Alemán becomes the headquarters of the warehouse, which is located at a distance of 29.6 km from the first seat municipality, that is, Fortín de las Flores.



While Figure 2 shows the regionalization of the state in the study, it should be remembered that in previous paragraphs, the state was zoned, based on this regionalization. The regions that make up the central area subject to study are the Capital region (1) and the High Mountains Region (2).

The location of each of the selected municipalities is shown in Figure 2: 1. Emiliano Zapata; 2. Fortín de las Flores; 3. Atoyac, Miguel Alemán, and 4. Paso del Macho.

- Inventory levels.

The levels of inventory of food, equipment, water, and medicines grouped in kits are shown in Table 3. It shows the economic order quantities, the reorder point, the safety inventory, and the number of people not supplied.

Table 3: Inventory levels

TYPE OF DEMAND	Kits	Economic order quantity ( $Q_j^p$ )	Re-order point ( $R_j^p$ )	Safety stock	None-supplied
Demand 1 $D_j^{p-}$ (Early childhood)	Food	15732	6545	653	1
	Equipment	15521	6545	653	1
Demand 2 $D_j^{p-}$ (childhood)	Food	32793	13999	1443	3
	Equipment	32793	13999	1443	3
Demand 3 $D_j^p$ (adolescence)	Food	48396	20353	1781	3
	Equipment	48396	20353	1781	3
Demand 4 $D_j^p$ (adulthood)	Food	37000	15460	1275	2
	Equipment	37000	15460	1275	2
Demand 5 $D_j^{p-}$ (old age)	Food	9078	3755	327	1
	Equipment	9050	3755	327	1
Demand 6 $D_j^{p-}$ (General)	Water	227931	60014	5379	8
Demand 6 $D_j^p$ (General)	Medicine	60014	28324	5379	8
Inventory Cost			640,501,050 MXN		

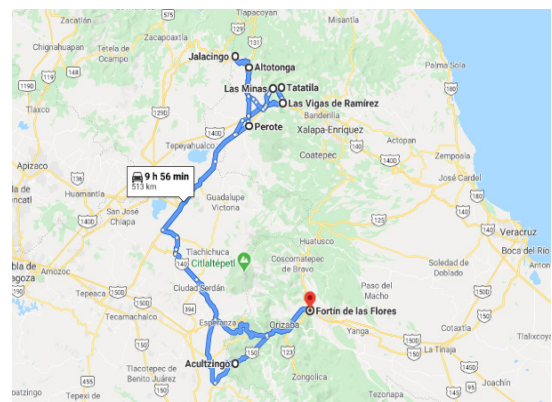


Figure 2: Municipalities headquarters of the pre-positioned warehouse [42]



• Delivery routes

Table 4 shows 4 of the 12 delivery routes obtained with a pre-positioned warehouse within the municipality of Fortín de las Flores. The overall results in the said municipality showed a total of 1,964.68 km traveled for delivery, an arithmetic mean of 163.72 km traveled per route, and arithmetic means of supplying victims of 6,291.45. The delivery of products can be carried out in a minimum time, under the assumption of not affecting the road infrastructure.

Figures 3, 4, 5 y 6, They show the delivery routes in Table 4, in order to be able to observe the routes to travel. For this, the application of Google Maps®.

Table 4: Delivery routes

#	ROUTE								
1	Acutzingo	Perote	Las Vigas Ramírez	Tatitila	Las Minas	Villa Aldama	Altoatonga	Jalancingo	
5	Coatepec	Xalapa	Tlacolulan	Acajete	Teocelo	Xico	Cosautlan de Carvajal	Ixhuacan de los Reyes	Ayahualulco
7	Huatusco	Zentla	Sochiapa	Tlacotepec de Mejía	Tenampa	Totutla	Tlaltetela		
12	Tlilapan	San Andrés Tenejapan	Atlahuilco	Tlaquilpa	Xoxocotla	Astacinga	Tehuipango		

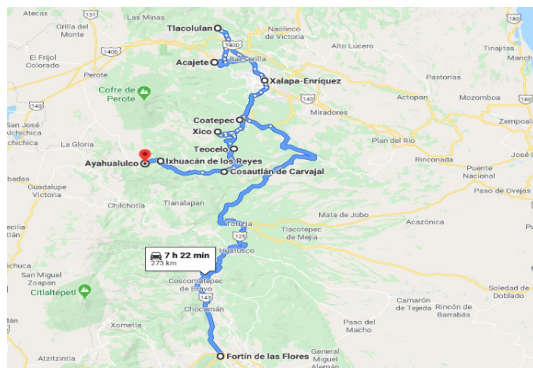


Figure 4: Route 5, 385 km traveled, nine municipalities supplied, extracted Google Maps ®

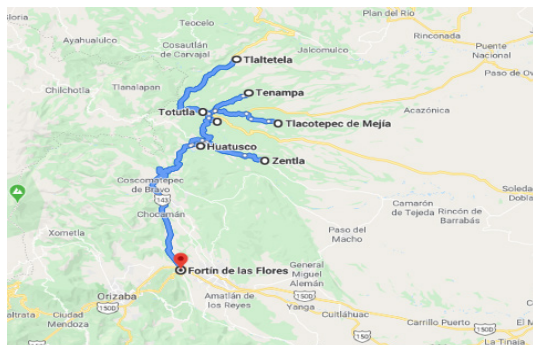


Figure 5: Route 7, 237 traveled, seven municipalities supplied, extracted from Google Maps ®

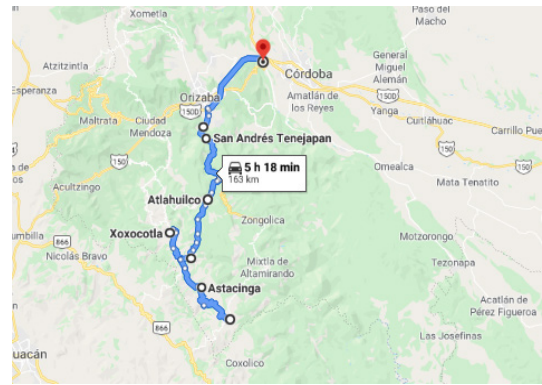


Figure 6: Route 7, 163 km traveled, seven municipalities supplied, extracted from Google Maps ®

4. Conclusions

Natural disasters have caused not only economic but also human losses. Mexico is not far from the effects of these events, much of its territory is exposed to the impact of natural phenomena of a hydrometeorological and geological nature, specifically the state of Veracruz, which has been considerably affected by hydrometeorological phenomena. This situation makes it imperative to locate facilities that can supply the needs, as mentioned above, to the victims in the shortest time.

The present document carries out the extension of a conference paper, evaluating the presented logistics model that has as its basis: a) the classic problem of the p-median for the location of a pre-positioning warehouse; b) an extension of the (q-R) model for calculating inventories of different products, according to different types of demand, and c) the problem of multiple vehicle routing with the capacity to establish delivery routes from warehouses to the affected municipalities. The evaluation is carried out in the same federal entity, with a more significant number of instances, in total 90, which are in the central area of the state, the results show that the location of the warehouse is assigned to the municipality of Fortín de las Flores, the delivery routes to the affected municipalities are presented, obtaining a total of 12 with a total of 1964 km traveled and an average travel distance of 163.72 km.

Similarly, the discarding of the municipality obtained is carried out, under the assumption that it has been affected by the natural event, obtaining another feasible location for the pre-positioned warehouse, this being the municipality of Emiliano Zapata, under the same. Of course, the latter municipality is discarded, obtaining Paso del Macho as the headquarters municipality of the warehouse. A fourth run of the program is carried out, in which the variable of positional weight and the previous municipalities are ruled out for having been affected, obtaining Atoya, Miguel Aleman, as the host municipality. In the same way, the increase in the number of products to be supplied was carried out. Therefore, the inventory levels obtained present a total cost of 640,501,050 Mexican pesos for the five demands: a) early childhood; b) childhood; c) adolescence; d) adulthood, and e) old age; with four product kits: 1. Food; 2. Equipment; 3. Health, and 4. Water.

It was carried out using the (q, R) model because the kits are supplied from 1, 2 or up to 3 times in specific time, according to the duration of the support bulletin's opening or, in accordance to

the need of the people. We did not consider the newsboy model because it considers a single purchase. Even so, it is suggested to look for some other model that can answer this problem.

People's lives are priceless, natural phenomena are not controllable, but actions can be carried out to reduce the damage that they may cause to people immersed in impacted areas at a particular time, and maintain a quality of life for them. Until his return to his healthy life. The state of Veracruz, year after year is hit by different phenomena, but derived from its location, those of the hydrometeorological type occurs with a specific frequency, even more so in the period covered between the months of June-November, which is why the research carried out, looks for feasible locations that allow the location of a pre-positioned warehouse in municipalities that meet conditions for its efficient management, and that this warehouse, in turn, contains food products, clothing, health, and clean water to ingest, at adequate inventory levels, with the aim of a resupply until the return to normality, with a delivery routing with minimum distances, so that attention is fulfilled in the shortest possible time in terms of food and health, the latter derived that diseases such as typhoid, tetanus, and lung can occur to name a few. In the literature, especially in studies carried out in this state, minimal research was visualized in the field of humanitarian logistics.

As future work, a extend model is proposed that can to locate multiple locations with their delivery routes and travel distance constraints.

## References

- [1] E. Barojas-Payán, D. Sánchez-Partida, V. Juárez-Rivera, R. Villafuerte-Díaz, J. Medina-Cervantes, "Plant location, inventory levels and supply of products to areas affected by a natural phenomenon" in 2019 IEEE International Conference on Engineering Veracruz (ICEV), Boca del Rio, Ver., Mex., 2019. <https://doi.org/10.1109/ICEV.2019.8920691>
- [2] United Nations. UN News. 2010 [Online]. Available at <https://news.un.org/es/story/2020/01/1467501> Last accessed February 1, 2020.
- [3] J.M. Rodríguez, C.M. Welsh, M.L. Romo, A.C. Travieso, Riesgo de desastres en México: eventos hidrometeorológicos y climáticos. Instituto Mexicano de Tecnología del Agua, 2018.
- [4] Cred-Usaid, Cred crunch, Brussels, Centre for Research on the Epidemiology of Disasters. United States Agency International Development. 2016. <https://www.emdat.be/publications?page=1>
- [5] Secretaría de Gobernación, Manual de organización y operación del sistema nacional de protección civil, 2018.
- [6] Secretaría de Hacienda y Crédito Público, Acuerdo por el que se emiten las reglas generales del fondo de desastres naturales, Capítulo I, Disposiciones generales, Artículo 1º, 2010.
- [7] Cámara de Diputados del H. Congreso de la Unión, Ley General de Protección Civil, Nueva Ley publicada en el Diario Oficial de la Federación el 6 de junio de 2012, 12(59), 1-6, 2018.
- [8] Cámara de Diputados del H. Congreso de la Unión, Ley General de Protección Civil, Nueva Ley publicada en el Diario Oficial de la Federación el 6 de junio de 2012, Capítulo XII, 2018.
- [9] Secretaría de Gobernación - Centro Nacional de Prevención de Desastres, Impacto socioeconómico de los principales desastres ocurridos en la República Mexicana en 2016, Dirección de Análisis y Gestión de Riesgos Subdirección de Estudios Económicos y Sociales México, 2018.
- [10] Instituto Nacional de Estadística y Geografía. Veracruz Ignacio de la Llave. 2019. <https://www.inegi.org.mx/app/areasgeograficas/?ag=30>
- [11] S. Saeyon, S. Pettit, I. Harris, A. Beresford, "The pre-positioning of warehouses at regional and local levels for humanitarian relief organization" International Journal of Production Economics, 170(1) Part B, 616-628, 2015.
- [12] C.G. Rawls, M.A. Turnquist, "Pre-positioning of emergency supplies for disaster response" Transportation Research Part B, 44(4), 521-534, 2009.
- [13] M. S. Daskin, Network and Discrete Location: Models, Algorithms, and Applications, John Wiley and Sons, Inc., 2013.
- [14] Veracruz, Gobierno del Estado. Ciclo de los desastres. 2020. <http://www.veracruz.gob.mx/proteccioncivil/ciclo-de-los-desastres/>
- [15] C. Boonmee, M. Arimura, T. Asada, "Facility location optimization model for emergency humanitarian logistics" International Journal of disaster risk Reduction, 24(2017), 485-498, 2017. <https://doi.org/10.1016/j.ijdr.2017.01.017>
- [16] A. Vermaa, G. Gaukler, "Pre-positioning disaster response facilities at safe locations: An evaluation of deterministic and stochastic modeling approaches" Computers & Operations Research, 62(2015), 197-209, 2015. <https://doi.org/10.1016/j.cor.2014.10.006>
- [17] T. Saputra, O. Pots, K.D. de Smidt-Destombes, "The impact of mean time between disasters on inventory pre-positioning strategy" Disaster Prevention and Management, 24(1), 115-131, 2015. <https://doi.org/10.1108/DPM-11-2013-0197>
- [18] I. Landa-Cruz, D. Sánchez-Partida, E. Barojas-Payán, J.L. Martínez-Flores, "Optimization model for relocation of distribution sites of economic support in areas of high marginalization" DYNA Management, 4(1), [15], 2016. <http://dx.doi.org/10.6036/MN8104>
- [19] M. Mostajbadeh, W. Gutjahr, F. Salman, "Inequity-averse shelter location for disaster preparedness." IISE Transactions, 51(8), 809-829, 2019. <https://doi.org/10.1080/24725854.2018.1496372>
- [20] C. Cornejo-Sánchez, J. Vargas-Florez, L. Aragón-Casas, V. Serpa-Oshiro, "Innovation in engineering, technology and education for competitiveness and prosperity" in Eleventh LACCEI Latin American and Caribbean Conference for Engineering and Technology (LACCEI'2013), Cancún, Q. Roo, Mex, 2013.
- [21] A. USLU, C. ÇETİNKAYA, SK. İŞLEYEN, "Vehicle routing problem in post-disaster humanitarian relief logistics: a case study in Ankara" Sigma J Eng & Nat Sci. 35(3), 481-499, 2017.
- [22] M.E. Bruni, S. Khodaparasti, "A fast heuristic for routing in post-disaster humanitarian relief logistics" Transportation Research Procedia, 30, 304-313, 2018. <https://doi.org/10.1016/j.trpro.2018.09.033>
- [23] M.A. Gökçe, E. Ercan, "Multi-period vehicle routing & replenishment problem of neighbourhood disaster stations for pre-disaster humanitarian relief logistics" IFAC-PapersOnLine, 52(13), 2614-2619, 2019. <https://doi.org/10.1016/j.ifacol.2019.11.601>
- [24] E. Barojas-Payán, D. Sánchez-Partida, D.E. Riaño, J.L. Martínez-Flores, W. Velázquez, S.O. Caballero, "Heuristic of the nearest neighbor in the delivery of supports within the state of Veracruz" International Journal of Combinatorial Optimization Problems and Informatics, 10(2), 26-38, 2018.
- [25] L. Bertazzi, M. Speranza, "Models and algorithms for the minimization of inventory and transportation costs: a survey" New Trends in Distribution Logistics, 480, 137-157, 1999.
- [26] C.H. Glock, "The joint economic lot size problem: A review" International Journal of Production Economics, 135(2), 671-686, 2012.
- [27] A. Sarmiento, R. Nagi, "A review of integrated analysis of production-distribution systems" IIE Transactions, 31(11), 1061-1074, 1999.
- [28] Y.B. Park, "An integrated approach for production and distribution planning in supply chain management" International Journal of Production Research, 43(6), 1205-1224, 2005.
- [29] G. Ghiani, A. Manni, E. Manni, M. Toraldo, "The impact of an efficient collection sites location on the zoning phase in municipal solid waste management" Waste Management, 34, 1949-1956, 2014.
- [30] V.C. Hemmelmayr, K.F. Doerner, R.F. Hartl, D. VIGO, "Models and algorithms for the integrated planning of bin allocation and vehicle routing in solid waste management" Transportation Science, 48, 103-120, 2014.
- [31] J. Gaytán-Iniestra, P.E. Arroyo-López, R. Enríquez-Colón, "A bi-criteria model to locate shelters as part of an evacuation plan in the case of floods" Revista Ingeniería Industrial, 2, 35-56, 2012.
- [32] Instituto Nacional para el Federalismo y Desarrollo Municipal, Secretaría de Gobernación. Sistema Nacional de Información Municipal: Municipios en cifras. 2017. <http://www.snim.rami.gob.mx/>
- [33] Secretaría de Comunicaciones y Transportes. Traza tu ruta. 2017. [http://app.sct.gob.mx/sibuac\\_internet/ControllerUI?action=cmdEscogeRuta](http://app.sct.gob.mx/sibuac_internet/ControllerUI?action=cmdEscogeRuta)
- [34] Secretaría de Desarrollo Social. Programa de abasto rural. 2017. <http://catalogo.datos.gob.mx/dataset/programa-de-abasto-rural-tiendas-2015>
- [35] Organización Mundial de la Salud. Biblioteca electrónica de documentación científica sobre medidas nutricionales (eLENA), "Ciclo de vida". 2017. [https://www.who.int/elena/life\\_course/es/](https://www.who.int/elena/life_course/es/)
- [36] Procuraduría Federal del consumidor. Quién es quién en los precios. 2018. <https://www.profeco.gob.mx/precios/canasta/home.aspx?th=1>
- [37] Servicio de Administración Tributaria. Salarios Mínimos. 2017. [www.sat.gob.mx/informacion\\_fiscal/tabla\\_indicadores/Paginas/salarios\\_minimos.aspx](http://www.sat.gob.mx/informacion_fiscal/tabla_indicadores/Paginas/salarios_minimos.aspx)
- [38] Secretaría del trabajo y previsión social. Salarios Mínimos. 2017. [www.gob.mx/cms/uploads/attachment/file/175865/Tabla\\_de\\_salarios\\_minimos\\_vigentes\\_a\\_partir\\_de\\_01\\_enero\\_2017.pdf](http://www.gob.mx/cms/uploads/attachment/file/175865/Tabla_de_salarios_minimos_vigentes_a_partir_de_01_enero_2017.pdf)

- [39] P. Toth, D. Vigo, *An Overview of Vehicle Routing Problems*, Monographs on Discrete Mathematics and Applications, SIAM, 2000.
- [40] J.P. Orrego, D. Ospina, E.M. Toro, “Solución al problema de ruteo de vehículos con capacidad limitada (CVRP) usando una técnica metaheurística” *Scientia Et Technica*, 21(3), 225-233, 2016. <https://doi.org/10.22517/23447214.9013>
- [41] Sistema Nacional de Protección Civil, México DDGR (FONDEN). Insumos autorizados por declaratoria de emergencia. 2016. <https://www.proteccioncivil.gob.mx/es/ProteccionCivil>
- [42] National Institute for Federalism and Municipal Development, SEGOB, Secretary of the Interior. National Municipal Information System: Municipalities in numbers. 2010. <http://www.snim.rami.gob.mx/>

## Distributed Microphone Arrays, Emerging Speech and Audio Signal Processing Platforms: A Review

Shahab Pasha<sup>\*</sup>,<sup>1</sup>, Jan Lundgren<sup>1</sup>, Christian Ritz<sup>2</sup>, Yuexian Zou<sup>3</sup>

<sup>1</sup>STC Research Centre, Mid-Sweden University, Sundsvall 85230, Sweden

<sup>2</sup>School of Electrical, Computer and Telecommunication Engineering, University of Wollongong, NSW 2500, Australia

<sup>3</sup>ADSPLAB/Intelligent Lab, School of ECE, Peking University, Shenzhen 518055, China

### ARTICLE INFO

#### Article history:

Received: 14 April, 2020

Accepted: 03 June, 2020

Online: 28 July, 2020

#### Keywords:

Distributed signal processing

Machine learning

Spatial selectivity

Speech enhancement

### ABSTRACT

Given ubiquitous digital devices with recording capability, distributed microphone arrays are emerging recording tools for hands-free communications and spontaneous tele-conferencings. However, the analysis of signals recorded with diverse sampling rates, time delays, and qualities by distributed microphone arrays is not straightforward and entails important considerations. The crucial challenges include the unknown/changeable geometry of distributed arrays, asynchronous recording, sampling rate mismatch, and gain inconsistency. Researchers have recently proposed solutions to these problems for applications such as source localization and dereverberation, though there is less literature on real-time practical issues. This article reviews recent research on distributed signal processing techniques and applications. New applications benefitting from the wide coverage of distributed microphones are reviewed and their limitations are discussed. This survey does not cover partially or fully connected wireless acoustic sensor networks.

## 1. Introduction

With new portable devices, such as smartphones and tablets, conventional microphone arrays are no longer the main signal and speech acquisition platform; rather, distributed microphone arrays (also called ad hoc microphone arrays) formed by the joint analysis of randomly located independent recording devices such as laptops and cell phones are emerging recording platforms for various applications [1]. Conventional compact microphone arrays and recording devices are now processed within ad hoc arrays and are jointly analyzed with other such devices.

For this reason, distributed microphones are popular and have been used in a wide range of applications, such as speaker tracking and speech recognition systems [2]. Currently, there is substantial potential for applications that use digital recording devices collaboratively as a virtual array [3]. By positioning such devices at random locations within the acoustic scene, the array geometry is no longer limited to standard structures. Distributed

microphone nodes provide wide spatial coverage and are ideal for capturing multiple speakers located meters away from one another. This advantage in the means of signal (i.e., sound and other cues derived from it) acquisition allows different audio scenes to be captured to give a more accurate picture of the user environment. To enable such ubiquitous and flexible teleconferencing and multimedia applications within the distributed signal processing context, several important technical and theoretical problems should be addressed. Some of the main challenges are the unknown/changeable array structure, inconsistent sampling frequencies, varying gains (due to varying source-to-microphone distances), and unsynchronized recordings. Although speech and signal processing applications are well studied and straightforward [4] in the context of known-geometry compact microphone arrays, existing methods are not directly applicable to distributed scenarios.

This survey article reviews the strengths and limitations of recently proposed distributed signal processing methods. The

<sup>\*</sup>Corresponding Author: Shahab Pasha, Email: shahab.pasha@miun.se



focus is on scenarios in which the microphones are independent and do not communicate within the array.

## 2. Literature Selection Methodology

The authors have reviewed the most recent theoretical and practical literature on distributed signal processing and related, overlapping fields of study. More than 250 research papers, theses, patents, and online tools and resources were studied and compared over a five-year period. One hundred peer-reviewed published works were chosen for final comparison and investigation according to the novelty of the work and the credibility of the journal/conference. The current set of references covers the literature and online resources as of 2020, though most of the main methods were proposed between 2013 and 2017.

## 3. Distributed Signal Processing

Ad hoc microphone arrays [1, 3] consist of a set of recording devices (referred to as nodes [5]) randomly distributed in an acoustic environment to record an unknown acoustic scene with wide spatial coverage (Figure 1). The nodes can be identical [5] or different [6] in terms of their structure and number of elements [7, 8] (Table 1). Ad hoc arrays eliminate the restrictions on microphone and source placement at fixed locations and facilitate dynamic and flexible recording experiences.

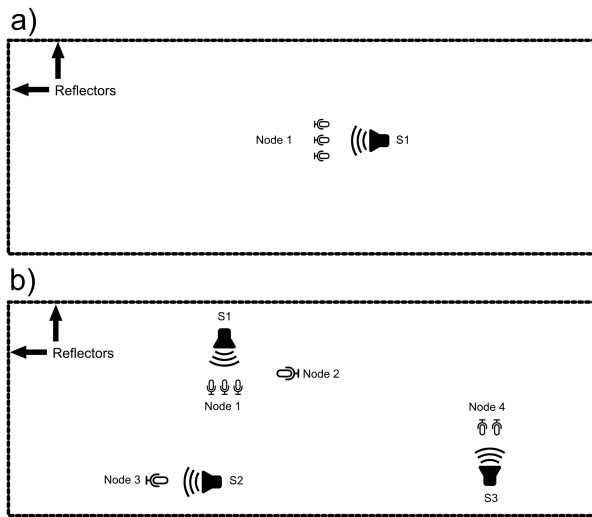


Figure 1: a) A three-element linear microphone array recording a single source; b) a distributed acoustic scene comprising three sound sources, two single-channel nodes, and two multi-channel linear arrays.

Table 1: Ad hoc microphone arrays compared with compact arrays

<i>Distributed</i>	<i>Compact</i>
Unknown structure [9]	Known structure
Changeable microphone locations [10]	Fixed topology
Unknown inter-channel time delays [11]	Known inter-channel time delays
Inconsistent gain within the array [12]	Consistent gain
Uncertain direction of arrival (DOA) definition [13]	Straightforward DOA definition
Large phase differences (i.e., spatial aliasing)	Negligible spatial differences
Inconsistent signal quality [5]	Consistent signal quality

### 3.1. Definition

$M$  recording devices (both single- and multi-channel devices) located randomly to capture  $N$  sound sources form an ad hoc microphone array. Node  $m$  contains  $M_m$  channels and the signal picked up by the  $i$ th channel forming the  $m$ th node is modelled as

$$x_{m,i}(n) = \sum_j \sum_t s_j(n) * h_{m,i,j}(n) + v_{m,i}(n) + w_{m,i}(n), \quad (1)$$

where  $s_j(n)$  is the  $j$ th sound source,  $h_{m,i,j}(n)$  represents the room acoustic response between the  $i$ th channel forming the  $m$ th node and source  $j$ , and  $v_{m,i}(n)$  is the additive noise;  $w_{m,i}(n)$  is the non-coherent component and represents reverberation and diffuse noise. Each channel location is modelled as

$$\mathbf{r}_{m,i} = [x_{m,i}, y_{m,i}, z_{m,i}], \quad (2)$$

and the  $j$ th source location is

$$\mathbf{r}_{s,j} = [x_j, y_j, z_j]. \quad (3)$$

in the Cartesian coordinate system.  $h_{m,i}(n)$ ,  $\mathbf{r}_{m,i}$ , and  $\mathbf{r}_{s,j}$  are assumed unknowns for all the values of  $m$ ,  $i$ ,  $j$ , and  $n$ . The total number of channels in the array is

$$M_{ch} = \sum_{m=1}^M M_m. \quad (4)$$

For ad hoc scenarios where all the nodes are single channel

$$M_{ch} = M. \quad (5)$$

The truncated RIRs ( $\mathbf{h}_{m,i,j}$ ) of length  $L$  are modelled as follows:

$$h_m(n) = \sum_{k=0}^L a_{m,k} \delta(n - \tau_{m,k}), \quad (6)$$

with time delays,  $\tau_{m,k}$ , and amplitudes,  $a_{m,k}$ .  $L$  is chosen based on the application and the reverberation time [14].  $\tau_{m,0}$  represents the time of arrival (TOA) at node  $m$ .

Assuming that the distances between the channels at each node are relatively small and that each node forms a compact microphone array [5, 7], each node can deliver a single-channel output, so there will be only  $M$  distributed recordings [15]. Having only one active source during the short frames simplifies (1) to

$$x_m(n) = s(n) * h_m(n) + v_m(n) + w_m(n). \quad (7)$$

### 3.2. Significance and applications

Tablets, smartphones, sound recorders, and other portable and wearable digital devices are becoming prevalent in workplaces, homes, and lecture halls, redefining how we communicate and record our communications. Consequently, these devices are becoming key tools for daily activities, including teleconferencing and hands-free speech communication [16]; ad hoc signal processing is therefore inevitable.

Advances in ad hoc signal processing technologies can also relax the highly demanding constraints on network layer design [17]. Ad hoc arrays also improve the quality of speech communication, acoustic scene analysis, and speech recognition (Figure 2) due to the multiple observations they make within a wider area [18].

### 3.2.1. Improved Distributed Meetings

Meetings are an important part of everyday life for many types of educational and professional workgroups. The main application of ad hoc microphone arrays is for distributed meetings (DMs) [19], as such meetings are recorded by microphones at unknown and varying environments with different signal qualities. DM systems enable the high-quality broadcasting and recording of meetings by utilizing the flexible spatial coverage of ad hoc arrays. Distributed meetings using signal processing techniques are not necessarily online meetings, but could be group meetings [20]. Ad hoc arrays are more suitable recording tools for distributed meetings than are compact microphone arrays, as the recording devices can be spread out in the meeting room. Joint analysis of the distributed microphone signals yields more accurate results for signal processing applications such as active source detection (i.e., localization) [21].

### 3.2.2. Hearing Aids

Reduction of interference and noise is important in hearing aids (HAs) to provide intelligible speech signals in noisy environments [22]. Using an array of microphones, it is possible to exploit the

spatial characteristics of the acoustic scenario to obtain more information about the target scene. Although the scenario investigated by Bertrand and Moonen [22] was a fully connected binaural network, the unknown geometry of the array and the random recording setup (i.e., the source locations) form an ad hoc acoustic scene.

### 3.2.3. Hands-Free Communication

Wearable recording nodes are essential to hands-free communication systems, and the node structure can also vary. As the node and source locations in such systems are unknown and changeable, classic array-processing methods cannot be applied. Ad hoc signal processing offers a solution for the joint analysis of the nodes in hands-free voice communication systems for improved speech enhancement and source localization applications. The general scenario of hands-free audio recording and voice recognition with distributed nodes has been described and investigated by Jia et al. [23].

### 3.2.4. Ambient Intelligence and Smart Homes

Ambient intelligence advances are based on advances in sensors, pervasive computing, and artificial intelligence [24]. Very little can be done by an ambient intelligent system without detecting the user's presence and movements. The ambient intelligence system must be aware of the users' locations in each time period. One way to do this is by tracking the sound sources and acoustic cues. Spatially distributed microphone arrays are effective tools for monitoring user movements [25]. The insights gained provide important clues as to the type of activities the user is engaged in, and make the system responsive to the user's

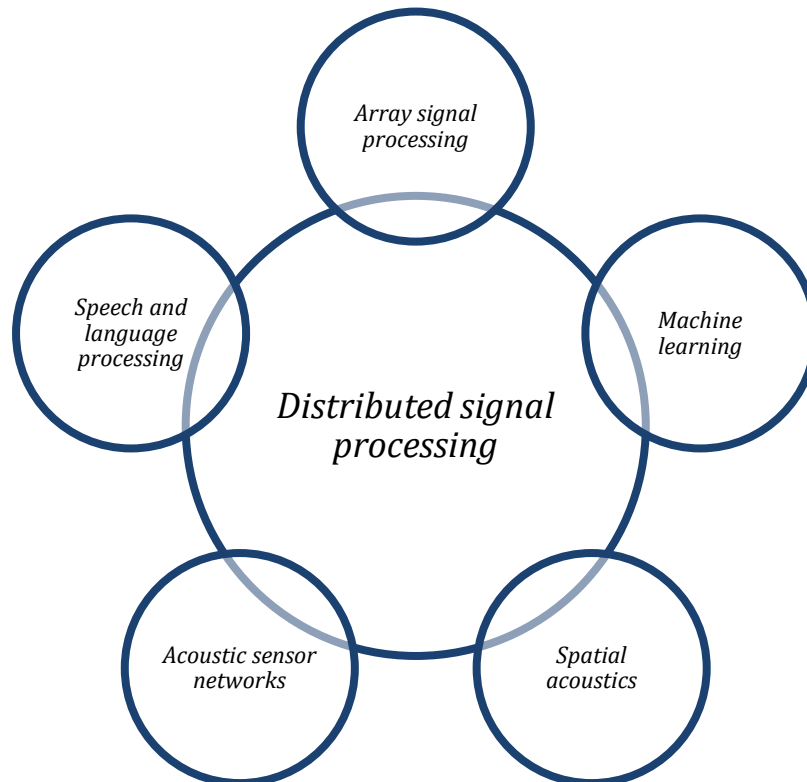


Figure 2: Distributed signal processing and overlapping fields of study

location [26]. Speaker verification and identification by ad hoc microphone arrays and speech recognition for automated human-computer interaction [18] are important fields of research that can help customize ambient intelligence applications.

3.2.5. *Monitoring*

Some recent research focuses on analyzing the environment by means of acoustic sensing due to its unobtrusive nature. Acoustic monitoring covers a wide range of applications, from security tasks (e.g., intrusion detection, malicious activity early detection, and traffic monitoring [27]) to whale migration tracking [28] and bird behavior studies [29].

3.2.6. *Medical Signal Processing*

Unconventional distributed microphone arrays have been used in medical applications for accurate and high-quality chest sound pick-up [30] and vital sound separation [31]. Compared with conventional compact arrays, ad hoc microphone arrays provide more accurate recordings of closely located organs such as the lungs and heart. The design and development of ad hoc acoustic sensors and microphones are essential for medical and E-health research.

3.3. *Room Acoustics and Distributed recording*

Unlike conventional compact microphone arrays, in which the noise and reverberation levels are consistent, in distributed signal processing each microphone has its own unique reverberation level and room acoustic response. Researchers have recently shown that the unique RIR and echo pattern at each ad hoc

microphone location contain location and distance information even if the recording setup is unknown [32, 33]. This idea is applicable to ad hoc microphone arrays for microphone clustering [11], room geometry reconstruction [33], and microphone localization [34].

The framework for recording using distributed microphone nodes was formulated by Tavakoli et al. [16] (1), covering both reverberation and noise. The signals recorded by all  $M_{ch}$  channels (4) within an ad hoc array are modeled as

$$\mathbf{x}(n) = [x_1(n), \dots, x_M(n)]^T. \tag{8}$$

As each ad hoc microphone receives its own unique distorted version of the source signal, the RIR and reverberation at each microphone location contain important information [11, 35]. Each RIR ( $h_m(n)$ ) (6) can be represented by sets of reverberation times and reverberation amplitudes [11], as follows:

$$[a_{m,0}, \dots, a_{m,L}], \tag{9}$$

where  $a_0$  is the direct path impulse amplitude, and

$$[\tau_{m,0}, \dots, \tau_{m,L}]. \tag{10}$$

Researchers have divided the RIR (1) into two segments: early and late echoes [36]. This segmentation is specifically important in the context of ad hoc microphone arrays and is the basis for defining discriminative features. The clarity feature is used for sound source localization by ad hoc microphones [37]. Table 2 summarizes the acoustic features applied to signal processing applications.

Table 2: Room acoustics features applied to distributed signal processing

<i>Method/features</i>	<i>Application</i>	<i>Reference(s)</i>	<i>Year</i>
<i>Energy</i>	Localization	Liu et al. [12]	2007
<i>Noise coherence</i>	Clustered beamforming	Himawan et al. [49]	2011
<i>Voice activity detection</i>	Speech enhancement	Sakanashi et al. [50]	2013
<i>Sparsity analysis</i>	Localization	Asaei et al. [21]	2014
<i>Time delay and sound level</i>	Speech separation	Souden et al. [7]	2014
<i>Signal power</i>	Traffic monitoring	Toyoda et al. [27]	2014
<i>Euclidian distance matrix (EDM)</i>	Microphone localization	Dokmanic et al. [51]	2015
<i>Kurtosis of the linear prediction residuals</i>	Clustering	Pasha et al. [8]	2015
<i>Generalized cross-correlation</i>	Speaker tracking	Tian et al. [52]	2015
<i>Reverberation and echoes</i>	Clustering	Pasha et al. [11]	2015
<i>Non-negative matrix factorization (NMF)</i>	Calibration	Asaei et al. [53]	2015
<i>Pseudo coherence vector</i>	Node selection for speech enhancement	Tavakoli et al. [5]	2015
<i>Reverberation</i>	DOA estimation	Pasha et al. [54]	2015
<i>Cepstral features</i>	Interference suppression	Gregen et al. [9]	2016
<i>C<sub>50</sub> (short-time reverberation)</i>	Multi-talk detection	Pasha et al. [37]	2016
<i>Magnitude-squared coherence (MSC)</i>	Crosstalk and multi-talk detection	Pasha et al. [37]	2016
<i>Echoes</i>	Room geometry reconstruction	Dokmanic et al. [33, 55]	2016
<i>Coherent-to-diffuse ratio</i>	Multi-talk and crosstalk detection	Pasha et al. [56]	2017
<i>Power spectral density (PSD)</i>	Spotforming	Habet et al. [57]	2017
<i>Signal to interference and noise ratio</i>	Noise cancellation	Tavakoli et al. [20]	2017
<i>Time difference</i>	Microphone localization	Woźniak et al. [58]	2019
<i>Distributed unscented Kalman particle filter (DUKPF)</i>	Source localization	Zhang et al. [59]	2020

### 3.4. Challenges and Limitations

Large scale distributed arrays are inherently asynchronous [38]. Inconsistent sampling rates [39], gain differences [12], and different signal-to-noise ratios (SNRs) at different locations [5, 12] are challenges with distributed signal processing. The main differences between ad hoc array and compact array signal processing are summarized in Table 1.

## 4. Distributed Signal Processing

The most recent existing distributed signal processing techniques proposed in the literature are reviewed here in terms of their target scenarios and requirements. Distributed signal processing overlaps with array processing [4, 40], wireless sensor networks [41] (not reviewed here), feature extraction and machine learning [42], hands-free speech communication [43], and guided/informed signal and speech processing [44, 45].

### 4.1. Microphone Calibration

Microphone calibration is especially important in the context of large arrays [46, 47] and distributed microphones, as the area covered by the microphones can be large [48]. These large distances and attendant time delays should be considered when time aligning the signals for beamforming and speech enhancement applications. Representing the gain of the array by  $\mathbf{g} = \{g_1, \dots, g_M\}$ , (1) can be rewritten as

$$x_m(n) = s(n) * g_m \bar{h}_m(n) + v_m(n), \quad (11)$$

for one active source, where  $g_m$  and  $\bar{h}_{m,i}(n)$  are the gain and normalized RIR of microphone  $m$ , respectively. Calibration often consists of estimating the distances between the pairs of microphones and reconstructing the array geometry given all the pairwise distances [53, 60]. Microphone array calibration in general suffers from reverberation, noise, and complicated mathematical computations (making calibration infeasible for real-time applications); specifically, sampling frequency

mismatch, inconsistent microphone gain, and non-stationary array geometry make calibration an ongoing process throughout the recording session. Due to calibration difficulties, some methods prefer to avoid microphone

calibration altogether, if possible applying methods inherently robust to microphone placement and steering error [61].

For microphone array calibration, some advanced mathematical methods use joint source and microphone localization methods [62] and incorporate matrix completion constrained by Euclidean space properties [63, 64]. Such methods require partial knowledge of pairwise microphone distances [65]. It has been shown that using sound emissions for self-calibration can result in a calibration method more robust to sampling frequency mismatch [66]; however, the method is only applicable to devices that have both recording and sound emitting capabilities.

### 4.2. Signal Synchronization

Signal synchronization is an essential task for more advanced applications such as dereverberation and speech enhancement. Improperly addressing the synchronization issue leads to poor dereverberation performance. The different types of delays, including microphone internal, TOA, and onset time delays (Figure 3), have been investigated and explained [67, 68]. The lack of a reference channel and inconsistent sampling frequencies [69] in ad hoc arrays lead to critical phase differences [70] between the channels, fundamentally challenging most speech and audio processing applications. Existing solutions to clock drift [71] require limiting assumptions, such as unmoving sources and stationary amplitudes.

If the array topology is available, the time difference of arrival (TDOA) (13) between any two microphones in the array can be calculated as

$$TOA_{s,m} = \frac{|\mathbf{r}_s - \mathbf{r}_m|}{c} + \delta_m \quad (12)$$

$$TDOA_{m,\hat{m}} = \frac{|\mathbf{r}_s - \mathbf{r}_m|}{c} - \frac{|\mathbf{r}_s - \mathbf{r}_{\hat{m}}|}{c} + (\delta_m - \delta_{\hat{m}}) + (T_{om} - T_{om\hat{m}}), \quad (13)$$

where  $\delta_m$  and  $T_{om}$  represent the internal delay and onset time of microphone  $m$ , respectively (Figure 3) [67, 68], and  $\mathbf{r}_s =$

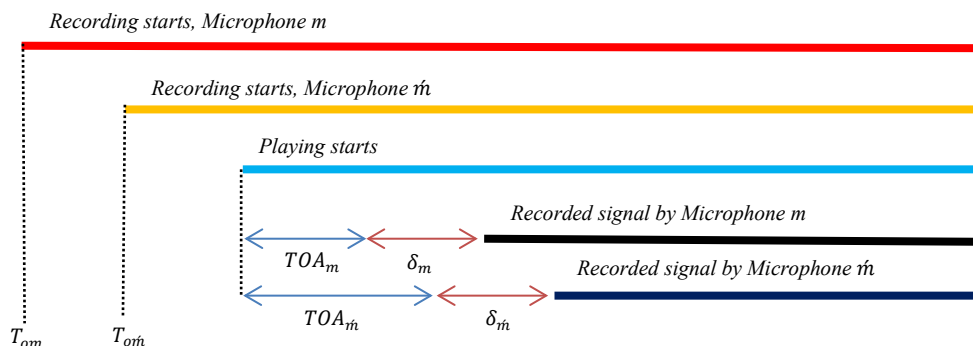


Figure 3: Time delays



$[x_s, y_s, z_s]^T$ ,  $\mathbf{r}_m = [x_m, y_m, z_m]^T$ , and  $\mathbf{r}_{\hat{m}} = [x_{\hat{m}}, y_{\hat{m}}, z_{\hat{m}}]^T$  are the source, microphone  $m$ , and microphone  $\hat{m}$  locations in Cartesian coordinate space, respectively. The objective of synchronization is estimating the overall time delay between every pair of microphones. The time delays are given as

$$\boldsymbol{\tau} = \begin{bmatrix} \tau_{11} & \cdots & \tau_{1M} \\ \vdots & \ddots & \vdots \\ \tau_{M1} & \cdots & \tau_{MM} \end{bmatrix}, \quad (14)$$

where  $\tau_{mm} = 0$  for  $m = 1$  to  $M$  and  $\tau_{m\hat{m}} = \tau_{\hat{m}m}$  for all  $m$  and  $\hat{m}$  values.

Researchers have used the time alignment of ad hoc channels for source localization by means of generalized cross correlation (GCC) [17] and for defining the parametric squared errors of time differences [72].

Some advanced techniques use the least squares for the temporal offset estimation [73] and for audio fingerprinting [38]. These methods are developed based on the clustering and synchronizing methods applied to unorganized multi-camera videos [74] which are based on matching the time-frequency landmarks between two channels.

The effect of synchronization on blind source separation (BSS) by a wireless acoustic sensor networks (WSAN) was investigated by Lienhart et al. [75], who concluded that full synchronization increases the BSS cost function by an average of 4 dB.

Most of the proposed synchronization methods can time-align the signals accurately and calculate the TDOA with an error of 1–10 milliseconds. The other important factor is the computational cost. The watermark-based algorithms [38] are shown to be more efficient than GCC methods [74].

### 4.3. Spatial Multi-Channel Linear Prediction (LP)

Multi-channel LP was developed for compact arrays and has applications, such as dereverberation [76] and compression [77, 78].

Using  $x_m$  from (1), the autocorrelation  $r_m(c)$  is obtained for channel  $m$  from

$$r_m(c) = E(x_m(n)x_m(n+c)), \quad c = 0, 1, 2, \dots \quad (15)$$

where  $E$  is the mathematical expectation.

$$\bar{r}(c) = \frac{1}{M} \times \sum_{m=1}^M r_m(c). \quad (16)$$

The baseline autocorrelation function (16), can be formulated in the more general form of a weighted average autocorrelation ( $\bar{r}_w(c)$ ) (17). Assuming that the applied weights are  $\boldsymbol{\beta} = \{\beta_1, \dots, \beta_M\}$ , the weighted average autocorrelation function is calculated as

$$\bar{r}_w(c) = \frac{1}{\sum_{m=1}^M \beta_m} \times \sum_{m=1}^M \beta_m r_m(c), \quad (17)$$

where  $\beta_m$  is the weights given to  $r_m(c)$ .

$$\mathbf{w}_s = \begin{bmatrix} \bar{r}_w(0) & \cdots & \bar{r}_w(P_{short}-1) \\ \vdots & \ddots & \vdots \\ \bar{r}_w(P_{short}-1) & \cdots & \bar{r}_w(0) \end{bmatrix}^{-1} \times \begin{bmatrix} \bar{r}_w(1) \\ \vdots \\ \bar{r}_w(P_{short}) \end{bmatrix}, \quad (18)$$

and the pre-whitened signal is

$$\tilde{e}_m(n) = x_m(n) - \sum_{k=1}^{P_{short}} w_{s,k} x_m(n-k), \quad (19)$$

where  $w_s = \{w_{s,1}, \dots, w_{s,P_{short}}\}$  [13]. Having the source-to-microphone distances  $\{q_{1,s}, \dots, q_{M,s}\}$ , the ideal distance weights are  $\mathbf{q} = \{\frac{1}{q_{1,s}}, \dots, \frac{1}{q_{M,s}}\}$ . It is observed that using  $\mathbf{q}$  as the weights significantly improves the autocorrelation function estimation in (18).

### 4.4. Beamforming

Beamforming, i.e., the process of focusing on a specific signal (based on the DOA or other characteristics), is widely used as part of multi-channel speech enhancement methods [6, 79]. Three beamforming techniques (listed below) have been applied to ad hoc microphone arrays. Generally, delay and sum beamforming (DSB) [8, 49] is more flexible and does not require limiting requirements, whereas more advanced beamforming methods assume some prior knowledge, which might not be the case for general scenarios.

#### 4.4.1. Delay and Sum Beamforming (DSB)

This beamforming technique has been successfully applied to ad hoc microphone arrays [49]. Using  $\mathbf{y}(n)$  (8), the DSB output is calculated as

$$\bar{x}_{DSB} = \sum_{m=1}^M x_m(n - \tau_{m,ref}), \quad (20)$$

where  $\tau_{i,ref}$  (14) is the time delay between channel  $i$  and the reference channel. The beamformer filter coefficients are obtained by:

$$\hat{\mathbf{W}} = \arg \min_{\mathbf{w}} \mathbf{w}^H \boldsymbol{\varphi}_x \mathbf{w}, \quad (21)$$

where  $\boldsymbol{\varphi}_x = E\{xx^H\}$  is the covariance. The solution is

$$\hat{\mathbf{W}} = \frac{\boldsymbol{\varphi}_x^{-1} \mathbf{h}}{\mathbf{h}^H \boldsymbol{\varphi}_x^{-1} \mathbf{h}}. \quad (22)$$

#### 4.4.2. Minimum Variance Distortionless Response beamforming (MVDR)

An optimization method for the MVDR beamformer using the pseudo-coherence model of the array (24), based on the coherence function (20), is proposed and successfully tested by Tavakoli et al. [5]:

$$P_{x_m x_{ref}} = \frac{E[X_m X_{ref}^*]}{E[|X_{ref}|^2]}, \quad (23)$$

$$X_m = \sum_{p=1}^P P_{x_m x_{ref}}^p X_{ref}^p, \quad (24)$$

where  $X_m$  is the frequency domain signal of  $\mathbf{x}_m$  (1) and  $*$  is the complex conjugate. The MVDR beamformer requires knowledge of the source DOA and of the steering vector of the array when applied to compact microphone arrays; however, under certain assumptions, it is possible to modify the MVDR beamformer and apply it to the distributed scenarios. The assumptions include connection between the channels [80] and an ad hoc array with nodes of known geometry [5].

#### 4.4.3. Linearly Constrained Minimum Variance (LCMV)

The LCMV was applied to distributed scenarios by Wood et al. [81] in experimental setups covering a wide range of random meeting scenarios. However, the LCMV beamformer requires knowledge of the RIR at each microphone location for each source in the meeting room. Himawan [82] proposed and successfully tested a clustered approach to blind beamforming for ad hoc microphone arrays, the applied features being the coherence between the diffuse noise and TDOA. The fact that the noise coherence between two microphones depends on the inter-microphone distance is exploited to estimate how close two microphones are. It is also well known that microphones located near each other have lower TDOAs, whereas microphones located farther away (i.e., metres) from each other have larger TDOAs.

#### 4.5. Speech Enhancement

Speech enhancement can cover applications, such as noise cancellation [22, 83], beamforming [5], and echo cancellation [8]. These applications can be used separately or jointly as a combined speech enhancement method. The state-of-the-art speech enhancement techniques proposed for conventional arrays of known geometry are inapplicable to ad hoc microphone arrays, and existing approaches are confined to basic beamforming techniques [5, 49]. Some basic techniques apply centralized multi-channel Wiener filters and the so-called distributed adaptive node-specific signal estimation (DANSE) algorithm [22] to remove noise in distributed hearing aid systems. These methods assume that the channels can communicate and transmit time stamps. Other speech enhancement methods proposed for ad hoc arrays require limiting supervision requirements, such as user identification of the target speech [50].

A general scenario with  $M$  ad hoc microphones is rewritten as

$$\mathbf{x}(n) = \mathbf{h}(n) * s(n) + \mathbf{v}(n), \quad (25)$$

$$\mathbf{h}(n) = \begin{bmatrix} h_1(n) \\ \vdots \\ h_M(n) \end{bmatrix}, \quad (26)$$

where  $\mathbf{x}(n) = [x_1(n), \dots, x_M(n)]^T$  (8) contains the multi-channel recording of all  $M$  microphones in the array,  $\mathbf{h}(n)$  is the RIR matrix at each microphone's location for source  $j$ , and  $\mathbf{v}(n)$  is the diffuse noise. The goal is to retrieve  $s(n)$  from  $\mathbf{x}(n)$ (25). A clustered speech enhancement approach based on beamforming and auto-regressive (AR) modelling of the speech signal was proposed and tested by Pasha and Ritz [8]. They showed that removing the microphone located far from the active source and exclusively applying the multi-channel dereverberation method for the microphones located nearer the source improved the speech enhancement performance.

$$e_{x_m}(n) = x_m(n) - \sum_{k=1}^p b_k x_m(n-k). \quad (27)$$

where  $x_m(n)$  (7) is the single-channel recording of an ad hoc microphone array and  $b_k$  is the LPC coefficient of order  $P$ . The kurtosis of the LP residual signals is then calculated as

$$k_{x_m}(n) = \frac{E\{e_{x_m}^4(n)\}}{E^2\{e_{x_m}^2(n)\}} - 3 \quad (28)$$

where  $E\{\}$  denotes the mathematical expectation and  $k_{x_m}(n)$  is the kurtosis of the LP residual signals.

#### 4.6. Source Localization and DOA Estimation

DOA estimation refers to one-dimensional source localization (i.e., angle of arrival) [84], whereas source localization can have a more general meaning, such as pinpointing the source location in a room (i.e., two- or three-dimensional localization) [85]. In this section, DOA estimation and source localization are reviewed together as they both require location feature estimation [21].

Source localization methods proposed for microphone arrays are based on extracting location features from the recorded signals and analyzing them to localize the active source [7]. This approach has limiting assumptions when applied to ad hoc arrays; for instance, if the relative distance between the microphones is unknown, extracted features such as amplitude attenuation and time delays cannot be accurately translated to location features. Researchers have tried to address this issue for the gain feature by proposing a relative attenuation feature [12, 62] that can localize collocated sources and microphones when the microphones have different gains. It has also been shown that if some sources (i.e., three out of seven) are not collocated with any microphones, the method can still localize all the sources and microphones accurately.

It has been shown that arbitrarily arranged sensors (forming a network of acoustic sensors) can be effectively applied for source DOA estimation [13]. The results indicate that the proposed method detects the source angle of arrival in degrees with less than 2% error.

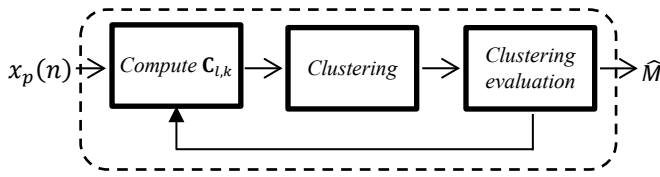


Figure 4: The proposed source counting system

Features derived from RIRs, such as attenuation [54], and the clarity feature ( $C_{50}$ ) [37], are also applied for the two-dimensional localization of sources.

#### 4.7. Source Counting and crosstalk detection

Speech processing methods use voice activity detection (VAD) to detect the periods with an active speaker. In scenarios with more than one microphone, VAD can be applied in source counting (Figure 4) and multi-talk detection applications as well [9].

Inspired by VAD algorithms, researchers have proposed multi-talk detectors in which the source and microphone locations are not available (Figure 4). Moonen and Bertrand [86] suggested a multi-speaker voice activity detection method that tracks the power of multiple simultaneous speakers. Coherent-to-diffuse ratio (CDR) values (32) calculated or estimated at dual microphone node locations are also applied for source counting [56].

For two-element nodes where  $\tilde{x}_{n,p}(t)$  (29) represents the signals recorded by the two channels,  $p \in \{1, 2\}$ , at node  $n \in \{1, \dots, N\}$ ,  $\tilde{s}(t)$  is the source signal and  $\tilde{h}_{n,p}(t)$  is the RIR at the  $n$ th node as in (6). The CDR features are calculated using the following:

$$\tilde{x}_{n,p}(t) = \tilde{s}(t) * \tilde{h}_{n,p}(t) + u_{n,p}(t), \quad (29)$$

$$C_{v_n}(f) = \frac{|\varphi_{v_{n,1}|v_{n,2}}(f)|^2}{\varphi_{v_{n,1}|v_{n,1}}(f) \varphi_{v_{n,2}|v_{n,2}}(f)}, \quad (30)$$

$$C_{x_n}(l, f) = \frac{|\varphi_{x_{n,1}|x_{n,2}}(l, f)|^2}{\varphi_{x_{n,1}|x_{n,1}}(l, f) \varphi_{x_{n,2}|x_{n,2}}(l, f)}, \quad (31)$$

where  $C_{v_n}(f)$  and  $C_{x_n}(f)$  are the noise and the signal magnitude squared coherence (MSC) values respectively and  $l$  represents the frame index. The CDR is

$$CDR_n(l, f) = \frac{C_{u_n}(f) - C_{x_n}(l, f)}{C_{x_n}(l, f) - C_s(l, f)}, \quad (32)$$

from which the use of the average CDR over the entire frequency band and  $L$  frames is given by

$$\overline{CDR}_n = \frac{1}{L(f_B - f_0)} \int_{f=f_0}^{f_B} \sum_{l=1}^L CDR_n(l, f) df. \quad (33)$$

The main limitation of the method proposed by Pasha et al. [56] is that all the nodes must be of the same structure, which

limits the method's applicability. The MSC is found using the cross-power spectral density (CPSD) as presented by Pasha et al. [87] (Figure 4):

$$c(l, f) = \frac{|\varphi_{x_1|x_2}(l, f)|^2}{\varphi_{x_1|x_1}(l, f) \varphi_{x_2|x_2}(l, f)}, \quad (34)$$

where  $f \in \{1, \dots, F\}$  is the frequency index of  $F$  total frequencies. The CPSD function used in (34) is defined as

$$\varphi_{x_1|x_2}(l, f) \triangleq \frac{1}{AK} \sum_{a,b} (x_1 * x_2)(Al + a, b) e^{-\frac{j2\pi kb}{K}}, \quad (35)$$

where  $a \in \{1, \dots, A\}$  is a frame index and  $j = \sqrt{-1}$  represents the imaginary unit. The cross-correlation is calculated by:

$$(x_1 * x_2)(\cdot, b) \triangleq \sum_n x_1(\cdot, n) x_2(\cdot, n + b), \quad (36)$$

where  $b$  is the displacement and  $x_p(n)$  framed is  $x_p(Al + a, n)$ .

#### 4.8. Source Separation

Crosstalk and speaker overlap decrease the signal quality and intelligibility in scenarios such as teleconferencing and meetings [88, 89].

The problem is mathematically formulated for  $M$  microphones and  $N$  sources (1) as

$$y_m(n) = \sum_j s_j(n) * h_{m,i,j}(n) \quad (37)$$

where  $y_m(n)$  is the speech mixture recorded by the  $m$ th microphone from the array,  $s_i(n)$  is the  $j$ th-source speech signal, and  $h_{m,i,j}(n)$  is the RIR for the  $m$ th microphone and the  $j$ th source (1). The goal is to obtain  $\mathbf{s}(n) = \{s_1(n), \dots, s_N(n)\}$  from  $\mathbf{y}(n) = \{y_1(n), \dots, y_M(n)\}$ .

Independent component analysis (ICA) was applied to blind speech separation in the online teleconferencing applications by Dmochowsky et al. [90]. Although the proposed ICA method is formulated for a general scenario, the experimental setup does not cover various scenarios and is limited to a two-element array. The noise and reverberation levels are low in the experimental setup and challenging scenarios (e.g. reverberation times higher than 800ms and very low SNRs) are not investigated.

More advanced sound source separation methods take into account the spatial coverage of the distributed arrays [91]. The signals obtained by the sub-arrays are then filtered by a geometric filter to achieve the highest output SIR. It is concluded that the proposed method can suppress (reject) interference by up to 40 dB in a reverberant environment. The novelty of this method is its use of passing and rejecting masks in the time-frequency domain to partition the microphones based on their power spectral density (PSD). The experimental setup is confined to one scenario with three sources and three microphones located near each other in pairs.

#### 4.9. Speech Recognition

Speech recognition as a main aspect of human-machine interaction has attracted significant attention in recent years. Distributed microphone arrays have significant advantages over compact microphone arrays as they provide unobtrusive and spatially flexible interaction between humans and personal devices spread out within a room. A series of ad hoc signal processing techniques for speech recognition applications, such as spatial directivity, beamforming, and speech feature extraction, was discussed by Himawan [18], the main focus being on beamforming for speech enhancement. Generalized side lobe cancelling techniques [92] and linear prediction (LP)-based speech enhancement [79] have proven to be successful speech recognition methods in ad hoc scenarios.

### 5. Machine Learning Applied to Distributed Scenarios

Machine learning techniques have been widely used in different areas of speech and audio signal processing, such as emotion recognition and source localization [93]. Machine learning and data mining techniques have been shown to be effective for learning and predicting nonlinear patterns. They have been widely used for beamforming via support vector machines (SVMs), source localization via neural networks, and other applications. As machine learning techniques are highly sensitive to the training set and parameters, using them in the flexible and uncertain distributed scenarios is very challenging. However, researchers have managed to define informative discriminative features for clustering and classifying microphones and signals, features that are independent of any specific setup [8, 91] and can discriminate among the microphones within an ad hoc array regardless of array topology. These methods flexibly exclude a subset of the microphones (nodes) from the multi-channel process (e.g., multi-channel speech enhancement [8]) and are based on certain predefined selection criteria [5].

#### 5.1. Microphone Clustering

Clustering is an unsupervised machine learning technique the goal of which is to assign objects (e.g., microphones) to groups with small intra-group differences and large inter-group differences [94]. The problem of clustering microphones based on their spatial locations was investigated by Gregen et al. [95]. It is important to cluster the microphones based on their spatial distances to select an optimal subset of microphones. Clustered speech processing approaches are applied to take advantage of the spatial selectivity of beamforming [49], dereverberation [8], and interference suppression [9]. The mel-frequency cepstral coefficients (MFCCs) [95], the coherence feature [3], the Legendre polynomial-based cepstral modulation ratio regression (LP-CMRARE) [95] (38)–(39), and the kurtosis of the LP residual [8] are used as features in various clustering methods reported in the literature. Along with the discriminative clustering feature, the clustering technique is important as it determines the limitations of the overall clustering method. For instance, the fuzzy clustering method

presented by Gregen et al. [95] requires prior knowledge of the number of sources, whereas a flexible codebook-based clustering method based on RIRs was applied by Pasha et al. [11]:

$$\hat{X}_c = |F(X_c)| \quad (38)$$

where  $X_c$  is the recorded signal in the cepstral domain and  $F$  denotes the Fourier transform. The average magnitude over the window length ( $C_T$ ) of the modulation spectrum is calculated using

$$\bar{X}_c = \frac{1}{C_T} \sum_{c=0}^{C_T-1} |\hat{X}_c|, \quad (39)$$

as the clustering feature. A practical clustering method in which the numbers of microphones and clusters are not known was proposed by Pasha and Ritz [8], who applied clustering to exclude highly distorted and reverberant microphones from the dereverberation process. It is concluded that the clustered approach improves the direct-to-reverberant ratio (DRR) by 5 dB. Although it is impossible to evaluate the clustering methods decisively [95] when there is no ground truth, researchers have proposed clustering performance measures such as the success rate (SR) [11] and purity [49].

#### 5.2. Signal Classification

Supervised machine learning methods such as classification are highly sensitive to the training set and setup, and are not always applicable to uncertain changeable scenarios such as ad hoc microphone arrays and meetings [96]. However, it has been shown that under certain assumptions (e.g., the availability of a clean training set) [95], it is possible to classify the recorded signals based on certain predefined classes (e.g., speech, noise, and music). It is concluded that cepstral features such as MFCC and LP-CMRARE are reliable features with which to discriminate speech, music, and noise signals [95]. The microphone clustering method uses MFCC and LP\_CMRARE as the features and divides microphones into two (i.e., the number of sources, which is assumed to be known) clusters. The signals recorded by the clustered microphones are then classified based on the predefined classes. It is also assumed that clean training data for each class are available [97].

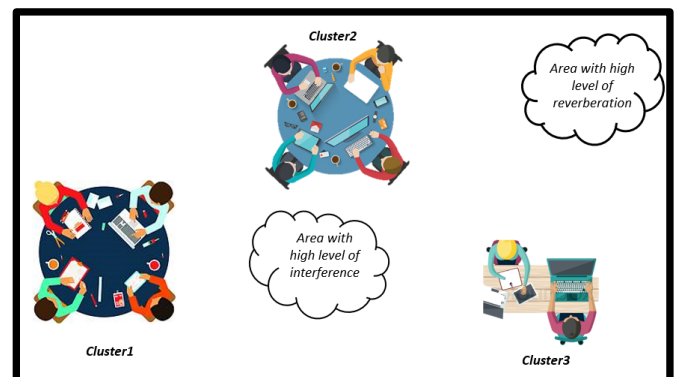


Figure 5: Microphone clusters [3]



## 6. Distributed Signal Processing Resources

An extensive database recorded using ad hoc microphones was reported and applied by Wood et al. [81] for real-world beamforming experiments. Twenty-four microphones were positioned in various locations on a central table in a reverberant room, and their outputs were recorded while four target talkers seated at the table read some text or had natural conversations. The recorded speech signals have been made publicly available [98]. Distributed beamforming tools and tutorials covering distributed speech enhancement and random microphone deployment have been made available [99]. Materials and resources associated with the distributed speech recognition (DSR) research conducted by IBM have been archived and sourced [100]. A coherence-based (31) source counting method for distributed scenarios has been made available by Donley [101]; the method counts the number of active speakers (up to eight) in a spontaneous meeting in reverberant environments.

University of Illinois have provided a dataset which facilitates distributed source separation and augmented listening research [102]. The dataset is recorded using 10 speech sources and 160 microphones in a large, reverberant conference room. The applied microphone array includes wearable sensors and microphones connected to tablets.

## 7. Conclusion

This review paper discussed recent advances in the context of distributed microphone arrays and signal processing. Standard dereverberation, speech separation, and source counting methods have been successfully adapted to the context of distributed signal processing using novel features and machine learning. Most existing ad hoc beamforming methods suffer from limiting assumptions that make them niche applications. Issues such as real-time source localization and DOA estimation are still challenging.

## References

- [1] S. Pasha, C. Ritz and J. Lundgren, "A Survey on Ad Hoc Signal Processing: Applications, Challenges and State-of-the-Art Techniques," in *International Symposium on Signal Processing and Information Technology (ISSPIT)*, Ajman, United Arab Emirates, 2019, DOI:10.1109/ISSPIT47144.2019.9001860.
- [2] M. J. Taghizadeh, Enabling speech applications using ad hoc microphone arrays: PhD dissertation, Lausanne: EPFL, 2015, [https://publications.idiap.ch/downloads/papers/2015/Taghizadeh\\_THESIS\\_2015.pdf](https://publications.idiap.ch/downloads/papers/2015/Taghizadeh_THESIS_2015.pdf).
- [3] S. Pasha, Analysis and Enhancement of Spatial Sound Scenes Recorded using Ad-Hoc Microphone Arrays, Doctor of Philosophy thesis, School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, 2017, <https://ro.uow.edu.au/theses1/450>.
- [4] J. Benesty, J. Chen and J. Huang, *Microphone Array Signal Processing*, Verlag Berlin Heidelberg: Springer, 2008, DOI: 10.1007/978-3-540-78612-2.
- [5] V. M. Tavakoli, J. R. Jensen, M. Christensen and J. Benesty, "Pseudo-coherence-based MVDR beamformer for speech enhancement with ad hoc microphone arrays," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, South Brisbane, 2015, DOI: 10.1109/ICASSP.2015.7178453.
- [6] N. D. Gaubitch, J. Martinez, W. B. Kleijn and R. Heusdens, "On near-field beamforming with smartphone-based ad-hoc microphone arrays," in *14th International Workshop on Acoustic Signal Enhancement (IWAENC)*, Juan-les-Pins, 2014, DOI: 10.1109/IWAENC.2014.6953345.
- [7] M. Souden, K. Kinoshita, M. Delcroix and T. Nakatani, "Location Feature Integration for Clustering-Based Speech Separation in Distributed Microphone Arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(2), 354-367, 2014, DOI: 10.1109/TASLP.2013.2292308.
- [8] S. Pasha and C. Ritz, "Clustered multi-channel dereverberation for ad-hoc microphone arrays," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, Hong Kong, 2015, DOI: 10.1109/APSIPA.2015.7415519.
- [9] S. Gregen and R. Martin, "Estimating Source Dominated Microphone Clusters in Ad-Hoc Microphone Arrays by Fuzzy Clustering in the Feature Space," in *Speech Communication; 12. ITG Symposium*, Paderborn, Germany, 2016, Print ISBN: 978-3-8007-4275-2.
- [10] C. Evers, Y. Dorfan, S. Gannot and P. Naylor, "Source tracking using moving microphone arrays for robot audition," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, 2017, DOI: 10.1109/ICASSP.2017.7953337.
- [11] S. Pasha, Y. Zou and C. Ritz, "Forming ad-hoc microphone arrays through clustering of acoustic room impulse responses," in *China Summit and International Conference on Signal and Information Processing (ChinaSIP)*, Chengdu, 2015, DOI: 10.1109/ChinaSIP.2015.7230367.
- [12] Z. Liu, Z. Zhang, L. W. He and P. Chou, "Energy-Based Sound Source Localization and Gain Normalization for Ad Hoc Microphone Arrays," in *International Conference on Acoustics, Speech and Signal Processing*, Honolulu, 2007, DOI: 10.1109/ICASSP.2007.366347.
- [13] S. Araki, H. Swada, R. Mukai and S. Makino, "DOA Estimation for Multiple Sparse Sources with Arbitrarily Arranged Multiple Sensors," *Journal of Signal Processing Systems*, 63(3), 265-275, 2011, <https://doi.org/10.1007/s11265-009-0413-9>.
- [14] L. Krishnan, P. Teal and T. Betlehem, "A robust sparse approach to acoustic impulse response shaping," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, South Brisbane, QLD, 2015, DOI: 10.1109/ICASSP.2015.7178067.
- [15] T. v. Waterschoot, "Distributed estimation of cross-correlation functions in ad-hoc microphone arrays," in *23rd European Signal Processing Conference (EUSIPCO)*, Nice, 2015, DOI: 10.1109/EUSIPCO.2015.7362385.
- [16] V. M. Tavakoli, J. R. Jensen, M. G. Christensen and J. Benesty, "A Framework for Speech Enhancement With Ad Hoc Microphone Arrays," *ACM Transactions on Audio, Speech, and Language Processing*, 24(6), 1038-1051, June 2016, DOI: 10.1109/TASLP.2016.2537202.
- [17] A. Bertrand, "Applications and trends in wireless acoustic sensor networks: A signal processing perspective," in *18th Symposium on Communications and Vehicular Technology in the Benelux (SCVT)*, Ghent, 2011, DOI: 10.1109/SCVT.2011.6101302.
- [18] I. Himawan, Speech recognition using ad-hoc microphone arrays, PhD dissertation, Queensland university of technology, 2010, <https://pdfs.semanticscholar.org/e95b/ad880c47be6e146dd75b82bc72e51e97db2f.pdf>.
- [19] R. Cutler, Y. Rui, A. Gupta, J. Cadiz, I. Tashev, L. He and A. Colburn, "Distributed meetings: A meeting capture and broadcasting system," in *ACM Multimedia*, Juan-les-Pins, France, 2002, DOI: 10.1109/ICASSP.2003.1202753.
- [20] V. Tavakoli, J. Jensen, R. Heusdens, J. Benesti and M. Christensen, "Distributed max-SINR speech enhancement with ad hoc microphone arrays," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, 2017, DOI: 10.1109/ICASSP.2017.7952136.
- [21] A. Asaei, H. Bourlard, M. Taghizadeh and V. Cehver, "Model-based sparse component analysis for reverberant speech localization," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, 2014, DOI: 10.1109/ICASSP.2014.6853835.
- [22] A. Bertrand and M. Moonen, "Robust distributed noise reduction in hearing aids with external acoustic sensor nodes," *EURASIP Journal on*

- Advances in Signal Processing*(14, 2009, <https://doi.org/10.1155/2009/530435>.
- [23] Y. Jia, L. Yu and I. Kozintsev, "Distributed Microphone Arrays for Digital Home and Office," in *IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, Toulouse, 2006, DOI: 10.1109/ICASSP.2006.1661463.
- [24] Diane J. Cooka, Juan C. Augustob and Vikramadit, "Ambient intelligence: Technologies, applications, and opportunities," *Pervasive and Mobile Computing*, 5, 277-298, August 2009, <https://doi.org/10.1016/j.pmcj.2009.04.001>.
- [25] A. Brutti, M. Omologo and P. Svaizer, "Oriented global coherence field for the estimation of the head orientation in smart rooms equipped with distributed microphone arrays," in *NTERSPEECH 2005 - Eurospeech, 9th European Conference on Speech Communication and Technology*, Lisbon, Portugal, 2005, Corpus ID: 2927411.
- [26] J. Lee, C. H. Lee, D. W. Kim and B. Y. Kang, "Smartphone-Assisted Pronunciation Learning Technique for Ambient Intelligence," *IEEE Access*, 2016, DOI: 10.1109/ACCESS.2016.2641474.
- [27] T. Toyoda, N. Ono, S. Miyabe, T. Yamada and S. Makino, "Traffic monitoring with ad-hoc microphone array," in *14th International Workshop on Acoustic Signal Enhancement (IWAENC)*, Juan-les-Pins, 2014, DOI: 10.1109/IWAENC.2014.6954310.
- [28] K. Tsuji, "Migration monitoring of fin whales in the southern Chukchi Sea with acoustic methods during 2012-2015," in *Techno-Ocean (Techno-Ocean)*, Kobe, 2016, DOI: 10.1109/Techno-Ocean.2016.7890746.
- [29] D. Stowell, E. Benetos and L. Gill, "On-Bird Sound Recordings: Automatic Acoustic Recognition of Activities and Contexts," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(6), 1193-1206, June 2017, DOI: 10.1109/TASLP.2017.2690565.
- [30] A. M. McKee and R. A. Goubran, "Chest sound pick-up using a multisensor array," in *SENSORS*, Irvine, CA, 2005, DOI: 10.1109/ICSENS.2005.1597816.
- [31] J. Chien, M. Huang, Y. Lin and F. Chong, "A Study of Heart Sound and Lung Sound Separation by Independent Component Analysis Technique," in *International Conference of the IEEE Engineering in Medicine and Biology Society*, New York, NY, 2006, DOI: 10.1109/IEMBS.2006.260223.
- [32] M. krekovic, I. Dokmanic and M. Vetterli, "EchoSLAM: Simultaneous localization and mapping with acoustic echoes," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, 2016, DOI: 10.1109/ICASSP.2016.7471627.
- [33] I. Dokmanic, L. Daudet and M. Vetterli, "From acoustic room reconstruction to slam," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, 2016, DOI: 10.1109/ICASSP.2016.7472898.
- [34] I. Dokmanic, L. Daudet and M. Vetterli, "How to localize ten microphones in one finger snap," in *22nd European Signal Processing Conference (EUSIPCO)*, Lisbon, 2014, Electronic ISBN: 978-0-9928-6261-9.
- [35] F. Antonacci, "Inference of Room Geometry From Acoustic Impulse Responses," *IEEE Transactions on Audio, Speech, and Language Processing*, 20(10), 2683-2695, 2012, DOI: 10.1109/TASL.2012.2210877.
- [36] P. Peso Parada, D. Sharma, J. Lainez, D. Barreda and P. Naylor, "A Single-Channel Non-Intrusive C50 Estimator Correlated With Speech Recognition Performance," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(4), 719-732, April 2016, DOI: 10.1109/TASLP.2016.2521486.
- [37] S. Pasha, C. Ritz and Y. X. Zou, "Detecting multiple, simultaneous talkers through localising speech recorded by ad-hoc microphone arrays," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, Jeju, 2016, DOI: 10.1109/APSIPA.2016.7820873.
- [38] T. K. Hon, L. Wang, J. D. Reiss and A. Cavallaro, "Fine landmark-based synchronization of ad-hoc microphone arrays," in *23rd European Signal Processing Conference (EUSIPCO)*, Nice, 2015, DOI: 10.1109/EUSIPCO.2015.7362600.
- [39] S. Araki, N. Ono, K. Kinoshita and M. Delcroix, "Estimation of Sampling Frequency Mismatch between Distributed Asynchronous Microphones under Existence of Source Movements with Stationary Time Periods Detection," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, United Kingdom, 2019), 785-789, DOI: 10.1109/ICASSP.2019.8683192.
- [40] S. Pazors, M. Hurtado and C. Muravchik, "On Sparse Methods for Array Signal Processing in the Presence of Interference," *Antennas and Wireless Propagation Letters*, 14, 1165-1168, 2015, DOI: 10.1109/LAWP.2015.2394233.
- [41] M. Taseska, S. Markovich-golan, E. Habets and S. Gannot, "Near-field source extraction using speech presence probabilities for ad hoc microphone arrays," in *14th International Workshop on Acoustic Signal Enhancement (IWAENC)*, Juan-les-Pins, 2014, DOI: 10.1109/IWAENC.2014.6954000.
- [42] A. D. Sarwate and K. Chaudhuri, "Signal Processing and Machine Learning with Differential Privacy: Algorithms and Challenges for Continuous Data," *IEEE Signal Processing Magazine*, 30(5), 86-94, Sept. 2013, DOI: 10.1109/MSP.2013.2259911.
- [43] R. Wehrmann, "Concepts of improving hands-free speech communication," in *IEEE International Symposium on Circuits and Systems*, San Diego, CA, 1992, DOI: 10.1109/ISCAS.1992.230435.
- [44] P. Chevalier and A. Maurice, "Blind and informed cyclic array processing for cyclostationary signals," in *9th European Signal Processing Conference (EUSIPCO 1998)*, Rhodes, 1998, Print ISBN: 978-960-7620-06-4.
- [45] E. Vincent, N. Bertin, R. Gribouval and F. Bimbot, "From Blind to Guided Audio Source Separation: How models and side information can improve the separation of sound," *EEE Signal Processing Magazine*, 31(3), 107-111, May 2014, DOI: 10.1109/MSP.2013.2297440.
- [46] C. Vanwynsberhea, P. Challande, J. Marchal, R. marchiano and F. Ollivier, "A robust and passive method for geometric calibration of large arrays," *The Journal of the Acoustical Society of America*, 139, p. 1252, 2016, <https://doi.org/10.1121/1.4944566>.
- [47] I. McCowan, M. Lincoln and I. Himawan, "Microphone array shape calibration in diffuse noise fields," *Transactions on Audio, Speech and Language Processing*, 16(3), 666-670, 2008, DOI: 10.1109/TASL.2007.911428.
- [48] D. McCarthy and F. Boland, "A method for source-microphone range estimation, using arrays of unknown geometry, in reverberant room environments," in *15th European Signal Processing Conference*, Poznan, 2007, Print ISBN: 978-839-2134-04-6.
- [49] I. Himawan, I. McCowan and S. Sridhan, "Clustered Blind Beamforming From Ad-Hoc Microphone Arrays," *IEEE Transactions on Audio, Speech, and Language Processing*, 19(4), 661-676, May 2011, DOI: 10.1109/TASL.2010.2055560.
- [50] R. Sakanashi, N. Ono, S. Miyabe, T. Yamada and S. Makino, "Speech enhancement with ad-hoc microphone array using single source activity," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, Kaohsiung, 2013, DOI: 10.1109/APSIPA.2013.6694323.
- [51] I. Dokmanic, J. Raineri and M. Vetterli, "Relax and unfold: Microphone localization with Euclidean distance matrices," in *23rd European Signal Processing Conference (EUSIPCO)*, Nice, 2015, DOI: 10.1109/EUSIPCO.2015.7362386.
- [52] Y. Tian, Z. Chen and F. Yin, "Distributed Kalman filter-based speaker tracking in microphone array networks," *Applied Acoustics*, 89), 71-77, 2015, <https://doi.org/10.1016/j.apacoust.2014.09.004>.
- [53] A. Asaei, N. Mohamadiha, M. J. Taghizadeh, S. Doclo and H. Bourlard, "On application of non-negative matrix factorization for ad hoc microphone array calibration from incomplete noisy distances," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, South Brisbane, 2015, DOI: 10.1109/ICASSP.2015.7178460.
- [54] S. Pasha and C. Ritz, "Informed source location and DOA estimation using acoustic room impulse response parameters," in *International Symposium on Signal Processing and Information Technology (ISSPIT)*, Abu Dhabi, 2015, DOI: 10.1109/ISSPIT.2015.7394316.

- [55] I. Dokmanic, Y. Lu and M. Vetterli, "Can one hear the shape of a room: The 2-D polygonal case," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, 2011, DOI: 10.1109/ICASSP.2011.5946405.
- [56] S. Pasha, J. Donley, C. Ritz and X. Y. Zou, "Towards real-time source counting by estimation of coherent-to-diffuse ratios from ad-hoc microphone array recordings," in *Hands-free Speech Communication and Microphone Arrays (HSCMA)*, San Francisco, 2017, DOI: 10.1109/HSCMA.2017.7895582.
- [57] M. Taseska and A. P. Habets, "Spotforming using distributed microphone arrays," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, 2013, DOI: 10.1109/WASPAA.2013.6701876.
- [58] S. Woźniak and K. Kowalczyk, "Passive Joint Localization and Synchronization of Distributed Microphone Arrays," *Signal Processing Letters*, 26(2), 292-296, Feb. 2019, DOI: 10.1109/LSP.2018.2889438.
- [59] Q. Zhang, W. Zhang, J. Feng and R. Tang, "Distributed Acoustic Source Tracking in Noisy and Reverberant Environments With Distributed Microphone Networks," *IEEE Access*, 8), 9913-9927, 2020, DOI: 10.1109/ACCESS.2020.2965210.
- [60] S. Vesa, "Binaural Sound Source Distance Learning in Rooms," *Transactions on Audio, Speech, and Language Processing*, 17(8), 1498-1507, 2009, DOI: 10.1109/TASL.2009.2022001.
- [61] I. Himawan, S. Sridharan and I. McCowan, "Dealing with uncertainty in microphone placement in a microphone array speech recognition system," in *International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, NV, 2008, DOI: 10.1109/ICASSP.2008.4517922.
- [62] M. Chen, Z. Liu, W. He, P. Chou and Z. Zhang, "Energy-based position estimation of microphones and speakers for ad hoc microphone arrays," in *Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, New York, 2007, DOI: 10.1109/ASPAA.2007.4393035.
- [63] M. J. Taghizadeh, R. Parhizkar, P. N. Garner and H. Bourlard, "Euclidean distance matrix completion for ad-hoc microphone array calibration," in *18th International Conference on Digital Signal Processing (DSP)*, Fira, 2013, DOI: 10.1109/HSCMA.2014.6843239.
- [64] M. J. Taghizadeh, N. Carner, H. Bourlard and A. Asaei, "Ad Hoc Microphone Array Calibration: Euclidean Distance Matrix Completion Algorithm and Theoretical Guarantees," *Signal Processing*, p. <https://doi.org/10.1016/j.sigpro.2014.07.016>, 2014.
- [65] M. Taghizadeh, A. Asaei, P. Garner and h. Bourlard, "Ad-hoc microphone array calibration from partial distance measurements," in *4th Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA)*, Villers-les-Nancy, 2014, DOI: 10.1109/HSCMA.2014.6843239.
- [66] N. Ono, K. Shibata and H. Kameoka, "Self-localization and channel synchronization of smartphone arrays using sound emissions," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, Jeju, 2016, DOI: 10.1109/APSIPA.2016.7820778.
- [67] L. Wang, T. K. Hon and A. Cavallaro, "Self-Localization of Ad-Hoc Arrays Using Time Difference of Arrivals," *IEEE Transactions on Signal Processing*, 64(4), 1018-1033, Feb.15, 2016, DOI: 10.1109/TSP.2015.2498130.
- [68] N. D. Gaubitch, W. B. Kleijn and R. Heusdens, "Auto-localization in ad-hoc microphone arrays," in *International Conference on Acoustics, Speech and Signal Processing*, Vancouver, BC, 2013, DOI: 10.1109/ICASSP.2013.6637618.
- [69] E. Robledo, T. S. Wada and B. Juang, "On Dealing with Sampling Rate Mismatches in Blind Source Separation and Acoustic Echo Cancellation," in *Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, USA, 2007, DOI: 10.1109/ASPAA.2007.4393044.
- [70] H. Chiba, N. Ono, S. Miyabe, Y. Takashaahi, T. Yamada and S. Makino, "Amplitude-based speech enhancement with nonnegative matrix factorization for asynchronous distributed recording," in *14th International Workshop on Acoustic Signal Enhancement (IWAENC)*, Juan les pins, France, 2014, DOI: 10.1109/IWAENC.2014.6954007.
- [71] S. Miyabe, N. Ono and S. Makino, "Blind compensation of interchannel sampling frequency mismatch for ad hoc microphone array based on maximum likelihood estimation," *Signal Processing*, p. 185-196, 2015, <https://doi.org/10.1016/j.sigpro.2014.09.015>.
- [72] N. Ono, H. Kohno, N. Ito and S. Sagayama, "Blind alignment of asynchronously recorded signals for distributed microphone array," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, 2009, DOI: 10.1109/ASPAA.2009.5346505.
- [73] P. Pertila, M. S. Hamalainen and M. Mieskolainen, "Passive Temporal Offset Estimation of Multichannel Recordings of an Ad-Hoc Microphone Array," *IEEE Transactions on Audio, Speech, and Language Processing*, 21(11), 2393-2402, 2013, DOI: 10.1109/TASLP.2013.2286921.
- [74] N. J. Bryan, P. Smargadis and G. J. Mysore, "Clustering and synchronizing multi-camera video via landmark cross-correlation," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, 2012, DOI: 10.1109/ICASSP.2012.6288396.
- [75] R. Lienhart, I. Kozintsev, S. Wher and M. Yeung, "On the importance of exact synchronization for distributed audio signal processing," in *International Conference on Acoustics, Speech and Signal Processing*, Hong Kong, 2003, DOI: 10.1109/ASPAA.2003.1285842.
- [76] T. Nakatani, T. Yoshioka, K. Kinoshita and M. Miyoshi, "Blind speech dereverberation with multi-channel linear prediction based on short time fourier transform representation," in *International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, NV, USA, 2008, DOI: 10.1109/ICASSP.2008.4517552.
- [77] S. Pasha, C. Ritz and Y. X. Zou, "Spatial multi-channel linear prediction for dereverberation of ad-hoc microphones," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Kuala Lumpur, 2017, DOI: 10.1109/APSIPA.2017.8282306.
- [78] S. Pasha and J. Lundgren, "Multi-Channel Compression and Coding of Reverberant Ad-Hoc Recordings Through Spatial Autoregressive Modelling," in *30th Irish Signals and Systems Conference (ISSC)*, Maynooth, Ireland, 2019, DOI: 10.1109/ISSC.2019.8904918.
- [79] M. Delcroix, T. Yosghioka, A. Ogawa and Y. Kubo, "Linear prediction-based dereverberation with advanced speech enhancement and recognition technologies for the REVERB challenge," *Proceedings of REVERB Workshop*, 2014.
- [80] V. M. Tavakoli, J. R. Jensen, R. Heusdens, J. Benesty and M. G. Christensen, "Ad hoc microphone array beamforming using the primal-dual method of multipliers," in *24th European Signal Processing Conference (EUSIPCO)*, Budapest, 2016, DOI: 10.1109/EUSIPCO.2016.7760416.
- [81] W. S. Wood, E. Hadad, I. Merks, B. Xu and S. Gannot, "A real-world recording database for ad hoc microphone arrays," in *Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, 2015, DOI: 10.1109/WASPAA.2015.7336915.
- [82] I. Himawan, I. McCowan and S. Sridharan, "Clustering of ad-hoc microphone arrays for robust blind beamforming," in *International Conference on Acoustics, Speech and Signal Processing*, Dallas, TX, 2010, DOI: 10.1109/ICASSP.2010.5496201.
- [83] T. C. Lawin, S. Stenzel, J. Freudenberger and S. Doclo, "Generalized Multichannel Wiener Filter for Spatially Distributed Microphones," in *Speech Communication; 11. ITG Symposium*, Erlangen, Germany, 2014, Print ISBN: 978-3-8007-3640-9.
- [84] T. Ballal and C. J. Bleakley, "DOA estimation of multiple sparse sources using three widely-spaced sensors," in *17th European Signal Processing Conference*, Glasgow, 2009, Print ISBN: 978-161-7388-76-7.
- [85] M. H. Hennecke and G. A. Fink, "Towards acoustic self-localization of ad hoc smartphone arrays," in *Joint Workshop on Hands-free Speech Communication and Microphone Arrays*, Edinburgh, 2011, DOI: 10.1109/HSCMA.2011.5942378.
- [86] A. Moonen and M. Bertrand, "Energy-based multi-speaker voice activity detection with an ad hoc microphone array," in *International Conference on Acoustics, Speech and Signal Processing*, Dallas, TX, 2010, DOI: 10.1109/ICASSP.2010.5496183.



- [87] S. Pasha, J. Donley and C. Ritz, "Blind speaker counting in highly reverberant environments by clustering coherence features," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Kuala Lumpur, 2017, DOI: 10.1109/APSIPA.2017.8282303.
- [88] Z. Liu, "Sound source separation with distributed microphone phone arrays in the presence of clock synchronization errors," in *Proceedings of the International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Seattle, 2008.
- [89] S. Araki, N. Ono, K. Kinoshita and M. Delcroix, "Projection Back onto Filtered Observations for Speech Separation with Distributed Microphone Array," in *8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Le Gosier, Guadeloupe, 2019, DOI: 10.1109/CAMSAP45676.2019.9022666.
- [90] J. Dmochowsky, Z. Liu and P. Chou, "Blind source separation in a distributed microphone meeting environment for improved teleconferencing," in *International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, Las Vegas, NV, USA, 2008, DOI: 10.1109/ICASSP.2008.4517553.
- [91] V. M. Tavakoli, J. R. Jensen, J. Benesty and M. G. Christensen, "A partitioned approach to signal separation with microphone ad hoc arrays," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, 2016, DOI: 10.1109/ICASSP.2016.7472272.
- [92] S. Golan, S. Gannot and I. Cohen, "Distributed GSC beamforming using the relative transfer function," in *Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, Bucharest, 2012, Print ISBN: 978-1-4673-1068-0.
- [93] D. Salvati, C. Drioli and L. Foresti, "On the use of machine learning in microphone array beamforming for far-field sound source localization," in *IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, Vietri sul Mare, 2016, DOI: 10.1109/MLSP.2016.7738899.
- [94] B. Clarke, E. Fokoue and H. Zhang, *Principles and Theory for Data Mining and Machine Learning*, Springer-Verlag New York Inc, 1st August 2009, <https://doi.org/10.1007/978-0-387-98135-2>.
- [95] S. Gregen, A. Nagathil and R. Martin, "Audio signal classification in reverberant environments based on fuzzy-clustered ad-hoc microphone arrays," in *International Conference on Acoustics, Speech and Signal Processing, ICASSP*, Vancouver, BC, 2013, DOI: 10.1109/ICASSP.2013.6638347.
- [96] "IEEE DCASE challenge," [Online]. Available: <http://dcase.community/>. [Accessed 30 3 2020].
- [97] S. Gergen and R. Martin, "Linear Combining of Audio Features for Signal Classification in Ad-hoc Microphone Arrays," in *Speech Communication; 11. ITG Symposium*, Erlangen, Germany, 2014.
- [98] "Ad-hoc recording signals," <http://www.eng.biu.ac.il/gannot/speech-enhancement>
- [99] "Prof. Gannot speech enhancement tutorial," <http://www.eng.biu.ac.il/gannot/tutorials-and-keynote-addresses/introduction-to-distributed-speech-enhancement-algorithms-for-ad-hoc-microphone-arrays-wireless-acoustic-sensor-networks/>.
- [100] IBM, "Distributed Speech Recognition," <https://www.ibm.com/blogs/research/2019/10/asr-deep-learning/>.
- [101] J. Donley. <https://github.com/JacobD10/CoherenceBasedSourceCounter>.
- [102] R. M. Corey, M. D. Skarha and A. C. Singer, *Composers, Massive Distributed Microphone Array Dataset*. [Sound Recording]. University of Illinois at Urbana-Champaign. [https://doi.org/10.13012/B2IDB-6216881\\_V1](https://doi.org/10.13012/B2IDB-6216881_V1). 2019.



## Computational Intelligence and Statistical Learning Performances on Predicting Dengue Incidence using Remote Sensing Data

Nittaya Kerdprasop<sup>\*1</sup>, Kittisak Kerdprasop<sup>1</sup>, Paradee Chuaybamroong<sup>2</sup>

<sup>1</sup>School of Computer Engineering, Suranaree University of Technology, Nakhon Ratchasima, 30000, Thailand

<sup>2</sup>Department of Environmental Science, Thammasat University, Rangsit Campus, Pathum Thani, 12120, Thailand

### ARTICLE INFO

Article history:

Received: 02 June, 2020

Accepted: 13 July, 2020

Online: 28 July, 2020

Keywords:

Dengue incidence

Data modeling

Computational intelligence

Statistical learning

Satellite-based index

Climate variability

### ABSTRACT

Dengue is a viral infection disease transmitted to people through the bite of specific mosquito species living in a tropical zone. According to the World Health Organization, dengue has been listed among the top-ten diseases for 2019 as it makes 3.9 billion people in 128 countries be at risk of infection. One major cause of substantial dengue widespread is the globally warm climate that accelerates rapid growth of mosquito vectors. In this research, we aim to build data-driven models to predict dengue cases using satellite index data to represent temperature, humidity, and greenness over the surface area of Bangkok, which is our target area of dengue prediction because of its high infection cases. Oceanic Niño Index is also used as a predictor variable to represent climate variability. The modeling methods employ seven algorithms from two broad schemes of the machine learning field. Artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) are algorithms from the subfield of computational intelligence, whereas multiple linear regression (MLR), generalized linear model (GLM), support vector regression (SVR), classification and regression tree (CART), and chi-squared automatic interaction detection (CHAID) are from the statistical learning subfield. Performances of these algorithms are evaluated on the same set of out-of-sample data. The results are that ANFIS is the best model for predicting dengue outbreak in the capital city of Thailand.

### 1. Introduction

Dengue is a communicable disease that can indirectly transmit from an infected person to another through the bite by a female mosquito in the two major *Aedes* species, *Ae. aegypti* and *Ae. albopictus*. These small mosquitoes have white stripes on their back and legs. They bite during day-time and live in and around human habitats. These mosquitoes can transmit not only dengue virus, but also chikungunya, Zika, and yellow fever [1]. Illness from dengue infection is caused by one of the four types of viruses [2]: DENV1, DENV2, DENV3, and DENV4. Patients recovered from infection by a specific type of dengue virus are immune to that virus type for life. Unfortunately, they are still vulnerable to the other three types of dengue viruses. The many types of viruses is just one reason for a widespread of dengue fever. The other causes for a rapid spread of dengue are global warming phenomenon and poorly or unplanned urbanization. The rise in temperature globally makes mosquitoes spreading from a tropical

region into subtropical zone such as China, Japan, U.S.A., and many countries in Europe [3]. A fast expansion of big cities is also another factor accelerating dengue spread as the *Ae. aegypti* mosquitoes can breed and lay eggs in small water holders such as vessels and abandoned containers.

In 1950, there were only 908 dengue cases reported [4] to the World Health Organization (WHO). The endemic areas are mostly countries in the Southeast Asia including Cambodia, Philippines, Thailand, and Vietnam. In 2016, the number of dengue infection cases reported to WHO has been substantially jumped to over 3.34 million cases. The latest reports [5] during September to November 2019 of the European Centre for Disease Prevention and Control (ECDC) list countries with most dengue cases in descending order as Brazil, Philippines, Mexico, Nicaragua, Thailand, Malaysia, and Colombia. The current number of dengue infections in Brazil is over two million cases which is almost ten times higher than the same period in the previous year. In Thailand as of November 2019, more than 120,000 cases have been reported and 126 cases are fatal [5]. The highest number of dengue

\*Corresponding Author Nittaya Kerdprasop, nittaya@sut.ac.th

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050440>

infections in Thailand is in the Bangkok metropolitan area, which is the capital city with massive urbanization and warm atmosphere.

Based on the dengue case classification criteria released by WHO in 1997 [6], clinical infections can be broadly categorized into two distinct classes: dengue fever (DF) and dengue haemorrhagic fever (DHF)/dengue shock syndrome (DSS). DF is a classic dengue fever that infected person may either show no signs of illness, or having mild illness with symptoms such as abdominal pain and mucosal bleeding [7]. DHF is a type of severe dengue with manifestation of vascular permeability resulting in plasma leakage, a decrease in platelets in the blood, and an increase of red blood cells in relative to the plasma that may cause the fatal shock syndrome, DSS [7]. The classification of dengue infection as DF/DHF/DSS is based on historical cases in Thailand that infections normally occur in children younger than 15 years old [7]. At the present decade that dengue infections have extended to older age groups and spatially invade to different zones, clinical symptoms of infected patients show some variations from the WHO's descriptions [8]. Therefore, in 2009 WHO developed new criteria and guideline for dengue case diagnosis and treatment. The new criteria [8] classify dengue infection into three main groups based on the ascending levels of severity as dengue without warning signs, dengue with warning signs such as abdominal pain and mucosal bleeding, and severe dengue that includes severe plasma leakage leading to shock, severe bleeding, and severe organ failure.

Fatal cases of severe dengue in Thailand are increasing from 36 deaths in 2014 to 127 deceases in 2019 [9]. The rise in dengue mortality in Thailand is more or less coincident with the occurrence of medium to strong level of El Niño events. This relation can be noticed from the high number of dengue fatality in El Niño years of 2015 and 2019. The coincidence of El Niño and dengue endemic, especially around the area of capital city, is of particular interest to our research team. We thus plan our modeling method to employ climate variability through the Oceanic Niño Index (ONI) as one independent factor to predict the number of dengue cases in Bangkok. Other independent variables are amount of rainfall and remote sensing indices to represent temperature, humidity, brightness, and greenness of vegetation. The modeling algorithms are those from the computational intelligence (artificial neural network -- ANN, adaptive neuro-fuzzy inference system -- ANFIS) and statistical learning schemes (multiple linear regression -- MLR, support vector regression -- SVR, generalized linear model -- GLM, classification and regression tree -- CART, chi-squared automatic interaction detection -- CHAID).

A review of these and other learning techniques applied to predict dengue incidence is presented in Section 2. Details of dataset and our modeling method are explained in Section 3. Experimental results of model performance comparison are illustrated in Section 4. The conclusion is in Section 5.

## **2. Related Work**

Historically, dengue has been identified as an endemic locally found in the Southeast Asian countries. After World War II, severe type of dengue (DHF) firstly appeared in Philippines in 1954 and later emerged in Thailand in 1958 and became an outbreak in Malaysia, Cambodia, Singapore, and Vietnam in the 1960s [9]. Currently, dengue expands its boundary from a local

communicable disease to become a global epidemic. Disease prevention and efficient mechanisms to monitor and control infection are urgent actions needed to be implemented instantly by governments and health organizations worldwide. Scientists and researchers attempt to estimate in advance the outbreak of dengue infection with the purpose of precise warning leading to proper preparation and management.

Research teams from Singapore estimate dengue incidence with nonlinear regression model [10], time-series Poisson regression model [11], and neural network model [12]. Scientists in Malaysia apply satellite imagery to study the pattern of dengue incidence distribution with correlation analysis against land-use patterns in the geographic information system (GIS) [13]. To predict dengue cases in Malaysia, variations of network modeling methods such as ANFIS [14], self-organizing map [15], and a hybrid of genetic algorithm and ANN [16] are applied. Various research groups in Philippines study dengue incidence patterns through regression analysis [17], ANN [18], fuzzy association rule mining [19], and many descriptive studies [20]. The sophisticated methods such as a combination of flower pollination algorithm and least squares support vector machines (FPA-LSSVM) [21] and ANFIS with various parameter tuning methods [22] are applied to forecast dengue incidence in Indonesia. Modeling dengue incidence in Thailand includes a wide range of techniques such as exploratory analysis through the remote sensing GIS [23], Poisson autoregressive [24], cellular automata with the inclusion of genetic algorithm for chromosome weighting and ANN for determining probability in state change [25], SVR [26], Bayesian network [27], and generalized additive model (GAM) [28].

Besides the Southeast Asia, dengue has become a public health threat in more than hundred countries worldwide. Such global pandemic urgently requires collaboration among researchers and practitioners at both regional and global scales. The awareness regarding an attempt to monitor and control dengue in Taiwan can be noticed through many proposals adopting various state-of-the-art machine intelligence techniques including fuzzy inference system [29], multivariate logistic regression analysis [30], and decision tree induction [31]. To forecast dengue outbreak in southern China, researchers apply GAM [32], SVR as well as variations of regression modeling [33], and the factor-detector modeling module of the ArgGIS software [34]. Researchers in Sri Lanka adopt fuzzy system with trapezoidal membership function to predict the developing from classic dengue fever to the severe level using clinical evidences as input data [35].

In Latin America that dengue has spread massively during this decade, researchers apply numerous machine learning techniques to build an efficient predictive model. The applied techniques include SVR and k-nearest neighbor regression to predict cases in the northwest city of Argentina [36]. SVR and ANN are adopted to predict dengue infection in Paraguay and ANN shows the best result [37]. In Brazil which confronts the most serious dengue outbreak in 2019, GAM [38] and Pearson correlation analysis [39] are the main methods applied to predict dengue cases in Sao Paulo, whereas nearest-neighbor regression analysis [40] is the technique that researchers use for predicting cases in Rio de Janeiro. To project the spread of dengue virus globally towards the year 2020, 2050 and 2080, an international team of researchers [41] has applied the ensemble of boosted regression tree to show the trend

in dengue transmission that around 60% of the global population living in Asia, Africa and Americas, in descending order, are at risk of infection.

It can be concluded from the literature review that dengue modeling techniques based on the mature machine learning methodology include regression analysis, ANN, ANFIS, SVR, GAM, and regression tree. We thus apply these main methods to derive prediction models to estimate number of dengue infection in the capital and also the most populous city of Thailand. The main focus is to observe performances of these popular modeling methods, and then select the most accurate model for predicting dengue outbreak in Bangkok.

### 3. Data and Modeling Method

#### 3.1. Data Sources

In our modeling process, the target for prediction is number of patients in Bangkok metropolitan area that had been diagnosed as infected with dengue virus. These historical data are monthly dengue cases in Bangkok during the years 2003-2017. Data between 2003-2015 are used for training the models, whilst data of 2016-2017 are held out for testing performance of the models.

Independent variables to be used as predictors are time of the year represented by number of month (1 to 12) and satellite-based data obtained from the National Oceanic and Atmospheric Administration (NOAA), U.S.A. Rainfall data are from the Centers for Environmental Information [42]. Satellite indices to reflect vegetation health are obtained from the NOAA's global vegetation health products [43]. The indices to be used as predictors in this research include smoothed and normalized difference vegetation index (SMN, or greenness), smoothed brightness temperature index (SMT, a thermal condition), vegetation condition index (VCI, a proxy for greenness anomaly), temperature condition index (TCI), and vegetation health index (VHI, a combined moisture and thermal conditions).

Another satellite-based index to reflect climate variability is also used in our modeling process. This climate index is Oceanic Niño Index (ONI) that can be accessed from the NOAA's Weather Service [44]. The ONI is computed from the anomaly of sea surface temperature (SST) at the Niño3.4 region in the Pacific Ocean at the equator line near Indonesia and north Australia. The SST is monthly collected to compute ONI by averaging from the 3-month periods: the current one, the previous, and the following month. This 3-month average value is then compared to a thirty-year average to check temperature deviation. If the ONI values in five consecutive months are +0.5 °C or higher than the normal value, then the El Niño (warm event) phenomenon is announced. On the contrary, if the ONI values in five consecutive periods are -0.5 °C or lower than the normal value, the opposite phenomenon called La Niña (the cold event) is announced.

#### 3.2. A Framework of Modeling Method

Main steps in our data modeling process are shown in Figure 1. In data preparation and integration step, remote sensing indices have to be transformed from weekly to be monthly by averaging weekly values of each month. We then select only the attributes contributing the most to accurate prediction in step 2. This is done by evaluating predictor importance value. Step 3 is the train-test

separation. Training data are used in the model building step. Test data are used in the model evaluation step.

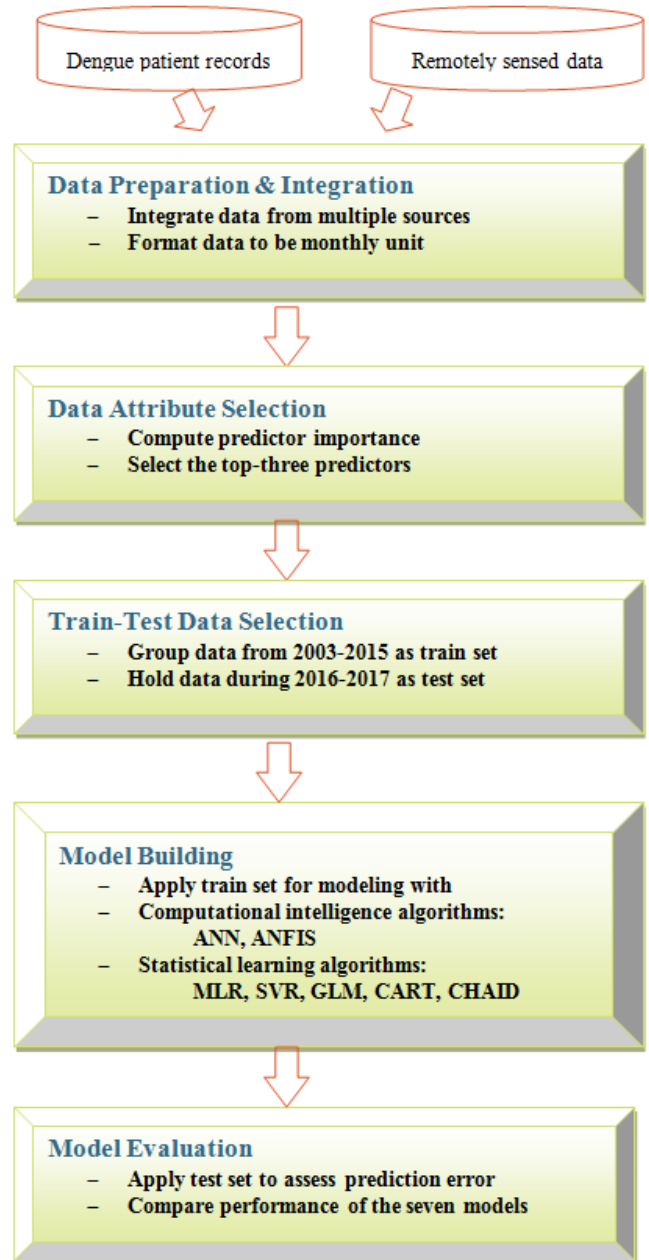


Figure 1: Steps in Data Modeling for Dengue Incidence Prediction

### 4. Performance Comparison Results

We perform experiments following the steps shown in Figure 1. Software tool used for running ANFIS algorithm is MATLAB, whereas other learning algorithms are run with IBM SPSS Modeler. All algorithms are tested with the same set of testing data that are unseen by the algorithms during the training phase to build model.

Data attributes used as model predictors to estimate number of dengue cases in Bangkok include the satellite-based indices (ONI, SMN, SMT, VCI, TCI, VHI), amount of rainfall, and number of month (1 to 12, in which 1 is January and 12 is December) in each year. Representation of each satellite-based index is as follows:

ONI -- climate variability, SMN -- surface greenness, SMT -- thermal condition, VCI -- greenness anomaly, TCI -- temperature condition, VHI -- moisture and thermal conditions. Number of month can represent season in that months 5 to 8 are rainy season in Thailand. Correlations among the eight predictors are presented in Table 1. Strong correlations are highlighted with bold font.

Prior to building the model with training data set, we analyze importance of each predictor to select only the top-three ones contributing most toward the final predictive model. To compute the importance value [45], each independent variable is assessed through sensitivity analysis to observe the reduction in variance of the target field, which is the dengue incidence in this specific domain. The computed sensitivity values are then ranked to evaluate importance of each predictor; the higher the value, the more important the variable. The ranking result of all seven predictors is shown in Figure 2. The attribute SMT is missing from the figure because the coefficient of variation in this attribute is below the threshold. The top-three variables, which are ONI,

month of the year, and SMN as a proxy of greenness, are then selected as important features to be used further in the modeling phase.

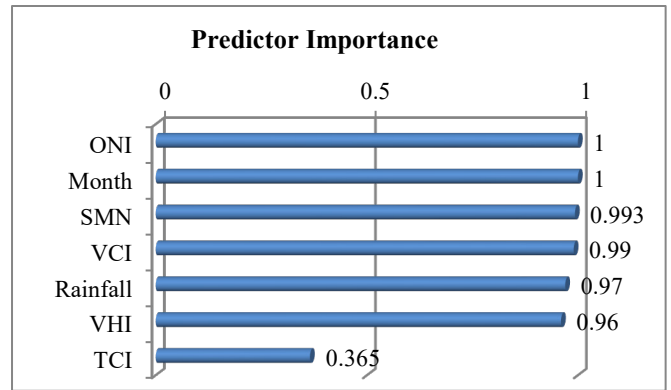


Figure 2: Importance Value of Independent Predictors

Table 1: Pearson Correlation among Predictor Attributes

	ONI	SMN	SMT	VCI	TCI	VHI	Month
<b>SMN</b>	0.078						
<b>SMT</b>	0.118	<b>0.735</b>					
<b>VCI</b>	0.135	<b>0.656</b>	0.152				
<b>TCI</b>	<b>-0.362</b>	-0.107	-0.143	<b>-0.331</b>			
<b>VHI</b>	<b>-0.282</b>	<b>0.304</b>	-0.046	<b>0.289</b>	<b>0.807</b>		
<b>Month</b>	-0.046	<b>-0.251</b>	<b>-0.460</b>	0.020	<b>-0.287</b>	<b>-0.281</b>	
<b>Rainfall</b>	-0.115	0.120	<b>0.398</b>	-0.131	<b>0.236</b>	0.161	<b>-0.439</b>

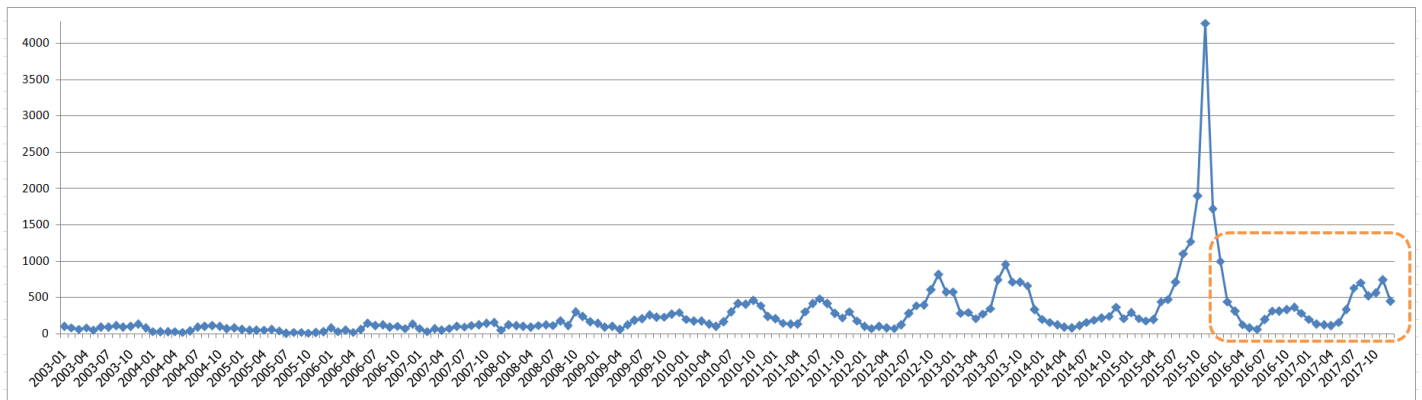


Figure 3: Distribution of Dengue Infection in Bangkok During 2003 to 2017 in which 2016-2017 are Test Data

Table 2: Performance of Learning Algorithms

Algorithm	Mean Absolute Error (MAE)	Root Mean Square Error (RMSE)	Model Correlation	Model Building Time (seconds)
<b>Computation Intelligence</b>				
ANFIS	151.511	216.536	0.835	4
ANN	211.296	308.277	0.827	8
<b>Statistical Learning</b>				
MLR	200.681	236.613	0.501	4
GLM (log function)	217.706	280.076	0.501	4
CHAID	217.018	292.826	0.515	3
SVR	236.732	325.958	0.501	4
CART	278.939	381.517	0.815	3



After data attribute selection step, we prepare two subsets of data: training data and test data. Training data are the records during the years 2003-2015, containing 156 records. Test data are cases from 2016-2017, comprising of 24 records. Dengue case distribution in Bangkok between 2003-2017 is displayed in Figure 3. Vertical axis is the number of dengue infected patients. There exists an extreme outbreak in November 2015, which is the month that ONI value is as high as 2.5. ONI normally ranges between +2.5 to -2.5. The ONI value of -2.5 represent the strong El Nino event that makes the weather in Bangkok hotter and drier than normal years. The dashed area in the figure marks distribution of dengue cases in the test data.

The training data are used as input for each of the learning algorithms to build predictive model. The models are then tested for their performance using the out-of-sample test data of the years 2016-2017. Performance evaluation results are illustrated in Table 2. The measurements for performance evaluation of each model include the mean absolute error (MAE), root mean square error (RMSE), correlation between predictors and the target of the model, and time to build model. Consider from predictive errors, ANFIS shows the best performance with the minimum values in both error metrics. MLR shows the second best performance with errors a little bit higher than ANFIS. On the metric of model correlation, ANFIS is also the best one, while ANN comes as the second best model and CART is in the third place. CART and CHAID, which are the tree-based modeling methods, take less time to build model than the other methods. Based on the time aspect, ANN is the worst one.

Figure 4 shows comparative performance of the seven models based on four measurements in which the scales are normalized to be between 0 to 1. The three measurements (MAE, RMSE, time) seek for the lowest one to be the best, while correlation requires the highest one to outperform the others. Therefore, we transform the normalized correlation metric to be the reverse one by applying the formula  $(1 - \text{normalized\_correlation})$  to make this metric in accordance with the others. Thus, model with all measurements closest to zero is the best one. From this criteria, we can rank the models from the best to the worst ones as follows: (1) ANFIS, (2) MLR, (3) CHAID, (4) ANN, (5) CART, (6) GLM, (7) SVR.

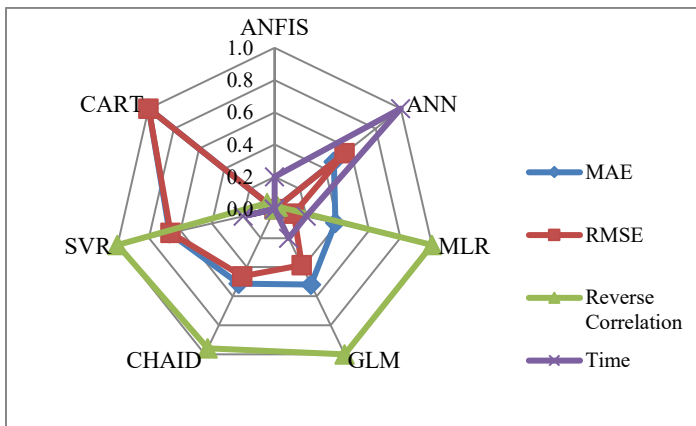


Figure 4: Comparison of Model Performance using Normalized Measurements

The objective of our modeling process is to predict the number of dengue cases in a specific region. It is thus a macro analysis aiming at dengue outbreak monitoring rather than a micro scope of

classifying patients as infected with dengue virus or not. Moreover, our target variable is continuous data. Therefore, traditional measurements such as sensitivity, specificity, F-measure, recall, precision, and so on, cannot be applied. The most suitable metric is the error made by the model to predict number of dengue incidence. The plot in Figure 5 shows the difference between actual number of dengue cases (circle) against the predicted number of cases forecasted by ANFIS (star). It can be noticed that at extreme level of dengue infected cases higher than 700, ANFIS forecasts quite inaccurate. For the average level of infection around 100 to 300 cases, ANFIS can make a good prediction.

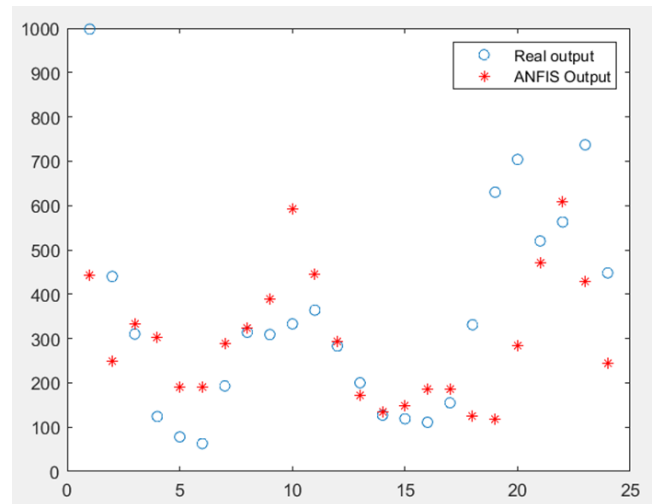


Figure 5: Predicting Performance of ANFIS on Test Data

## 5. Conclusion

Dengue is a fatal communicable viral disease threatening life of population in Thailand as well as many countries in the tropical zone with warm and humid weather providing a good condition for the spread of *Ae. aegypti* and *Ae. albopictus* mosquitoes that are the main carriers of dengue virus. The outbreak of dengue is seasonal in that infection is abnormally high in every two or three years. Many researchers observed that this spasmodic infection is coincident to the rise of sea surface temperature in the equatorial Pacific region. We thus include Oceanic Niño Index (ONI) as one main factor to derive a model to predict dengue cases in Bangkok. Other predictors used in our modeling phase are smoothed and normalized difference vegetation index (SMN), which is a proxy for vegetation greenness, and month of the year (1 to 12) that is used to capture season as dengue fever in Thailand normally widespread during months 6 to 11. The satellite features (SMN and ONI) and month attribute are selected from the original seven variables through the computation of predictor importance to select only the top-three features showing the best variance reduction against the target attribute, which is the number of dengue infected patients.

The selected features (ONI, SMN, month) in addition to the target field (dengue cases) are then used for the next step of data-driven model building. Modeling algorithms applied in this work are of two categories: computational intelligence and statistical learning. Computational intelligence algorithms are ANN and ANFIS. The other five algorithms (multiple linear regression, generalized linear model, support vector regression, regression

trees -- CART and CHAID) are in a group of statistical learning. We evaluate performance of all algorithms using the same set of out-of-sample data for a fair comparison. The experimental results reveal that ANFIS is the best predictor model. Multiple linear regression is the second best one. In the future, we plan to further our empirical study to cover every province of Thailand.

**Conflict of Interest**

The authors declare no conflict of interest.

**Acknowledgment**

This work has been supported by grant from Suranaree University of Technology through the funding of Data and Knowledge Engineering Research Unit in which the first and second authors are principal researchers.

**References**

[1] WHO, World Health Organization, "Dengue and severe dengue" Fact Sheet, updated 4 November 2019. <https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue>

[2] CDC, Centers for Disease Control and Prevention, "Dengue" Access 10 May 2020. <https://cdc.gov/dengue/about/index.html>

[3] N.E.A. Murray, M.B. Quam, A. Wilder-Smith, "Epidemiology of dengue: past, present and future prospects" *Clinical Epidemiology*, 5, 299-309, 2013.

[4] N. Gyawali, R.S. Bradbury, A.W. Taylor-Robinson, "The epidemiology of dengue infection: harnessing past experience and current knowledge to support implementation of future control strategies" *Journal of Vector Borne Diseases*, 53, 293-304, 2016.

[5] ECDC, European Centre for Disease Prevention and Control, "Dengue worldwide overview" Access 21 April 2020. <https://ecdc.europa/en/dengue-monthly>

[6] S. Bandyopadhyay, L. Lum, A. Kroeger, "Classifying dengue: a review of the difficulties in using the WHO case classification for dengue haemorrhagic fever" *Tropical Medicine & International Health*, 11(8), 1238-1255, 2006.

[7] S.R.S. Hadinegoro, "The revised WHO dengue classification: does the system need to be modified?" *Paediatrics and International Child Health*, 32(S1), 33-38, 2012.

[8] B.A. Ajlan, M.M. Alafif, M.M. Alawi, N.A. Akbar, E.K. Aldigs, T.A. Madani, "Assessment of the new World Health Organization's dengue classification for predicting severity of illness and level of healthcare required" *PLoS Neglected Tropical Diseases*, 13(8), e0007144, 2019. <https://doi.org/10.1371/journal.pntd.0007144>

[9] Vector Borne Disease Division, Ministry of Public Health, "Dengue infection (in Thai)" Access 10 May 2020. <https://ddc.moph.go.th/index.php>

[10] B. Loh, R.J. Song, "Modeling dengue cluster size as a function of *Aedes aegypti* population and climate in Singapore" *Dengue Bulletin*, 25, 74-78, 2001.

[11] Y. Hii, J. Rocklöv, N. Ng, C. Tang, F. Pang, R. Sauerborn, "Climate variability and increase in intensity and magnitude of dengue incidence in Singapore" *Global Health Action*, 2, 124-132, 2009.

[12] H.M. Aburas, B.G. Cetiner, M. Sari, "Dengue confirmed-cases prediction: a neural network model" *Expert Systems with Applications*, 37, 4256-4260, 2010.

[13] N.C. Dom, A.H. Ahmad, Z.A. Latif, R. Ismail, B. Pradhan, "Coupling of remote sensing data and environmental-related parameters for dengue transmission risk assessment in Subang Jaya, Malaysia" *Geocarto International*, 28(3), 258-272, 2013.

[14] T. Faisal, M.N. Taib, F. Ibrahim, "Adaptive neuro-fuzzy inference system for diagnosis risk in dengue patients" *Expert Systems with Applications*, 39, 4483-4495, 2012.

[15] F. Ibrahim, T.H.G. Thio, To Faisal, M. Neuman, "The application of biomedical engineering techniques to the diagnosis and management of tropical diseases: a review" *Sensors*, 15, 6947-6995, 2015.

[16] N.A. Husin, N. Mustapha, M.N. Sulaiman, R. Yaacob, H. Hamdan, M. Hussin, "Performance of hybrid GANN in comparison with other standalone models on dengue outbreak prediction" *Journal of Computer Sciences*, 12(6), 300-306, 2016.

[17] G.L. Su, "Correlation of climatic factors and dengue incidence in metro Manila, Philippines" *AMBIO: A Journal of the Human Environment*, 37(4), 292-294, 2008.

[18] H.I. Datoc, R. Caparas, J. Caro, "Forecasting and data visualization of dengue spread in the Philippines Visayas island group" in 7th Int. Conf. on Information, Intelligence, System & Applications (IISA), Chalkidiki, Greece, 2016. <https://doi.org/10.1109/IISA.2016.7785420>

[19] A.L. Buczak, B. Baugher, S.M. Babin, L.C. Ramac-Thomas, E. Guven, Y. Elbert, P.T. Koshute, J.M.S. Velasco, V.G. Roque Jr, E.A. Tayag et al., "Prediction of high incidence of dengue in the Philippines" *PLoS Neglected Tropical Diseases*, 8(4), e2771, 2014. <https://doi.org/10.1371/journal.pntd.0002771>

[20] K.A. Agrupis, M. Ylade, J. Aldaba, A.L. Lopez, J. Deen, "Trends in dengue research in the Philippines: a systematic review" *PLoS Neglected Tropical Diseases*, 13(4), e0007280, 2019. <https://doi.org/10.1371/journal.pntd.0007280>

[21] Z. Mustafa, M.H. Sulaiman, F. Ernawan, Y. Yusof, M.F.M. Mohsin, "Dengue outbreak prediction: hybrid meta-heuristic model" in 19th IEEE/ACIS Int. Conf. on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), Busan, South Korea, pp. 271-274, 2018.

[22] B. Rachmat, O.D. Nurhayati, Suryono, "Prediction the number of patients at dengue hemorrhagic fever cases using adaptive neural fuzzy inference system (ANFIS)" *International Journal of Innovative Research in Advanced Engineering*, 3(4), 23-28, 2017.

[23] A. Chaikoolvatana, P. Singhasivanon, P. Haddawy, "Utilization of a geographical information system for surveillance of *Aedes aegypti* and dengue haemorrhagic fever in north-eastern Thailand" *Dengue Bulletin*, 31, 75-82, 2007.

[24] M. Tipayamongkhogul, C. Fang, S. Klinchan, C. Liu, C. King, "Effects of the El Niño-Southern Oscillation on dengue epidemics in Thailand, 1996-2005" *BMC Public Health*, 9, article 422, 2009. <https://doi.org/10.1186/1471-2458-9-422>

[25] T. Soemsap, S. Wongthanavas, W. Satimai, "Forecasting number of dengue patients using cellular automata model" in 2014 Int. Electrical Engineering Congress (IIECON), Chonburi, Thailand, 2014. <https://doi.org/10.1109/IIECON.2014.6925876>

[26] N. Sumanasinghe, A.R. Mikler, J. Muthukudage, C. Tiwari, R. Quiroz, "Data driven prediction of dengue incidence in Thailand" in Int. Conf. on Computing and Information Technology (IC2IT 2017), Bangkok, Thailand, pp. 95-107, 2017.

[27] C. Ruangudomsakul, A. Duangsin, K. Kerdprasop, N. Kerdprasop, "Application of remote sensing data for dengue outbreak estimation using Bayesian network" *International Journal of Machine Learning and Computing*, 8(4), 394-398, 2018.

[28] R. Jain, S. Sontisirikit, S. Iamsirithaworn, H. Prendinger, "Prediction of dengue outbreaks based on disease surveillance, meteorological and socio-economic data" *BMC Infectious Diseases*, 19, article 272, 2019. <https://doi.org/10.1186/s12879-019-3874-x>

[29] D. Saikia, J.D. Dutta, "Early diagnosis of dengue disease using fuzzy inference system" in 2016 Int. Conf. on Microelectronics, Computing and Communications (MicroCom), Durgapur, India, 2016. <https://doi.org/10.1109/MicroCom.2016.7522513>

[30] C. Huang, C. Hsu, H. Guo, S. Su, H. Lin, "Dengue fever mortality score: a novel decision rule to predict death from dengue fever" *Journal of Infection*, 75, 532-540, 2017.

[31] C. Wu, S. Kao, C. Shih, M. Kan, "Open data mining for Taiwan's dengue epidemic" *Acta Tropica*, 183, 1-7, 2018.

[32] X. Qi, Y. Wang, Y. Li, Y. Meng, Q. Chen, J. Ma, G.F. Gao, "The effects of socioeconomic and environmental factors on the incidence of dengue fever in the Pearl River Delta, China, 2013" *PLoS Neglected Tropical Diseases*, 9(10), e0004159, 2015. <https://doi.org/10.1371/journal.pntd.0004159>

[33] P. Guo, T. Liu, Q. Zhang, L. Wang, J. Xiao, Q. Zhang, G. Luo, Z. Li, J. He, Y. Zhang, W. Ma, "Developing a dengue forecast model using machine learning: a case study in China" *PLoS Neglected Tropical Diseases*, 11(10), e0005973, 2017. <https://doi.org/10.1371/journal.pntd.0005973>

[34] Y. Chen, Z. Zhao, Z. Li, W. Li, Z. Li, R. Guo, Z. Yuan, "Spatiotemporal transmission patterns and determinants of dengue fever: a case study of Guangzhou, China" *International Journal of Environmental Research and Public Health*, 16(14), article 2486, 2019. <https://doi.org/10.3390/ijerph16142486>

[35] S.D.P. Jayasundara, S.S.N. Perera, G.N. Malavige, S. Jayasinghe, "Mathematical modelling and a systems science approach to describe the role of cytokines in the evolution of severe dengue" *BMC Systems Biology*, 11, article 34, 2017. <https://doi.org/10.1186/s12918-017-0415-3>

[36] J.M. Scavuzzo, F. Trucco, M. Espinosa, C.B. Tauro, M. Abril, C.M. Scavuzzo, A.C. Frery, "Modeling dengue vector population using remotely sensed data and machine learning" *Acta Tropica*, 185, 167-175, 2018.

[37] J.D. Mello-Roman, J.C. Mello-Roman, S. Gomez-Guerrero, M. Garcia-Torres, "Predictive models for the medical diagnosis of dengue: a case study in Paraguay" *Computational and Mathematical Methods in Medicine*, 2019, article 7307803, 2019. <https://doi.org/10.1155/2019/7307803>

- [38] E.C. Farinelli, O.S. Baquero, C. Stephan, F. Chiaravalloti-Neto, "Low socioeconomic condition and the risk of dengue fever: a direct relationship" *Acta Tropica*, 180, 47-57, 2018.
- [39] I. Ogashawara, L. Li, M.J. Moreno-Madrinan, "Spatial-temporal assessment of environmental factors related to dengue outbreaks in Sao Paulo, Brazil" *GeoHealth*, 3(8), 202-217, 2019.
- [40] A.P.R. Dalvi, J.U. Braga, "Spatial diffusion of the 2015-2016 Zika, dengue and chikungunya epidemics in Rio de Janeiro Municipality, Brazil" *Epidemiology & Infection*, 147, article e237, 2019. <https://doi.org/10.1017/s0950268819001250>
- [41] J.P. Messina, O.J. Brady, N. Golding, M.U.G. Kraemer, G.R. William Wint, S.E. Ray, D.M. Pigott, F.M. Shearer, K. Johnson, L. Earl et al., "The current and future global distribution and population at risk of dengue" *Nature Microbiology*, 4(9), 1508-1515, 2019.
- [42] NOAA, National Centers for Environmental Information, "Climate Data Online" Access 11 January 2020. <https://www.ncdc.noaa.gov/cdo-web/>
- [43] NOAA STAR, Center for Satellite Applications and Research, "Global Vegetation Health Products" Access 11 January 2020. [https://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh\\_browseByCountry\\_province.php](https://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_browseByCountry_province.php)
- [44] NOAA National Weather Service, Climate Prediction Center, "Cold & Warm Episodes by Season" Access 11 January 2020. [http://origin.cpc.ncep.noaa.gov/products/analysis\\_monitoring/ensostuff/ONI\\_v5.php](http://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php)
- [45] IBM, IBM SPSS Modeler 17 Algorithms Guide, IBM Corp, 2015.

## Design and Optimization of a Three Stage Electromechanical Power Unit using Numerical Methods

Yashwant Kolluru<sup>1,\*</sup>, Rolando Doelling<sup>1</sup>, Lars Hedrich<sup>2</sup>

<sup>1</sup>Robert Bosch GmbH, eBikes Department, Kusterdingen, 72770, Germany

<sup>2</sup>Johann Wolfgang Goethe-University, Institute of Informatics, Frankfurt am Main, 60054, Germany

### ARTICLE INFO

Article history:

Received: 15 May, 2020

Accepted: 17 July, 2020

Online: 28 July, 2020

Keywords:

Vibro-acoustics

Numerical Methods

Compliance

Stress

### ABSTRACT

*The advent of electric vehicles has changed the face of the automobile industry. The drive system properties of vehicles such as eBikes or electric cars differ fundamentally from those of a diesel engine. The lack of a conventional internal combustion engine has made the vehicles considerably silent. Nevertheless, previously hidden sources of vibration and noise have become more dominant. In addition to these emissions, other structural properties such as compliance and deformation also appear as relevant factors for the original equipment manufacturer. Usually, deterioration of these variables affects the efficiency of the power unit. In this paper, a simulation template is created to understand and analyze these properties of the drive unit. Furthermore, new enhancements to improve the key indicators, such as strain energy, natural frequencies, etc., are shown, thereby creating a potential method flow to develop better performing drive units. Numerical optimization tools are used to simulate structures with complex shapes that exactly meet the mechanical constraints and use as little material as possible. In this work, two optimized variants of electromechanical drives are presented. The first scenario illustrates the optimized model with an objective of minimizing the strain energy of the structures, whereas the second task aids in the development of a variant with superior dynamic properties than the current drive units. Ultimately, several numerical calculations are validated using experiments.*

### 1. Introduction

This paper is an extension of work that was originally presented in ISSE 2019 conference [1]. Originally, paper [1] focused on creation of a multi-body model for analyzing the vibro-acoustic characters inside the eBike. The current work illustrates the extension of the previous setups to the optimization domain using the finite element techniques.

In the current pedelec industry, the core problems faced by manufacturers are the structural behaviors, such as vibration and deformation, due to loading. Lately, researchers have mainly focused on the development of the electric motors, batteries and power electronics, which has laid a solid foundation for building advanced electric powertrains [2]-[4]. Improving the energy efficiency of the entire electromechanical drives was the essential parameter, while areas related to vehicle dynamics and structural properties, such as vibration, noise, stiffness, compliance, etc., are not extensively examined. However, several findings show that the

vibrations of electric drives can lead to several problems in regard to stability and driving comfort of the vehicle [5, 6]. In addition, the presence of stresses in structures beyond the yield limits leads to larger deflections and thereby reduce the efficiency of the components [7]. Understanding the static and dynamic behavior of the electromechanical drives therefore remains an essential task for the Original Equipment Manufacturers (OEMs). This paper attempts to develop a method that can be used to capture structural behavior early in the product to avoid unnecessary costs in the later stages of the product lifecycle. The method was developed largely with the help of the Finite Element Modeling techniques (FEM).

The numerical discretization of known structures is a popular topic in computational mechanics. The end user uses the software iteratively to optimize a structure. It is up to him/her to improve the design in order to identify a good geometry [8]. In contrast, topology optimization software outputs the optimal geometry using structured mesh, directly after the optimization task has been defined. Structural topology optimization helps the designer in modeling the type of structure which best fits to satisfy the operating conditions for the problem [9]. It changes the general

\*Corresponding Author: Yashwant Kolluru, Gerhard-Kindler-Str. 3, Kusterdingen 72770, Germany. +49-17656830708, yashwant.kolluru@de.bosch.com



component layout to find designs with goals such as minimization compliance or maximization of natural frequencies. Stress and compliance optimization are key indicators in the structural design for a range of engineering applications. Optimization helps in the design of mechanical components with lower stress concentrations [10]. Once the results are extracted from the optimization module, they are enhanced using the smoothing functions provided by the optimization tool. It can be seen as a procedure to optimize the rational arrangement of the available elements in the design volume and to eliminate the unnecessary elements.

In a similar way, the structure borne emission from housings must also be optimized. Vibrations are undesirable movements in the powertrain and are responsible for the problematic noise and energy losses in the system. In general, they are generated from excitations at structural components and would later affect the overall performance of the Drive Unit (DU). In order to maintain the efficiency of the drive, vibration and noise emission should be reduced below the threshold values. One way of achieving this is through the dynamic optimization technique [11]. As part of this approach, the simulation template is created to analyze the surface acceleration and acoustic quantities. Later, a dynamic optimization module is developed to strengthen the structure and reduce the vibrations in the drivetrain.

## 2. State of the Art

### 2.1. Stress Analysis in Drivetrain Components

Structural analysis is generally performed to check the durability of certain parts for a particular load and support condition. For the component to be structurally safe, stresses in the body should not be more than the yield strength of the material. Considering the possibility of fatigue failure, the component must be optimized such that the Von-mises stress generated does not exceed the endurance strength of the material. There are several articles that generally describe the stress and deformation of various components such as gears, motors, etc. For instance, paper [12] written by Mao discusses the importance of transmission error in gears and its influence in powertrain structures. Similarly, Chung published his results in paper [13] that focused on stress analysis of helical gears and their deformation profile for different load cases in normal drive trains. Electromechanical drives have distinct structural characteristics. Most of these papers principally illustrate the gear forces in static load cases or influences in the performance of normal drivetrains due to errors at the gears. However, there are no proper references available for analyzing electromechanical drives as entire unit, for example, considering the influences of multiple domains such as axial and electromagnetic forces, static and linear dynamic cases together, or the impact of excitations, forces at the bearings and the structural behavior of the housing combined. Moreover, adding optimization tasks makes the existing scenario more complex. This paper tries to solve this problem by examining and simulating each juncture responsible for deformations, critical stresses etc. In this way, a holistic simulation template is provided with which the entire drivetrain can be analyzed and a future variant with optimized properties can be developed.

### 2.2. Modal Analysis of Complex Structures

Vibrations are nothing but mechanical oscillations of a system/component at an equilibrium point. A statically balanced

object oscillates at a certain frequency depending on its properties such as mass and stiffness. This frequency is referred to as the natural frequency of that structure [14]. The calculation of these frequencies and their respective mode shapes is essential for researchers to understand the vibro-acoustic properties of the components. There are usually two sets of models used to analyze the vibrations of structures, continuous and discrete. The continuous models are mathematical models applied to continuous data and are represented by the Partial Differential Equations (PDE). The coefficients of these PDE's depend on the geometrical and material properties. Solutions to these equations are obtained using numerical methods such as Dunkerley's method, Rayleigh-Ritz method, Newton-Raphson method, Galerkin approach, etc., [15 - 17]. The discrete models are those models with finite degrees of freedom and are generally described by ordinary differential equations. Two main techniques that deal with the discrete models are finite element and rigid element methods.

The structure of the DU is complex and consists of more than 60 components and 800000 Degrees Of Freedom (DOF), which makes it difficult to analyze either with the continuous models or the rigid element method. FEM is used as a simulation technique considering the factors such as complexity (large number of DOF's), modeling (more flexibility to include the effects like contact, hybrid modeling, etc.), adaptability and accuracy (better precision) for calculating the vibrations. FEM generates the physical response of the system at any given position, including some that may have been neglected in an analytical approach. In addition, FEM has an advantage of modeling noise using acoustic finite, infinite elements and the impedance boundaries [18]. The Eigen forms are calculated using the solution technique described by Grimes, Lewis, and Simon in paper [19].

### 2.3. Relevance of Optimization

Researchers have developed numerous solutions with regard to the structural topology optimization problems. Below is a list of few approaches developed within the topology optimization domain [8]:

- Ground structure approach discussed in the Shiehin's paper [20].
- Homogenization method described by the Bendsoe and Suzuki in papers [11, 21], respectively.
- Bubble method shown by the Eschenauer's paper [9].
- Fully stressed design technique depicted in Xiein's paper [10].

The first three approaches listed above have some characteristics in common. These are optimization methods with an objective function and a set of design variables and design constraints. They solve the optimization problem, either by a sequential, quadratic programming algorithm (approach 1) or by an optimality criterion concept (approaches 2 and 3). In contrast, the fully stressed design technique, although not an optimization algorithm in the conventional sense, proceeds by removing inefficient material and thereby optimizes the use of the remaining material in the structure, in a step-by-step process.

Within this paper, the methods majorly from approach 1 are used as guidelines for the design of drivetrain structures to improve the stiffness, dynamic characteristics and to reduce the weight of the components. In addition, the optimization design can be

regarded as the topology optimization problem of continuum structures. Structural optimization attempts to find the best path for transferring all types of loads.

This paper deals with the optimization of drivetrain structures for both static and dynamic load cases. Using a topology optimization approach to design new drivetrains should take less time than the traditional trial-and-error experiments.

#### 2.4. Definition of the Optimization Task

Optimization can be used to shorten the development phase by upgrading insight and intuition of the designers through an automated process. To optimize the powertrain model, the variable that needs to be optimized must first be determined. Optimizations to minimize stress or to maximize the Eigen frequency are simply not sufficient. The task must be more specific, for example minimizing the maximum node stresses that occur during a static load case with a limitation of the existing weight, or similarly maximizing the sum of the first natural frequencies with limit on volume of the components. The goal of the optimization task is referred to as the objective function. In addition, certain conditions can be enforced during the optimization, for instance, the displacement of a given node can be curtailed to a certain value. This enforced condition is often called a constraint.

In addition, the structural domain called the reference domain can be divided into the design and non-design areas. The non-design domain includes key regions such as supports, boundaries and areas where loads are applied. Therefore, these regions cannot be modified throughout the entire optimization process. With the constraint on weight, objective functions, such as optimizing stress, stiffness, etc., are used for the optimization of the static load case. In linear dynamic analysis, the visco-elastic material definitions are used to simulate the drivetrain. Here, the goal of optimization is to curtail the surface vibrations and acoustic noise below the limits.

Overall an optimization task is represented using three characteristics. Below is a list of the characteristics, including concerning questions:

- Design variables - How can the structure be modified? Values, such as frequencies, weight, volume and stiffness, can be altered to obtain better vibrations.
- Design constraints – Which restrictions must be implemented in the model? Variables, such as mass and volume, can be limited to certain ranges.
- Design objective functions - What should be minimized, maximized, optimized, etc.? Function of compliance with a goal in increasing the stiffness and reducing the stress. Function of natural frequencies (variable) to push the critical spectrum domains, thereby avoiding vibrations.

### 2.5. Mathematical formulations

#### 2.5.1. Gearmesh Excitations

The parametric formulation for gears characterized as lumped mass is shown in (1) (1 – gear1, 2 –gear2). It represents the dynamic Equation of Motion (EOM) of one of the gear with to gearmesh excitations. Here the term  $e(\psi(t))$  includes the effects of

parametric excitation. It is developed as sum of mean excitation plus sine curves (considering the harmonics of the tooth cycle) with corresponding amplitudes.  $\psi_i(t)$  ( $i = 1, i = 2 - \text{gear2}$ ) and its time derivatives represent the rotation angle, angular velocity and angular acceleration respectively.  $M, J, r, z, \psi, e, k_z$  and  $d_z$  represent torque, inertia, radius of pitch circle, number of tooth, rotation angle, excitation, stiffness and damping respectively [1].

$$J_1 \ddot{\psi}_1 = M_1(t) - r_1 k_z (\psi_1(t)) [r_1 \psi_1(t) + r_2 \psi_2(t) + e(\psi_1(t))] - r_1 d_z (\dot{\psi}_1(t)) [r_1 \dot{\psi}_1(t) + r_2 \dot{\psi}_2(t) + \dot{e}(\psi_1(t))] \quad (1)$$

Furthermore, the gear forces are described in (2), (3) and (4). Here,  $\bar{F}$  consists of three components, namely the tangential, radial and axial force of the gear.  $\bar{F}_n, T_\beta, T_{an}, c$  and  $s$  describe the normal force, rotation matrix for the helix angle, rotation matrix for the normal pressure angle, cosine angle and sine angle, respectively [1]. These gear forces cause the shafts and bearings to vibrate, which in turn sets the structure borne emission in rest structures.

$$T_{an} = \begin{bmatrix} c(an) & s(an) & 0 \\ -s(an) & c(an) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$$T_\beta = \begin{bmatrix} c(\beta) & 0 & -s(\beta) \\ 0 & 1 & 0 \\ s(\beta) & 0 & c(\beta) \end{bmatrix} \quad (3)$$

$$\bar{F} = T_\beta T_{an} \bar{F}_n \quad (4)$$

#### 2.5.2. Fluid Structure Coupling

The housing-air coupling occurs at the interface at which the acoustic medium interacts with the structure. This is called the fluid structure interface. This coupling between the housing structure and air medium is modeled by conservation of linear momentum.

Equation (5) shows the general formulation of the housing and surrounding fluid at its interface. The terms  $u, M, M_f, M_{fs}, K, K_f, K_{fs}, F_s$  and  $F_a$  indicate the displacement of housing, mass of the housing, mass of the fluid/ air, mass of the fluid structure interface, stiffness of the structure, stiffness matrix of fluid, stiffness matrix of fluid structure interface, force applied by the structure and force exerted by the air respectively. Detailed information about the coupling finite element models can be found in papers [22, 23].

$$\begin{bmatrix} M & 0 \\ M_{fs} & M_f \end{bmatrix} \begin{bmatrix} \ddot{u} \\ \ddot{p} \end{bmatrix} + \begin{bmatrix} K & K_{fs} \\ 0 & K_f \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} F_s \\ F_a \end{bmatrix} \quad (5)$$

#### 2.5.3. Optimization formulations

Furthermore, (6) represents the mathematical form of the optimization task. The objectives are defined as minimizing the compliance (C) and the surface accelerations (A). Volume is considered as a design constraint (in this example) and is limited to  $V_{max}$ .  $\rho$  is a density vector containing the element densities  $\rho_e$ . When  $\rho_e$  is 1 we consider an element to be filled whereas an element with  $\rho_e$  as 0 is considered to be a void element.  $\Omega$  is considered as the design volume of the optimization task. The volume constraint prevents the optimized structure from ending with the full design volume when searching for its maximum structural stiffness.

$$\begin{aligned} & \text{Minimize } C(\rho) \text{ or } A(\rho) \\ & \text{Subject to } 0 \leq \rho \leq 1 \text{ and} \\ & \int_{\Omega} \rho \, d\Omega - V_{max} \leq 0 \end{aligned} \quad (6)$$

One way to maximize a structure's global stiffness is to minimize its compliance. The compliance is therefore defined as the equivalent strain energy of the FE solution, which yields higher stiffness when minimized. The compliance is defined in (7) and  $u$  is obtained by solving the equilibrium equation (8).

$$C(\rho) = f^T u \quad (7)$$

$$K(\rho)u = f, \text{ with } K(\rho) = \sum_{e=1}^{nel} \rho_e K_e^0 \quad (8)$$

$K_e^0$  is the initial elemental stiffness matrix. Using the gradient based approach, derivatives with respect to  $C(\rho)$  are evaluated. Finally, the case of multiple loads as (9) can be included in the structural optimization task by using weights and objectives that are subjected to a particular load case with the index  $k$ .  $M$  is the total amount of load cases. Equations (7) and (8) can also be further modified for dynamic optimization of surface accelerations [24].

$$f = \sum_{k=1}^M f_k w_k \quad (9)$$

In previous models, optimality criterion, a controller based algorithm was utilized for the optimization. This is faster than the sensitivity based algorithm since no gradient information is required [11, 24]. However, the controller based algorithm is based on stresses and is limited to optimizing stiffness under a volume fraction constraint. Sensitivities are not evaluated since the controller uses strain energy and stresses as an input. In this paper, the optimization tasks are executed using sensitivity based Method of Moving Asymptotes (MMA) approach. It solves non-linear optimization problems in the sense that it uses sequences with sub-problems that are approximations of the original problem [25]. For MMA, these sub-problems are constructed by gradient information and it is also assumed that these approximations are convex. Different design responses can also be combined, for example if the objective is to minimize compliance, both the volume and effective stress can be included as constraints.

The primary focus of the paper lies in the development and optimization of a three stage transmission. The modeling and simulation section describes the architectural layout, intricacies involved in the development of the static simulation template, the forces in the linear dynamic step, coupling between the vibration and acoustic domains, etc., are described. In the optimization section, housing weights are used as design variables. The vehicle performance and indicators, such as compliance, acoustic pressure, etc., are considered as design objectives.

The design procedure is defined in three steps. First, the FE model is used to analyze the original structure. The stiffness and Von-mises stress obtained from the analysis are then treated as the design variables in the topology optimization. The constraint of the optimization task is the reduction in total mass while improving the performance parameter such as vibration in comparison to the original. Eventually, two different optimized variants are developed: one for the static scenario and the other for the linear dynamic load case.

### 3. Development of the Numerical Model

Before the optimization task is illustrated, it is essential to create a simulation template that takes into account all the characteristics, such as excitation mechanisms, loads at the bearings, deformations of the critical components, etc. Considering the necessary complexities within the modeling technique enables the user to build a model that is close to reality. Detailed information regarding the simulation templates with analyses at multiple domains, such as the Lumped Parameter Model (LPM), Multi-Body Dynamics (MBD), etc., can be found in papers [1, 23].

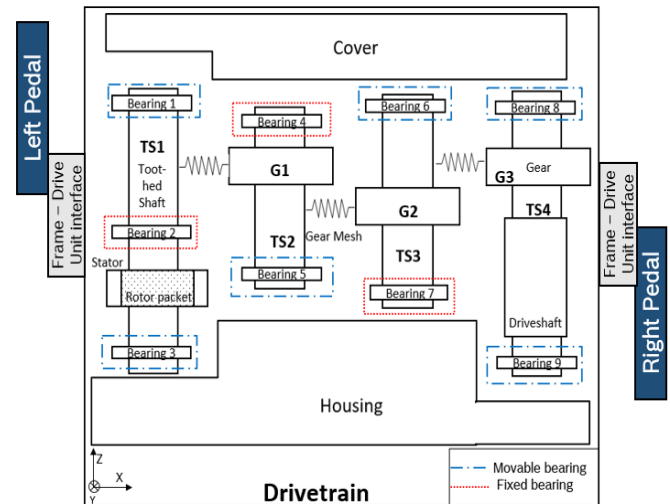


Figure 1: Architectural layout of housings, gears, toothed shafts and bearings of the drivetrain

Figure 1 describes the layout of simulated drivetrain instance to check for critical stresses, deformations and Eigen frequencies. The rotation of the toothed shafts and gears would cause transmission errors that lead to vibrations in the structure. The torque at the toothed shafts and axial load at the bearings would in turn cause the housing structures to deform. Along with the load, boundaries are also applied in the simulation template. For instance, the DOF of bearing nodes are either fixed, thus the z-axis rotation is free, or movable, thus the z-axis translation and rotation are free. Furthermore, the drivetrain is connected to pedals at the frame bracket interface; depending on the load case scenario, the model is enforced with loads at the left pedal, right pedal, shaft axles, gear toothmesh, etc.

Vibrations are influenced by structures having heavier mass or greater stiffness. Therefore, the components that contribute to the drive train vibrations must first be determined. Table 1 describes the scaled masses, Young's modulus and the yield limits of different structures available in the drivetrain. Together, these structures represent about 92 % of the total mass of the power unit. Hence, the structural validation of these components individually and in combination would help to understand the deformations and vibrations of the entire drivetrain in later stages.

The FE software Abaqus and the optimization tool Tosca from Dassault Systems are used to implement the simulation and optimization models [26].

Table 1: Modeling information of the individual components of the drivetrain (scaled)

Component	Mass (Kg)	Young's modulus (Mpa)	Plasticity limit (MPa)
Housing	0.113	0.119	0.14
Cover	0.077	0.119	0.14
Stator	0.268	0.952	0.651
TS4	0.065	0.938	1
G3	0.063	1	1
TS3	0.065	1	1
G2	0.019	1	1
TS2	0.01	1	1
G1	0.015	1	1
TS1	0.023	1	1
Driveshaft	0.044	0.933	0.94
Rotor packet	0.079	0.881	0.417
Frames	0.045	1	0.515

### 3.1. Static Analysis Template

This section describes the model parameters associated with simulation of the static load case. It illustrates a simulation template for analyzing the electromechanical drive for its stresses and deformations. In this scenario, the load is applied in multiple steps: first at the left pedal and later at the right pedal.

To test the mechanical strength of the power unit, it must be viewed in its most critical state, i.e., when it is subjected to the maximum forces during its use. For this reason, the bike is considered at a position with a maximum force of 1800 N on the pedals. According to the German standard DIN EN ISO 4210-8 for "bicycles-safety-technical requirements for bicycles", two cases have to be considered in order to validate the strength of a system [27]. These cases are called "uphill" and "downhill".

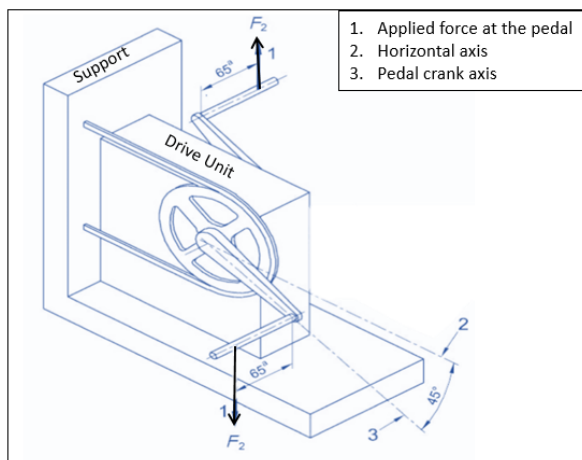


Figure 2: Schematic view of the static load case

Figure 2 corresponds to the uphill load case, where the force on the left and right pedals act in the downward and upward directions, respectively. This standard defines the position of the

pedals in relation to the axis of the chain, as well as the forces and boundary conditions acting on them.

Furthermore, the Abaqus simulation model developed must comply with these standards. The boundary conditions apply to the chain and frame. In the uphill position, the simulation consists of two steps, in which a force of 1800 N is applied first to the right pedal and then to the left pedal. The mesh is then created and refined so that the simulation converges and areas with high stress are re-meshed finely for better accuracy. Table 2 describes the general information regarding the FE model.

Table 2: FE model information for static load case

Count of nodes	Count of elements	Load step 1	Load step 2
872148	525712	Right pedal 1800 N	Left pedal -1800N

A specific magnesium alloy with a yield limit of 140 MPa is used as the material for the housing structures. The fracture occurs at a plastic elongation of more than 6 %. With a safety factor of 2, the limit value at 3 % is considered further for the simulations.

The maximum stresses of both uphill and downhill simulation are comparable for their respective cases and hence do not allow the most critical case to be determined. In order to circumvent this problem, a life cycle study is performed to determine the most critical position. This is not the subject of this paper, but is carried out to find the critical load case. Uphill load case with specific orientation (not disclosed) is found out to be the worst in comparison with downhill and other orientations, therefore, it is considered for the further part of the work.

Table 3 corresponds to results obtained for the uphill simulation. It can be noted that PEEQ (equivalent plastic strain) of the two housing bodies is below the tolerable limit. The maximum stresses exceeds the yield limit. However, these are concentrated in small, highly localized areas. The highest stresses occur at the screw holes where the two parts are connected and where they are in contact with the brackets that connect them to the frame. These areas therefore need to be finely meshed [28]. The contour plots are not illustrated due to the non-disclosure agreement.

Table 3: Stress, deformation and strain energy obtained from the housing and cover

Component	Stress (MPa)	PEEQ (%)	Strain Energy (N.mm)
Housing	218.3	1.57	14210
Cover	211.2	1.32	9821

When a load is applied and the material deforms, energy is introduced into the structure. The energy introduced into the material due to the loading is referred to as the strain energy. Strain energy density gives an overview to which elements/areas of the model are contributing more in supporting the load (absorbing). As a result, an idea as to, where to strengthen the part and where to remove the material, is obtained. The strain energy is later used in optimization of housing structures, where the load-paths are identified.



### 3.2. Vibro-acoustic Template

The modular diagram in Figure 3 aids with the development of a correct model for curtailing the Noise Vibration and Harshness (NVH) below the threshold limits. The dynamic characteristics obtained from the vibration and acoustic module are transferred to the optimization module. Later, these values are analyzed and altered according to the changes made in the design variables.

This section gives an overview of the first step, i.e., building an simulation template, to calculate the vibration and noise. Vibrations originating at the gears and shafts need to be well understood and later contained at the desired boundaries. For this reason, modal analyses for the DU components are developed. This aids in evaluating the vibration behavior of the individual structures. The drivetrain consists of four main components: gears, motor, shafts and housings. At the development stage, essential parameters for vehicle vibration such as mass, inertia, stiffness and damping for these structures are evaluated.

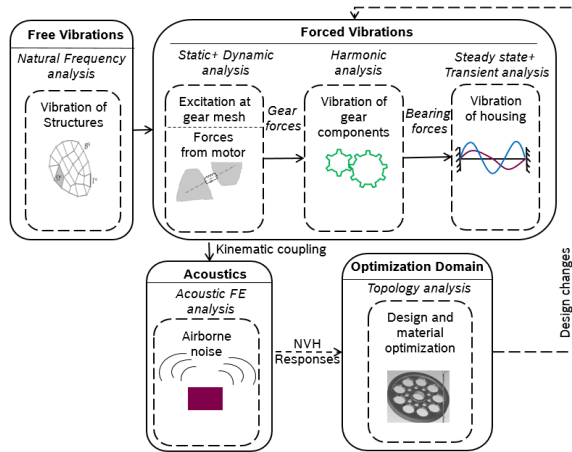


Figure 3: Flow of the FE model for dynamic simulations

Damping is the diminution, restriction or prevention of the vibrations. To achieve this, damping elements are introduced to the systems; the addition of these make the power unit more efficient and provide a better ride comfort [23]. When a damped structure, which is initially at rest, is excited with a harmonic load, it has a transient response at first which then fades out quickly. Furthermore, the structure attains a steady state which is characterized by a harmonic response with the same frequency as that of the applied load. The analysis is performed across a frequency sweep by applying the loading at a series of different frequencies and logging its response. This procedure provides solutions to the linear equations of motion when the loading is harmonic. As part of this section, several submodules of the vibro-acoustic template, such as excitations and acoustic influences, are discussed further.

#### 3.2.1. Excitations from the Electric Motor

Electromagnetic (EM) phenomena in drive units can generate undesirable effects such as vibrations. There are, already, a few simulation methods that have been developed for understanding this effect. For instance, paper [29] from Li describes the electromagnetic simulations developed to study the characteristics of flux and EM force under different conditions using FEM, thereby it is utilized to analyze the vibration behavior of the

structures. Similarly, paper [30] published by Dupont, focuses on excitations due to electromagnetic phenomena using an electromagnetic FE solver. This excitation is then projected onto the structure mesh of the stator in order to calculate the dynamic response. However, they neglected the effects of excitation due to the transmission errors. Nevertheless, through this work, the radial, tangential and axial forces arising at the stator due to the EM forces are coupled with the mechanical excitations for analyzing the vibration characteristics of the drivetrain.

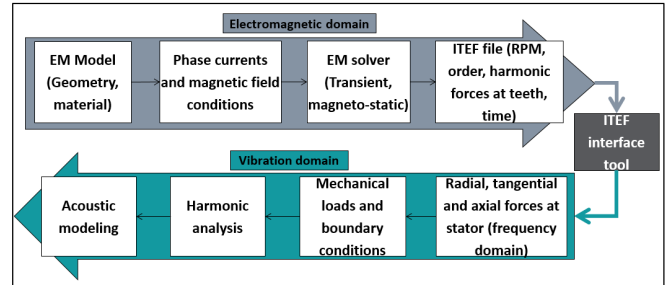


Figure 4: Flow for simulating vibration characteristics using EM forces

Figure 4 describes the flowchart for the calculation of the vibration behavior for the drivetrain based on the EM characteristics. The EM model is prepared using motor geometries and appropriate definitions, such as the phase current. Apart from the calculation of flux and field lines, the simulation also generates an ITEF (Interface for Transfer of Excitation Forces) file. This file consists of information, including RPM, orders and forces at each individual tooth, across discrete time points. The ITEF tool acts as an interface to transfer the forces onto the stator surfaces with radial, tangential and axial components along the frequency spectrum of interest. Later, these forces are used for simulating the harmonics of the drivetrain structures.

#### 3.2.2. Excitations from the Bearings

For further analysis, the gear model with bearings and shafts is simulated. Here, the axial and radial forces at the bearings and other critical areas are calculated. The reaction forces in the simulation are logged in order to retrieve the complete profile of bearing forces. Figure 5 describes the profile of axial forces at the crankshaft after converting into the frequency domain (as an example). The other axes also have analogous profiles with their respective orders and amplitudes. These are not described in detail as the idea was to introduce the user to force amplitude vs. order graphs that are later used in the harmonic excitation step. Information about the simulation model encompassing axial forces and housing structures are explained in the papers [1, 31]. The scale on the y-axis is normalized for non-disclosure purposes.

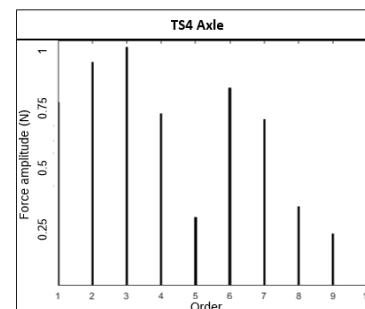


Figure 5: Forces at the TS4 axle represented in the frequency domain

### 3.2.3. Acoustic Mesh

The surface velocities of the housing surfaces create pressure fluctuations of the fluid around them [32]. These disturbances result in noise around the drivetrain. The acoustic pressures and intensities across different distances are estimated with the help of an acoustic FE mesh.

Acoustic elements model the propagation of the acoustic waves and are active only in dynamic analysis procedures. Figure 6 represents the schematic view of the acoustic mesh developed around the drivetrain structure. The terms  $n_s$  and  $n_a$  indicate the normal vectors of the structural and acoustic mesh, respectively. Furthermore, information corresponding to the coupling between these domains is explained by Sandberg in book [33]. The length of the acoustic elements should be able to capture the maximum frequency of interest. In (10),  $c$ ,  $n$  and  $h$  correspond to the speed of sound, number of elements required to capture the wave (generally 5-8) and wavelength, respectively [31].

$$freq_{max} = \frac{c}{n \cdot h} \tag{10}$$

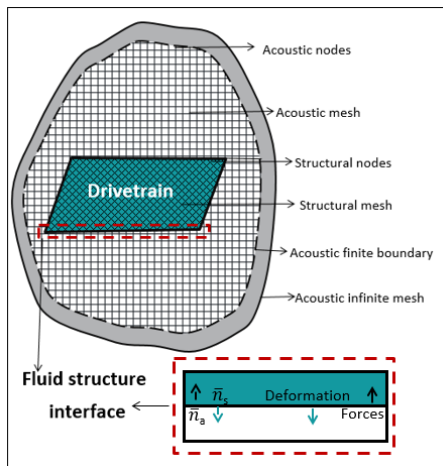


Figure 6: Structural, acoustic finite and infinite element model [20]

Exterior surfaces of the structure are tied to the inner surface of the finite acoustic mesh, hence, velocities at the structural surfaces are transferred to the acoustic nodes. Furthermore, the acoustic velocities of nodes can be altered by introducing impedances between the surfaces. Later, the acoustic finite mesh is attached to the acoustic infinite mesh for estimating the pressures at farther distances. The loss effects within the acoustic medium are modeled via volumetric drag and material impedance parameters.

Table 4: Step information for harmonic simulation

Frequency range	300 – 10000 Hz
Applied load	Axial forces (section Excitations from the bearings) and motor forces (section Excitations from the Electric Motor)
Damping factor	0.03 - 0.24 (frequency dependent) (exact values not disclosed)
Step type	Steady state dynamics - modal

### 3.2.4. Vibration Results

This section provides a brief overview of the acceleration magnitudes for a drive unit simulation. Table 4 indicates the step information such as the frequency range, applied load and damping constants used for calculation of harmonic responses at the housing structure.

Figure 7 describes the surface acceleration recorded at the housing structure element that has the maximum displacement for constant damping factor across the spectrum. The x-axis and y-axis represent the frequency in Hertz and amplitude in Decibels, respectively. Note that these values are scaled appropriately for non-disclosure reasons. The graph shows the influence of the damping ratio on the acceleration magnitude. For lesser damping factors, the peaks tend to be very sharp and for higher damping factors, the amplitudes are reduced and smoothed out. Therefore, the models with damping structures along the transfer path have the potential to create an impedance jump across the intersections, thereby reducing the vibrations. The horizontal line indicates the threshold limit defined using the specification data. The specification sheet including the vibration, sound limits, weight restrictions etc. are obtained from the requirements engineering team, who study the present market situation and analyze the various factors associated with customer satisfaction.

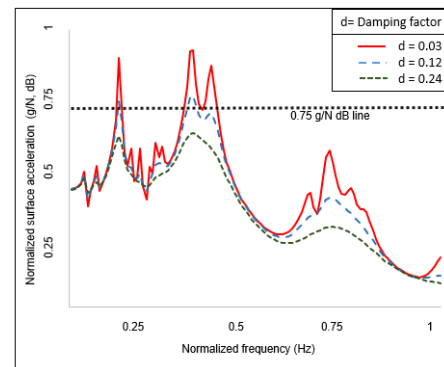


Figure 7: Acceleration simulated at a critical node with varying damping factors

Further descriptions of the results are illustrated in section 6, after describing the test setups and connected software.

## 4. Topology Optimization

Topology optimization starts with an initial model and determines an optimum design by modifying the properties of the material in selected areas, effectively removing elements from the analysis. The numerical methods developed to achieve structural optimization can be classified into two categories: the deterministic method and the probabilistic method. The former uses mathematical programming and a special gradient-based optimizer to perform the optimization. The latter uses heuristic algorithms which are nature-inspired and are developed based on the successful evolutionary behavior of natural systems [34, 35].

For this paper, mostly mathematical programming along with an optimal criteria approach are used in the optimization tasks. The mathematical programming performs iterative changes of the initial design for improving the objective function and fulfilling constraints in each optimization iteration, while the optimal criteria uses mathematical formulations of conditions that characterize the

optimum. The design variables are redesigned so that they fulfill the optimal criteria. The main advantages of this method are that it has faster convergence and the speed of this convergence is independent of the number of design variables. Figure 8 shows the flow of data within this optimization module for reducing the stress at critical housing areas.

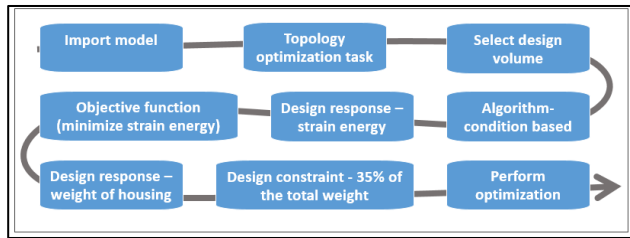


Figure 8: Optimization flow for minimizing strain energy

Table 5 describes the variables defined in the Tosca parameter file, which are used in order to optimize the drive unit for the strain energy. A design constraint with 48 % and 27 % of total weight for the housing and cover, respectively, is imposed. The area near the motor, bracket interface, other critical boundaries, etc., must not be changed. Therefore, the model is also given the frozen area for both structures. Minimization of the strain energy of both parts is considered as the objective function of the optimization task. The convergence of results is obtained at optimization iteration number 32. Later, using the Tosca GUI’s Smooth module, visualization of the smoothed, optimized part and the downside of it, i.e., the removed material, is determined.

Table 5: Topology optimization definition for minimizing strain energy

Optimization task	Topology
Objective function	Minimize strain energy / Minimize compliance
Design constraint	Reduction in weight of housing and cover
Strategy	Topology sensitivity
Iteration stop	50

Figure 9 describes the reduction in strain energy through the optimization simulation. The smoothed results from the Tosca are revised in the CAD software and finally new housing and cover parts with 55 % and 31 % of the original parts are obtained.

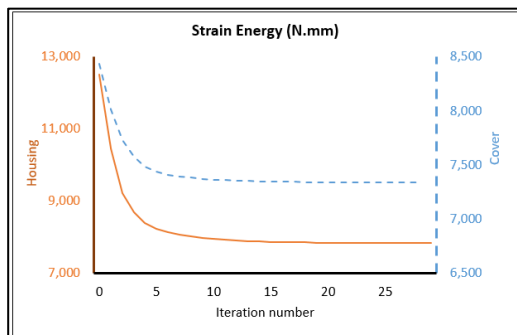


Figure 9: Strain energy vs. optimization iteration

Once the optimization task has been accomplished, new components with a mixture of the two materials can be defined. It is understood that the existing structures, i.e., the resulting

elements of the optimization, have the maximum stress and strain energy, therefore, these elements were retained during the optimization. These structures are referred to as skeletons in the further sections of the work. This mechanical optimization shows that it is possible to simplify the geometry of the housing and cover whilst maintaining the stiffness of the structure. The metal skeleton can then be provided with a plastic insert in order to close and seal the drive unit.

In the next step, removed elements (negative volume) are replaced with different plastic materials to obtain the weight reduction in comparison with the original parts and, at the same time, curtail the stress limits below the thresholds.

Table 6: Maximum stress, drivetrain weight and percentage mass in comparison with original drive unit

Material	Max stress (MPa)	Drivetrain weight (gm)	Optimized drivetrain mass in %
Mg Skeleton	1530	243	47
EP43	380.6	425	82
Akulon	581.2	425	82
Delrin	361.9	467	90
Polyfort	553.4	396	76

The results in Table 6 show that the addition of plastic significantly reduces the stresses in the parts compared to the pure metal skeleton. Only a few areas of the housing have stresses that exceed the yield limit. In addition, these areas are thin and very localized. The geometry of these zones must be revised in CAD software in order to bring their maximum stress below the fracture limit.

Among the various plastics used, Delrin offers the lowest maximum stress at 361.9 MPa, while the addition of Polyfort enables the best weight savings. In fact, the magnesium skeleton supplemented with Polyfort accounts for 76 % of the mass of the original DU. EP43 is a compromise between these two cases and represents the best optimization solution. Indeed, when added to the magnesium skeleton, this allows a maximum tension of 380.6 MPa to be obtained, whilst the end part (skeleton and plastic) comprises 82 % of the original drivetrain mass. This weight saving of almost 20 % represents a potential improvement in the drivetrain. The geometry of this new structure must be revised for it to be constructed. In addition, thermal aspects must also be considered. Indeed, the structure must allow the thermal dissipation of the heat generated by the engine and transmission during the operation of the DU. The connection between metal and plastic must adapt to the different thermal expansions of the materials to ensure a perfect seal during operation.

## 5. Dynamic Optimization

Developing a sensible combination of objective functions and constraints is one of the most challenging steps in applying topology optimization. For the design of dynamic systems, structural vibration control is the most essential consideration. The objective of this type of optimization is to achieve the optimal material layout for the load-bearing components.

Figure 10 describes the flow of the optimization module to optimize the structural dynamics of the housing and cover, thereby improving the drivetrain. The two basic parts (cover and housing) are covered with an envelope volume (design volume). The internal forces, the electromagnetic forces of the motor and the transmission from section 3.2 are applied. The displacements of frame interface points are measured, which enables reaction forces to be calculated. The aim of the dynamic optimization is then to minimize the amplitude of the reaction forces in order to reduce the noise generated by the drive unit during its operation. For this purpose, the optimization software removes material from the design volume to form the optimal ribs. The complete design volume (from both housing and cover) weighed 1.25 kg at the beginning of the simulation. The restriction introduced via the volume constraint corresponds to a final weight corresponding to 30 % of the initial weight. This optimization is carried out in 42 iterations over 38 hours to give a final weight of 312 gm (i.e., approximately 26 % of its starting weight).

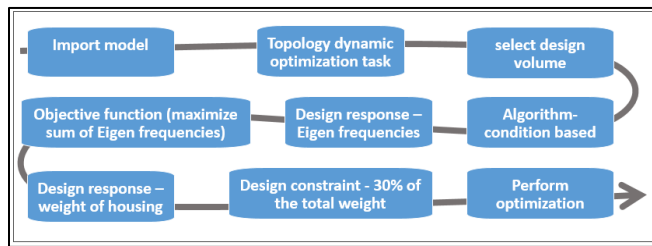


Figure 10: Optimization flow for linear dynamics

Table 7 illustrates the variables defined in the Tosca parameter file that tries to optimize the drive unit for the dynamic characteristics.

Table 7: Topology optimization definition for reducing the noise level

Optimization task	Topology
Objective function	Minimize the vibration and sound levels
Design constraint	Reduction in weight and stress of housing and cover
Strategy	Topology sensitivity
Iteration stop	80

The Frequency Response Function (FRF) is a frequency-based function used to identify the resonant frequencies, damping and mode shapes of the structure. This expresses the frequency domain relationship between an input (force) and output (acceleration) of a linear, time-invariant system. Figure 11 describes a FRF graph for the mean acceleration measured at the critical interface points. This optimization shows a decrease in the maximum acceleration by 4.2 dB. This reduction is greater at higher frequencies and reaches 12 dB compared to the initial drive unit. Overall, the acceleration of the optimized model lies well within the thresholds (limits shown in Figure 11).

Similarly, Figure 12 represents the octave spectrum simulated at 1000 mm distance from the main housing structure. This figure compares the sound pressure levels of the optimized and original drivetrain structures. Optimization shows a decrease in the sound pressure level by 6.2 dB at the frequency spectrum band with the maximum amplitude. This difference is greater at higher

frequencies and reaches 13.4 dB compared to the original drivetrain. Overall, the acoustic values of the optimized model lie well within the thresholds.

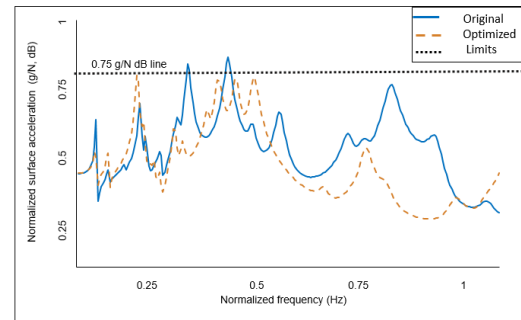


Figure 11: Vibration plot of optimized vs. original drivetrain structure

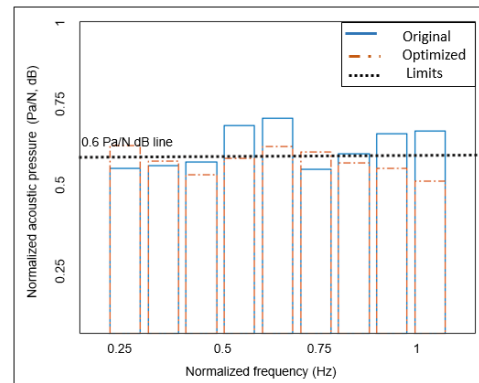


Figure 12: Octave (acoustic) plot of optimized vs. original drivetrain structure

## 6. Method Verification

This part of the paper describes the validation of the methodology generated using the simulation techniques in section 3.2. Initially, the experimental setup developed for a DU variant is illustrated. Later, the FE model verification is shown using the values of surface acceleration and acoustic pressure.

### 6.1. Experimental Setup

Experiments are performed to verify the surrounding pressure and surface velocities of the current drivetrain variant obtained from the numerical calculations. Figure 13 illustrates the schematic experimental setup of the drivetrain and corresponding measuring devices. The absorber box is used to isolate the powertrain vibrations from the surroundings. The different forms of the dashed lines indicate the surface velocities and acoustic fluctuations respectively.

RPM and motor torque profiles are considered as inputs for simulations and experimental runs. The Polytec laser vibrometer is used for the measurement of exterior vibrations of the housing and cover surfaces [36]. Acceleration sensors are mounted to obtain accelerations at precise locations. The sound pressure level is measured using the acoustic camera (from Head Acoustics) and microphones installed at a distance of 1000 mm [37]. LMS software serves as a central node to send the force reference signals and also to record the vibrations and noise [38]. Finally, after the analysis of multiple DU variants for various torque and RPM profiles, the Campbell and spectral plots are generated.



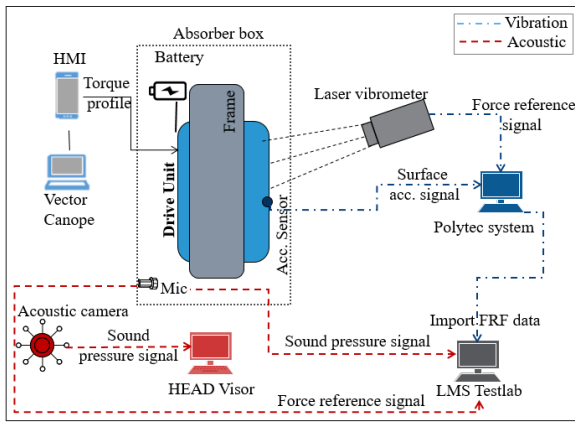


Figure 13: Experimental model for measuring noise and vibration [1]

## 6.2. Numerical Validation

This section attempts to establish the relationship between the experimental results and the FE values. As a part of it, the first step is to validate the model components for their natural frequencies. It aids in verifying properties like damping factors, Eigen forms and frequencies of the components (not shown in this paper). The comparison of Eigen forms between the Experiment Modal Analysis (EMA) and FE frequency analysis helped to determine the variables like loss coefficients for the spectrum of interest. This data is further used for simulating the harmonic step using the steady state dynamic solution technique. Validating the complete drivetrain at the first go can be extremely hard due to complexities associated in the system. At the start, individual structures are verified and later for the combined structures. Validating various combination of structures enabled for studying the properties of damping associated with contact.

The presence of permanent deformations can alter the dynamic characteristics. Hence, before performing the harmonic analysis of complete drivetrain, the structures are checked for plastic deformations for the static load case. In the end, the damping and material values obtained from the combined structures are used for the simulation and validation of the forced excitation (Figure 14) of the complete drivetrain. Here, an excitation with a torque (2.5 N.m), RPM (0-2500) and appropriate boundaries are used.

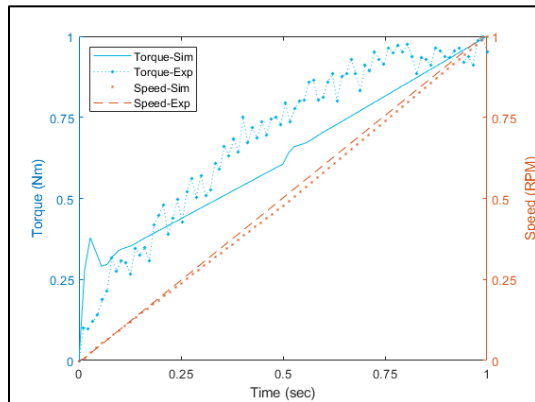


Figure 14: Torque and RPM profiles vs. time

Figure 15 shows the FRF plot of the simulation and experimental results for the vibrations of a DU variant (using the template discussed in section 3.2). The x-axis corresponds to the

frequency spectrum. The y-axis represents the surface velocity of a critical node on the exterior surface of the drivetrain (g/N, dB). The numbers shown on the graph correspond to the resonance frequencies. These peaks indicate the presence of Eigen frequencies for the drivetrain components. Figure 16 describes the octave plot (1/3rd) of the acoustic quantity. The y-axis of Figure 16 represents the acoustic pressures (Pa/N, dB). The simulation and experimental values shown in Figure 15 and Figure 16 have a deviation of around 8.78 % and 4.38 % (on average), respectively.

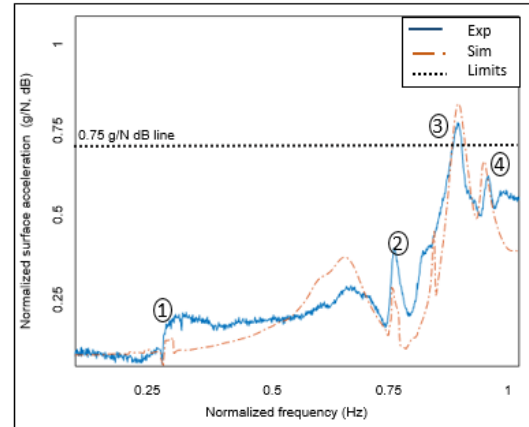


Figure 15: FRF vibration plot - Experimental vs. Simulated

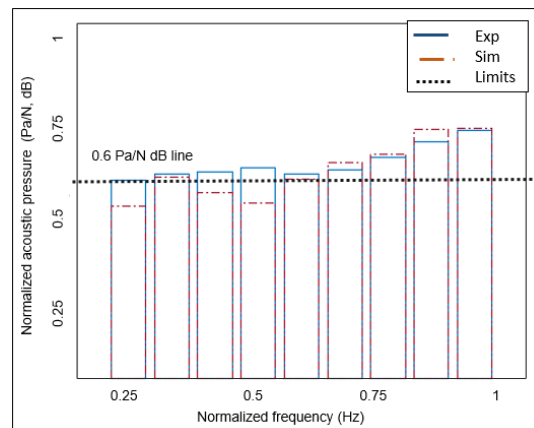


Figure 16: Sound octave spectrum - Experimental vs. Simulated

## 7. Conclusion

A numerical modeling method for the drivetrain with the aim of understanding the structural properties, such as stress, strain energy, compliance, surface velocities and acoustic pressure is proposed. First, a static model template with its appropriate forces and boundaries is shown. The model is then simulated for the worst load case, i.e., the uphill condition. The vibro-acoustic template introduces the user to topics related to electromagnetic influences, axial forces and structural fluid coupling. Furthermore, the harmonic forces emanating from the bearings and the motor are included into the template. Harmonic analysis is used to calculate the surface acceleration at the housings and the sound pressure in the surroundings. The next section describes two different optimized models with targets of minimizing strain energy and surface velocities. These two tasks cannot be encompassed into one single model as the solver for both the cases is different. The first scenario dealt with a static non-linear model and the second

focused on the structural dynamic characteristics. The validations are illustrated by comparing the frequency curves of simulations and experiments.

The dynamic responses described in this paper are considerably improved in comparison with the multi-body models that are illustrated in the state of the art [1]. Although the results from physical modeling approach [1] are promising, it failed in providing the information like the locations with higher vibration amplitudes in the structures. The FE method helped to realize the critical areas of the structures and to develop high-level models for calculating the acoustic reactions at precise distances. The other prior art approaches such as [6, 8, 9, 17] dealt with a single facet of the problem such as the analysis of the transmission vibrations or just a static optimization or FE modeling etc., for the traditional drive trains. The approach described in this paper aided in building a complete flow analyzing the excitations from several sources on the gears, the motor, the bearings and the housing structures for the electromechanical drives. The method also helped in connecting the static and dynamic domain to build a model with minimized deformations and vibrations simultaneously (Figure 3).

Although the simulation results obtained do not fully complement with the experiments, a substantial 78 % of results are decently fitted. The deviations between results are caused by reasons such as device calibrations, simulation errors and complexities in representing the physical significance of drivetrain parts. Improvements such as better contact, acoustic modeling in simulations and more robust test setups like anechoic chambers can aid in mitigating these differences. At the fluid structure boundary, the simulation setups transfer all the velocities from structural to acoustic domain, while in reality there is energy dissipation due to experimental conditions. Although the losses can be included using impedance boundaries in the FE, a detailed study needs to be performed to define the exact admittance and volumetric drag coefficients. Overall, the system method developed appears to have the potential to predict the NVH characteristics and can be considered for analyzing the optimized drivetrain models.

## 8. Outlook

Since the FE values obtained for the original drivetrain model are close enough to the experimental results, the next major step would be to build the real parts of the optimized housing and cover. It would be essential to perform validation on the improvised designs obtained from Tosca.

Other aspects such as thermal fluctuations inside drivetrain can also influence the stresses and linear dynamic characteristics of the structures. The developed simulation templates can thus be further integrated into other areas such as the thermal, control systems and the electrical system model, via heterogeneous modeling techniques.

Also, wear among the gears could affect the noise of the drive train. For example, the methods described in paper [39] can be used as a basis for understanding the dependence of the abrasion mechanism on the noise emission of the drivetrain.

## Conflict of Interest

The authors declare no conflict of interest.

## References

- [1] Y. Kolluru, R. Doelling, L. Hedrich, "Numerical simulations of vibro acoustic behaviors related to drive train assemblies", in 2019 International Symposium on Systems Engineering (ISSE), 2019. <https://doi.org/10.1109/ISSE46696.2019.8984458>
- [2] T. Kato, R. Mizutani, H. Matsumoto, K. Yamamoto, "Advanced technologies of traction motor for automobile", in 2013 IEEE ECCE Asia Downunder, IEEE, 147–152, 2013. <https://doi.org/10.1109/ECCE-Asia.2013.6579088>
- [3] J. de Santiago, H. Bernho, B. Ekegard, S. Eriksson, S. Ferhatovic, R. Waters, M. Leijon, "Electrical motor drivelines in commercial all-electric vehicles: A review", IEEE Transactions on Vehicular Technology – IEEE TRANS VEH TECHNOL, 61, 475–484, 2012. <https://doi.org/10.1109/TVT.2011.2177873>
- [4] E. Rask, M. Duoba, H. Lohse-Busch, Recent hybrid electric vehicle trends and technologies, in 2011 IEEE Vehicle Power and Propulsion Conference, IEEE, 1–6, 2011. <https://doi.org/10.1109/VPPC.2011.6043172>
- [5] J. M. Rodriguez, R. Meneses, J. Orus, "Active vibration control for electric vehicle compliant drivetrains, in IECON 2013-39th Annual Conference of the IEEE Industrial Electronics Society", IEEE, 2590–2595, 2013. <https://doi.org/10.1109/IECON.2013.6699539>
- [6] Y. Ito, S. Tomura, K. Moriya, "Vibration-reducing motor control for hybrid vehicles", RD Review of Toyota CRDL, 37–43, 2005. [https://www.tytlabs.com/english/review/rev402epdf/e402\\_037ito.pdf](https://www.tytlabs.com/english/review/rev402epdf/e402_037ito.pdf), last accessed 2020/21/07
- [7] K. Jensen, Optimising stress constrained structural optimisation, in 23rd International Meshing Roundtable (IMR), Elsevier, 2014. <https://imr.sandia.gov/papers/abstracts/Je751.html>, last accessed 2020/21/07
- [8] E. Hinton, J. Siens, fully stressed topological design of structures using an evolutionary procedure, Engineering Computations: Int J for Computer-Aided Engineering 12, 229–244, 1993. <https://doi.org/10.1108/02644409510799578>
- [9] H. Eschenauer, A. Schumacher, T. Vietor, Decision makings for initial designs made of advanced materials, in Topology design of structures, Springer, 469–480, 1993. [https://doi.org/10.1007/978-94-011-1804-0\\_33](https://doi.org/10.1007/978-94-011-1804-0_33)
- [10] Y. M. Xie, G. P. Steven, "A simple evolutionary procedure for structural optimization", Computers & structures 49, 885–896, 1993. [https://doi.org/10.1016/0045-7949\(93\)90035-C](https://doi.org/10.1016/0045-7949(93)90035-C)
- [11] K. Suzuki, N. Kikuchi, Layout optimization using the homogenization method, in Optimization of large structural systems, Springer, 157–175, 1993. [https://doi.org/10.1007/978-94-010-9577-8\\_7](https://doi.org/10.1007/978-94-010-9577-8_7)
- [12] K. Mao, "An approach for powertrain gear transmission error prediction using the non-linear finite element method, Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering 220, 1455–1463, 2006. <https://doi.org/10.1243/09544070JAUTO251>
- [13] Y.-C. Chen, C.-B. Tsay, Stress analysis of a helical gear set with localized bearing contact, Finite Elements in Analysis and Design 38, 707–723, 2008. [https://doi.org/10.1016/S0168-874X\(01\)00100-7](https://doi.org/10.1016/S0168-874X(01)00100-7)
- [14] D. A. Crolla, D. Cao, The impact of hybrid and electric powertrains on vehicle dynamics, control systems and energy regeneration, Vehicle system dynamics 50, 95–109, 2012. <https://doi.org/10.1080/00423114.2012.676651>
- [15] R. Anderl, Simulations with NX: Kinematics, FEM, CFD, EM and Data Management. with Numerous Examples of NX 9, Hanser eLibrary, Hanser Publications, 2014.
- [16] S. Iyengar, R. N. Elements of Mechanical Vibration. India, I.K. International Publishing House Pvt. Limited, 2010.
- [17] J. A. Morgan, M. R. Dhulipudi, R. Y. Yakoub, A. D. Lewis, Gear mesh excitation models for assessing gear rattle and gear whine of torque transmission systems with planetary gear sets, SAE Transactions, 1780–1789, 2007. <https://doi.org/10.4271/2007-01-2245>
- [18] H. Allik, Accuracy of infinite elements in structural acoustics modeling, The Journal of the Acoustical Society of America 77, 1985. <https://doi.org/10.1121/1.2022426>
- [19] R.G. Grimes, J. G. Lewis, and H. D. Simon, A Shifted Block Lanczos Algorithm for Solving Sparse Symmetric Generalized Eigenproblems, SIAM Journal on Matrix Analysis and Application, 228–272, 1994. <https://doi.org/10.1137/S0895479888151111>
- [20] R. C. Shieh, Massively parallel structural design using stochastic optimization and mixed neural net/finite element analysis methods, Computing Systems in Engineering 5, 455–467, 1994. [https://doi.org/10.1016/0956-0521\(94\)90026-4](https://doi.org/10.1016/0956-0521(94)90026-4)
- [21] M. P. Bendsoe, N. Kikuchi, Generating optimal topologies in structural design using a homogenization method, Computer Methods in Applied Mechanics and Engineering, 197–224, 1988. [https://doi.org/10.1016/0045-7825\(88\)90086-2](https://doi.org/10.1016/0045-7825(88)90086-2)
- [22] W. Desmet and D. Vandepitte, Numerical Acoustics: Finite Element Modeling for Acoustics, ISAAC 13- International Seminar on Applied Acoustics, 37–85, 2005.

- [23] Y. Kolluru, R. Doelling, L. Hedrich, Multi domain modeling of NVH for electro-mechanical drives - accepted for publication, in 2020 Integrating Seamlessly NVH, 2020. <https://saemobilus.sae.org/content/2020-01-1584>, last accessed 2020/21/07
- [24] M.P. Bendsoe and O. Sigmund, Topology optimization: Theory, methods and applications, Springer-Verlag, Berlin Heidelberg, 2003.
- [25] K. Svanberg, The method of moving asymptotes - A new method for structural optimization. International Journal for Numerical Methods in Engineering, 359-373, 1987. <https://doi.org/10.1002/nme.1620240207>
- [26] Simulia, Simulia User Assistance Guide 2018, About Abaqus and Tosca, Dassault Systems Simulia Corp, 2018. [http://si0vm1337.de.bosch.com:8081/EstProdDocs2018/English/DSSIMULIA\\_Established.htm?show=SIMULIA\\_Established\\_FrontmatterMap/DSDoc\\_Abaqus.htm](http://si0vm1337.de.bosch.com:8081/EstProdDocs2018/English/DSSIMULIA_Established.htm?show=SIMULIA_Established_FrontmatterMap/DSDoc_Abaqus.htm), last accessed 2020/21/07
- [27] I. 4210-8:2014, Cycles - safety requirements for bicycles - part 8: Pedal and drive system test methods (iso 4210- 8:2014), Online Browsing Platform 1, 2014. <https://www.iso.org/standard/59915.html>, last accessed 2020/21/07
- [28] P. Wriggers, Contact mechanics, Prentice Hall, New Jersey, 2007.
- [29] R. Li, X. Peng, W. Wei, Electromagnetic vibration simulation of a 250 mw large hydropower generator with rotor eccentricity and rotor deformation, Energies 10, 2017. <https://doi.org/10.3390/en10122155>
- [30] J.-B. Dupont, P. Bouvet, Multiphysics modelling to simulate the noise of an automotive electric motor, in 2012 Integrating Seamlessly NVH, 2012. <https://doi.org/10.4271/2012-01-1520>
- [31] Y. Kolluru, R. Doelling, L. Hedrich, Numerical Analysis of the Influences of Wear on the Vibrations of Power Units - accepted for publication, in 2020 Integrating Seamlessly NVH, 1–12, 2020. <https://saemobilus.sae.org/content/2020-01-1506>, last accessed 2020/21/07
- [32] L. Kinsler, Fundamentals of Acoustics, John Wiley & Sons Australia, Limited, 2000.
- [33] G. Sandberg, R. Ohayon, Computational Aspects of Structural Acoustics and Vibration, CISM International Centre for Mechanical Sciences, Springer Vienna, 2009.
- [34] R. Picelli, S. Townsend, C. Brampton, J. Norato, H. Kim, Stress based shape and topology optimization with the level set method, Computer Methods in Applied Mechanics and Engineering 329, 1-23, 2018. <https://doi.org/10.1016/j.cma.2017.09.001>
- [35] C. Mitropoulou, Y. Fourkiotis, N. Lagaros, M. Karlaftis, Evolution Strategies Based Metaheuristics in Structural Design Optimization, 79–102, 2013. <https://doi.org/B978-0-12-398364-0.00004-8>
- [36] P. G. Technology, Vibrometer measurements methods, 2019. <https://www.polytec.com/eu/vibrometry/technology/>, last accessed 2020/21/07
- [37] H. acoustics, Acoustic camera / head visor(code /5000\_) handbook., 2016. [https://www.head-acoustics.com/eng/nvh\\_head-visor.htm](https://www.head-acoustics.com/eng/nvh_head-visor.htm), last accessed 2020/21/07
- [38] Siemens, Lms test.lab, “Analysis and structural design manual”, Siemens Industry Software, 2015. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.372.6925&rep=rep1&type=pdf>, last accessed 2020/21/07
- [39] Y.Kolluru, R.Doelling, L. Hedrich, “Wear estimation using FEM and machine learning techniques”, in 2019 ECOTRIB conference, 176–178, 2019. <https://ecotrib2019.oetg.at/program2>.

## Bilateral Communication Device for Deaf-Mute and Normal People

Raven Carlos Tabiongan\*

College of Engineering, Samar State University, Catbalogan, 6700, Philippines

---

### ARTICLE INFO

Article history:

Received: 07 May, 2020

Accepted: 09 July, 2020

Online: 28 July, 2020

---

Keywords:

Two-way

Communication device

Deaf-mute

Full duplex

RX-TX

---

### ABSTRACT

Communication is a bilateral process and being understood by the person you are talking to is a must. Without the ability to talk nor hear, a person would endure such handicap. Given that hearing and speech are missing, many have ventured to open new communication methods for them through sign language. This bilateral communication device can be utilized by both non-sign language users and Deaf-mute together in a single system. Shaped as a box (8in x 8in) with two multi-touch capable displays on both ends, the contraption has several microcontrollers and touch boards within. The latter has the technology of twelve interactive capacity touch and proximity electrode pads that react when tapped, producing quick response phrases audible via speaker or earphone. These touch boards are equipped with an MP3 decoder, MIDI synthesizer, 3.5mm audio jack and a 128MB microSD card. The touch screen modules mounted on top of the microcontrollers transfer data to and from each other in real-time via receiver-transmitter (RX-TX) full duplex UART serial communication protocol. The device is lightweight weighing at about 3 lbs. The prototype device was piloted in an academic institution of special education for deaf-mute students. Participants were 75 normal and 75 Deaf-mute people aged between 18 and 30 years. The experimental results show the overall rating of the device is 90.6%. The device is designed to promote the face-to-face socialization aspect of the Deaf-mute users to the normal users and vice versa. Several third-party applications were utilized to validate the accuracy and reliability of the device thru metrics of consistency, timing and delay, data transmission, touch response, and screen refresh rates.

---

### 1. Introduction

This paper is an extension of work originally presented in the 2019 8th International Symposium on Next Generation Electronics (ISNE) [1]. Talking and hearing play a very important part of our daily lives. It helps in finishing up tasks, from the simplest to the most complex. It is inherent in humans to rely on other senses when deprived of hearing or speaking. Most people having these disabilities try to learn the art of sign language due to the urge of communicating independently. Unfortunately, being able to acquire such skills has its own sets of fallbacks, such as the inability to converse with the non-sign language, speaking, and hearing-capable people [2]. Communication is a bilateral process and being understood by the person talking to is a must [3]. Without the ability to talk nor hear, a person would be such in a handicap, socially and emotionally. Cases of these handicap conditions are commonly termed as Deaf-mute. In deaf

community, the word deaf is spelled in two separate ways. The small “d”, deaf represents a person’s level of hearing through audiology and not being associated with other members of the deaf community whereas the capital “D”, Deaf indicates the culturally Deaf people who use sign language for communication [4]. Research indicated [5] that Deaf people, especially Deaf children, have high rates of behavioral and emotional issues in relation to different methods of communication. Most people with such disabilities become introverts and resist social connectivity and face-to-face socialization. Imagine the depression and discomfort experienced when one is unable to express their thoughts, for sure frustration arises every now and then.

People rely on words and sound from the environment for them to grasp what is happening. Other alternatives to remedy the said language barrier is to acquire a sign language interpreter, but practically speaking it is not the optimal solution. An interpreter is an inconvenience basically due to the 24/7 limitation. Not everyone around can check what they are saying or is intending to

---

\*Corresponding Author: Raven Carlos Tabiongan, Philippines, +639758694325 & raven.tabiongan@ssu.edu.ph



express. Most Deaf-mute just limit themselves to performing simple tasks to avoid irritation to others and to themselves. Persons who are Deaf-mute normally suffer when performing typical day-to-day tasks. Given that hearing and speech are missing, many have tried to open new communication medium for them such as sign language.

The above-mentioned problems can be solved by integrating a bilateral communication device capable of sending and receiving text and audio responses via full duplex serial communication protocol in real-time. Moreover, the device is equipped with an operating system that converts sign language into text, images, and audio for better communication between the Deaf-mute and the normal users [6], [7].

The main purpose of this research paper is to provide a user experience setting wherein the Deaf-mute and normal users can communicate and chat in a close-proximity, face-to-face setting. Another purpose of this system is to provide a simple and cost-effective solution that can be utilized by both Deaf-mute and non-sign language people together simultaneously in a single system.

## 2. Related Works

The Deaf community is not a monolithic group; it has adversity of groups which are as follows [8, 9]: (1) Hard-of-hearing people: they are neither fully deaf nor fully hearing, also known as culturally marginal people [10]. They can obtain some useful linguistic information from speech; (2) Culturally deaf people: they might belong to deaf families and use sign language as the primary source of communication. Their voice (speech clarity) may be disrupted. (3) Congenital or prelingual deaf people: they are deaf by birth or become deaf before they learn to talk and are not affiliated with Deaf culture. They might or might not use sign language-based communication; (4) Orally educated or post lingual deaf people: they have been deafened in their childhood but developed the speaking skills. (5) Late-deafened adults: they have had the opportunity to adjust their communication techniques as their progressive hearing losses.

### 2.1 Sensor Module Technology Approach.

Sensors and touch screen technology can be integrated in a system to bridge the communication gap between the Deaf-mute and normal people with or without knowledge in sign-language. Sharma et al. used wearable sensor gloves for detecting the hand gestures of sign language [11].

### 2.2 Visual Module Technology Approach.

Many vision-based technologies interventions are used to recognize the sign languages of Deaf people. For example, Soltani et al. developed a gesture-based game for Deaf-mutes by using Microsoft Kinect which recognizes the gesture command and converts it into text so that they can enjoy the interactive environment [12]. Voice for the mute (VOM) system was developed to take input in the form of fingerspelling and convert into corresponding speech [13]. The images of finger spelling signs are retrieved from the camera. After performing noise removal and image processing, the finger spelling signs are matched from the trained dataset. Processed signs are linked to appropriate text and convert this text into required speech. Nagori and Malode [14]

[www.astesj.com](http://www.astesj.com)

proposed the communication platform by extracting images from the video and converting these images into corresponding speech. Sood and Mishra [15] presented the system that takes images of sign language as input and displays speech as output. The features used in vision-based approaches for speech processing are also used in different object recognition-based applications [16]-[22].

### 2.3 Product Design and Development Approach.

The device targets to improve the current way Deaf-mute communicates with the normal people, and vice versa. Thus, a bilateral communication device with the intent to promote proximity and face-to-face socialization between the parties involved is developed [1].

- (1) To design and develop a portable device that enables the Deaf-mute and normal people communicate via multiple modes in a contraption that is designed for proximity and face-to-face communication.
- (2) To establish a stable RX/TX full duplex serial UART communication protocol between the microcontrollers and modules within the device.
- (3) To construct a sturdy packaging for the device suitable for indoor and outdoor portability and use.
- (4) To test the system and device using third-party software. (Terminal Monitor, Balabolka (TTS), Arduino IDE)
  - a. Accuracy in transmitting data
  - b. Consistency in transmitting data
  - c. Speed (Timing and Delay)
  - d. Screen Refresh Rates
  - e. Touch Responses
- (5) To test the system and device using a Five-Point Scale questionnaire given to 150 respondents: (75 Deaf-mute and 75 normal people).
  - a. Value-Added
  - b. Aesthetic Design
  - c. Hardware Components
  - d. Built-in Features
  - e. User-Friendliness
  - f. Reliability
  - g. Accuracy

### 2.4 Mobile Application Technology Approach.

Many of the new smartphones are furnished with advanced sensors high processors, and high-resolution cameras [23]. A real-time emergency assistant "iHelp" [24] was proposed for Deaf-mute people where they can report any kind of emergency. The current location of the user is accessed through built-in GPS system in a smartphone. The information about the emergency is sent to the management through SMS and then passed on to the closest suitable rescue units, and hence the user can get rescue using iHelp. MonoVoix [25] is an Android application that also acts as a sign language interpreter. It captures the signs from a mobile phone camera and then converts them into corresponding speech. Sahaaya [26] is an Android application for Deaf-mute people. It uses sign language to communicate with normal people. The speech-to-sign and sign-to-speech technology are used. For a hearing person to interact with Deaf-mute, the text-to-speech (TTS) technology inputs the speech signal, and a corresponding

sign language video is played against that input through which the mute can easily understand. Bragg et al. [27] proposed a sound detector. The app is used to detect alert sounds and alert the deaf-mute person by vibrating and showing a popup notification.

Improvement in health care access among Deaf people possible by providing the sign language supported visual communication and implementation of communication technologies for healthcare professionals. Some of the implemented technology-based approaches for facilitating Deaf-mutes with easy-to-use services are as follows.

### 3. Materials and Methods

The Arduino Mega 2560 is a microcontroller board based on the ATmega2560. It has 54 digital input/output pins, which can be used as PWM outputs, 16 analog inputs, 4 UARTs hardware serial ports, a 16MHz crystal oscillator, a USB connection, a power jack, an ICSP header, and a reset button. It contains everything needed to support the microcontroller; simply connect it to a computer with a USB cable or power it with a AC-to-DC adapter or battery. The Mega 2560 board is compatible with most shields designed.



Figure 1: Arduino Mega 2560 Microcontroller

Table 1: Arduino Mega 2560 Product Specification

Microcontroller -	ATmega2560
Input Voltage - (recommended)	7-12V
Input Voltage (limit) -	6-20V
Digital I/O Pins	54 (of which 15 provide PWM output)
Analog Input Pins -	16
DC Current per I/O Pin -	20 mA
DC Current for 3.3V Pin -	50 mA
Flash Memory -	256 KB of which 8 KB used by bootloader
SRAM -	8 KB
EEPROM -	4 KB
Clock Speed -	16 MHz
LED BUILTIN -	13
Length -	101.52 mm
Width -	53.3 mm
Weight -	37 g

The Touch Board is a microcontroller board with dedicated capacitive touch and MP3 decoder ICs. It has a headphone socket and micro SD card holder (for file storage), as well as having 12

capacitive touch electrodes. It is based around the ATmega32U4 and runs at 16MHz from 5V. It has a micro USB connector, a JST connector for an external lithium polymer (LiPo) cell, a power switch, and a reset button. It is like the Arduino Leonardo board and can be programmed using the Arduino IDE. The ATmega32U4 can appear to a connected computer as a mouse or a keyboard, (HID) serial port (CDC) or USB MIDI device.

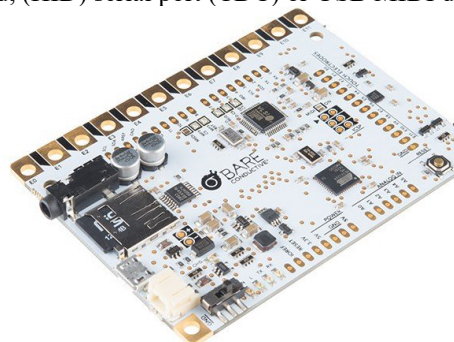


Figure 2: Bare Conductive Touch Board

Table 2: Bare Conductive Touch Board Product Specification

Microcontroller - Touch IC -	Microchip ATmega32U4 Resurgent Semiconductor MPR121
MP3 Decoder IC -	VLSI Solution VS1053b
Audio Output -	15mW into 32Ω via 3.5mm stereo socket
Removable Storage -	Up to 32GB via micro SD card
Input Voltage -	3.0V DC – 5.5V DC
Operating Voltage -	5V DC
Max. Output Current - (5V Rail)	400mA (100mA at startup)
Max. Output Current - (3.3V Rail)	300mA
LiPo Cell Connector -	2-way JST PH series – pin 1 +ve, pin 2 -ve
LiPo Charge Current -	200mA
Capacitive Touch Electrodes-	12 (of which 8 can be configured as digital I/O)
Digital I/O Pins -	20 (of which 3 are used for the MPR121 and 5 are used for the VS1053b-the latter can be unlinked via solder blobs)
PWM Channels -	7 (shared with digital I/O pins)
Analogue Input Channels -	12 (shared with digital I/O pins)
Flash Memory -	32kB (ATmega32U4) of which 4kB used by bootloader
SRAM -	2.5kB (ATmega32U4)
EEPROM -	1kB (ATmega32U4)
Clock Speed -	16MHz (ATmega32U4), 12.288MHz (VS1053b)
DC Current per I/O Pin -	40mA sink and source (ATmega32U4), 12mA source / 1.2mA sink (MPR121)
Analogue Input Resistance -	100MΩ typical (ATmega32U4)

Table 3: Bare Conductive Touch Board Input and Output Pins

Touch Electrodes - E0-E11	Connect to the MPR121 and provide capacitive touch or proximity sensing.
Serial - Pins 0 (RX) and 1 (TX)	Used to receive (RX) and transmit (TX) TTL serial data using ATmega32U4 UART.
TWI (I2C) - Pins 2 (SDA) and 3 (SCL)	Data and clock pins used to communicate with the MPR121.
IRQ - Pin 4	Used to detect interrupt events from the MPR121.
SD-CS - Pin 5	Used to select the micro SD card on the SPI bus.
D-CS - Pin 6	Used to select data input on the VS1053b.
DREQ -	Used to detect data request events from the VS1053b.
MP3-RST -	Used to reset the VS1053b.
MP3-CS -	Used to select the instruction input on the VS1053b
MIDI IN - Pin 10	Used to pass MIDI data to the VS1053b and have it behaved as a MIDI synthesizer
Headphone Output - AGND, R, L	Provide the headphone output from the VS1053b on 0.1" / 2.54mm pitch pads as an alternative to the 3.5mm socket.
External Interrupts - Pins 0, 1, 2, 3, 7	Configured to trigger an interrupt on a low value, a rising or falling edge, or a change in value.

This TFT display is big (2.8" diagonal) bright (4 white-LED backlight) and colorful (18-bit 262,000 different shades)! 240x320 pixels with individual pixel control. It has more resolution than a black and white 128x64 display. As a bonus, this display has a resistive touchscreen attached to it already. It can detect finger presses anywhere on the screen. It also includes an SPI touchscreen controller thus it only needs one additional pin to add a high-quality touchscreen controller. This display shield has a controller built into it with RAM buffering, so that almost no work is done by the microcontroller. This shield needs fewer pins than our v1 shield, in order to connect more sensors, buttons and LEDs: 5 SPI pins for the display, another pin or two for the touchscreen controller and another pin for microSD card to read images off of it.



Figure 3: TFT Touch Screen Display (2.8" diagonal)

Table 4: TFT Screen Pins

Digital #13 or ICSP SCLK -	This pin is used for the TFT, microSD and resistive touch screen data clock
Digital #12 or ICSP MISO -	This pin is used for the TFT, microSD and resistive touch screen data
Digital #11 or ICSP MOSI -	This pin is used for the TFT, microSD and resistive touch screen data.
Digital #10 -	This is the TFT CS (chip select pin). It is used by the Arduino to tell the TFT that it wants to send/receive data from the TFT only.
Digital #9 -	This is the TFT DC (data/command select) pin. It is used by the Arduino to tell the TFT whether it wants to send data or commands

Table 5: Resistive Touch Controller Pins

Digital #13 or ICSP SCLK -	This pin is used for the TFT, microSD and resistive touch screen data clock.
Digital #12 or ICSP MISO -	This pin is used for the TFT, microSD and resistive touch screen data.
Digital #11 or ICSP MOSI -	This pin is used for the TFT, microSD and resistive touch screen data.
Digital #8 -	This is the STMPE610 Resistive Touch CS (chip select pin). It is used by the Arduino to tell the Resistive controller that it wants to send/receive data from the STMPE610 only.

Table 6: Capacitive Touch Pins

SDA -	This is the I2C data pin used by the FT6206 capacitive touch controller chip. It can be shared with other I2C devices. On UNO's this pin is also known as Analog 4.
SCL -	This is the I2C clock pin used by the FT6206 capacitive touch controller chip. It can be shared with other I2C devices. On UNO's this pin is also known as Analog 5.

Table 7: microSD Card Pins

Digital #13 or ICSP SCLK -	This pin is used for the TFT, microSD and resistive touch screen data clock
Digital #12 or ICSP MISO -	This pin is used for the TFT, microSD and resistive touch screen data
Digital #11 or ICSP MOSI -	This pin is used for the TFT, microSD and resistive touch screen data
Digital #4 -	This is the microSD CS (chip select pin). It is used by the Arduino to tell the microSD that it wants to send/receive data from the microSD only



### 3.1 Power Supply Design

The two microcontrollers used in the study are the Bare Conductive Touch Board and Arduino Mega 2560 Rev3. As shown in Figure 5, both microcontrollers are directly powered. Arduino Mega 2560 Rev3 was powered by a 9V Battery with power snap for easier insertion. While the Touch Board was operated with a 3.7V 3000mAh Lithium Polymer Battery. Both power supplies are rechargeable, ensuring energy efficiency. Meanwhile, Bare Conductive Touch Board is powered via rechargeable 3.7V Li-Po Battery, which is directly inserted onto the boards LiPo socket.

The main reason why there are two designated power supplies for each microcontroller is to ensure that there is a sufficient supply independent from one another.

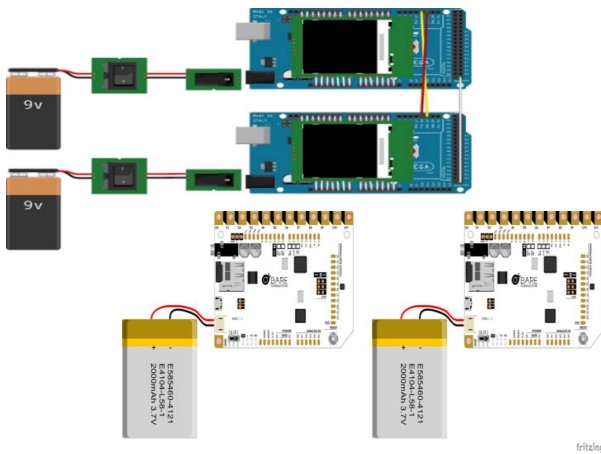


Figure 4: Illustrates how the LCD Touch Shield is mounted to Arduino Mega 2560 Rev3 with 9V power supply (top) and the Bare Conductive Touch Boards is powered thru 3.7V, 3000mah LiPo battery (bottom).

### 3.2 Software Application – Terminal Emulation Program

Terminal Emulation Program is a simple serial port (COM) monitoring program that is a very useful tool in debugging serial communication applications.

### 3.3 Software Application – Arduino Integrated Development Environment (IDE)

Figure 6 shows the real-time process happening inside the touch boards. Using the Arduino IDE 1.6.7. serial monitor, the reading indicates an accurate response, with no delay or whatsoever. From the instance that it registers a physical touch a pre-programmed audio response is played instantly. The Arduino Integrated Development Environment is a cross-platform application that is written in functions from C and C++. It is used to write and upload programs to Arduino compatible boards, but also, with the help of 3rd party cores, other vendor development boards. Furthermore, contains a text editor for writing code, a message area, a text console, a toolbar with buttons for common functions and a series of menus.

### 3.4 Hardware Design

The indicated connections are the I/O pins that when attached together fits in a mount position. These are the exact pin outs to where the touch screens should be connected to the

microcontrollers, to execute the desired output. Basic RX-TX communication was utilized to interact and link the two pairs as shown in Figure 7.

Programming for Arduino is done on Arduino IDE which is based on C Language. It's neither completely C nor embedded C. The code written on the IDE will be passed directly to a C/C++ compiler (avr-g++).

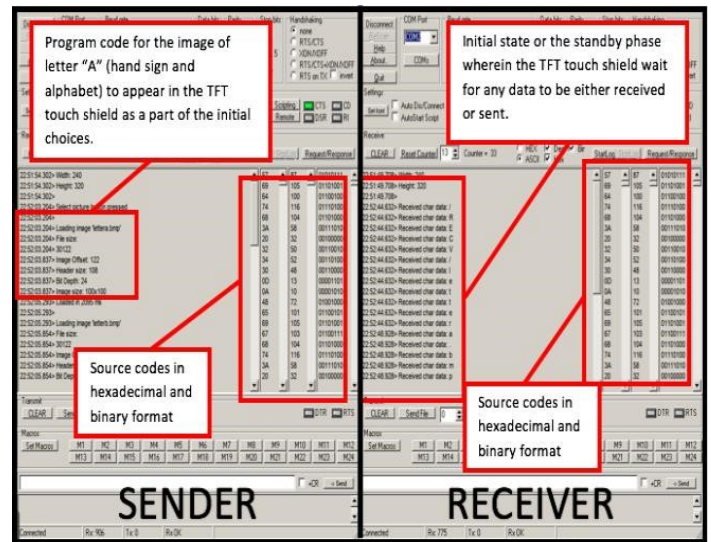


Figure 5: Actual data acquired using the Terminal Emulation Program. What is shown are the source codes in hexadecimal and binary format in real time. This was utilized to ensure accuracy and reliability in transmitting data from one shield to the other and vice-versa.

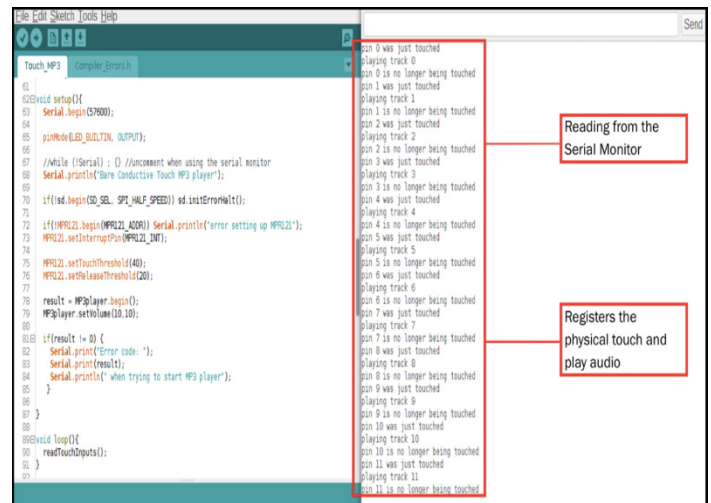


Figure 6. As indicated, for every physical tap a user made on the solder pads, it will play the preloaded phrase track. A total of twelve tracks, one for each pad.

### 3.5 Software Application – Balabolka Text-to-Speech (TTS)

One of the twelve preloaded phrase tracks is “Hello! How are you? I am a Bilateral Communication Device for Hearing and speech Impaired. A communication device that pertains to people suffering from hearing and speaking disabilities most referred to as deaf, mute, and deaf-mute individuals. I am capable of being used both indoor and outdoor environment. I am dedicated to bridge the language gap between the impaired subjects to those of the normal, talking, non-sign language people. Have a good day.”



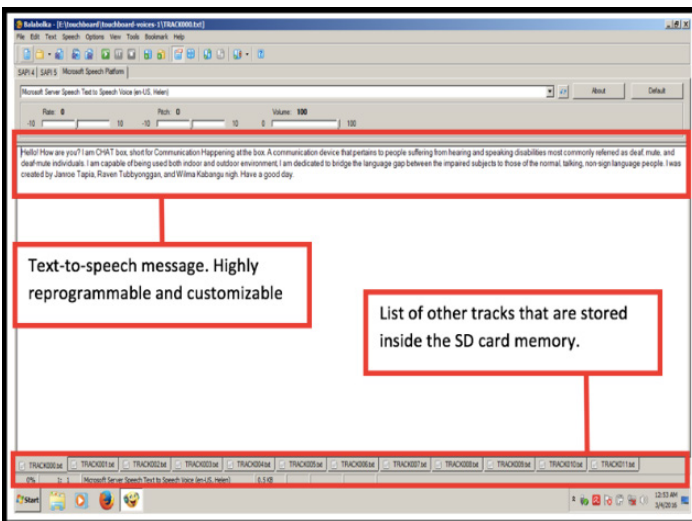
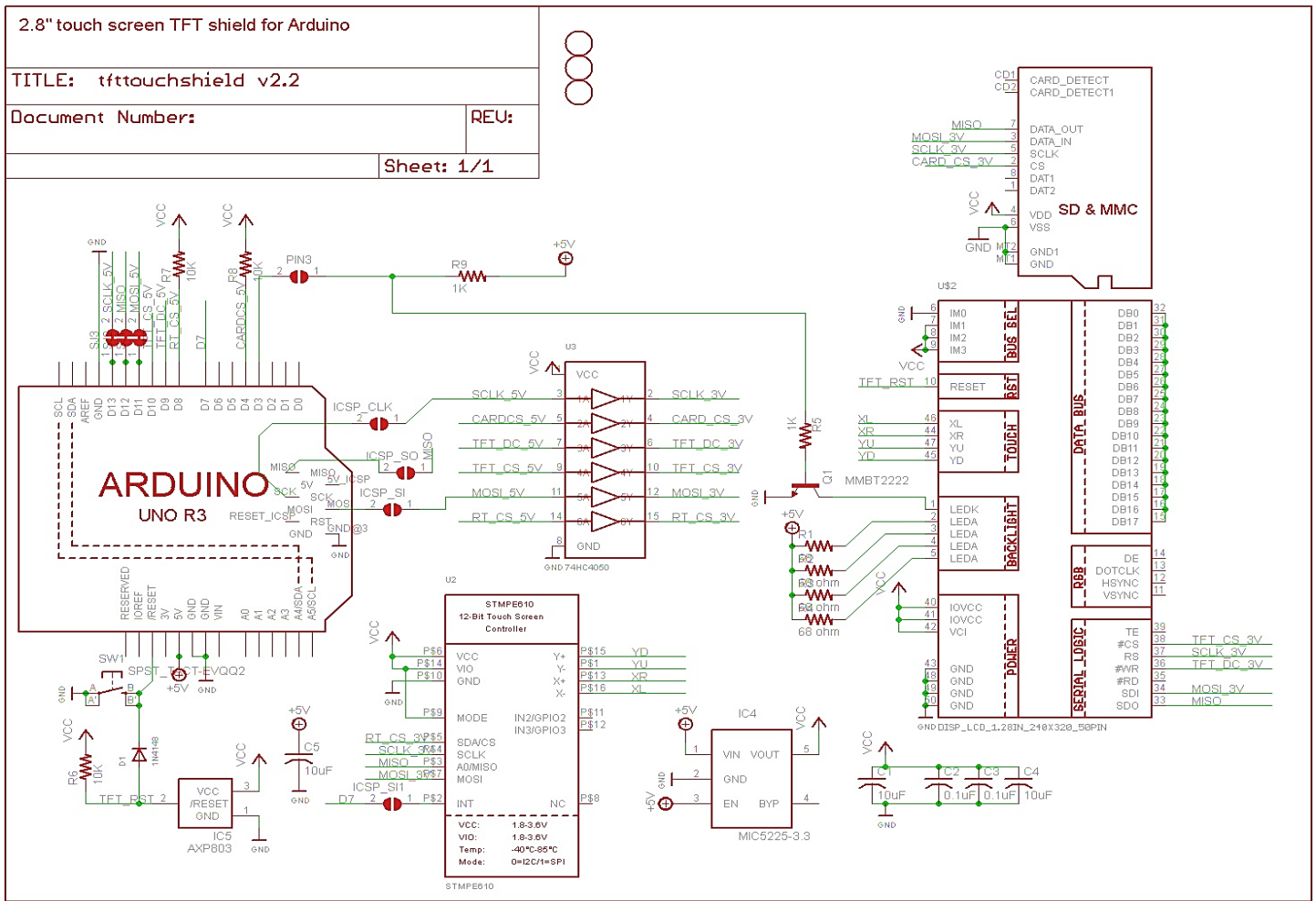


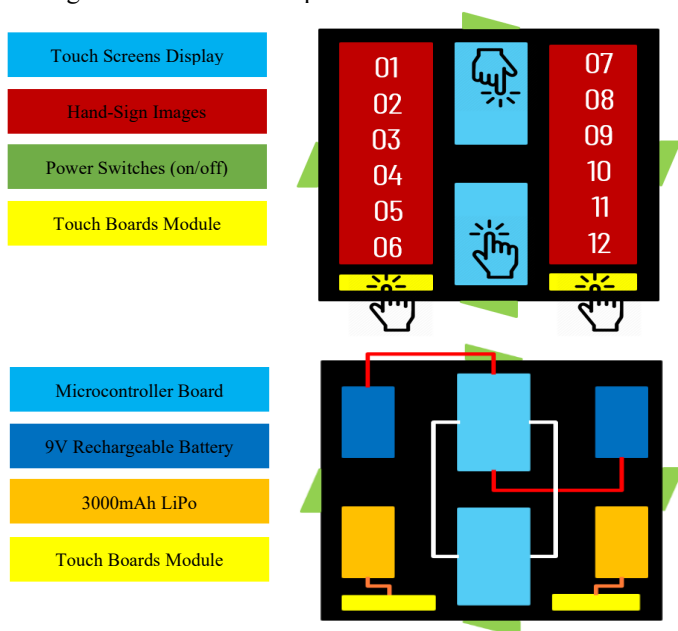
Figure 8: Actual data seen using a third-party program named Balabolka. What shown is the raw text file to be imported as .mp3 format to the micro SD Card which will then be inserted to the Touch Board's compartment.

The system is composed of two microcontroller and two touch boards. It can receive and send sign language images in .bmp

format via touch screens through the microcontrollers' RX-TX serial communication. The system comprises four microcontrollers, one electric paint, two earphones, two rechargeable 9V batteries and two 3.7V lithium-ion rechargeable batteries with 3000mAh capacity and a twin barrel battery charger. (1) A surface consists of twelve common words and/or phrases based from the 1000 most common phrases used); (2) Two touch screens consisting of preloaded standard English alphabet and phrases for the non-sign language users. The impaired subjects have two options either to tap the surface with the signed image of their choice, correspondingly the Touch Board will read the message out loud via built-in speaker or to communicate using the touch screens like texting.

It is a bilateral communication device comprising of the following processes and components: (1) two microcontrollers, two touch boards, configured to act as with powerful processing power with a built-in Li-Po and Li-Ion battery charger. Standard audio jack for speakers and earphones. Composed of twelve electrode solder pads which produce audio when touched or tapped: (2) Innovative User Interface. Two microcontrollers with two mounted screen display module and touch boards Solder Pads; (3) Stable UART Communication. An RX-TX serial connection was utilized to send and receive data from one touch screens to another: (4) On

Top Guide. A surface consisting of twelve hand sign images with their corresponding translations; (5) Text-to-Speech Capable. A third-party text-to-speech program (TTS) called Balabolka, enabling the converted audio in an .mp3 format to be exported within the Touch Board's SD card memory; (6) Lightweight and Sturdy Casing. A lightweight packaging made from stainless alloy was constructed to form a rigid box for indoor and outdoor usage; (7) Accessories Integrated. Two earphones attached on the two touch boards respectively for audio feature and One twin barrel battery charger: (8) Long-lasting Battery Consumption. Two rechargeable 9V 220mAh battery supply for the microcontrollers; Two rechargeable 3.7V 3000mAh li-ion battery; (9) Convenient On & Off Preferences. Two separate rocker switches (on/off) states for the touch screens; Two separate built-in switches (on/off) states found in the touch boards for interface purposes; (10) Non-Water-Resistant. Any nearby water reservoir could damage the device when exposed to.



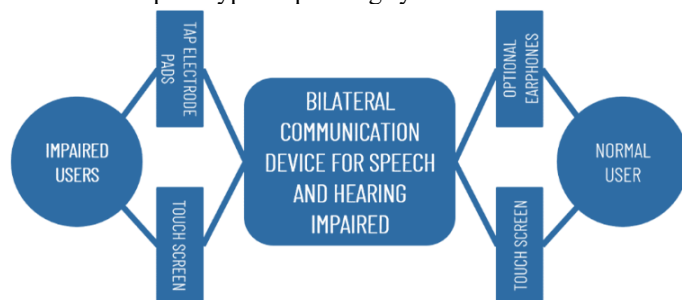
## 5. Perspective Renderings of the Device

Based on design it has an eight by eight (8x8) inch aluminum casing, with four differently sized compartments dedicated for the two 2.8-inch touch shields comprising the thirty percent (30%), two touch boards comprising the ten percent (10%), and display of hand sign images with English phrase translation comprising the sixty percent (60%) of the space. Additionally, there are four (4) separate power switches, one for each microcontrollers and boards, respectively. The device is approximately weighed at around 2.8 lbs. to 3.0 lbs.

## 6. Proposed Methodology

Distress caused by a verbal barrier which hinders the subjects from being able to communicate freely and easily to sign language incapable individuals and vice-versa is the focal point of concern this project aims to resolve. This communication device consists of two major parts: (1) a surface consists of two Bare Conductive Touch Boards preloaded with the twelve phrases and words with

its corresponding display of hand sign images converted to its English translation and (2) two TFT Touch Shields mounted in Arduino Mega 2560 Rev3 respectively, preloaded with the standard English alphabet and American Sign Language symbols. The Deaf-mute have two options either to tap the desired electrode pads on the touch board or to communicate using the TFT touch shield like a multi-touch screen. In this regard, we proposed and developed a system that converts and interprets American sign-language (ASL) hand-sign image to its corresponding English translation and vice versa. Aside from that, it can also produce speech or audio responses via interactive capacity touch and proximity electrode pads. The bilateral communication device promotes two-way communication between Deaf-mute and normal person in a close. The normal person refers to the one who has no hearing or vocal impairment or disability. The main features of the prototype's operating system are listed below.



Users communicate with one another using the device and its peripherals.

### 6.1 Normal to Deaf-Mute Person Communication.

This module takes text message of a normal person as an input and outputs an image file in a .bmp format, that displays sign language images for the Deaf-mute person. Each image file is given tags and indexing. The steps of normal to Deaf-mute person communication are as follows:

- The application takes text of normal person as an input.
- The application converts the text message of a normal person into an image by using the text-to-image conversion program.
- The program matches the text to any of the image tags and index associated with the file and displays the corresponding sign for the Deaf-mute.

### 6.2 Deaf-Mute to Normal Person Communication.

Not everyone has knowledge of sign language. The Deaf-mute have two options either to tap the desired electrode pads preloaded with the most commonly used English words and phrases located on the touch board or to communicate using the TFT touch shield like a multi-touch screen, swiping and tapping hand-sign and illustrative images of English words in alphabetical order by default. A search option is also integrated to the system to find images on the screen faster.

### 6.3 System Flowchart of the Operating System.

It is straightforward bilateral communication device with simple, yet sharp responses made possible through the layer of process it undergoes. As shown in Figure 11, the device starts with tapping

(choosing the words), followed by chatting (forming syntax and semantics), then a process of sending or receiving (decision-making), after that two alternative options follow either to ignore (no) or to reply (yes). To ignore means ending the system flow while replying resumes the entire process creating a loop from tapping to responding, until the user ends it. The reply stage involves the sending and receiving process which constitutes conversion of text message to picture message (normal user to Deaf-mute) or conversion of picture message to text message (Deaf-mute to normal user).

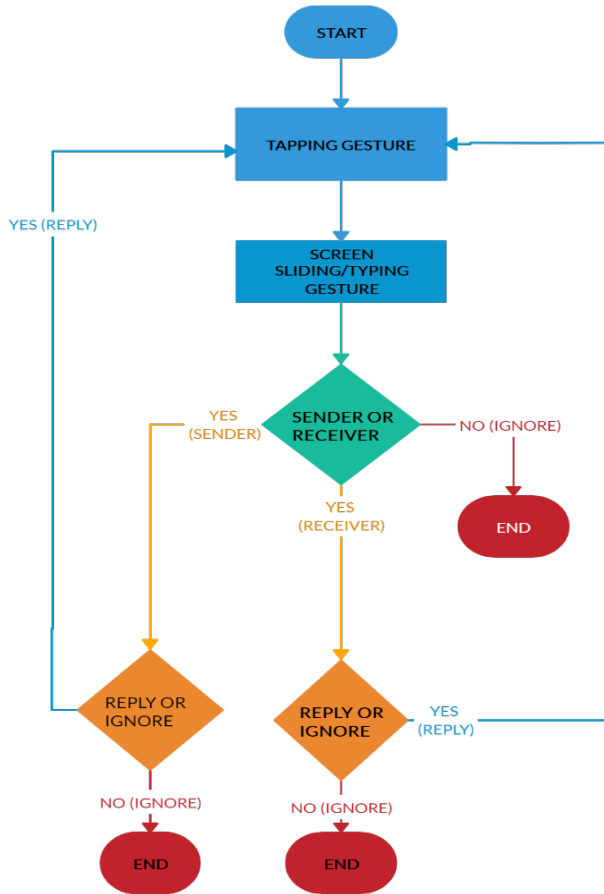


Figure 11: Presents the System Flowchart: it shows the internal logic of the entire system, thus providing the backbone of the how the project works (start and end)

#### 6.4 Framing.

Framing process is used to split the pre-emphasized audio files into short segments. The Bare Conductive Touch Boards consists of capacitive and proximity electrodes that is responsive to touch. The said board also stores audio files via microSD card. The audio files is represented by  $N$  frame samples and the interframe distance or frameshift is  $M$  ( $M < N$ ). In the program, the frame sample size ( $N$ ) = 256 and frameshift ( $M$ ) = 100. The frame size and frameshift in milliseconds are calculated as follows:

$$\text{FrameSize(ms)} = f_n = \frac{1}{N * M} = 25.6 \text{ milliseconds} \quad (1)$$

$$\text{FrameShift} = 10 \text{ milliseconds}$$

#### 6.5 Statistical Treatment of Data.

Weighted Mean is used for calculating the weighted mean statistics for the given data set. If all the values are equal, then the weighted mean is equal to the arithmetic mean. It is a kind of average, wherein instead of each data point contributing equally to the final mean, some data points contribute more weight than the others. If all the weights are equal, then the weighted mean equals the arithmetic mean.

Formula:

$$\bar{x} = \frac{\sum_{i=1}^n (x_i w_i)}{\sum_{i=1}^n (w_i)} \quad (2)$$

where,

- $\Sigma$  = the sum of
- w = the weights
- x = the value

### 7. Experimental Results and Discussion

#### 7.1 Experimental Setup

This communication device consists of two major parts: (1) a surface consists of two Bare Conductive Touch Boards preloaded with the twelve phrases and words with its corresponding display of hand sign images converted to its English translation and (2) two TFT Touch Shields mounted in Arduino Mega 2560 Rev3 respectively, preloaded with the standard English alphabet and American Sign Language symbols. The Deaf-mute have two options either to tap the desired electrode pads on the touch board or to communicate using the TFT touch shield like a multi-touch screen. The device is powered via two separate power supply. Rechargeable 9V DC supply charges up the Arduino microcontrollers and a different rechargeable 3.7V, 3000mAh LiPo powers up the touch boards. Sufficient power supply is vital to ensure the device functionality and features work.

#### 7.2 Message Conversion and Transmission Phase

The working operating system of the bilateral communication device. (a) Write/Text Message Mode allows normal user to type-in English words and sent it to Deaf-mute. (b) Picture Message Mode converts the previously sent text message, into its translation in hand-sign images. (c) Picture Message Mode allows the Deaf-mute to select sign language image sets and sent to normal user. (d) Write/Text Message converts the previously sent picture message, into its English word translation. The operating system treats each image stored in the memory card of touch screen displays, as indexed then call each .bmp image file once its match word is chosen. The same protocol works the other way around from text messages to picture message.

### 7.3 Qualitative Feedback

Researchers formalized questionnaire survey to evaluate the overall performance device through several criteria. The demographics of the respondents includes fifty percent (50%) non sign language users aged 18 to 30 years old and fifty percent (50%) Deaf-mute aged 18 to 30 years old. The researcher divides the population into strata based off the participant’s racial demographics. Here, there are two strata, one for each of the two racial categories. Using this information, a disproportional stratified sample was further implemented. The basis for rating are the following parameters: (1) Value Added – The device was advantageous to the speech and hearing impaired people; (2) Aesthetic Design – The packaging of the device prove to be stable and sturdy; (3) Hardware Components - device’s hardware is properly placed and well organized; (4) Built-in Features – The device has innovative user interfaces like the touch shield making it more interesting; (5) User-Friendliness - The device is easy to use, very convenient, and innovative; (6) Reliability – The data sent and received by the device are consistent and correct; and (7) Accuracy – The data sent and received by the device are exact and true on what was intended to.

### 7.4 Application Interface

Initial bootup of the screens from both ends of the device differs from each other. Normal users and Deaf-mute have their designated position to where they should be at. Display screen of the Deaf-mute contains picture message tab showing the multiple sign language images with English translation at the bottom, in grid view. These images can be a word, a letter, or a phrase. These are arranged alphabetically, in which an additional search bar is added for faster image searches. Meanwhile, display screen of the normal users contains a standard QWERTY keyboard layout. Both screens have sent button, wherein the conversion and transmission of data is initialized and completed once data is received from one end to the other

In this example (a), the message contains the word “hello”. From this point, the normal user will tap “Send Message”, then this will be sent and converted to the corresponding hand sign images on the second TFT Touch Shield for the Deaf-mute users. Accordingly, the previously sent “hello” word was converted into hand sign images as shown at (b).

In this example (c), the image that has the word “all done” was tapped. This is how the Deaf-mute will communicate with the normal user. Consequently, the image that has the word “all done” was received by the TFT Touch Shields from the normal user in an English-based text form (d).

### 7.5 Questionnaire Design

Corresponding interpretation includes: Excellent (4.51-5.00); Very Good (3.51-4.00); Average (2.51-3.50); Fair (1.51-2.50); Poor (1.00-1.50); Very Poor (0.00-0.99). Likert surveys are also

quick, efficient, and inexpensive methods for data collection. They have high versatility and can be sent out fast. In answering the questionnaire, the respondents are group in pairs of normal users and Deaf-mute, for a total of 75 pairs.

The respondents rated the device based on the following criteria:

- a. Value Added
- b. Aesthetic Design
- c. Hardware Components
- d. Built-in Features
- e. User-Friendliness
- f. Reliability
- g. Accuracy

Test cases involved:

- a. Accuracy in transmitting data (back and forth sending and receiving 10 times)
- b. Consistency in transmitting data (back and forth sending and receiving 10 times)
- c. Speed (timing and delay in seconds)
- d. Screen Refresh and Transitions Rates (min. 2 seconds)
- e. Touch Responsiveness (min. 2 seconds)
- f. User Experience (easy-to-use)
- g. User Interface (easy-to-learn)

Table 8: Respondents’ Scale

Range	Interpretation
5	Excellent (4.51-5.00)
4	Very Good (3.51-4.50)
3	Average (2.51-3.50)
2	Fair (1.51-2.50)
1	Poor (1.00-1.50)
0	Very Poor (0.00-0.99)

Table 9: Talled Results based on the Respondents’ Scale

Metrics	Users’ Feedbacks						Weighted Mean	Interpretation
	0	1	2	3	4	5		
Value Added	0	130	10	10	0	0	4.80	Excellent
Aesthetic Design	0	50	100	0	0	0	4.33	Very Good
Hardware Components	0	70	80	0	0	0	4.47	Very Good
Built-in Features	0	90	40	20	0	0	4.47	Very Good
User-Friendliness	0	100	40	0	10	0	4.53	Excellent
Reliability	0	100	30	20	0	0	4.53	Excellent
Accuracy	0	110	20	20	0	0	4.60	Excellent
Overall Rating							4.53	Excellent



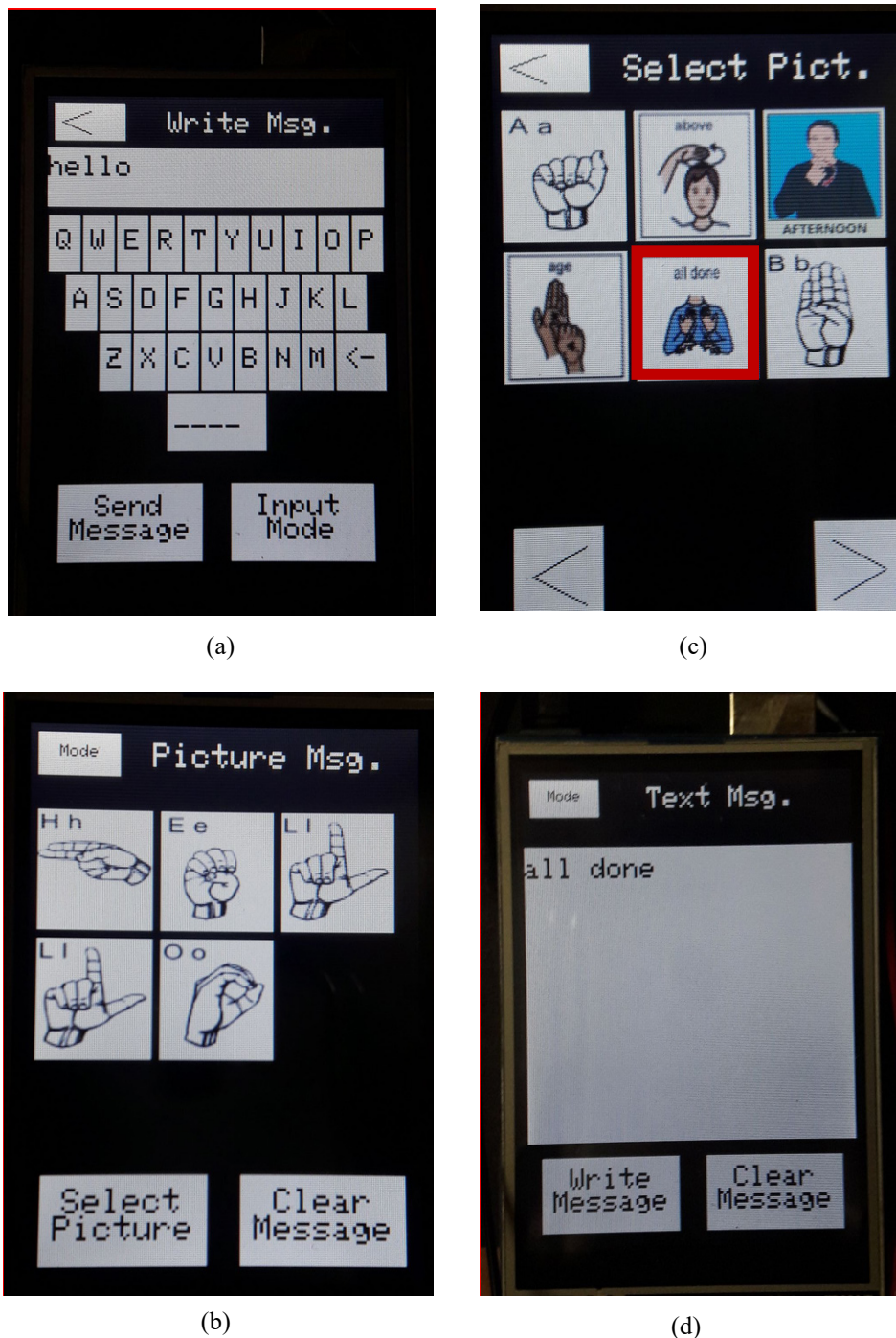


Figure 12: The working operating system of the bilateral communication device. (a) Write/Text Message Mode allows normal user to type-in English words and sent it to Deaf-mute. (b) Picture Message Mode converts the previously sent text message, into its translation in hand-sign images. (c) Picture Message Mode allows the Deaf-mute to select sign language image sets and sent to normal user. (d) Write/Text Message converts the previously sent picture message, into its English word translation.

## 8. Conclusion

Communication is a bilateral process and being understood by the person you are talking to is a must. Without the ability to talk nor hear, a person would endure such handicap. Given that hearing and speech are missing, many have ventured to open new communication methods for them through sign language. This bilateral communication device can be utilized by both non-sign language users and Deaf-mute together in a single system. Shaped as a box (8in x 8in) with two multi-touch capable displays on both ends, the contraption has several microcontrollers and touch

boards within. The latter has the technology of twelve interactive capacity touch and proximity electrode pads that react when tapped, producing quick response phrases audible via speaker or earphone. These touch boards are equipped with an MP3 decoder, MIDI synthesizer, 3.5mm audio jack and a 128MB microSD card. The touch screen modules mounted on top of the microcontrollers transfer data to and from each other in real-time via receiver-transmitter (RX-TX) full duplex UART serial communication protocol. The device is lightweight weighing at about 3 lbs. The prototype device was piloted in an academic institution of special

education for deaf-mute students. Participants were 75 normal and 75 Deaf-mute people aged between 18 and 30 years. The experimental results show the overall rating of the device is 90.6%. The device is designed to promote the face-to-face socialization aspect of the Deaf-mute users to the normal users and vice versa. Several third-party applications were utilized to validate the accuracy and reliability of the device thru metrics of consistency, timing and delay, data transmission, touch response, and screen refresh rates.

### Conflict of Interest

The author declares no conflict of interest.

### Acknowledgment

The proponent would like to acknowledge the Samar State University particularly the College of Engineering and the Center for Engineering, Science and Technology Innovation for the continuous support, resources, and unwavering encouragements in order to achieve milestones and deliverables set for this research undertaking.

### References

- [1] R. Tabiongan, "Chatbox: two-way schemed communication device for speech and hearing impairment," in 2019 8<sup>th</sup> International Symposium on Next Generation Electronics (ISNE), Zhengzhou China, CN, 2019. doi: 10.1109/ISNE.2019.8896577
- [2] T. Humphries and J. Humphries, "Deaf in the time of the cochlea," *Journal of Deaf Studies and Deaf Education*, 16(2), 53-163, 2010. doi: 10.1093/deafen/enq054
- [3] V. Kimmelman, "List of sign languages and their abbreviations," *Information Structure in Sign Languages*, 23-26, 2019. doi: 10.1515/9781501510045-206
- [4] I.W. Leigh and J.F. Andrews, "Deaf people and society: psychological, sociological and educational perspectives", Psychology Press, 2016. doi: 10.4324/9781315473819
- [5] P. Vostanis, M. Hayes, M. Du Feu, and J. Warren, "Detection of behavioural and emotional problems in deaf children and adolescents: Comparison of two rating scales," *Child: Care, Health and Development*, 23(3), 233–246, 1997. doi: 10.1111/j.1365-2214.1997.tb00966.x
- [6] J. Gugenheimer, K. Plaumann, F. Schaub et al., "The impact of assistive technology on communication quality between deaf and hearing individuals," in Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW 2017, 669–682, March 2017. doi: 10.1145/2998181.2998203
- [7] S. Kaur, and K.S. Dhinda, "Design and Development of Android Based Mobile Application for Specially Abled People". *Wireless Pers Commun* 111, 2353-2367, 2020. doi: 10.1007/s11277-019-06990-y
- [8] J.L. Prayandi. K. Jordan, "The deaf community and culture at a crossroads: Issues and challenges," *Journal of Social Work in Disability and Rehabilitation*, 9, 2, 168–193, 2010. doi: 10.1080/1536710X.2010.493486
- [9] S. Barnett, "Communication with deaf and hard-of-hearing people: A guide for medical education," *Academic Medicine: Journal of the Association of American Medical Colleges*, 77(7), 694–700, 2002. doi: 10.1097/00001888-200207000-00009
- [10] E. Rothe and A. Pumariega, "Culture, identity, and mental health: conclusions and future directions" *Oxford Medicine Online*, February 2020. doi: 10.1093/med/9780190661700.001.0001
- [11] M.V. Sharma, N.V. Kumar, S.C. Masaguppi, S. Mn, and D.R. Ambika, "Virtual talk for deaf, mute, blind and normal humans," in Proceedings of the 2013 1st Texas Instruments India Educators' Conference, TIIEC 2013, 316–320, April 2013. doi: 10.1109/TIIEC.2013.63
- [12] F. Soltani, F. Eskandari, and S. Golestan, "Developing a gesture based game for deaf/mute people Using microsoft kinect," in Proceedings of the 2012 6<sup>th</sup> International Conference on Complex, Intelligent, and Software Intensive Systems, ICCISIS 2012, 491–495, July 2012. doi: 10.1109/CISIS.2012.55
- [13] A. K. Tripathy, D. Jadhav, S. A. Barreto, D. Rasquinha, and S.S. Mathew, "Voice for the mute", in Proceedings of the 2015 International Conference on Technologies for Sustainable Development, ICTSD 2015, February 2015. doi: 10.1109/ICTSD.2015.7095846
- [14] N. P. Nagori and V. Malode, "Communication Interface for Deaf-Mute People using Microsoft Kinect," in Proceedings of the 1st International Conference on Automatic Control and Dynamic Optimization Techniques, ICACDOT2016, 640–644, September 2016. doi: 10.1109/ICACDOT.2016.7877664
- [15] A. Sood and A. Mishra, "AAWAAZ: A communication system for deaf and dumb," in Proceedings of the 5<sup>th</sup> International Conference on Reliability, Infocom Technologies and Optimization, ICRITO 2016, 620–624, September 2016. doi: 10.1109/ICRITO.2016.7785029
- [16] M. Yousuf, Z. Mehmood, H.A. Habib et al., "A Novel Technique Based on Visual Words Fusion Analysis of Sparse Features for Effective Content-Based Image Retrieval," *Mathematical Problems in Engineering*, 2018, Article ID 2134395, 13 pages, 2018. doi: 10.1155/2018/2134395
- [17] D. Zhong and I. Defee, "Visual retrieval based on combination of histograms of ac block patterns and block neighborhood," *International Conference on Image Processing*, Atlanta, GA, 2006, 1481-1484, 2006. doi: 10.1109/ICIP.2006.312562
- [18] Z. Mehmood, T. Mahmood, and M. A. Javid, "Content-based image retrieval and semantic automatic image annotation based on the weighted average of triangular histograms using support vector machine," *Applied Intelligence*, 48(1), 166–181, 2018. doi: 10.1007/s10489-017-0957-5
- [19] N. Ali, K.B. Bajwa, R. Sablatnig, and Z. Mehmood, "Image retrieval by addition of spatial information based on histograms of triangular regions," *Computers & Electrical Engineering*, 54, 539–550, 2016. doi: 10.1016/j.compeleceng.2016.04.002
- [20] Z. Mehmood, S.M. Anwar, N. Ali, H.A. Habib, and M. Rashid, "A Novel image retrieval based on a combination of local and global histograms of visual words," *Mathematical Problems in Engineering*, 2016, Article ID 8217250, 1-12, 2016. doi: 10.1155/2016/8217250
- [21] Z. Mehmood, F. Abbas, T. Mahmood, M.A. Javid, A. Rehman, and T. Nawaz, "Content-based image retrieval based on visual words fusion versus features fusion of local and global features," *Arabian Journal for Science and Engineering*, 1–20, 2018. doi: 10.1007/s13369-018-3062-0
- [22] S. Jabeen ,Z. Mehmood, T. Mahmood, T. Saba, A. Rehman, and M. T. Mahmood, "An effective content-based image retrieval technique for image visuals representation based on the bag-of-visual-words model," *PLoS ONE*, 13, 4, e0194526, 2018. doi: 10.1371/journal.pone.0194526
- [23] S. Ghanem, C. Conly, and V. Athitsos, "A Survey on Sign Language Recognition Using Smartphones," in Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments, 171–176, ACM, June 2017. doi: 10.1145/3056540.3056549
- [24] L.B. Chen, C.W. Tsai, W.J. Chang, Y. M. Cheng, and K.S.M. Li, "A real-time mobile emergency assistance system for helping deaf-mute people/elderly singletons," in Proceedings of the IEEE International Conference on Consumer Electronics, ICCE 2016, 45-46, January 2016. doi: 10.1109/icce.2016.7430516
- [25] R. Kamat, A. Danoji, A. Dhage, P. Puranik, and S. Sengupta, "MonVoix-An Android Application for the acoustically challenged people," *Journal of Communications Technology, Electronics and Computer Science*, 8, 24–28, 2016. doi: 10.22385/jctecs.v8i0.123
- [26] A. Kaur and P. Madan, "Sahaaya: Gesture Recognising System to Provide Effective Communication for Specially-Abled People," *2012 IEEE Global Humanitarian Technology Conference*, Seattle, WA, 2012, 93-97, doi: 10.1109/GHTC.2012.65
- [27] D. Bragg, N. Huynh, and R. E. Ladner, "A personalizable mobile sound detector app design for deaf and hard-of-hearing users," in Proceedings of the 18<sup>th</sup> International ACM SIG ACCESS Conference on Computers and Accessibility, ASSETS2016, 3–13, 2016. doi: 10.1145/2982142.2982171

## Survey Analysis: Enhancing the Security of Vectorization by Using word2vec and CryptDB

Hana Yousuf<sup>1</sup>, Said Salloum<sup>1,2,\*</sup>

<sup>1</sup>Faculty of Engineering &IT, The British University in Dubai, 345015, United Arab Emirates

<sup>2</sup>Research Institute of Sciences & Engineering, University of Sharjah, 27272, United Arab Emirates

### ARTICLE INFO

Article history:

Received: 24 June, 2020

Accepted: 17 July, 2020

Online: 28 July, 2020

Keywords:

Natural Language Processing

Vectorisation

Word2vec

CryptDB

### ABSTRACT

Vectorization is extracting data from strings through Natural Language Processing by using different approaches; one of the best approaches used in vectorization is word2vec. To make the vectorized data secure, we must apply a security method, which will be CryptDB. The paper is analyzing the survey, which is created to interview security engineers through the SPSS tool. By analyzing the responses from software security engineers, it is seen that both word2vec and CryptDB works significantly. Word2vec is an effective vectorization approach, while CryptDB is an effective, secure database. In future work, we will be developing a secure vectorization using both approaches.

## 1. Introduction

A vector is a set of real numbers that is converted from a string. This is performed to extract useful data from a specific work using Natural Language Processing (NLP) techniques through deep learning and machine learning techniques [1]–[6]. This conversion is known as vectorization or word embedding. It converts multiple words or an entire phrase to the corresponding vector format of real numbers that are utilized to predict the words and similarities. This would be helpful for classifying texts, identifying similar words, clustering the documents, or extracting the features [7]. On the whole, vectorization is a term referring to the conversion of scalar instructions to vector instructions. The data is stored in vector format, where the loop is reconfigured in such a way that multiple elements can be processed in place of a single array. Vectorisation is used for executing systematic computation where large amounts of data must be processed effectively. Since the size of the code is reduced, the number of bugs is also reduced. It is also easier to substitute the values in mathematical expressions, thereby making the calculation easier and also easier to read and the data can be stored in compact manner.

Technically speaking, in word embeddings, the individual words are characterized as vectors with real value in a predefined cell [8]. The individual words are mapped to their respective vectors. The values of these vectors are trained similarly to neural networks; hence they are often grouped under deep learning

techniques. Representations are provided based on the similarities between the words, thereby understanding their meanings. An embedding layer is a part of the work embedding which is trained jointly with neural network models on specific NLP like classifying the documents or language modelling [9]. The document containing the lexical data must be cleaned so that the individual words are not encoded. They are initially initialized with smaller random numbers. The method of training embedding layers needs a large amount of training data. This approach of learning an embedding layer requires a lot of training data and can be slow, but will learn an embedding both targeted to the specific text data and the NLP task.

There are many word embedding approaches where the models have been pre-trained. Some of them are word2vec, Glove, and Fast text from Google, Stanford, and Facebook, respectively [10]. Word2vec was developed by Toma Mikolov and other researchers from Google in 2013 [11]. It is necessary to understand why word2vec was created. The other NLP systems at that period used to consider the words as individual atomic units. There were no similarities between the words, and the models were not effective for smaller datasets. Hence, larger datasets with complicated models have used neural network approaches for training complicated data models and works better for larger datasets with lots of vocabulary. Therefore, since the neural network is used in word2vec, it works effectively in measuring the quality of the resulting vectorized contents. Hence, the degree of similarity is considered for identifying similar words and grouping them.

\*Corresponding Author: Said Salloum, University of Sharjah, UAE. Tel: +971507679647 Email: ssalloum@sharjah.ac.ae



In order to improve the architecture for the word representations, the accuracy must be improved, and the computational complexity must be minimized. Hence, neural network models like Recurrent and Feedforward neural networks are linked to Net Language Model (NLM) [12]. They can be trained through Backpropagation or Stochastic gradient descent. In FFNNLM, it contains the layers like input, hidden, and output layers like any other work. In addition, it also contains the projection layer. Since the computation between hidden layers and projection layers are complicated, the data in the projection layer becomes denser [13]. In RNNLM, the model has more complicated patterns and layers; hence projection layer is not required in this model. Only the three basic layers of the deep network are sufficient.

There are two different models for learning that can be part of the word2vec technique for learning the word embeddings. They are:

- Continuous Bag-of-Words (CBOW)
- Continuous Skip-Gram

The former gets trained based on the prediction of the present word based on its context, while the latter gets trained by learning the surrounding words, which will lead to the present word. Both of them focus on learning the words concerning their context and usage and also the configurable parameters of the models. The major benefit of the technique is that higher-quality word embeddings may be trained effectively with reduced space and time requirements. This allows bigger embeddings to get trained with higher magnitudes and more lexical data. The architecture of the word2vec algorithm is given in Figure 1.

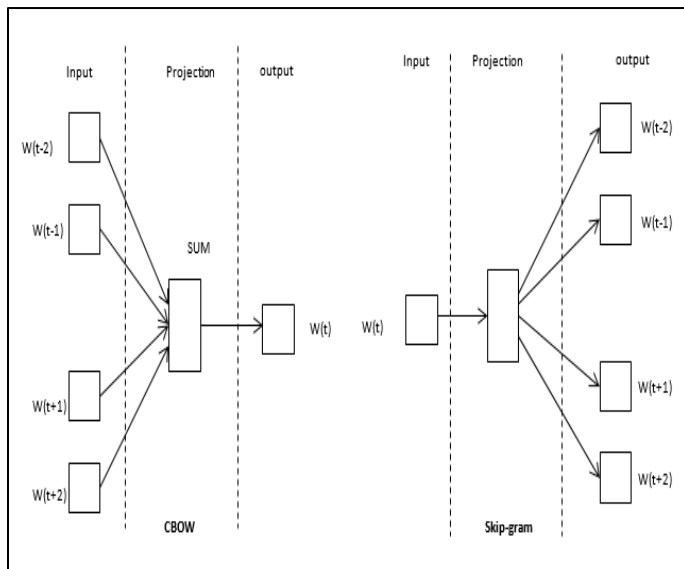


Figure 1: Architecture of word2vec

Network Functions Virtualisation (NFV) is a network architecture that utilizes IT virtualization to virtualize the entire classes of network node functions into structural blocks [14], [15]. Firewalls and network intrusions detections are necessary to provide security to the hardware. Hence these must be utilized in the application layer of the system [16]. Vectorisation is gradually taking center stage, which is being utilized in hardware equipment [17]. Newer processors are altering their architectures to improve

the security for vectorization [18]. In these CPUs, it is obligatory to achieve vectorization for attaining higher performance [19]. When vectorized query execution is utilized, the processes are streamlined by processing multiple blocks at the same time. This may usually be 1024 blocks. Within the block, every column is saved as a single array, i.e., vector. Simple arithmetic operations and comparisons are performed by rapid iterations of vectors in a tight loop with very few function calls or conditional branches in the loop. These loops compile in a streamlined manner with fewer instructions and complete the instructions with a reduced clock cycle by utilizing the processors and cache memory.

The databases may be vulnerable to sensitive information or theft and can be attacked since the attackers can take advantage of the bugs in order to gain access to private information. Also, the attackers may seize and leak the data to those who require them. Hence, the confidentiality and piracy of the database must be safeguarded. CryptDB provides confidentiality against the attackers through SQL databases [20]. The query processes in the SQL are encrypted by collecting effective SQL aware encryption schemes. The process can also be used for encrypting the credential keys so that an item can be decrypted with merely a password. Since the processes and data are encrypted, the database administrator will not be able to view the data. In case of an attack on the server, the invader would not be able to gain access to the decrypted data. The database is capable of handling multiple queries simultaneously with fewer maintenance costs.

## 2. Literature Review

A study by [21] has assessed security risks through both quantitative and qualitative analysis. A critical review of the existing standards has been performed [22], [23], and then the results of the quantitative data are compared to qualitative data. From the analysis, it has been identified that long term security was preferred by the security engineers, and the financial impact of the information security is difficult to calculate quantitatively. A similar analysis has been performed by [24] on the performance of information security and security awareness of the engineers. Interviews have been performed in an IT company for exploring the personal role of information security workers. It was learned that most of the workers did not have any individual security responsibilities; higher workload increases the conflict between security and performance. They must be involved more into the processes to have a better-balanced improvement in security and performance.

The different types of databases have been discussed in [25] and have focussed on the CryptDB database. The type that provides the most secure among the different types of CryptDB database is Random (RND). This type is attributed to providing security against under a probabilistic adaptive plain text attack. The available data has been split and then mapped into the respective ciphertexts using the vectorized process. Ransom makes sure that the calculations are conducted directly on the ciphertexts. Encryption approaches like AES and Blowfish are performed simultaneously in RND with a random initialization vector. AES is used more than Blowfish for 64-bit block sizes since AES has 128-bit blocks that cause the ciphertexts to be very long.

Researchers of [26] has proposed the vectorization in the form of cryptographic security and digital map in this implementation.



Data encryption has been used to provide security to the network in this work. An iterative encryption organization has been proposed for compressing the vector domain in copy protection and retrieving the control of vector data. The cipher size is reduced by encrypting the various processes through Minimum Coding Object (MCO). The mean points of the data are permuted with position-based encryption, and then the secondary data is fit inside.

A Ciphertext policy with consistent size has been introduced by [27] known as Cipher-text Policy Hidden Vector Encryption (CP-HVE). This type of encryption is unique, where the IBE gets anonymous and utilizes the identity as the main parameters where the attributes linked with the ciphertext or secret keys contain wild-cards. Different arrangements have been utilized, one with a composite ordered bi-linear group while it is a prime ordered bilinear group. These schemes offer higher security by distinguishing the plaintext and ciphertext. This vectorized encryption can have a uniform size than the other HVE methods.

Scholars of [28] has performed quantitative analysis using word2vec to form semantic forms of Naturalistic Driving Data (NDD). Since large amounts of data must be compressed and create compact versions without any loss, the performance has been carried out. Quantitative analysis has been performed to identify the semantic link between the sequences in words. [29] has proposed an automatic intrusion detection system known as MalDozer for the Android operating system. It uses deep learning techniques for classifying the sequences and then word2vec has been used for API for detecting the patterns and malware in the system. The proposed approach has been implemented on servers and evaluated on multiple devices. The result shows the proposed approach and showed high accuracy

According to a study by [30] has also performed quantitative analysis for sentiment analysis. Word2vec method has been used for extracting the features, and then Principal Component Analysis (PCA) was used to identify the important elements and structure. Various classifiers have been compared for performing the work and evaluating the results. In [31] has conducted research for the holistic management of cybersecurity by both quantitative and qualitative methodologies. From the analysis, it was concluded that small and medium-sized businesses had no structural mechanisms to reduce security risks. Instead, they applied the solutions after the breach takes place. Hence, there are not equipped to handle the risks and threats beforehand. The main elements of the framework are identified as external factors, assessing the risks and organizational behaviors.

A study by [32] has proposed a one-shot learning approach for the rapid detection of the attacks. The implementation has been performed for computer vision for recognizing images and texts. Memory Augmented Neural Network has been used in this work, combined with malware detection. Word2vec was used for converting the Application Programming Interface (API) sequences into numeric vectors before getting fed into the one-shot learning network. High accuracy was obtained due to the vectorization using word2vec.

In [33] has proposed a malware detection approach, namely DroidVecDeep, that uses a deep learning approach. Different features are extracted and then ranked by utilizing the Mean Decrease Impurity approach. The features are then transformed

into compact vectors on the basis of word2vec. A deep learning model is used to train the classifier, and various malware detection methods are compared. From the results, it is seen that the proposed approach works better than other detection techniques. Hence, vectorization using word2vec is an effective algorithm with deep learning.

A study by [34] has performed group encryption with lattice assumptions through the relation between the matrix vectors. The implementation has been performed through the proposed model, and it was seen that the proposed model provided higher security with certain assumptions. Public key encryption has been used to demonstrate the processes of the novel methodology. However, only generic data has been vectorized rather than the processes.

In [35] has also utilized word2vec for constructing vectors with context sentences and sense the definition vectors giving scores to individual words. WordNet database has been used for retrieving the words and structure and get trained. Scores are based on specific thresholds, and they will be combined with the probability of SEMCOR, which is a sense tagged corpus. It is seen from the results that the performance is better without the probability of distribution. A study by [36] has developed a framework for improving network security. The related studies that focus on network security have been studied with respect to changes to ensure security, passive defenses, the formation of strategy, and overall trend prediction. A detailed critical analysis has been performed by comparing the related articles. The process has included representations, analysis of the solution, predicting the situation, and acquiring the factors.

Encryption is usually performed in such a way that there is very little loss in the vectorized data. A study by [37] has implemented word2vec with Long Short-Term Memory (LSTM) approach for provided security from the attacks. The opcodes have been analyzed in the executable files and maybe clustered using the LSTM network. Word2vec was used for converting the names of the API through one hot encoding results in higher dimensional vectors since the individual cases are represented using the individual dimensions. The proposed approach had better performance than a one-hot encoding-based approach. The issues with data encryption and decryption have been analyzed by [38] through a questionnaire-based methodology. An overview of different schemes for single and multiple secrets of image data has been performed. The different factors that are taken into consideration for the encryption are contrast, capacity, number of shares, type of shares, accuracy, security, and the image format. From these factors, the issues have been identified and have been suggested to address them through novel methods.

Authors of [39] has analyzed the trends in blockchain technology by using word2vec in combination with Latent Semantic Analysis (LSA). Most existing research work has used effort demanding a complete textual investigation and traditional bibliometric approaches. Hence, k means clustering approach has been used for capturing the lexical context. Annual trend analysis has been used for various chains of blockchain. It is seen that the proposed approach has higher accuracy with respect to quantitative methodology. A qualitative approach has also been performed.

From the above review of literature, it can be seen that there are not many studies performing quantitative analysis through

interviews and questionnaires. Most of the studies have directly improved the work by trying to implement the algorithms; however, there are hardly any studies that analyze the management problem by interviewing the public on this. Hence, a comprehensive methodology is required to directly interview the concerned people and collect the data to understand the problem faced and create an effective solution.

### 3. Methodology

In this work, the quantitative methodology is used to collect the data and analyze the problem. The quantitative methodology can be defined as a type of analysis that uses data that can be easily quantifiable to perform mathematical, computational, and statistical approaches [40]–[52]. Data is collected from the potential sample population through any of the sampling methods in the form of questionnaires [53]. These sets of questions may be asked directly or through the internet in the form of polls or mails. The responses are always converted to the numerical form in such a way that they are quantifiable. The implementation uses interpretivism to understand the problem and gain solutions. Interpretivism is a type of methodology that the researcher implements for synthesizing the facts that are derived from the sources. These facts depend on lots of factors that cannot be measured physically, or they are difficult to measure. The factors may fall under social, cultural, or economic categories. This is analyzed to realize whether the database or security engineers know how to provide security to the database that is vectorized.

The quantitative strategy is applicable for this research since the link between the various variables can be set by interpreting them [54]. This requires interviews to be conducted through questionnaires. The questionnaire addressed the software engineer, who is the most suitable discipline to respond and address the issues in dealing with wrd2vec and CryptDB. Questionnaires are selected for this analysis since they are more reliable and can collect data from lots of respondents quickly and effectively. This is essentially true for larger projects with thousands of respondents. The major disadvantage, however, is that the format is fixed for the entire population and cannot be customized depending on the respondent and this removes the likelihood of in-depth analysis. However, this work does not require variable responses, and hence questionnaires are sufficient. A survey contains various questions relating to the knowledge of vectorization, and security is distributed to the engineers of different companies. They are selected based on the different available companies identified online and their engineers are contacted through email. The responses are captured through online forms and surveys and then compiled together. A total of 30 responses are captured in this way. No sampling strategy was followed here, and only one group of respondents is identified. The questionnaire consists of eleven questions divided into three groups. The groups are dedicated to experience, security, and time. The experience gets from the data related to the educational background and work experience of the respondents. Security aspect relates to the knowledge of the correspondence on security of word2vec and vectorization. Time measures the ability of the security engineers to manage time in performing any changes to the design of the database. The results from the questionnaire are compiled into a single file. Since the number of respondents is very small, detailed analysis is not

required. In this work, SPSS software is used for the quantitative analysis.

### 4. Data Analysis

From the collected results, the analysis performed through SPSS tool is discussed in this section. Table 1 gives details of the qualification of the respondents. A major portion of them is undergraduates at 53.5%, followed by postgraduates at 30%, doctorates at 10%, and certification courses at 6.7%. Table 2 explains the work experience of the respondents. Most of the respondents have experienced between 4 and 8 years, while only 2 respondents had experienced less than four years, suggesting that most of the respondents are well experienced. This is because they are in a senior position. Around 30% of the respondent had more than nine years' experience, out of which 10% had more than 15 years of experience. The respondents who have completed certificate courses have a very high experience, which conveys that only graduates are allowed to become security engineers in the recent decade.

Table 1: Frequency of Qualification

Qualification	Number of respondents	Percent
Computer certification course	2	6.7
Undergraduate	16	53.3
Postgraduate	9	30.0
Doctorate	3	10.0
Total	30	100.0

Table 2: Frequency of Experience

Experience	Number of respondents	Percent
0-3 years	2	6.7
4-8 years	19	63.3
9-15 years	6	20.0
More than 15 years	3	10.0
Total	30	100.0

The ability to solve the security problems in vectorization is given in Table 3. It is seen that most of the respondents are able to solve security issues, with 60% of them pertaining to certain problems, while 33.3% were confident of solving all the problems. Two of the respondents said that they are unable to solve security issues. While more than 90% of the respondents are able to solve security problems, only 66.7% of them have used the CryptDB database for security, as shown in Table 4. The data is also visualized in Figure 2.

Table 3: Ability to solve the security issues during vectorization

Ability to solve the security issues during vectorization	Number of respondents	Percent
Unable to solve the problems	2	6.7
Can solve certain problems	18	60.0
Can effectively solve all the problems	10	33.3
Total	30	100.0

Table 4: Frequency of use of CryptDB database

Use CryptDB database	Number of respondents	Percent
Yes	20	66.7
No	10	33.3
Total	30	100.0

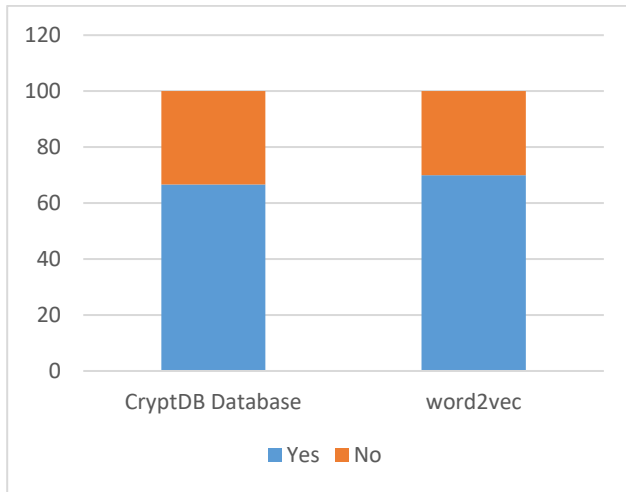


Figure 2: Use of CryptDB and word2vec

A large portion of the respondents has used word2vec embedding for vectorization. Around 70% of them have used them, while the other 30% have either used other types of embedding techniques or never used vectorization as shown in Table 5.

Table 5: Frequency of use of Word2ved embedding

Do you use Word2vec embedding?	Number of respondents	Percent
Yes	21	70.0
No	9	30.0
Total	30	100.0

Since word2vec can be performed easily, it is implemented; however it takes time for embedding them along with the secure database. Only a few of the respondents could complete the implementation within a day, while almost three quarters of the respondents will take from 2 to 7 days. 13.3% of the respondents will take more than a week to complete as shown in Table 6.

Table 6: Time taken for Word2ved embedding

Time is taken for Word2ved embedding	Number of respondents	Percent
Within a day	4	13.3
Takes about a week	22	73.3
Takes more than a week	4	13.3

Once the database is secured, then there might be possibilities of making changes to the design or contents in the work. This might be easy for some minor modifications, where the contents can be modified easily. However, for adding a new component or embed a new element, the design has to be modified, which would take some time. Most of the respondents at around 86.7% of them

responded that it would take a long time for them to make changes in the design, only 13.3% of them said it could be done quickly as shown in Table 7.

Table 7: Frequency of changes be made in the design quickly if required

Can any changes be made in the design quickly if required	Number of respondents	Percent
Yes, design can be changed quickly	4	13.3
It takes a long time to make design changes	26	86.7
Total	30	100.0

Even though most respondents have used CryptDB and word2vec, they do not use it in their everyday results. Table 8 shows whether the respondents use CryptDB and word2vec regularly.

Table 8: Frequency of the usage of the concepts

Which one do you use regularly?	Number of respondents	Percent
CryptDB	10	33.3
Word2vec	9	30.0
Both	11	36.7

The mean and the standard deviation for the security is also given in Table 9 for the CryptDB database and word2vec embedding. The p values are seen to be at 0.002, which is considered to be a significant value.

Table 9: Difference in mean security between types of database

	CryptDB database	Word2vec	Both	p-value
	Mean ± SD			
Security	2.73±0.90	2.63±0.51	3.64±0.35	0.002**

Table 10 gives the opinion of the engineers related to database security. Most of the engineers strongly agree that their database is secure. 40% of the respondents strongly agree, followed by 26.7% agreeing that their database is secure. On the other hand, 32% of the respondents do not agree to this and feel that security must be improved. The majority of the respondents also feel that CryptDB would improve security in the database. 66.7% of the respondents agree, while the other 33.3% disagree that CryptDB improves the security features. While 40% of the respondents strongly agree, only 3.3% of them strongly disagree with CryptDB. Most of the respondents also agree that a secure database must be used for secure vectorizations. 86.7% of the respondents agree that a secure database must be used, while 13.3% of them disagree that security is not required.

The descriptive statistics for the database security are shown in Table 11. Out of the 4-point scale from strongly disagree to strongly agree, a mean of 3.3 is seen along with a standard deviation of 0.76 for the 30 respondents. This is more inclined towards the right with more people agreeing to the conditions in table 10. Since the mean is near 3, more people just agree to the

points while some of them strongly agree. The frequency of disagreeing and strongly disagree is also less.

Table 10: Database security

		N	%
My database is secure	Strongly Disagree	5	16.7
	Disagree	5	16.7
	Agree	8	26.7
	Strongly Agree	12	40.0
CryptDB database improve the security features	Strongly Disagree	1	3.3
	Disagree	9	30.0
	Agree	8	26.7
	Strongly Agree	12	40.0
Secure database must be used for vectorization?	Strongly Disagree	1	3.3
	Disagree	3	10.0
	Agree	16	53.3
	Strongly Agree	10	33.3

Table 11: Descriptive statistics for security

	N	Minimum	Maximum	Mean	SD
Security	30	1.33	4.00	3.03	0.76

## 5. Conclusions and Recommendations

From the above analysis, it is identified that most of the scientists think that security is necessary for the databases. Even though there are a lot of secure databases and word embedding options, however, around 63% of the interviewed engineers use CryptDB for security purposes, and around 66% use word2vec for embedding options. While many scientists agree that security must be improved in the databases, most scientists completely agree that CryptDB can definitely improve security. Also, most of the engineers accept that word2vec is an effective way to completely improve the vectorization. This work has presented a quantitative methodology for understanding the link between the word vectorization approach word2vec and providing security to the database by using CryptDB. Initially, a brief introduction to both these terminologies concerning vectorization was studied, and then related studies have been studied in the literature review. A questionnaire was created to perform interviews on security engineers. Data is collected in the form of online forms, and then, the data is analyzed through SPSS tool. From the data, it is identified that both word2vec and CryptDB is recommended for embedding the words and security, respectively. The simultaneous use of both of them is suggested as an effective combination for secure vectorization. Hence, the implementation of both of them together is considered as future scope of this work.

## 6. Limitation and Caveats

The gap identified is that even though many researchers have individually implemented them both. They haven't been

implemented together. Also, most of the studies have implemented them technically, but haven't performed interview-based analysis. The limitation of this work was due to the lockdown that happened in March due to COVID-19, and the analysis was conducted on a small scale group of security engineers, the number of correspondents considered in the study was little. In future work, the study will consider both quantitative and qualitative methods. In this study, there are no caveats discussed in this paper.

## References

- [1] H. Zamani and W. B. Croft, "Relevance-based word embedding," in *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2017, 505–514. <https://doi.org/10.1145/3077136.3080831>
- [2] S. A. Salloum, M. Alshurideh, A. Elnagar, and K. Shaalan, "Mining in Educational Data: Review and Future Directions," in *Joint European-US Workshop on Applications of Invariance in Computer Vision*, 2020, 92–102. [https://doi.org/10.1007/978-3-030-44289-7\\_9](https://doi.org/10.1007/978-3-030-44289-7_9)
- [3] S. A. Salloum, R. Khan, and K. Shaalan, "A Survey of Semantic Analysis Approaches," in *Joint European-US Workshop on Applications of Invariance in Computer Vision*, 2020, 61–70. [https://doi.org/10.1007/978-3-030-44289-7\\_6](https://doi.org/10.1007/978-3-030-44289-7_6)
- [4] M. Al-Emran, S. Zaza, and K. Shaalan, "Parsing modern standard Arabic using Treebank resources," in *2015 International Conference on Information and Communication Technology Research, ICTRC 2015*, 2015. [10.1109/ICTRC.2015.7156426](https://doi.org/10.1109/ICTRC.2015.7156426)
- [5] C. Mhamdi, M. Al-Emran, and S. A. Salloum, *Text mining and analytics: A case study from news channels posts on Facebook*, vol. 740. 2018. [https://doi.org/10.1007/978-3-319-67056-0\\_19](https://doi.org/10.1007/978-3-319-67056-0_19)
- [6] M. Al-Emran, "Hierarchical Reinforcement Learning: A Survey," *Int. J. Comput. Digit. Syst.*, 4(2), 137–143, 2015. <https://doi.org/10.12785/IJCD/040207>
- [7] H. Zamani and W. B. Croft, "Estimating embedding vectors for queries," in *Proceedings of the 2016 ACM International Conference on the Theory of Information Retrieval*, 2016, 123–132. <https://doi.org/10.1145/2970398.2970403>
- [8] O. Press and L. Wolf, "Using the output embedding to improve language models," *arXiv Prepr. arXiv1608.05859*, 2016. <https://arxiv.org/abs/1608.05859>
- [9] K. Hashimoto, C. Xiong, Y. Tsuruoka, and R. Socher, "A joint many-task model: Growing a neural network for multiple nlp tasks," *arXiv Prepr. arXiv1611.01587*, 2016. <https://arxiv.org/abs/1611.01587>
- [10] D. Sarkar, "The Promise of Deep Learning," in *Text Analytics with Python*, Springer, 2019, 631–659. [https://doi.org/10.1007/978-1-4842-4354-1\\_10](https://doi.org/10.1007/978-1-4842-4354-1_10)
- [11] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv Prepr. arXiv1301.3781*, 2013. <https://arxiv.org/abs/1301.3781>
- [12] S. Ghosh, M. Chollet, E. Laksana, L.-P. Morency, and S. Scherer, "Affectlm: A neural language model for customizable affective text generation," *arXiv Prepr. arXiv1704.06851*, 2017. <https://arxiv.org/abs/1704.06851>
- [13] T.-H. Yang, T.-H. Tseng, and C.-P. Chen, "Recurrent neural network-based language models with variation in net topology, language, and granularity," in *2016 International Conference on Asian Language Processing (IALP)*, 2016, 71–74. [10.1109/IALP.2016.7875937](https://doi.org/10.1109/IALP.2016.7875937)
- [14] Y. Li and M. Chen, "Software-defined network function virtualization: A survey," *IEEE Access*, 3, 2542–2553, 2015. [10.1109/ACCESS.2015.2499271](https://doi.org/10.1109/ACCESS.2015.2499271)
- [15] R. Mijumbi, J. Serrat, J.-L. Gorricho, N. Bouten, F. De Turck, and R. Boutaba, "Network function virtualization: State-of-the-art and research challenges," *IEEE Commun. Surv. tutorials*, 18(1), 236–262, 2015. [10.1109/COMST.2015.2477041](https://doi.org/10.1109/COMST.2015.2477041)
- [16] J. F. Kurose, *Computer networking: A top-down approach featuring the internet*, 3/E. Pearson Education India, 2005.
- [17] "Intel Corporation (2015). Intel Vectorization Tools. [Online].," 2015.
- [18] "Intel (2013). Intel® Xeon Phi™ Processors. [Online].," 2013.
- [19] C. Stylianopoulos, L. Johansson, O. Olsson, and M. Almgren, "CLort: High Throughput and Low Energy Network Intrusion Detection on IoT Devices with Embedded GPUs," in *Nordic Conference on Secure IT Systems*, 2018, 187–202. [https://doi.org/10.1007/978-3-030-03638-6\\_12](https://doi.org/10.1007/978-3-030-03638-6_12)
- [20] R. A. Popa, C. M. S. Redfield, N. Zeldovich, and H. Balakrishnan, "CryptDB: processing queries on an encrypted database," *Commun. ACM*, vol. 55, no.



- 9, 103–111, 2012.
- [21] A. Munteanu, "Information security risk assessment: The qualitative versus quantitative dilemma," in *Managing Information in the Digital Economy: Issues & Solutions-Proceedings of the 6th International Business Information Management Association (IBIMA) Conference*, 2006, 227–232. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=917767](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=917767)
- [22] S. A. Salloum, M. Alshurideh, A. Elnagar, and K. Shaalan, "Machine Learning and Deep Learning Techniques for Cybersecurity: A Review," in *Joint European-US Workshop on Applications of Invariance in Computer Vision*, 2020, 50–57. [https://doi.org/10.1007/978-3-030-44289-7\\_5](https://doi.org/10.1007/978-3-030-44289-7_5)
- [23] S. K. Yousuf H., Lahzi M., Salloum S.A., "Systematic Review on Fully Homomorphic Encryption Scheme and Its Application.," *Al-Emran M., Shaalan K., Hassantien A. Recent Adv. Intell. Syst. Smart Appl. Stud. Syst. Decis. Control. vol 295. Springer, Cham*, 2021. [https://doi.org/10.1007/978-3-030-47411-9\\_29](https://doi.org/10.1007/978-3-030-47411-9_29)
- [24] E. Albrechtsen, "A qualitative study of users' view on information security," *Comput. Secur.*, 26(4), 276–289, 2007. <https://doi.org/10.1016/j.cose.2006.11.004>
- [25] R. A. Popa, C. M. S. Redfield, N. Zeldovich, and H. Balakrishnan, "CryptDB: protecting confidentiality with encrypted query processing," in *Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles*, 2011, 85–100. <https://doi.org/10.1145/2043556.2043566>
- [26] B.-J. Jang, S.-H. Lee, and K.-R. Kwon, "Perceptual encryption with compression for secure vector map data processing," *Digit. Signal Process.*, vol. 25, 224–243, 2014. <https://doi.org/10.1016/j.dsp.2013.09.013>
- [27] T. V. X. Phuong, G. Yang, and W. Susilo, "Efficient hidden vector encryption with constant-size ciphertext," in *European Symposium on Research in Computer Security*, 2014, 472–487. [https://doi.org/10.1007/978-3-319-11203-9\\_27](https://doi.org/10.1007/978-3-319-11203-9_27)
- [28] Y. Fuchida, T. Taniguchi, T. Takano, T. Mori, K. Takenaka, and T. Bando, "Driving word2vec: Distributed semantic vector representation for symbolized naturalistic driving data," in *2016 IEEE Intelligent Vehicles Symposium (IV)*, 2016, 1313–1320. [10.1109/IVS.2016.7535560](https://doi.org/10.1109/IVS.2016.7535560)
- [29] E. B. Karbab, M. Debbabi, A. Derhab, and D. Mouheb, "Android malware detection using deep learning on API method sequences," *arXiv Prepr. arXiv:1712.08996*, 2017. <https://arxiv.org/abs/1712.08996>
- [30] X. Ge, X. Jin, and Y. Xu, "Research on sentiment analysis of multiple classifiers based on word2vec," in *2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 2018, vol. 2, 230–234. [10.1109/IHMSC.2018.10159](https://doi.org/10.1109/IHMSC.2018.10159)
- [31] J. Jung, "A Study of Cyber Security Management within." University of Portsmouth, 2018.
- [32] T. K. Tran, H. Sato, and M. Kubo, "One-shot Learning Approach for Unknown Malware Classification," in *2018 5th Asian Conference on Defense Technology (ACDT)*, 2018, 8–13. [10.1109/ACDT.2018.8593203](https://doi.org/10.1109/ACDT.2018.8593203)
- [33] T. Chen, Q. Mao, M. Lv, H. Cheng, and Y. Li, "DroidVecDeep: Android Malware Detection Based on Word2Vec and Deep Belief Network.," *TIIS*, 13(4), 2180–2197, 2019.
- [34] B. Libert, S. Ling, F. Mouhartem, K. Nguyen, and H. Wang, "Zero-knowledge arguments for matrix-vector relations and lattice-based group encryption," *Theor. Comput. Sci.*, 759, 72–97, 2019. [https://doi.org/10.1007/978-3-662-53890-6\\_4](https://doi.org/10.1007/978-3-662-53890-6_4)
- [35] K. Orkphol and W. Yang, "Word sense disambiguation using cosine similarity collaborates with Word2vec and WordNet," *Futur. Internet*, vol. 11, no. 5, p. 114, 2019. <https://doi.org/10.3390/fi11050114>
- [36] Y. Li, G. Huang, C. Wang, and Y. Li, "Analysis framework of network security situational awareness and comparison of implementation methods," *EURASIP J. Wirel. Commun. Netw.*, 2019(1), 205, 2019. <https://doi.org/10.1186/s13638-019-1506-1>
- [37] J. Kang, S. Jang, S. Li, Y.-S. Jeong, and Y. Sung, "Long short-term memory-based malware classification method for information security," *Comput. Electr. Eng.*, 77, 366–375, 2019. <https://doi.org/10.1016/j.compeleceng.2019.06.014>
- [38] L. R. Logeshwari, R. & Parvathy, "A Quantitative and Qualitative Reasoning of Various Visual Cryptographic Schemes to Uphold Secrecy. International Journal of Innovative Technology and Exploring Engineering. [Online]. 8 (11). 1148–1151.," 2019.
- [39] S. Kim, H. Park, and J. Lee, "Word2vec-based latent semantic analysis (W2V-LSA) for topic modeling: a study on blockchain technology trend analysis," *Expert Syst. Appl.*, 113401, 2020. <https://doi.org/10.1016/j.eswa.2020.113401>
- [40] C. Q. (2019). Dhir, R.K., de Brito, J., Silva, R. V. & Lye, "Methodology. In: Sustainable Construction Materials.," [Online]. Elsevier, 15–34, 2019.
- [41] S. A. S. Salloum and K. Shaalan, "Investigating students' acceptance of E-learning system in Higher Educational Environments in the UAE: Applying the Extended Technology Acceptance Model (TAM)." The British University in Dubai, 2018.
- [42] M. Habes, M. Alghizzawi, S. Ali, A. SalihAlnaser, and S. A. Salloum, "The Relation among Marketing ads, via Digital Media and mitigate (COVID-19) pandemic in Jordan.," *Int. J. Adv. Sci.*, 29(7), 2326–12348, 2020.
- [43] M. Habes, M. Alghizzawi, S. A. Salloum, and C. Mhamdi, "Effects of Facebook Personal News Sharing on Building Social Capital in Jordanian Universities," in *Recent Advances in Intelligent Systems and Smart Applications*, Springer, 2020, 653–670. [https://doi.org/10.1007/978-3-030-47411-9\\_35](https://doi.org/10.1007/978-3-030-47411-9_35)
- [44] S. S. A. Al-Marouf R.S., "An Integrated Model of Continuous Intention to Use of Google Classroom.," *Al-Emran M., Shaalan K., Hassantien A. Recent Adv. Intell. Syst. Smart Appl. Stud. Syst. Decis. Control. vol 295. Springer, Cham*, 2021. [https://doi.org/10.1007/978-3-030-47411-9\\_18](https://doi.org/10.1007/978-3-030-47411-9_18)
- [45] S. A. Salloum, A. Q. M. Alhamad, M. Al-Emran, A. A. Monem, and K. Shaalan, "Exploring Students' Acceptance of E-Learning Through the Development of a Comprehensive Technology Acceptance Model," *IEEE Access*, vol. 7, 128445–128462, 2019. [10.1109/ACCESS.2019.2939467](https://doi.org/10.1109/ACCESS.2019.2939467)
- [46] S. A. Salloum and M. Al-Emran, "Factors affecting the adoption of E-payment systems by university students: Extending the TAM with trust," *Int. J. Electron. Bus.*, 14(4), 371–390, 2018. <https://doi.org/10.1504/IJEB.2018.098130>
- [47] S. A. Salloum and K. Shaalan, "Factors affecting students' acceptance of e-learning system in higher education using UTAUT and structural equation modeling approaches," in *International Conference on Advanced Intelligent Systems and Informatics*, 2018, 469–480. [https://doi.org/10.1007/978-3-319-99010-1\\_43](https://doi.org/10.1007/978-3-319-99010-1_43)
- [48] S. A. Salloum and K. Shaalan, "Adoption of e-book for university students," in *International Conference on Advanced Intelligent Systems and Informatics*, 2018, 481–494. [https://doi.org/10.1007/978-3-319-99010-1\\_44](https://doi.org/10.1007/978-3-319-99010-1_44)
- [49] M. Al-Emran, I. Arpacı, and S. A. Salloum, "An empirical examination of continuous intention to use m-learning: An integrated model," *Educ. Inf. Technol.*, 2020. <https://doi.org/10.1007/s10639-019-10094-2>
- [50] S. A. Salloum, M. Al-Emran, R. Khalaf, M. Habes, and K. Shaalan, "An Innovative Study of E-Payment Systems Adoption in Higher Education: Theoretical Constructs and Empirical Analysis," *Int. J. Interact. Mob. Technol.*, 13(6), 2019. <https://onlinejour.journals.publicknowledgeproject.org/index.php/ijim/article/view/9875>
- [51] R. S. Al-Marouf, S. A. Salloum, A. Q. AlHamadand, and K. Shaalan, "Understanding an Extension Technology Acceptance Model of Google Translation: A Multi-Cultural Study in United Arab Emirates," *Int. J. Interact. Mob. Technol.*, 14(3), 157–178, 2020. <https://onlinejour.journals.publicknowledgeproject.org/index.php/ijim/article/view/11110>
- [52] M. Habes, S. A. Salloum, M. Alghizzawi, and C. Mhamdi, "The Relation Between Social Media and Students' Academic Performance in Jordan: YouTube Perspective," in *International Conference on Advanced Intelligent Systems and Informatics*, 2019, 382–392.
- [53] P. Galbacs, *The Friedman-Lucas Transition in Macroeconomics: A Structuralist Approach*. Academic Press, 2020.
- [54] M. Al-Emran, V. Mezhyuev, and A. Kamaludin, "PLS-SEM in Information Systems Research: A Comprehensive Methodological Reference," in *4th International Conference on Advanced Intelligent Systems and Informatics (AISI 2018)*, 2018, 644–653. [https://doi.org/10.1007/978-3-319-99010-1\\_59](https://doi.org/10.1007/978-3-319-99010-1_59)

## Development of Soil Moisture Monitoring by using IoT and UAV-SC for Smart Farming Application

Sarun Duangsuwan<sup>\*1</sup>, Chakree Teekapakvisit<sup>1</sup>, Myo Myint Maw<sup>2</sup>

<sup>1</sup>Information Engineering, Prince of Chumphon Campus, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 86160, Thailand

<sup>2</sup>Department of Computer Engineering and Information Technology (CEIT), Mandalay Technological University (MTU), Mandalay, 05011, Myanmar

### ARTICLE INFO

Article history:

Received: 19 June, 2020

Accepted: 17 July, 2020

Online: 28 July, 2020

Keywords:

UAV-SC

IoT

Soil moisture monitoring

Mobile application

Smart farming

### ABSTRACT

Soil moisture is a fundamental factor for smart farming that is used to control the water management system. In this paper, the unmanned aerial vehicle (UAV) small cell (UAV-SC) can provide Internet of things (IoT) as the hotspot mobility network, due to the minimum limitation energy of connected IoT. The development of ground sensor (GS) communicates to the UAV-SC called GS-UAV-SC model for soil moisture monitoring is proposed to smart farming. UAV-SC aims to fulfill the data collection task with a limitation of GSs power. In the experiment, the two case scenarios: Napier grass farm and Ruzi grass farm are implemented. The result of soil moisture status is demonstrated as an example of data in real time on a mobile application monitoring system. The proposed system is useful for users/farmers to know the soil moisture data quickly for smart farming applications.

## 1. Introduction

Recently, agricultural UAV or drones become one of the most useful agricultural instruments utilized in smart farming, especially in the ground sensing applications. The UAV is employed as a mobility hotspot in the area lacking of the communication network to collect the data of soil and crops in order to monitor and control the water management system for precision agriculture. The agricultural UAV with a combination of IoT sensors are preferred. At present, the IoT sensors detect temperature and moisture of the soil and send them to farmers by using the UAV. Therefore, the identification of locations that suffer from droughts, water scarcity, and dryness of soil profile is easily observed. It helps the farmers to take precautionary methods such as contouring, damming, draining the surface water, and curtailing for further irrigation. Small UAV, either copters or flat-winged models with a combination of IoT sensors are essential.

In [1], the authors firstly introduced UAV which plays a great role in the aerial imagery of soils, particularly soil characteristics. Drones are attractive because they can fly at a low attitude over the fields, as many times as requires in the short intervals to study the same location. The result in [1] showed that the UAV or drones are one of the common methods for obtaining high-resolution maps

showing spatial and temporal changes in soil moisture. However, the UAV imagery with low resolution is not used consistently to evaluate a field experiment. In [2], the authors have shown that the precision of soil moisture measurement with the frequency domain reflection (FDR) method was highly accurate. However, the soil moisture FDR sensor type is still expensive. Therefore, the low-cost capacitive sensor and very low-cost resistive sensors have been investigated in [3]. The authors have presented that low-cost capacitive sensors with soil-specific calibration matched the performance of the secondary standard and could possibly be used for water management with limited effects on irrigation efficiency. Nowadays, most of soil moisture application is based on IoT networking [4] or wireless sensor networking (WSN) [5]. The objective is to automatically obtain the agricultural parameters without requiring human-to-human or human-to-computer.

## 2. Related Work

### 2.1. UAV-SC

Most of the wireless IoT nodes utilized in the agricultural environment limit power energy which leads to the propagation problem in the surrounding area. Interestingly, B. Liu et al. [6] have proposed to solve the resource allocation problem in their UAV and IoT. The UAV works as power distribution by using a dynamic gamin-based model for wireless power transfer

<sup>\*</sup>Corresponding Author: Sarun Duangsuwan, KMITL Prince of Chumphon Campus, Tel: +66-99186-4411; Email: sarun.du@kmitl.ac.th

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050444>

applications. The authors in [6] have illustrated that the power energy transfer of UAV depends on the amount of the density of IoT wireless nodes and they can be controlled by the power by using the dynamic game method. On the other hand, the proposed measurement aided dynamic planning algorithm in [7] can solve the problem-based of multiple IoT connectivity network. This algorithm can help to control the drone small cell (DSC) base station as a self-configuration on the fly regarding the cell being located and channel being used, based on the real-time measurement of network throughput. In addition, more increasingly, an optional mobile backhaul link might use the drone or UAV aided emergency network [8]. However, the problem-based such as the inter-cell interference should be considered. In [9], the multiple DSCs used for the several based stations were optimized by using the equivalent uniform density plane entity (EUDPE) method. The results indicated that the increase of the transmission power of drone did not necessarily improve coverage performance for the IoT connectivity. Similar to the drone base stations (DBSs) in [10], the authors proposed a realistic mathematical model for the joint optimization problem of DBS placement and IoT users. Path loss measurement for an air-to-ground channel was investigated in [10]. For the data monitoring system based on WSN with the large-scale area such as a coast or river [11], the authors considered to maximize data collection with given energy budgets for the drone and sensors. DroneTank [11], the UAV can determine the flight path and speed to conduct data collection. As mentioned in the works of literature, the UAV-SC for ground sensing is very important in the smart farming applications.

## 2.2. Smart Farming

Recently, the proposed AgriLogger has been implemented in Italy [12]. The aim is to build precision agriculture in the areas lacking data communication networks that can collect data and store the long periods and transmitted agrometeorological data, such as temperature and relative humidity. The drone is equipped with a payload, such as a wireless transceiver, radio transmitter, and GNSS receiver. Additionally, a communication link provides the narrowband IoT (NB-IoT) network to collect the data via cloud management. AgriLogger was able to collect and store for long periods estimated for more than ten years, and consistently, agrometeorological parameters data in a successful way would be great profit for agriculture. As well-known that the WSN technology is useful for precision agriculture as well as the IoT solution. There are several works focused on using WSN [13,14]. In [15], the authors have presented the technique and practice of precision agriculture. They implied that precision agriculture is to provide decision support systems based on multiple parameters impacted to the crops plantation such as soil moisture, water management, wind speed, and so on, by using WSN-based or IoT-based cloud computing. In [16], a secure IoT-based WSN framework for smart agriculture was described. The simulation results show that the proposed was more efficient compared to the IoT communication reliabilities. The network throughput was increased to 13.5 % and 38.5 % for the packets drop ratio, 13.5 % for the network latency, 16 % for the energy consumption, and 26 % for the routing overhead of smart agriculture.

SmartFarmNet [17] platform is an example application for smart agriculture in Australia by using IoT platform. It provides a scalable sensor data acquisition, analysis, and visualization [www.astesj.com](http://www.astesj.com)

platform for smart farming applications, based on IoT. Interestingly, this platform measures soil moisture at particular multiple depths and steams data from the physical sensors to monitor a phenomenon. By using the SmartFarmNet gateway to process soil moisture data which is close to the source, the cloud computing performs real-time analysis on incoming sensor data streams. For a platform of precision agriculture, SmartFarmNet is an existing model of smart agriculture based on IoT-based platform. Considering the IoT for smart farming, the IoT platform can help the farmer to predict the disease of the strawberry farm in [18]. However, the problem is the false detection probability of IoT sensors to the gateway. Then, the authors in [19] have presented a solution called SensorTalk to automatically detect the potential sensor failures and semi-automatically calibrate the aging sensors. The solution showed that the false detection probability of IoT sensors to the gateway was reduced to less than 0.7 %.

In related work of smart irrigation, IoT framework for smart irrigation such as data monitoring, preprocessing, fusion, synchronization, storage, and irrigation management enriched by the prediction of soil moisture was presented in [20]. Smart&Green was used to preprocess soil moisture data with outlier removal criteria and techniques of Zscore, MZscore, GESD, and Chauvenet to provide a more precise irrigation water needed in irrigation management. While the propose of IoT for the water management interestingly was presented in [21], WSN, LoRa, and ESP32 modules were used. The system retrieved real-time data and used them to determine the correct amount of water to be used in a garden. With this solution [21], it was possible to save up to 34 % of water with sensor data from temperature, humidity, and soil moisture, or up to 26 % with only temperature inputs. At present, the application of IoT in smart farming is rapidly emerging. A Long-range wide area network (LoRaWAN) based on IoT platform has introduced and called LoRaFarM [22]. The LoRaFarM platform has evaluated on a real farm in Italy. This system can collect data such as air, soil, temperature, and humidity. The advantage of the LoRa network is the long-range communication network and life-long battery of sensors, but it is still needed more gateways in the field. In [22], the design for the soil moisture sensor kit was proposed. The LoRaFarM hardware was designed by using the 1W solar-cell, solar charger, soil sensor, LoRa module at 868 MHz, and LiPo battery.

## 3. Proposed Model and Implementation

The GS-UAV-SC model for smart farming applications is shown in Figure 1. It can be noted that the IoT soil moisture GS kits were installed in the field, then, were connected to the UAV-SC to transfer data streaming to the internet. After that, the data are stored and computed on the cloud platform. Additionally, the farmer/user can control the UAV around the field in order to collect all data from every sensor. User can monitor the data via the application program interface (API) on the mobile application. This proposed model aims to compensate for the low energy from the multiple GSs connectivity.

The structure model of the soil moisture GS prototype is shown in Figure 2. It can be noted that this platform would be used in case of sunlight. Herein, the capacitive soil moisture was used in this study. By using the capacitive soil moisture GS, the capacitance is varied based on the water content present in the soil. The capacitance converted into voltage level basic from 1.2 V to 3.0 V

maximum. The advantage of capacitive soil moisture GS which is made of corrosion-resistant material is that it gives long service

life. The structure model of the soil moisture GS prototype consists of five parts:

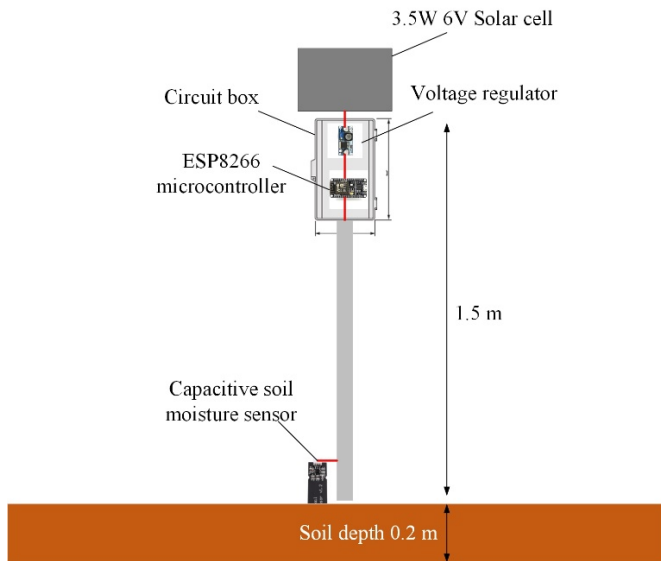
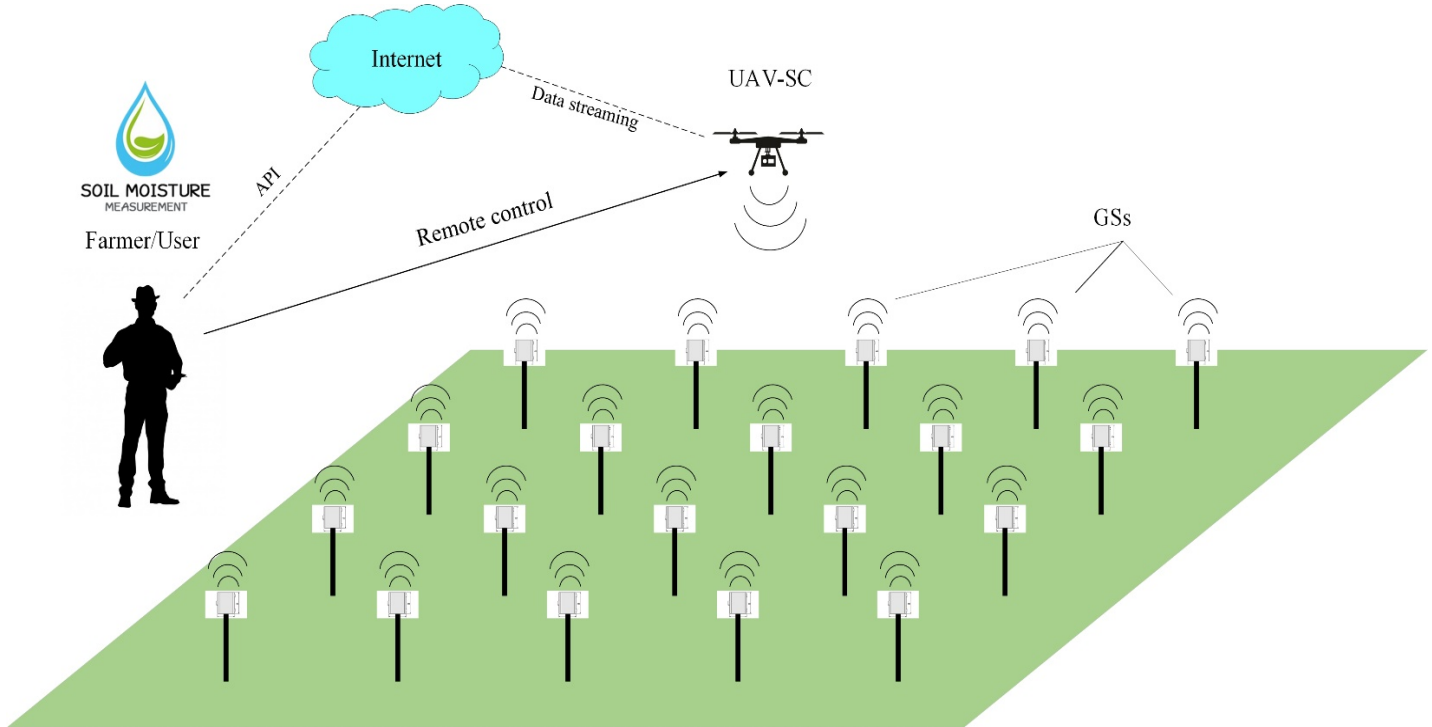


Figure 2: The structure model of soil moisture GS prototype.

2. Voltage regulator is DC to DC up/down LM2596 module, the maximum output current is 3 A, and the size is 50 mm wide and 37 mm length.
3. Microcontroller ESP8266 NodeMCU is used for controlling data from the soil moisture and transmitting to WiFi internet. The voltage is driven at 3.3 V and current consumption ranges from 70 mA to 240 mA. The transmitting power of microcontroller is 20 dBm, and it supports IEEE802.11 b/g/n.

4. Capacitive soil moisture sensor SKU: SEN0193, this module includes an on-board voltage regulator which gives an operating voltage range of 1.2 V to 3.0 V. The operating current is 5 mA. This module is an analog sensor which has 3-pins: A0 signal pin, positive pin +, and negative pin -. The depth of soil moisture sensor in the soil is 0.2 m according to the standard as shown Figure 2.
5. The material structure is made from aluminum pipe, and shield is inserted in the circuit box.

The implementation was conducted at the Tropical Animal Research Institute, Ramkhamhaeng University, Thailand. Figure 3(a) shows an example location of Napier grass farm and Ruzi grass farm in Figure 3(b) where the dimension of Napier grass farm is 25 m wide and 45 m length as shown in Figure 3(c) and Ruzi farm is 30 m wide and 50 m length in Figure 3(d). It is noted that these Napier and Ruzi grass are the main food of animals in the smart farm and 50 m of length in Figure 3(d). To note that, these Napier and Ruzi grass are mainly food of the animals within the smart farm.

### 3.1. Data Processing

The block diagram of data processing is shown in Figure 4. The quadcopter UAV is used for mobility and mounted with the WiFi transceiver. The WiFi transceiver radiates power to the GSs and is connected with the microcontroller ESP8266 NodeMCU module. After receiving the data from the capacitive soil moisture sensor, the ESP8266 NodeMCU transmits data to the internet, cloud computing, and cloud storage based on the Google firebase solution. Eventually, the soil moisture data is alerted to the mobile application.



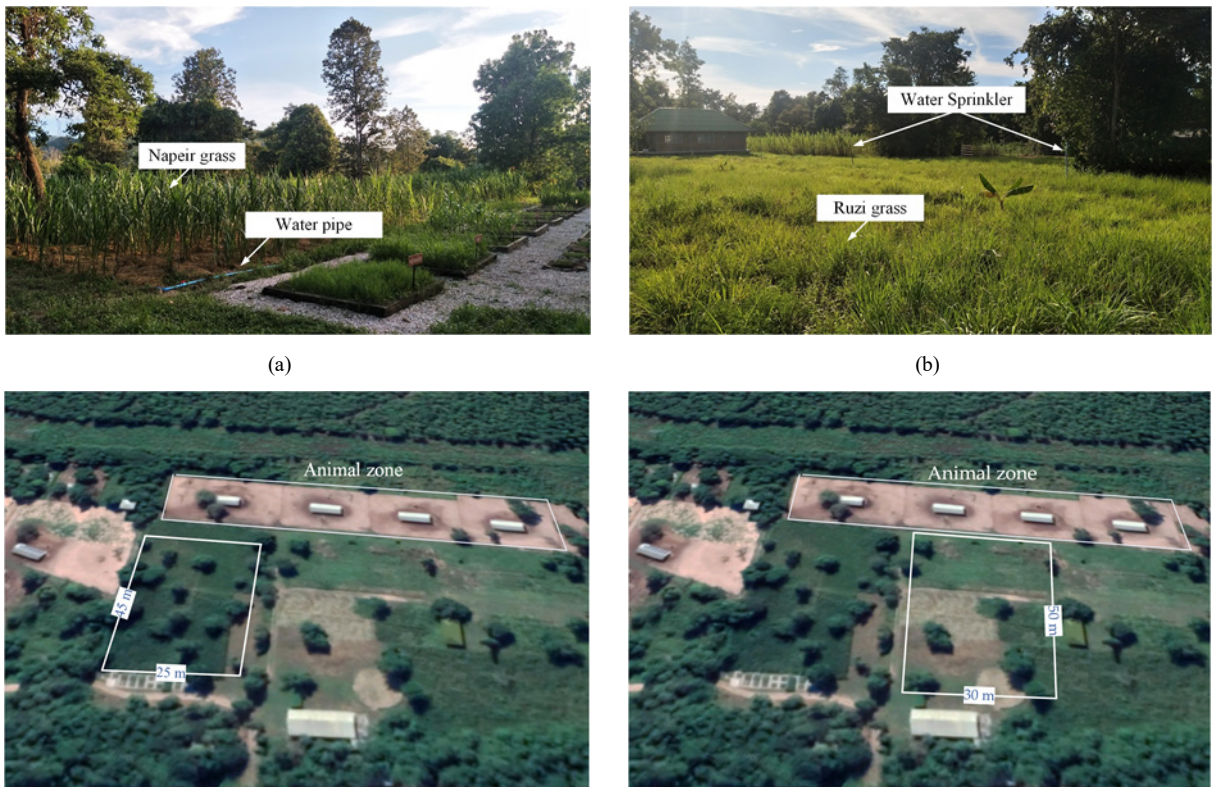


Figure 3: Smart farming of the Tropical Animal Research Institute, Ramkhamhaeng University: (a) Side view of Napier grass farm; (b) Side view of Ruzi grass farm; (c) Top view of Napier grass dimension; (d) Top view of Ruzi grass dimension.

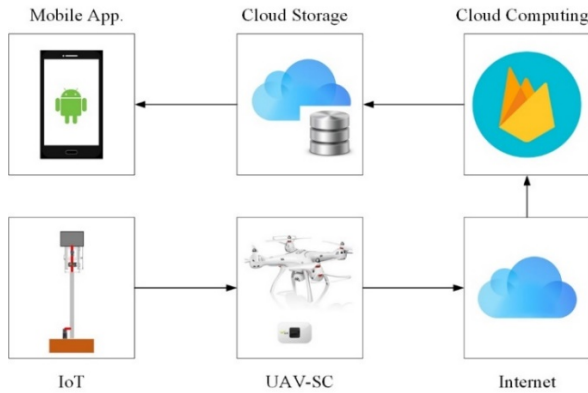


Figure 4: Block diagram of data processing.

The flowchart of the soil moisture monitoring process is shown in Figure 5. Under the condition of soil moisture, the level of the moisture is set to be alerted on the mobile application as follows:

- With the show “Dry” status, all sensors send the data as the range between 1 % to 45 %. It means that the soil is dry.
- With the show “Humid” status, all sensors send the data as the range from 46 % to 79 %. It means that the soil is humid.
- With the show “Wet” status, all sensors send the data as the range between 80 % and 100 % . It means that the soil is wet.

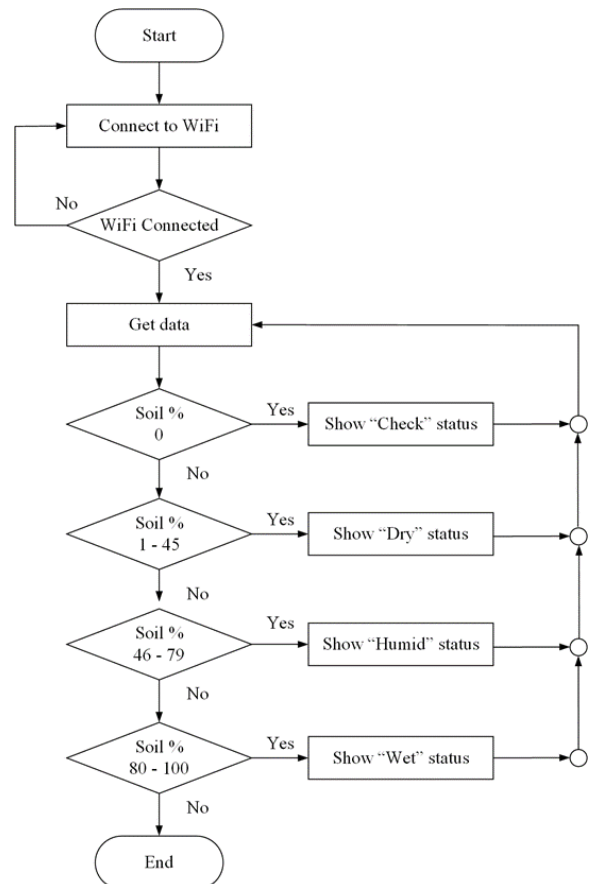


Figure 5: Flowchart of decision.

#### 4. Experimental Setup

This section demonstrates the experimental setup in both Napier farm and Ruzi farm.

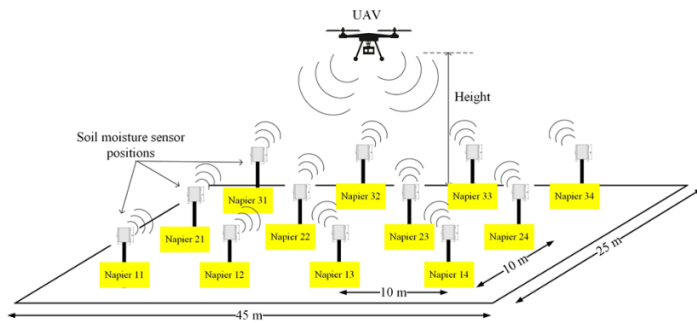
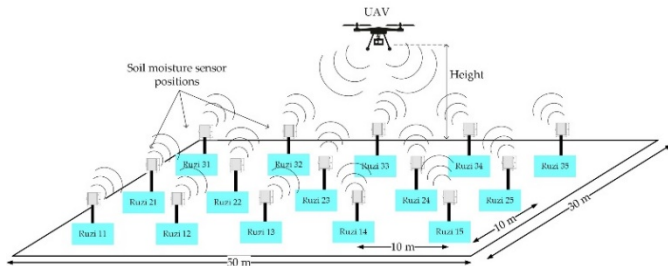


Figure 6: Experimental setup in Napier farm with 12 GSs.



: Real-time test in Napier farm.



The experimental setup in Napier farm is demonstrated in Figure 6. The GSs are installed along the line of drip irrigation water systems in the field. The separation distance of each sensor is 10 m, and the position names are Napier11 to Napier34. Real-time test in Napier farm is shown in Figure 7. The flight time of the UAV is approximately 20 minutes, and the altitude is varied from 1 m to 10 m. For the sensor depth, it is a pin into the soil with about 0.2 m depth. The tested soil moisture GS kits are performed by the power supply from the solar cell panel in the day time. Therefore, the farmer/user can monitor the data in real time..

shown in Figure 8 and real-time test in Figure 9. It demonstrates the position of sensors in the field in which the number of GS is 15 positions. The GS sets as Ruzi11 to Ruzi35. The sensors installed along the line of the water sprinkler in the field. At the experiment, the UAV altitude was also varied from 1 m to 10 m. Note that the average temperature of all days of these locations was 40 degrees. As a result, both the Napier and Ruzi farms need more water spraying

every day. Therefore, the GS-UAV-SC model is useful for the farmer/user to automatically control the on-off of the water management system in these agricultural farms.

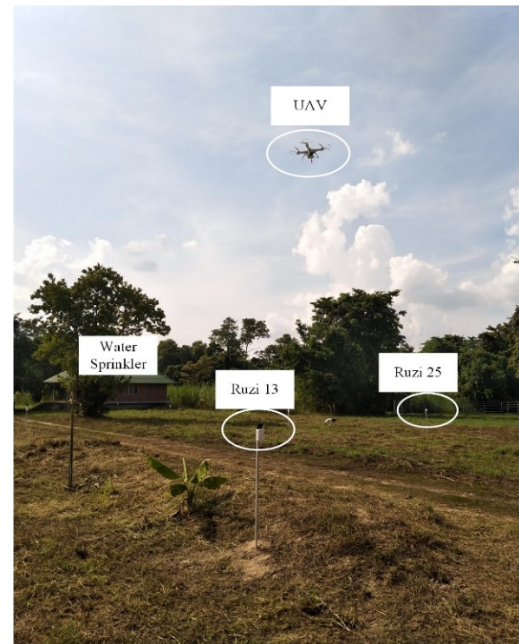


Figure 9: Real-time test in Ruzi farm.

#### 5. Results and Discussion

Table 1: In case: Before opening the water system in Napier farm.

Positions	Soil moisture monitoring (%)					
	2.00PM	2.15PM	2.30PM	2.45PM	3.00PM	3.15PM
Napier 11	23	23	24	25	25	25
Napier 12	32	32	34	33	34	34
Napier 13	37	37	38	38	38	39
Napier 14	43	43	44	45	45	43
Napier 21	40	41	42	41	41	42
Napier 22	22	23	25	24	23	25
Napier 23	28	30	29	30	28	28
Napier 24	42	41	41	40	42	42
Napier 31	32	31	31	31	32	32
Napier 32	32	33	34	33	33	34
Napier 33	30	30	32	32	31	30
Napier 34	26	26	28	28	27	28

Table 2: In case: After opening the water system in Napier farm.

Positions	Soil moisture monitoring (%)					
	4.00PM	4.15PM	4.30PM	4.45PM	5.00PM	5.15PM
Napier 11	45	47	52	56	61	59
Napier 12	48	49	55	60	62	61
Napier 13	50	53	57	61	64	63
Napier 14	51	55	61	64	66	65
Napier 21	45	47	54	55	64	64
Napier 22	52	53	56	60	63	63
Napier 23	50	52	55	61	62	61
Napier 24	51	53	58	64	65	62
Napier 31	64	66	67	67	68	65
Napier 32	44	45	55	58	59	56
Napier 33	64	65	67	69	69	67
Napier 34	47	47	48	50	51	49



Table 3: In case: Before opening the water system in Ruzi farm.

Positions	Soil moisture monitoring (%)					
	Time	2.00PM	2.15PM	2.30PM	2.45PM	3.00PM
Ruzi 11	21	20	20	18	19	20
Ruzi 12	21	20	19	20	19	21
Ruzi 13	25	25	26	27	26	25
Ruzi 14	31	30	30	31	30	31
Ruzi 15	32	32	33	34	33	34
Ruzi 21	16	16	15	16	16	18
Ruzi 22	13	14	15	16	17	17
Ruzi 23	25	26	26	27	28	28
Ruzi 24	31	30	31	30	30	31
Ruzi 25	33	34	33	34	34	33
Ruzi 31	12	13	14	14	15	16
Ruzi 32	15	14	15	18	17	18
Ruzi 33	22	23	22	23	23	25
Ruzi 34	0	0	0	0	0	0
Ruzi 35	0	0	0	0	0	0

Table 4: In case: After opening the water system in Ruzi farm.

Positions	Soil moisture monitoring (%)					
	Time	4.00PM	4.15PM	4.30PM	4.45PM	5.00PM
Ruzi 11	38	46	47	48	48	47
Ruzi 12	46	48	51	53	52	51
Ruzi 13	46	48	53	56	54	54
Ruzi 14	52	53	55	57	56	55
Ruzi 15	48	49	54	57	58	56
Ruzi 21	46	48	51	52	53	54
Ruzi 22	79	80	81	79	74	71
Ruzi 23	53	54	59	60	62	59
Ruzi 24	45	46	52	54	55	53
Ruzi 25	47	49	54	54	54	53
Ruzi 31	52	54	60	61	59	57
Ruzi 32	49	51	55	55	56	54
Ruzi 33	46	48	51	52	53	51
Ruzi 34	0	0	0	0	0	0
Ruzi 35	0	0	0	0	0	0

Before the opening of the drip irrigation water system in Napier grass farm, soil moisture sensors were tested from 2 PM to 3:15 PM. The results of soil moisture sensors show that the data is in the range from 22 % to 42 % of the real-time monitoring. Therefore, the status was Dry status. After the opening of the drip irrigation water system for about forty-five minutes, it is observed that the soil moisture values were rapidly increased in the range of 45 % to 69 % as shown in Table 2. Thus, the “Humid” status was shown on the mobile application.

In Ruzi farm, the soil moisture values in Ruzi farm were in the range from 12 % to 34 % in Table 3 before opening the water sprinkler system, Therefore, the status showed as Dry. After the opening water sprinkler system about forty-five minutes in Ruzi farm, the results of soil moisture values were in the range from 38 % to 81 %. As a result, the status shown on mobile application was Dry, Humid, Wet, and Check status. And then, the capacitive soil moisture sensors were removed from the soil. Both Ruzi 34 and Ruzi 35 were 0 % with the “Check” status. The experiment data in Ruzi farm were collected on April 3, 2020.

The examples of mobile application on screen as real-time monitoring were shown in Figure 10 (a)-(b). The data collected at 4 PM is shown in Table 2 and Table 4. Figure 10 (a) and 10 (b) show the status of soil moisture values in Napier farm and Ruzi farm. The application is easily used and can be applied to control the water management (on-off) system both in Napier and Ruzi farms. It took just to 15 minutes to collect the data for the proposed model.

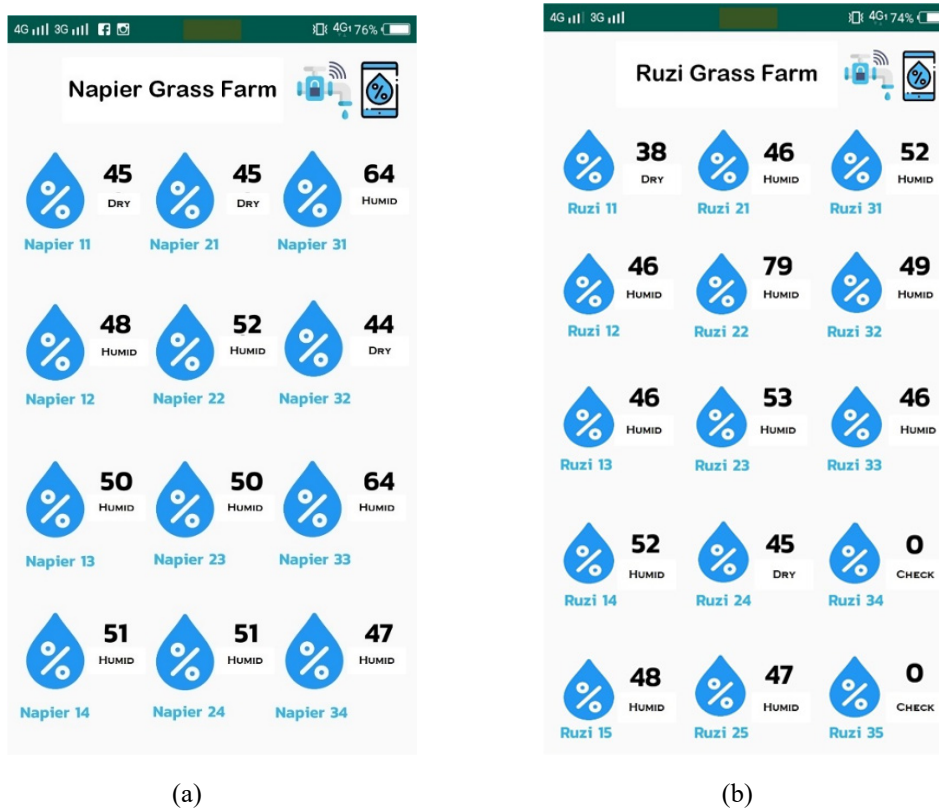


Figure 10: Example of real-time monitoring on mobile application at 4 PM:

(a) On the screen in Napier farm. (b) On the screen in Ruzi farm.

## 6. Conclusion

In this paper, the proposed GS-UAV-SC model based on IoT and wireless communication connected network in smart farming has been implemented for soil moisture monitoring. The low-cost multiple GSs were designed as the ground sensing to monitor the soil moisture in the daylight both for Napier grass farm and Ruzi grass farm. The UAV-SC is equipped with the WiFi transceiver similar to the hotspot mobility and connected with the multiple GSs. This proposed smart farming system is called the GS-to-UAV-SC communication system. The implementation and results confirm that the UAV-SC can be applied as mobility transceiver to compensate for the low energy of GSs, including the real-time data that are designed to monitor soil moisture values on the mobile application. In future work, the optimization of path loss model for smart farming (SF) system will be implemented by the machine learning algorithms.

## Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Acknowledgment

This research was financially supported by the Academic Melting Pot, KMITL research fund. The authors would like to thank the Tropical Animal Research Institute, Ramkhamhaeng University, in Thailand, which was cooperative research in this work.

## References

[1] C. M. Gevaert, J. Suomalainen, J. Tang, L. Kooistra, "Generation of Spectral-Temporal Response Surfaces by Combining Multispectral Satellite and Hyperspectral UAV Imaginary for Precision Agricultural Applications" *IEEE J. of Selected Topics in Applied Earth Observations and Remote Sensing*, **8**(6), 3140–3146, 2015. <https://doi.org/10.1109/JSTARS.2015.2406339>

[2] L. Chen, L. Zhangzhong, W. Zheng, J. Yu, Z. Wang, L. Wang, C. Huang "Data-Driven Calibration of Soil Moisture Sensor Considering Impacts of Temperature: A Case Study on FDR Sensors" *Sensors*, **19**(20), 1–11, 2019. <https://doi.org/10.3390/s19204381>

[3] S. Adla, N. K. Rai, S. H. Karumanchi, S. Tripathi, M. Disse, S. Pande "Laboratory calibration and performance evaluation of low-cost capacitive and very low-cost resistive soil moisture sensors" *Sensors*, **20**(2), 1–27, 2020. <https://doi.org/10.3390/s20020363>

[4] W. L. Chen, Y. B. Lin, Y. W. Lin, R. Chen, J. K. Liao, F. L. Ng, Y. Y. C. Liu, C. C. Wang, C. H. Chiu, T. H. Yen, "AgriTalk: IoT for Precision Soil Farming of Turmeric Cultivation" *IEEE Internet of Things J.*, **6**(3), 5209–5223, 2019. <https://doi.org/10.1109/JIOT.2019.2899128>

[5] F. Viani, M. Bertolli, A. Polo, "Low-Cost Wireless System for Agrochemical Dosage Reduction in Precision Farming" *IEEE Sensors J.*, **17**(1), 5–6, 2017. <https://doi.org/10.1109/JSEN.2016.2622244>

[6] B. Liu, H. Xu, X. Zhou, "Resource Allocation in Unmanned Aerial Vehicles (UAV)-Assisted Wireless-Powered Internet of Things" *Sensors*, **19**(8), 1–13, 2019. <https://doi.org/10.3390/s19081908>

[7] N. Lu, Y. Zhong, C. Shi, N. Cheng, B. Li, "Planning While Flying: A Measurement-Aided Dynamic Planning of Drone Small Cells" *IEEE Internet of Things J.*, **6**(2), 2693–2705, 2019. <https://doi.org/10.1109/JIOT.2018.2873772>

[8] G. Castellanos, M. Deruyck, L. Martens, W. Joseph, "Performance Evaluation of Direct-Link Backhaul for UAV-Aided Emergency Networks" *Sensors*, **19**(15), 1–16, 2019. <https://doi.org/10.3390/s19153342>

[9] C. Dong, J. Xie, H. Dai, Q. Wu, Z. Qin, Z. Feng "Optimal Deployment Density for Maximum Coverage of Drone Small Cells" *China Communications*, **15**(5), 25–40, 2018. <https://doi.org/10.1109/CC.2018.8387984>

[10] A. Ahmed, M. Awais, T. Akram, S. Kulac, M. Alhussein, K. Aurangzeb, "Joint Placement and Device Association of UAV Base Stations in IoT Networks" *Sensors*, **8**(6), 1–16, 2019. <https://doi.org/10.3390/s19092157>

[11] R. Xiong, F. Shan, "DroneTank: Planning UAVs Flight and Sensors Data Transmission under Energy Constraints" *Sensors*, **18**(9), 1–19, 2018. <https://doi.org/10.3390/s18092913>

[12] M. Idbella, M. Iadaresta, G. Gagliarde, A. Mennella, S. Mazzoleni, G. Bonanomi, "AgriLogger: A New Wireless Sensor for Monitoring Agrometeorological Data in Areas Lacking Communication Networks" *Sensors*, **20**(6), 1–13, 2020. <https://doi.org/10.3390/s20061589>

[13] M. Bacco, A. Berton, A. Gotta, L. Caviglione, "IEEE 802.15.4 Air-Ground UAV Communications in Smart Farming Scenarios" *IEEE Communication L.* **22**(9), 1910–1913, 2018. <https://doi.org/10.1109/LCOMM.2018.2855211>

[14] F. Viani, M. Bertolli, A. Polo, "Low-Cost Wireless System for Agrochemical Dosage Reduction in Precision Farming" *IEEE Sensors J.*, **17**(1), 5–6, 2017. <https://doi.org/10.1109/JSEN.2016.2622244>

[15] U. Shafi, R. Mumtaz, J. Garcia-Nieto, S. A. Hassan, S. A. Raza Zaidi, N. Iqbal, "Precision Agriculture Techniques and Practices: From Consideration to Applications" *Sensors*, **19**(17), 1–25, 2019. <https://doi.org/10.3390/s19173796>

[16] K. Haseeb, I. U. Din, A. Almogren, N. Islam, "An Energy Efficient and Secure IoT-based WSN Framework: An Application to Smart Agriculture" *Sensors*, **20**(7), 1–14, 2020. <https://doi.org/10.3390/s20072081>

[17] P. P. Jayaraman, A. Yavari, D. Geogakopoulos, A. Morshed, A. Zaslavsky, "Internet of Things Platform for Smart Farming: Experiences and Lessons Learnt" *Sensors*, **16**(11), 1–17, 2016. <https://doi.org/10.3390/s16111884>

[18] S. Kim, M. Lee, C. Shin, "IoT-based Strawberry Disease Prediction System for Smart Farming" *Sensors*, **18**(11), 1–17, 2018. <https://doi.org/10.3390/s18114051>

[19] Y. B. Lin, Y. W. Lin, J. Y. Lin, H. N. Hung, "SensorTalk: An IoT Device Failure Detection and Calibration Mechanism for Smart Farming" *Sensors*, **19**(21), 1–19, 2019. <https://doi.org/10.3390/s19214788>

[20] N. G. S. Campos, A. R. Rocha, R. Gondim, T. L. Coelho da Sliva, D. G. Gomes, "Smart & Green: An Internet-of-Things Framework for Smart Irrigation" *Sensors*, **20**(1), 1–25, 2020. <https://doi.org/10.3390/s20010190>

[21] A. Gloria, C. Dionisio, G. Simoes, J. Cardoso, P. Sebastiao, "Water Management for Sustainable Irrigation Systems using Internet-of-Things" *Sensors*, **20**(5), 1–14, 2020. <https://doi.org/10.3390/s20051402>

[22] G. Codelippi, A. Cilfone, L. Davoli, G. Ferrari, "LoRaFarM: a LoRaWAN-based Smart Farming Modular IoT Agriculture" *Sensors*, **20**(7), 1–24, 2020. <https://doi.org/10.3390/s20072028>



## The Role of KM in Enhancing AI Algorithms and Systems

Hani AlGhanem<sup>1</sup>, Mohammad Shanaa<sup>1</sup>, Said Salloum<sup>1,2,\*</sup>, Khaled Shaalan<sup>1</sup>

<sup>1</sup>Faculty of Engineering &IT, The British University in Dubai, 345015, United Arab Emirates

<sup>2</sup>Research Institute of Sciences & Engineering, University of Sharjah, 27272, United Arab Emirates

---

### ARTICLE INFO

Article history:

Received: 24 June, 2020

Accepted: 14 July, 2020

Online: 28 July, 2020

---

Keywords:

Knowledge management

Artificial intelligence

AI algorithms

---

### ABSTRACT

Knowledge Management processes present a vital role in improving AI systems and algorithms. Many studies and reviews were carried out to examine the relationship between KM processes and AI systems. However, studies were focusing on specific methods and the impact on some AI algorithms, neglecting the role of other KM processes and how it may affect the AI system to achieve the objective, which reduces the adoption in some organizations. The current study shows the relation between KM processes and AI systems from a higher perspective, giving different options to apply other KM processes for the same AI algorithm to reduce any implementation challenges and enhance the adoption level. The review looks into 16 studies collected from a different database from 2014 to 2019. The main finding of the research was the massive impact of some KM processes like knowledge acquisition and knowledge creation on the different types of AI systems and algorithms to give an additional option for organizations during the implementation. Additionally, the research finds that most of the studies agree on the positive relationship between knowledge management processes and the role-plays to enhance AI systems and algorithms. Finally, the study shows a decrease in the number of researches done for this topic in the selected databases, which can be enhanced by other researchers by examining other databases to increase results accuracy.

---

### 1. Introduction

Knowledge plays an essential role in the maintenance of organization business [1], as it can be explained as one of the most critical factors that companies need to maintain their business in the market [2]–[4], also knowledge considered as a valuable intangible resource. It includes all the ideas, concepts, data, technologies stored in the human brain [5]. Knowledge has many properties and can be summarized as 1) can be created by personnel, 2) dilatable, 3) can be saved in individuals brain or any other document storage or computerized storage, 4) can be stored in specific orders for easy of retrieving it when needed, 5) can be shared with others, 6) forgettable and can be lost [6]. Recently, the attention of KM has been evolved dramatically as many pieces of research, consultants, and specialists focus on organizational knowledge management in order to overcome the challenges that faced organizations [7], [8].

Knowledge management is considered a new domain for government institutes, which used to coordinate between all of the

human resources, technological resources, organization structure, and environment to improve organization strategies [5]. Besides, KM exerted an effective role in educational practices [9]–[12]. While [13] collect the agreed definition of knowledge management as “creation, communication, and application of knowledge.” Moreover, consider a vital tool to improve organizational processes. The primary purpose of KM is described by [14] “to improve the systematic handling of knowledge and potential knowledge within the organization”. AI is the way to build a software program that can imitate the human brain in many mental procedures and processes, including learning, thinking, building perception, natural language processing, innovation, and complex issues resolution [15]. AI can be employed in many applications that required thinking and analysis, using AI tools such as statistical and mathematical methodologies to build computer understanding and machine learning in the same way as human thinking and acquiring knowledge. In their study of intelligent vehicles, they assert that the usefulness of feeding back the generated knowledge form the test result of smart vehicles to AI systems in order to improve it, so the AI system can learn more

---

\*Corresponding Author: Said Salloum, University of Sharjah, UAE.

Tel: +971507679647 Email: [ssalloum@sharjah.ac.ae](mailto:ssalloum@sharjah.ac.ae)

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050445>

and provide better results[10]. In [16] they assert on the importance of the data warehouse to store organizational knowledge, and these data warehouses play an essential role in the application of artificial intelligent in general and decision support systems in specific, furthermore, Data/open data is considered a booster for the Information technology industry [17].

This review study raises the following four research questions:

1. What are the primary knowledge management processes studied, and how it is impacting AI systems?
2. What are the leading AI algorithms and systems reviewed in the studies collected?
3. What research methods have been used in finding the relation between knowledge management processes and AI systems?
4. How active is the topic of knowledge management impact on AI during the last five years, across the research databases?

The remainder of the review is split as follows: Section 2 containing a literature review, research methodology discussed in section 3. Section 4 demonstrates the findings and results achieved from the in-depth analysis of the articles reviewed. Finally, the review is concluded in section 5.

## **2. Literature review**

Nobody could deny the role of information systems in our daily lives [18], [19]. Knowledge management plays a vital role in accelerating the wheels of information systems [20]–[22]. The initial show of knowledge management return to the 80s as came along with AI and expert systems, then it spread out between businessmen, academics, and consultants, in the last ten year big companies like telecommunication has many success factors, and one of the critical factors is the knowledge that gained from employees experience, stakeholders, customers, and competitors. This knowledge provides the companies with a strategic and sustainable competitive advantage by building an organizational culture having active knowledge management in general and knowledge sharing in specific [23].

A study by [24] explains that knowledge management tools are used to help in knowledge generation, documentation, and sharing. KM tools are used to make managing the knowledge is easier by implementing IT technologies to create, organize, and share knowledge, which helps to improve organization operation. KM tools can be classified into four groups; the first one is tools to discover knowledge and to build the knowledge structure, which concentrates on the current knowledge and the existing one, and to make available in a well-structured format and easy to access by linking the information, ideas, and experiences. The second one is tools for knowledge processing, which can be defined as the workflow of how to deal with information and knowledge by filter, store, and analyze it. The third one is tools used for share and transfers the knowledge between individuals, which included organization local intranet or internet; furthermore, it can be extended to public online forums, wiki which make the organization issues resolving the environment having. The fourth and final one is tools used to analyzing and applying the knowledge, which includes a wide range of applications, including intelligent agents, expert systems, and many others that help in decision making.

After reviewing 20 styles of KM frameworks in the period 1991 to 2015, In [25] conclude a unified KM framework which includes the five primary processes, the first one is knowledge transfer that focuses how the knowledge and information are being moved from one individual/organization to another using interaction between people or organizations in a written or verbal way through traditional ways of using technological way. The second one is storing the knowledge in a suitable storage device, and nowadays it is stored electronically as documents or into databases; storing the knowledge is essential; because knowledge can be lost, especially the tacit knowledge. The third one is knowledge application, is the way how to get benefit from the stored knowledge and employ it to improve organization innovation, processes, performance to build a continuous competitive advantage, and this can be done using knowledge application software like expert system, intelligent agents, decision-making systems. The fourth one is knowledge creation, which is come by converting explicit and tacit using different ways like internalization, combination, externalization, socialization. Also, knowledge can be created by interviewing the experts and knowledge exchange between individuals or by documenting the organizational work procedure. The fifth and last one is knowledge acquisition; this is the process of getting knowledge from outside of the organization and uses it within the organization's knowledge base, like knowledge gained from customers, competitors, or suppliers.

Many tools can be used and present AI, such as Intelligent Agents, Business intelligence, gadgets dashboard, and expert systems [15]. While [26] define the intelligent agents as “an autonomous system that can obtain synergy effects by combining a practical user interface, on the one hand, and an intelligent system based on Artificial Intelligence, Neural Networks, and fuzzy theory, on the other hand.”.

An intelligent agent is defined as a computer system that can understand the environment that its existence in and interact based on the received data and predefined rules [27]. Intelligent agents can be categorized into different groups; the first one is the collaborative or non-collaborative, the collaborative mean that intelligent agents do a sequence of operation that one agent starts after the previous one finish his task. The second one is the profitable economy agents, which focus on the stock market and online trading by doing automatic deals and decisions. The third one is online assistant agents that can learn from the user experience. While the fourth one is called mobile agents as it can be used over the internet, for example, it can play a role in load balancing on specific servers [26].

The concept of Business intelligence (BI) was first introduced in the 90s and nowadays become widely used mainly in business to support decision making (Watson, H.J. and Wixom, B.H., 2007). BI provides a user interface that enables users to analyze and drill down the data into the information to extract and useful knowledge and data that needed to help in tacking a decision using a friendly geographical interface including charts, aggregates, and statistics [28].

The expert system is known as the first kinds of AI Implementation and can be used on an organizational level or personality level. Usually, an expert system consists of a

knowledge base module which used to keep all information and data related to issues solving. While fact base module contains some categorized information about specific cases, usually it used to build the final result, explanation module, which provides the steps that used in order to reach a specific solution depending on facts/ knowledge base. Finally, a presentation module responsible for providing all data to end-users in a human-readable and structured format [15]. There are two significant modules of expert systems, which are 1) knowledge-base which can build using fuzzy or first-order logic. So expert systems can fit the electronic governments due to the nature of e-government problems [29]. So expert systems consider and an application of knowledge management as it helps in decision making and problem-solving [30].

Neural Networks, including ANN, RNN, and CNN, simulates the human neural nerves, and it uses learning from provided data and experiments to build the network model and to find the unseen patterns within these data. Neural Networks, including ANN, RNN, and CNN, considered a robust algorithm to predict and forecast cretin scenario cases [31]. Artificial neural networks and data mining work in parallel for the CRM system in order to generate knowledge about customers' behaviors and to predict their willingness to purchase specific items [32]. Additionally, in [33] study, they express the importance of using knowledge management in order to feed the AI system as it is essential for personal characterization, which used to improve public safety using the decision tree algorithm.

It is essential to having a KM feeding to AI systems, which can help to improve AI in call center domains by helping call center agents to find the best resolution for the caller and customer, but KM should have reliable data and source of information. Machine learning (ML) can be employed by reading the history of customer chat conversation and previously solved issues with customers to prepare suggested answers to be used by customer services representative with the ability to customize the prepared answers by ML[24].

According to a study by [34] use many AI algorithms and implement it in a knowledge management system specifically in knowledge extraction to predict Heart failure for patients, among nine classifiers, the results come by employing Random forests algorithm with a range between 10 to 100 trees. Also, in [35] within the same medical domain, they use various AI algorithms, and one of them is the random forests, which is defined as a group of random and parallel created decision trees. The output of this algorithm is considered the average of the output of all decision trees within the forest.

Based on the studies that exist in the literature, most of the reviews are focusing on the direct impact on the knowledge management process to AI systems and algorithms, despite the cost or the possibility of implementing the idea within the organizations. Ignoring other criteria like cost has a negative impact, as seen in one of the studies that, despite the positive effect of system output to enhance performance and data accuracy, the cost of implementing these systems was a significant obstacle to adopt it. These facts encourage us to study the researches available to highlight this missing point and to assist researchers in focusing more on the implementation and adoption of the relation.

Additionally, our study gives other available options for Knowledge Management processes, which can help in enhancing AI systems and algorithms.

### 3. Methodology

The systematic review provides a way to explorer all published papers, chapters, and articles to remove any biased conclusions [36]–[44]. While researches that do not consider a systematic way might lead to higher risk by selecting certain studies that support the author's point of view only and eliminate other points of view [45]. Also, Systematic review gives the reader the latest explored knowledge available in the entire literature for a specific research area or question [46]–[48].

A study by [49] summarize the process of systematic review in six steps. 1) Starting by specifying the research questions or the hypothesis that needs to be tested. 2) Then an inclusion/ exclusion criteria to be determined in order to filter the collected papers based on that criteria. 3) After that, a thorough search for all papers related to the research question to be collected from all available databases by defining the strategy and keywords for the search. 4) review all the retrieved papers from the previous step and evaluate the quality of the papers based on the title, abstract, or the conclusion; this step will eliminate all the un-related papers. 5) Data mining and extraction from the non-eliminated papers from the previous step using thorough reading and analysis of the paper according to research questions. 6) Finally, doing data analysis and answer the research questions providing shreds of evidence from the selected papers, and supported by literature review.

The systematic review is used to understand the impact of applying knowledge management processes in improving the decision making systems for health care [39]. The study utilizes a four steps approach to finalize the selected papers for the review. The steps used in the study succeeded in eliminating papers that were not related to the research questions. Additionally, the study explains in a table, the research methodology and findings from each of the selected papers. However, the table did not show the direction (positive/negative) for each paper and country of implementation as these details may impact the implementation of knowledge management in healthcare decision making systems if applied to countries with different rules and regulations.

Finally, the authors find that following the steps from [40], will help in achieving the paper goals.

Table 1: Inclusion and exclusion criteria.

Inclusion	Exclusion
Should be related to knowledge management Methodologies, processes, or lifecycle.	Articles published before 2014.
Should be related to artificial Intelligent Algorithms, or Techniques.	Related to Knowledge management, but not linked to Artificial Intelligent.
Should be related to the government domain or related to the general domain (without specific domain mentioned).	Related to Artificial intelligence but not linked to Knowledge management.

3.1. Inclusion/exclusion criteria

For the collected paper, it filtered in or out based on certain criteria, which are summarized in Table 1.

So, the above Table 1 shows the criteria that need to be considered when preparing the papers for the systematic review, so the result after applying the inclusion and exclusion criteria will be a list of papers that will be used in this study.

3.2. Data sources and search criteria

In order to collect the papers that are considered in these systematic reviews, a list of keywords was used or a combination of it using “And” operations. These keywords are “Knowledge management”, “Artificial intelligent”, government, “Artificial Intelligence”, “e-Government”. These keywords were used in different publication databases, including “ArticleFirst”, “Electronic Collections Online”, “Emerald Group Publishing Limited”, “Google Scholar”, and “Worldcat”. The total numbers of retrieved papers were 239 .the first exclusion based on the year of publication and all papers before 2014, so 109 papers were eliminated. Two papers were found as duplicated. The remaining papers were 130. After applying all criteria from the inclusion and exclusion table, the remaining number of papers is 16. and listed in Table 2, which contain the source of the paper, details about knowledge management process, artificially intelligent algorithms or system, the research method that used in the paper, the country that targeted by the study, the target domain, the database that collected from and brief and summary about the paper and the main idea behind it.

no additional sources of papers, from screening phase all duplicated papers and year criteria applied to filter out un-needed papers. The eligibility phase came to check the content of the papers if it fit and contain the research required info or not. Moreover, the final phase came to include the remaining papers in the study and analyze it

3.3. Quality assessment

In the systematic review, there is a concern about the quality of the collected papers, even inclusion and exclusion criteria applied, but it is not enough to check the quality of paper contents. A quality test should be conducted as strong evidence of good papers quality can tell how strong is the conclusion can be made [50]. So to check the quality of the papers, a quality assessment checklist was defined as shown in Table 2. For scoring purposes, there will be three values; if the answer to the question is yes, then the score will be 1; if the answer is no, then the score will be 0; if the answer is partially presented, then the score will be 0.5.

Table 2: A quality assessment checklist.

#	Question
1	Are the research aims clearly specified?
2	Are the Knowledge/KM considered by the study clearly specified?
3	Is the AI considered by the study clearly specified?
4	Do the results add to the literature?
5	Does the study add to your knowledge or understanding?

Based on the scoring criteria and Table 2, for each paper, a score between 0 to 5. By applying the quality assessment for all the 16 papers, a result shows in Table 3, P1 to P16 is a list of papers in the same order of papers list in table 4 main table. The results show that all papers are in excellent quality, so all of the 16 papers will be included in the analysis.

Table 3: Quality assessment result.

Source	Q1	Q2	Q3	Q4	Q5	Total	%
P1	1	0.5	1	0.5	1	4	80%
P2	1	1	1	1	1	5	100%
P3	1	1	0.5	1	1	4.5	90%
P4	1	0.5	1	1	1	4.5	90%
P5	1	1	1	1	1	5	100%
P6	1	1	0.5	1	1	4.5	90%
P7	1	1	1	1	1	5	100%
P8	0.5	1	1	1	1	4.5	90%
P9	1	1	0.5	1	0.5	4	80%
P10	0.5	0.5	1	1	1	4	80%
P11	1	1	1	1	1	5	100%
P12	1	1	0.5	1	1	4.5	90%
P13	1	1	1	1	1	5	100%
P14	1	1	1	1	1	5	100%
P15	1	1	1	1	1	5	100%
P16	1	1	1	1	1	5	100%

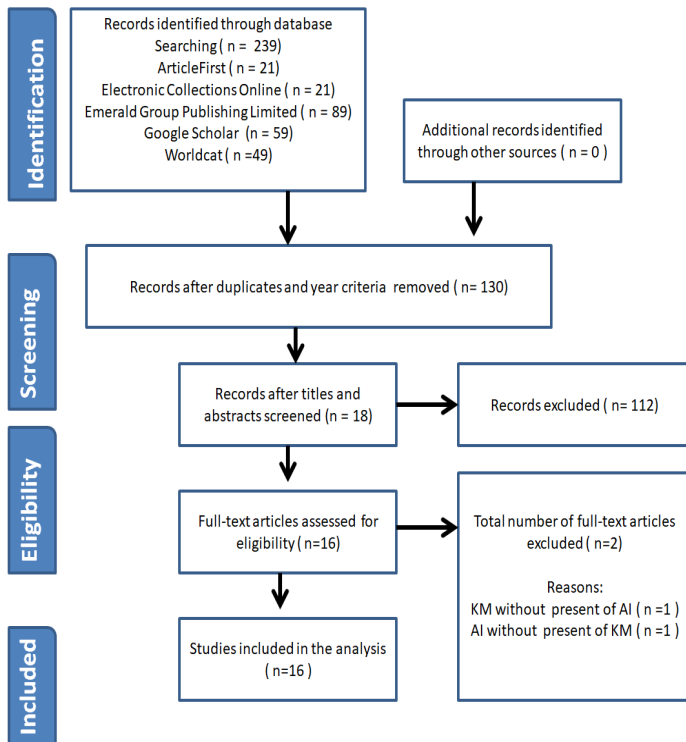


Figure 1: Process of selection papers.

Figure 1 shows the main phases of paper selection in systematic review starting by identifying the databases and collected numbers of each database; also, it explains that there are



3.4. Data coding and analysis

The analysis section depends on the six steps suggested by [51] for a systematic review. In the initial phase, this paper analyzed the relationship and the quality of coded characteristics. Initially, the authors list the main Knowledge management processes for the papers reviewed. In addition to AI Algorithms and Systems, Research database (e.g., WorldCat, Emerald, and others), study direction (Positive, Negative, Neutral), Target Organization (public, private), and country. During the first analysis phase, authors excluded all papers which do not directly describe the impact of knowledge management on AI algorithms and systems. The authors analyzed all documents by dividing all studies between the authors and manually examine the articles. Secondly, the authors reviewed the papers on a high level and start excluding duplicate papers, and papers not on direct relation to this paper topic based on title and abstract screening. After that, the authors deep analyze the remaining researches and start to exclude papers that are not eligible based on a full review and quality assessment questions defined. Finally, 16 papers were selected as they match the quality criteria. Table 4 present the selected papers with the coding characteristics for each paper. The table reflects the analysis done for each paper in rows, by specifying the KM process discussed in the paper, the AI algorithm or system used, and the source research database. Additionally, the table stating the type of research methodology used in each study, the country research used within, and target type audience. Finally, the table illustrated the direction for the paper (Negative/Positive) and added the findings from each study.

Table 4: Analysis of knowledge management research articles with regard to AI algorithms and systems

#	Ref.	KM Process	AI Algorithm / AI Systems	Method	Country
P1	[52]	Knowledge Application /Decision making	Neural Networks including ANN, RNN, and CNN	Survey	USA
P2	[15]	Knowledge Discovery	1) Expert Systems 2) Multi-agent systems 3) Business Intelligence 4) Performance Dashboards	Case Study	Poland
P3	[53]	Knowledge Creation	1) Natural Language Processing 2) Machine-learning techniques 3) Human intelligence	Experimental	General
P4	[31]	Knowledge Application /Decision making	Neural Networks including ANN, RNN, and CNN	Interviews and Experimental	Iran

P5	[54]	Knowledge Creation	AI in General	Experimental	China
P6	[55]	Knowledge Acquisition	Ontology and XML	Experimental	Italy
P7	[29]	Knowledge Creation	AI in General	Case Study	Bangladesh
P8	[33]	Knowledge Application /Decision making	Decision Trees C4.5	Experimental	Colombia
P9	[56]	Knowledge Application /Decision making	Chatbot	Case Study	General
P10	[57]	1) Knowledge-Creating, 2) Knowledge Acquisition 3) Knowledge Sharing	Expert Systems	Literature Review	KSA
P11	[58]	1) Knowledge Acquisition 2) Knowledge Sharing	1) Fuzzy Logic 2) Case-based reasoning	Case study	Taiwan
P12	[59]	Knowledge Acquisition	Fuzzy Logic	Literature Review	Brazil
P13	[60]	Knowledge Acquisition	1) Neural Networks including ANN, RNN, and CNN 2) Support Vector Machines 3) Genetic Algorithms 4) Rough sets 5) Fuzzy Systems	Literature Review	General
P14	[61]	Knowledge Acquisition	1) Expert System 2) Deep Learning	Case study	Australia
P15	[34]	Knowledge-Creating,	Prediction model	Experimental and comparison study	Greece
P16	[35]	Knowledge-Creating,	Prediction model	Experimental	Bangladesh

4. Results and discussion

This systematic review is summarizing 16 articles published between 2014 and 2019 in Google scholar, Emerald, and WorldCat databases on regards to the impact of knowledge management processes in improving AI systems. The findings of this review discussed the following five questions:

4.1. RQ1: What are the primary knowledge management processes studied, and how it is impacting AI systems?

Different studies were done to investigate the impact of knowledge management processes in enhancing the AI systems and algorithms. Table 5 describes the knowledge management processes involved in the study captured, and the frequency of using the method in each study. It is noticeable that knowledge acquisition is the most common processes used to enhance AI systems (N = 6), followed by both knowledge creation and knowledge application/decision-making process (N = 6), knowledge sharing (N = 2), and Knowledge discovery with the least number of studies (N = 1). Referring to Table 4, it seems that knowledge application and decision-making processes have a positive impact on Neural Networks, including ANN, RNN, and CNN [31], Decision Trees [33], and Chatbots [56]. On the other hand, [52] argued that knowledge management application and decision making could have some negative impact on Neural Networks, including ANN, RNN, and CNN, in some applications.

Table 5: Knowledge management processes in the studies reviewed.

Source	Knowledge Application /Decision making	Knowledge Discovery	Knowledge Creation	Knowledge Acquisition	Knowledge Sharing
[52]	X				
[15]		X			
[53]			X		
[31]	X				
[54]			X		
[55]				X	
[29]			X		
[33]	X				
[56]	X				
[57]			X	X	X
[58]				X	X
[59]				X	
[60]				X	
[61]				X	
[53]			X		
[31]			X		

Additionally, it has been noticed that the Knowledge creation process has a positive impact on Natural Language Processing, Machine Learning, Human Intelligence [54], on Expert Systems [57], and on AI algorithms in general [29], [62].

Furthermore, the most knowledge management process impacting the AI systems as per the collected studies are knowledge acquisition and knowledge creation, as it is affecting Fuzzy Logic systems [58]–[60], Expert systems [57], [61], Neural Networks including ANN, RNN, and CNN Support Vector

Machines, Genetic Algorithms, Rough sets [60], Deep Learning [61], Ontology and XML [55], and case-based reasoning [58]. Also, Knowledge Discovery seems to have a positive impact on Expert Systems, Multi-agent systems, Business Intelligence, Performance Dashboards [15]. Finally, the Knowledge Sharing process has a positive impact on Expert systems [57], Fuzzy Logic, and Case-based reasoning [58].

4.2. RQ2: What are the leading AI algorithms and systems reviewed in the studies collected?

Referring to Figure 2, it is noticeable that Expert systems, Neural Networks including ANN, RNN and CNN, and Fuzzy logic algorithms are taking the lead in the number of studies collected (N = 3), followed by general AI study (N = 2), and finally, other AI algorithms and Systems, Genetic Algorithms, Rough sets, Case-based reasoning, Human Intelligence, Machine Learning techniques, Chatbot, Decision Tree, Ontology and XML, Natural Language Processing, Performance Dashboards, Business Intelligence and Multi-language system (N = 1).

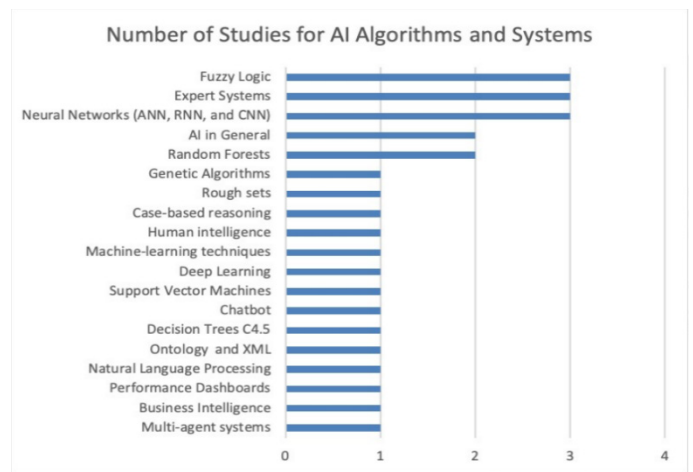


Figure 2: Number of Studies for AI Algorithms and systems.

4.3. RQ3: What research methods have been used in finding the relation between knowledge management processes and AI systems?

Analyzing Table 4 shows multiple research methods been used to find the relation between knowledge management processes, and AI algorithms and systems. Studies use Survey, Case Study, Experimental, Interviews, and Literature Review to perform the required analysis. Figure 3 presents the percentage and number of studies for each methodology. The experimental method takes the lead with seven studies with 44%. Followed by Case Study with five studies and 33%, followed by Literature Review with three studies and 18%, and finally, the lowest two methods used are surveys and interviews with one research each and 6%.

4.4. RQ4: How active is the topic of knowledge management impact on AI during the last five years, across the research databases?

Looking to the year of publication, and by analyzing Figure 4, we notice a drop in the number of researches studied the relationship between knowledge management and AI algorithms and systems since 2014. The highest number of papers was in 2014 (N = 5), then from 2015 to 2017 there with the same number of researches (N = 2), followed by 2018 (N = 2), and finally in 2019 increased to two pieces of research (N = 3).

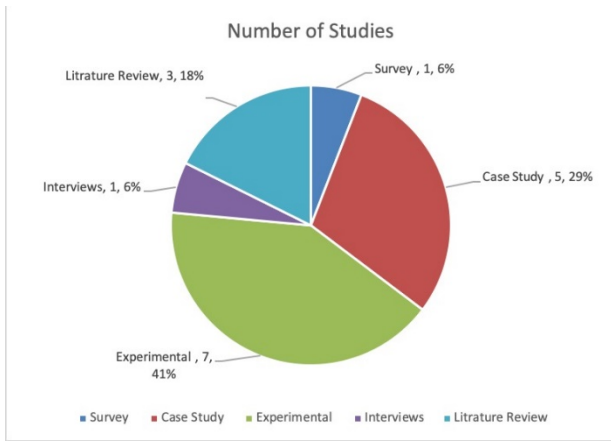


Figure 3: Research Methodology

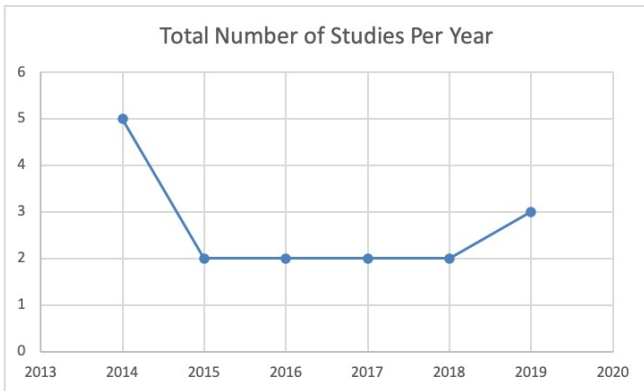


Figure 4: Total number of studies per year

Figure 5 shows how the Worldcat database is popular in having this topic as it has at least one paper every year from 2014 to 2019. The analysis indicates that Worldcat and Emerald have the top number of studies in 2014 (N = 2), followed by Google scholar (N = 1). In 2015, and 2016, Worldcat was the only database with studies related to the topic. In 2017, both Worldcat and Emerald had one research and appeared on both years. while in 2018, both Worldcat and google scholar database with one study for each related to the topic. Finally, in 2019, Worldcat, Emerald, and google Scholar had one research and appeared in each.

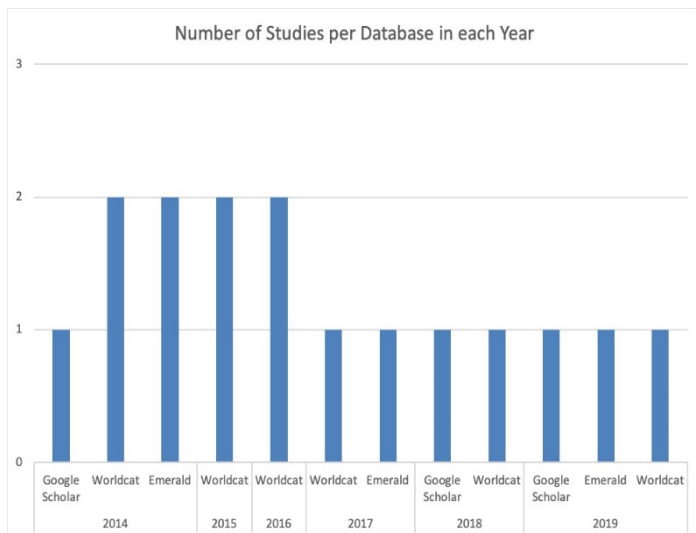
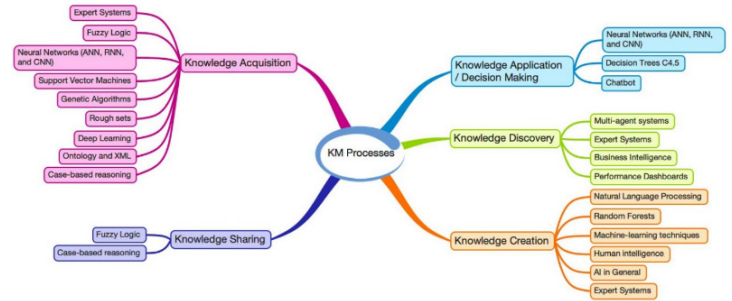


Figure 5: Number of studies per database in each year

Figure 6 describes the relationship between Knowledge management processes and the AI systems and algorithms that can be enhanced or impacted. This is also aligned with the previous discussion of the impact of knowledge acquisition on more number of AI systems and algorithms. Knowledge acquisition can passivity enhance Neural Networks, including ANN, RNN and CNN, Support Vector Machines, Genetic Algorithms, and rough sets, as also described in the study done by [60]. Both [58], [60] also agree that Knowledge acquisition can enhance Fuzzy Systems. Additionally, [61] added that knowledge acquisition could improve Expert systems and Deep learning by having a training way for the algorithms to enhance the outputs. Finally, based on Figure 4, we notice that the number of researches done for the relation between KM and AI systems is decreasing each year, comparing to the studies done in 2014.



Regarding managerial insight, this paper provides a clear understanding of how each stage of knowledge management plays a role in enhancing AI Algorithms. Furthermore, this paper provides Practitioners, including software solutions architectures, digital transformation framework designers, and knowledge management, a solid base to get benefits from their knowledge management and employee the right AI algorithms to increase the benefits by using the right algorithms at the right KM stage.

## 5. Conclusion

The role of Knowledge Management processes in improving AI algorithms and systems was discussed in multiple studies. The paper used systematic review as an approach to determine the relationship between KM and how it affects AI systems and algorithms. This paper followed the six steps summarized by [51], to apply for a systematic review of selected papers. As a result, 16 papers were selected as they fit the requirements and selection criteria defined by the authors. Applying the quality assessment questions on all reviewed papers helped in filtering only high-quality researches, which discussed the selected topic. However, and in some scenarios, authors need to be careful while putting the filtering questions as it could lead to bias in selecting papers based on the expected outcome, not the real results. On the other hand, this paper agreed that knowledge acquisition has a positive impact on enhancing Deep learning and Expert systems related to record-keeper knowledge management in Australia, which increases the job inside offices [61].

In reference to Table 5, an analysis was done to classify the studies referring to Knowledge Management processes in enhancing AI algorithms and systems. It is noticeable that knowledge acquisition is the most common process to improve AI systems. This is in alignment with the studies done by [60], [61], which agree on the positive impact of knowledge acquisition on different AI systems and algorithms. Previous Knowledge management process papers studied the relation with AI systems



and algorithms and focusing on specific AI system or algorithm. In this study, we tried to show the link between KM and AI systems looking from the higher picture, and showing that some KM processes can work not only for specific AI algorithms or methods, but it can also help in improving other AI algorithms. This can help researchers to identify other options to enhance an AI system by checking multiple options for a more cost-effective way if the current process used is having high implementation costs. Analysis finding illustrates the impact of knowledge management acquisition as the most used process to enhance the different type of AI systems and Algorithms. It also highlights that knowledge creation is the second option for improving AI systems. The analysis also shows that although some KM processes like knowledge application may have a positive impact on performance and data accuracy, it can have a negative impact on implementation due to high-cost requirements [52]. This research can suggest other KM processes as an option to achieve the required objectives, which may have less cost for implementations.

The limit of this study was on the implementation cost related to KM processes within AI systems as it will defer from an organization to another, based on the requirements and objectives required to achieve. Another limitation was due to the small number of studies done to find the relation between KM and AI systems, which limited the study outputs and may affect results accuracy. Future research may extend and improve the result of this research by increasing the number of databases that may have more studies related to the topic. Additionally, future researches can add a quality assessment question during the filtering steps to identify the degree or weight for the improvement in percentage; each KM process can give to AI algorithms and systems. For example, researches can highlight the results of how much improvement can KM process like knowledge acquisition give for AI Algorithm like Support Vector Machines. Furthermore, a detailed recommendation table can be created to highlight the best KM processes that can be used for each AI algorithm or system with expected improvement percentage and cost. Finally, this paper can be used as a foundation to prepare a comprehensive mapping table for the KM processes used to improve AI algorithms and systems.

It is beyond the scope of this paper to examine the actual implementation of the KM roles on AI algorithms, but it is recommended to make a real case study for an organization that has an existing knowledge management system or need to build a new knowledge management system. Moreover, to check how the theoretical results are supported in a practical case study.

### Acknowledgment

This work is a part of a project undertaken at the British University in Dubai.

### References

- [1] S. A. Salloum, M. Al-Emran, and K. Shaalan, "The Impact of Knowledge Sharing on Information Systems: A Review," in *13th International Conference, KMO 2018*, 2018. [https://doi.org/10.1007/978-3-319-95204-8\\_9](https://doi.org/10.1007/978-3-319-95204-8_9)
- [2] F. O. Omotayo and M. O. Salami, "Use of Social Media for Knowledge Sharing Among Students," *Asian J. Inf. Sci. Technol.*, 8(2), 65–75, 2018.
- [3] K. S. Ahlam Wahdan, Sendeyah Hantoobi, Said A. Salloum, "A systematic review of text classification research based on deep learning models in Arabic language," *Int. J. Electr. Comput. Eng.*, 10(6), 6629–6643, 2020. <http://doi.org/10.11591/ijece.v10i6.pp%25p>
- [4] S. S. B Al Kurdi, M Alshurideh, "Investigating a theoretical framework for learning technology acceptance," *Int. J. Electr. Comput. Eng.*, 10(6), 6484–6496, 2020. <http://doi.org/10.11591/ijece.v10i6.pp%25p>
- [5] P. López-Portillo, V. González, E. René, R. Hidalgo, and J. Alberto, "Knowledge management metrics for Public Organizations: A literature review-based proposal," *arXiv Prepr. arXiv1609.09541*, 2016. <https://arxiv.org/abs/1609.09541>
- [6] S. Satyanarayan and G. Azumah, "Operationalisation of knowledge management practices in r&d activity in multinational organisations," in *Proceedings of the 12th Management International Conference: Managing Sustainability, Portorož, Slovenia*, 23–26, 2011.
- [7] H. A. Nasiruzzaman, R. Qudaih, and D. Ahmad, "Project success and knowledge management (km) practices in Malaysian institution of higher learning (IHL)," *J. Educ. Vocat. Res.*, 4(5), 159–164, 2013. <https://doi.org/10.22610/jevrv.v4i5.114>
- [8] C. Rusulii, R. Tasmin, J. Takala, and H. Norazlin, "Knowledge Management Process At Malaysian University Libraries: A Review," *Int. J. Soc. Sci.*, 1–16., 2012
- [9] M. A. Arpaci, I. Al-Emran, M., Al-Sharafi, "The impact of knowledge management practices on the acceptance of Massive Open Online Courses (MOOCs) by engineering students: A cross-cultural comparison, Telematics and Informatics," 2020. <https://doi.org/10.1016/j.tele.2020.101468>
- [10] M. Al-Emran and V. Mezhuyev, "Examining the Effect of Knowledge Management Factors on Mobile Learning Adoption Through the Use of Importance-Performance Map Analysis (IPMA)," in *International Conference on Advanced Intelligent Systems and Informatics*, 449–458, 2019. [https://doi.org/10.1007/978-3-030-31129-2\\_41](https://doi.org/10.1007/978-3-030-31129-2_41)
- [11] M. Al-Emran, V. Mezhuyev, and A. Kamaludin, "An Innovative Approach of Applying Knowledge Management in M-Learning Application Development: A Pilot Study," *Int. J. Inf. Commun. Technol. Educ.*, 15(4), 94–112, 2019. DOI: [10.4018/IJICTE.2019100107](https://doi.org/10.4018/IJICTE.2019100107)
- [12] M. Al-Emran, V. Mezhuyev, and A. Kamaludin, "Towards a conceptual model for examining the impact of knowledge management factors on mobile learning acceptance," *Technol. Soc.*, 2020. <https://doi.org/10.1016/j.techsoc.2020.101247>
- [13] W. Bodrow and V. Magalashvili, "It-Based Purpose-Driven Knowledge Visualization," in *ICSOFT (PL/DPS/KE/MUSE)*, 194–197, 2007.
- [14] P. Heisig, "Harmonisation of knowledge management—comparing 160 KM frameworks around the globe," *J. Knowl. Manag.*, 13(4), 4–31, 2009. DOI: <https://doi.org/10.1108/13673270910971798>
- [15] P. Ziuziański, M. Furmankiewicz, and A. Sołtysik-Piorunkiewicz, "E-health artificial intelligence system implementation: case study of knowledge management dashboard of epidemiological data in Poland," *Int. J. Biol. Biomed. Eng.*, 8), 164–171, 2014.
- [16] H. R. Nemat, D. M. Steiger, L. S. Iyer, and R. T. Herschel, "Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing," *Decis. Support Syst.*, 33(2), 143–161, 2002. [https://doi.org/10.1016/S0167-9236\(01\)00141-5](https://doi.org/10.1016/S0167-9236(01)00141-5)
- [17] H. AlGhanem, A. Mustafa, and S. Abdallah, "Knowledge and Human Development Authority in Dubai (KHDA) Open Data: What Do Researchers Want?," in *European, Mediterranean, and Middle Eastern Conference on Information Systems*, 58–70, 2019,. [https://doi.org/10.1007/978-3-030-44322-1\\_5](https://doi.org/10.1007/978-3-030-44322-1_5)
- [18] C. Mhamdi, M. Al-Emran, and S. A. Salloum, *Text mining and analytics: A case study from news channels posts on Facebook*, 740. 2018. [https://doi.org/10.1007/978-3-319-67056-0\\_19](https://doi.org/10.1007/978-3-319-67056-0_19)
- [19] M. Al-Emran, S. Zaza, and K. Shaalan, "Parsing modern standard Arabic using Treebank resources," in *2015 International Conference on Information and Communication Technology Research, ICTRC 2015*, 2015. DOI: [10.1109/ICTRC.2015.7156426](https://doi.org/10.1109/ICTRC.2015.7156426)
- [20] M. Al-Emran and T. Teo, "Do knowledge acquisition and knowledge sharing really affect e-learning adoption? An empirical study," *Educ. Inf. Technol.*, 2019. <https://doi.org/10.1007/s10639-019-10062-w>
- [21] M. Al-Emran, V. Mezhuyev, A. Kamaludin, and M. AlSinani, "Development of M-learning Application based on Knowledge Management Processes," in *2018 7th International conference on Software and Computer Applications (ICSCA 2018)*, 2018. <https://doi.org/10.1145/3185089.3185120>
- [22] M. Al-Emran, V. Mezhuyev, and A. Kamaludin, "Students' Perceptions towards the Integration of Knowledge Management Processes in M-learning Systems: A Preliminary Study," *Int. J. Eng. Educ.*, 34(2), 371–380, 2018. <https://dialnet.unirioja.es/servlet/articulo?codigo=6872062>
- [23] S. Banerjee, "Knowledge Management and Organizational Performance: The Case of a Telecommunication Organization in India," 2018.
- [24] S. Al-Hawamdeh, "Knowledge management: re-thinking information management and facing the challenge of managing tacit knowledge," *Inf. Res.*, 8(1), 1–8, 2002.
- [25] M. M. Shongwe, "An analysis of knowledge management lifecycle



- frameworks: Towards a unified framework,” *Electron. J. Knowl. Manag.*, 14(3, p. 140, 2016.
- [26] D. S. Kim, C. S. Kim, and K. W. Rim, “Modeling and design of intelligent agent system,” *Int. J. Control. Autom. Syst.*, 1(2), 257–261, 2003.
- [27] D. Foster, C. McGregor, and S. El-Masri, “A survey of agent-based intelligent decision support systems to support clinical management and research,” in *proceedings of the 2nd international workshop on multi-agent systems for medicine, computational biology, and bioinformatics*, 16–34, 2005.
- [28] C. M. Olszak and E. Ziemba, “Approach to building and implementing business intelligence systems,” *Interdiscip. J. Information, Knowledge, Manag.*, 2(1), 135–148, 2007.
- [29] M. S. Hossain, P. Zander, M. S. Kamal, and L. Chowdhury, “Belief-rule-based expert systems for evaluation of e-government: a case study,” *Expert Syst.*, 32(5), 563–577, 2015. <https://doi.org/10.1111/exsy.12110>
- [30] H. Vacik, C. Torresan, T. Hujala, C. Khadka, and K. Reynolds, “The role of knowledge management tools in supporting sustainable forest management,” *For. Syst.*, 22(3), 442–455, 2013.
- [31] A. Azadeh, K. Darivandi Shoushtari, M. Saberi, and E. Teimoury, “An Integrated artificial neural network and system dynamics approach in support of the viable system model to enhance industrial intelligence: The case of a large broiler industry,” *Syst. Res. Behav. Sci.*, 31(2), 236–257, 2014. <https://doi.org/10.1002/sres.2199>
- [32] H. Y. Lam, G. T. S. Ho, C.-H. Wu, and K. L. Choy, “Customer relationship mining system for effective strategies formulation,” *Ind. Manag. Data Syst.*, 114(5), 711–733, 2014. <https://doi.org/10.1108/IMDS-08-2013-0329 Download as RIS>
- [33] V. H. Masias, M. A. Valle, J. J. Amar, M. Cervantes, G. Brunal, and F. A. Crespo, “Characterising the Personality of the Public Safety Offender and Non-offender using Decision Trees: The Case of Colombia,” *J. Investig. Psychol. Offender Profiling*, 13(3), 198–219, 2016. <https://doi.org/10.1002/jip.1451>
- [34] E. E. Tripoliti et al., “HEARTEN KMS—A knowledge management system targeting the management of patients with heart failure,” *J. Biomed. Inform.*, 94, p. 103203, 2019. <https://doi.org/10.1016/j.jbi.2019.103203>
- [35] N. Nahar and F. Ara, “Liver disease prediction by using different decision tree techniques,” *Int. J. Data Min. Knowl. Manag. Process.*, 8(2), 1–9, 2018.
- [36] S. A. S. Salloum and K. Shaalan, “Investigating students’ acceptance of E-learning system in Higher Educational Environments in the UAE: Applying the Extended Technology Acceptance Model (TAM).” The British University in Dubai, 2018.
- [37] S. A. Salloum, A. Qasim Mohammad Alhamad, M. Al-Emran, A. Abdel Monem, and K. Shaalan, “Exploring students’ acceptance of e-learning through the development of a comprehensive technology acceptance model,” *IEEE Access*, 7, 2019. DOI: [10.1109/ACCESS.2019.2939467](https://doi.org/10.1109/ACCESS.2019.2939467)
- [38] S. F. S. Alhashmi, M. Alshurideh, B. Al Kurdi, and S. A. Salloum, “A Systematic Review of the Factors Affecting the Artificial Intelligence Implementation in the Health Care Sector,” in *Joint European-US Workshop on Applications of Invariance in Computer Vision*, 37–49, 2020. [https://doi.org/10.1007/978-3-030-44289-7\\_4](https://doi.org/10.1007/978-3-030-44289-7_4)
- [39] S. K. Al Mansoori S., Salloum S.A., “The Impact of Artificial Intelligence and Information Technologies on the Efficiency of Knowledge Management at Modern Organizations: A Systematic Review,” *Al-Emran M., Shaalan K., Hassaniien A. Recent Adv. Intell. Syst. Smart Appl. Stud. Syst. Decis. Control. vol 295. Springer, Cham*, 2021. [https://doi.org/10.1007/978-3-030-47411-9\\_9](https://doi.org/10.1007/978-3-030-47411-9_9)
- [40] S. K. Areed S., Salloum S.A., “The Role of Knowledge Management Processes for Enhancing and Supporting Innovative Organizations: A Systematic Review,” *Al-Emran M., Shaalan K., Hassaniien A. Recent Adv. Intell. Syst. Smart Appl. Stud. Syst. Decis. Control. vol 295. Springer, Cham*, 2021. [https://doi.org/10.1007/978-3-030-47411-9\\_8](https://doi.org/10.1007/978-3-030-47411-9_8)
- [41] S. K. AlShuweih M., Salloum S.A., “Biomedical Corpora and Natural Language Processing on Clinical Text in Languages Other Than English: A Systematic Review,” *Al-Emran M., Shaalan K., Hassaniien A. Recent Adv. Intell. Syst. Smart Appl. Stud. Syst. Decis. Control. vol 295. Springer, Cham*, 2021. [https://doi.org/10.1007/978-3-030-47411-9\\_27](https://doi.org/10.1007/978-3-030-47411-9_27)
- [42] S. K. Almansoori A., AlShamsi M., Salloum S.A., “Critical Review of Knowledge Management in Healthcare,” *Al-Emran M., Shaalan K., Hassaniien A. Recent Adv. Intell. Syst. Smart Appl. Stud. Syst. Decis. Control. vol 295. Springer, Cham*, 2021. [https://doi.org/10.1007/978-3-030-47411-9\\_6](https://doi.org/10.1007/978-3-030-47411-9_6)
- [43] S. K. Habeh O., Thekrallah F., Salloum S.A., “Knowledge Sharing Challenges and Solutions Within Software Development Team: A Systematic Review,” *Al-Emran M., Shaalan K., Hassaniien A. Recent Adv. Intell. Syst. Smart Appl. Stud. Syst. Decis. Control. vol 295. Springer, Cham*, 2021. [https://doi.org/10.1007/978-3-030-47411-9\\_7](https://doi.org/10.1007/978-3-030-47411-9_7)
- [44] S. K. Yousuf H., Lahzi M., Salloum S.A., “Systematic Review on Fully Homomorphic Encryption Scheme and Its Application,” *Al-Emran M., Shaalan K., Hassaniien A. Recent Adv. Intell. Syst. Smart Appl. Stud. Syst. Decis. Control. vol 295. Springer, Cham*, 2021. [https://doi.org/10.1007/978-3-030-47411-9\\_29](https://doi.org/10.1007/978-3-030-47411-9_29)
- [45] R. French and P. Brocklehurst, “The effect of pregnancy on survival in women infected with HIV a systematic review of the literature and meta-analysis,” *BJOG An Int. J. Obstet. Gynaecol.*, 105(8), 827–835, 1998. <https://doi.org/10.1111/j.1471-0528.1998.tb10226.x>
- [46] D. M. Rousseau, J. Manning, and D. Denyer, “11 Evidence in management and organizational science: assembling the field’s full weight of scientific knowledge through syntheses,” *Acad. Manag. Ann.*, 2(1), 475–515, 2008. <https://doi.org/10.5465/19416520802211651>
- [47] N. Al-Qaysi, N. Mohamad-Nordin, and M. Al-Emran, “Employing the technology acceptance model in social media: A systematic review,” *Educ. Inf. Technol.*, 1–42, 2020. <https://doi.org/10.1007/s10639-020-10197-1>
- [48] M. Fatehah, V. Mezhujev, and M. Al-Emran, “A Systematic Review of Metamodelling in Software Engineering,” in *Recent Advances in Intelligent Systems and Smart Applications*, 3–27, Springer, 2021. [https://doi.org/10.1007/978-3-030-47411-9\\_1](https://doi.org/10.1007/978-3-030-47411-9_1)
- [49] F. F. Tránsito, M. R. Helena, J. G. P. Francisco, and A. M. V. José, “The systematic review of literature in LIS: An approach,” in *ACM International Conference Proceeding Series*, 02-04-(Nove), 291–298, 2016. <https://doi.org/10.1145/3012430.3012531>
- [50] L. Victor, “Systematic reviewing,” *Soc. Res. Updat.*, 54(1), 1–4, 2008.
- [51] T. Ferreras Fernández, H. Martín-Rodero, F. J. García-Peñalvo, and J. A. Merlo Vega, “The systematic review of literature in LIS: an approach,” 2016. <https://doi.org/10.1145/3012430.3012531>
- [52] O. Alsetoohy, B. Ayoun, S. Arous, F. Megahed, and G. Nabil, “Intelligent agent technology: what affects its adoption in hotel food supply chain management?,” *J. Hosp. Tour. Technol.*, 2019. <https://doi.org/10.1108/JHTT-01-2018-0005>
- [53] X. Liu, C. Guo, and L. Zhang, “Scholar metadata and knowledge generation with human and artificial intelligence,” *J. Assoc. Inf. Sci. Technol.*, 65(6), 1187–1201, 2014. <https://doi.org/10.1002/asi.23013>
- [54] L. Li et al., “Artificial intelligence test: a case study of intelligent vehicles,” *Artif. Intell. Rev.*, 50(3), 441–465, 2018. <https://doi.org/10.1007/s10462-018-9631-5>
- [55] G. Boella, L. Di Caro, L. Humphreys, L. Robaldo, P. Rossi, and L. van der Torre, “Eunomos, a legal document and knowledge management system for the web to provide relevant, reliable and up-to-date information on the law,” *Artif. Intell. Law*, 24(3), 245–283, 2016. <https://doi.org/10.1007/s10506-016-9184-3>
- [56] O. Smilansky, “The Real Benefits of Artificial Intelligence: Adding AI to the contact center enables rapid information retrieval, leading to a better experience for both agents and customers,” *Cust. Relatsh. Manag.*, 21 (11), 28–31, 2017.
- [57] A. Y. Al-Aama, “Technology knowledge management (TKM) taxonomy: using technology to manage knowledge in a Saudi municipality,” *Vine*, 44(1), 2–21, 2014. <https://doi.org/10.1108/VINE-12-2012-0052>
- [58] K. L. T. Choy et al., “An intelligent case-based knowledge management system for quality improvement in nursing homes,” *VINE J. Inf. Knowl. Manag. Syst.*, 48(1), 103–121, 2018. <https://doi.org/10.1108/VJIKMS-01-2017-0001>
- [59] A. Sheffer Corrêa, A. de Assis Mota, L. Toledo Moreira Mota, and P. Luiz Pizzigatti Corrêa, “A fuzzy rule-based system to assess e-government technical interoperability maturity level,” *Transform. Gov. People, Process Policy*, 8(3), 335–356, 2014. <https://doi.org/10.1108/TG-08-2013-0028>
- [60] H. M. Said and A.-B. M. Salem, “Exploiting Computational intelligence Paradigms in e-Technologies and Activities,” *Procedia Comput. Sci.*, 65), 396–405, 2015.
- [61] G. Rolan, G. Humphries, L. Jeffrey, E. Samaras, T. Antsoypova, and K. Stuart, “More human than human? Artificial intelligence in the archive,” *Arch. Manuscripts*, 47(2), 179–203, 2019. <https://doi.org/10.1080/01576895.2018.1502088>
- [62] Z. Li, W. M. Wang, G. Liu, L. Liu, J. He, and G. Q. Huang, “Toward open manufacturing: A cross-enterprises knowledge and services exchange framework based on blockchain and edge computing,” *Ind. Manag. Data Syst.*, 118(1), 303–320, 2018.

## Potential of Solar Energy in Residential Rooftop Surface Area in Semarang City, Indonesia

Djoko Adi Widodo<sup>1\*</sup>, Purwanto Purwanto<sup>1,2</sup>, Hermawan Hermawan<sup>1,2</sup>

<sup>1</sup>Doctorate Program of Environmental Science, School of Postgraduate Studies, Universitas Diponegoro, Semarang, 50241, Indonesia

<sup>2</sup>Department of Chemical Engineering, Faculty of Engineering, Universitas Diponegoro, Semarang, 50275, Indonesia

### ARTICLE INFO

Article history:

Received: 02 June, 2020

Accepted: 16 July, 2020

Online: 28 July, 2020

Keywords:

Geographic Information System

Residential Rooftop

Solar Energy Potential

### ABSTRACT

*This study examines the potential of solar energy in the urban environment with a case study in Semarang, Indonesia by analyzing the intensity of solar radiation and the residential rooftop area. The study aims to obtain a quantitative description of the potential for electricity production from rooftop solar photovoltaic systems in residential areas and estimate the mitigation potential of CO<sub>2</sub>. The estimation method has adopted the hierarchies assessment: estimation of physical, geographic, and technical potential. This study shows the residential roof area spread over 16 districts in the city of Semarang is 412,987.50 m<sup>2</sup> to 2,083,387 m<sup>2</sup> has the average potential to of solar energy every year of 44,051 - 222,222 MWh/year. Total the low-carbon electricity is equivalent to 40.87% of the total electricity consumption in 2018 at Semarang City and reduce 1,394 tonCO<sub>2</sub> in a year. Potential electricity production is proposed to set rules for the future empowerment of solar energy and analyze the potential at different time levels, such as monthly, weekly and daily.*

## 1. Introduction

Increasing concerns on the use of carbon-based or fossil fuel-based energy are apparent around the world. Approximately 70% of the global electricity supply is currently generated from fossil fuels [1]. It is estimated that if global energy demand grows from 2010 to 2030 at 1.5% per year, it could lead to an increase 22% in oil consumption, 42% in natural gas and 53% in coal during a 20-year time frame [2]. The use of fossil fuels for electricity generation releases greenhouse gases into the atmosphere, which causes climate change. The problems are considered not good from an environmental perspective, especially the utilization of electrical energy from unsustainable resources such as increases carbon dioxide concentration in atmospheric from greenhouse gas emission, environmental safety from energy production techniques, unstable prices of energy, and depletion of fuel reserves from carbon sources [3, 4]. Due to extensive exploitation, fossil fuels have gradually become depleted and the problem of global warming increases. Therefore, all the countries should develop alternative energy sources in order to reduce the potential risk of depletion of available fossil fuel energy supplies and to overcome degradation of environmental [5]. Nowadays, many nations are facing increasing challenges in order to diversify energy sources. Most countries that tend to the development of

low carbon sustainability pathways have made climate change mitigation a major concern and increased the research about analyzing various mitigation policies of climate change in the national context [6]. Renewable energy sources are very imperative in their potential to play a dual role in mitigating of global warming and assure long-term energy security. Renewable energy is a vital element for every sustainable solution [7, 8].

A power generation technology of renewable energy is a solution to the concerns of the use of fossil fuels [9]. Solar energy is a potential renewable energy resource for the world to recovery the environmental problems with clean energy [10]. The solar energy is widely regarded as a source of major renewable energy expected to contribute to the security of energy supply and environmental protection. Solar energy is readily available, the most abundant, and clean of all renewable energy resources to date, universal, noiseless and non-polluting [11, 12]. Solar energy resources have a significant role in future energy resources [13, 14]. Solar power technology is one of the main solutions to fulfill the increasing global energy demand [14]. There are two technologies used to harness solar energy: solar thermal and solar photovoltaic [15,16].

Photovoltaic technology has become a promising renewable energy supply from clean energy sources in all of scale production [17]. Its use can produce resources with zero emissions, zero noise, and reliability [17]. In the last

\*Corresponding Author: Djoko Adi Widodo, Email: [djoko.adiw@gmail.com](mailto:djoko.adiw@gmail.com)

decades, photovoltaic technology has been growing more rapidly than other renewable energy sources. Solar photovoltaic systems in rooftop are increasingly becoming fundamental system in the energy sector globally. It is reported that they contribute around 50% of the total installed solar energy capacity [21, 22]. Photovoltaic is seen as the main climate change mitigation technology. Achieving this potential requires large-scale photovoltaic installations, either on the roof of a house or as an array mounted on the ground [20].

Local renewable energy generation through photovoltaic systems mounted on a rooftop buildings in urban areas with great potential for mitigating emissions of greenhouse gas [21]. The potential of photovoltaic electricity generation on the rooftop depends from several global, local, temporal, and spatial variables [22]. The geographical potential of a rooftop solar photovoltaic system of an area can be determined by estimating the availability of the roof area in a certain percentage [23]. The intensity of the solar radiation in Indonesia is around 4.8 kWh/m<sup>2</sup> per day in all its regions. It means that every 1 m<sup>2</sup> with the solar module can produce 4.8 kWh of electricity every day [24]. The solar photovoltaic potential of a rooftop can be determined hierarchically through of different stages: (i) potential of physical: the total amount of solar energy that reaches the surface target, it can be called total solar radiation on the roof; (ii) potential of geographical: spatial availability of roofs in buildings, where solar energy can be obtained, it can be called to as the roof area available for solar photovoltaic installations; (iii) technical potential: the total amount of electricity that takes into technical characteristics of solar photovoltaic systems: module efficiency, it can be called electricity generation [25,26].

Based on the literature review, the primary energy combination for Indonesia's electricity generation is still largely dominated by fossil energy sources such as coal and petroleum, especially in big cities [27]. Clean energy from the energy source of solar radiation should be utilized. Therefore, the investigation of the potential of solar energy in urban areas is also required to determine the contribution of solar energy as a clean energy supply for domestic use and the reduction of CO<sub>2</sub>. Government policies regarding the use of PV systems in roof buildings should be set to support the National Energy Policy in 2025, which increases the role of new and renewable energy, by 36% in 2050. This study aims to obtain an estimation of the solar energy potential and mitigation of CO<sub>2</sub> potential based on photovoltaic roof systems in Semarang City as one of big city in Indonesia. The rooftop solar photovoltaic system in Semarang City is currently not widely installed, even though geography and meteorology have great solar energy potential. Until 2019, a rooftop solar photovoltaic system has only installed 95 kWp capacity in 3 government buildings and is connected to the national grid. Specifically, studies have not yet been carried out on the potential of local area solar energy for electricity production in the roof space of buildings in a comprehensive manner. In this study, the evaluation of solar energy potential focuses on the roofs of residential areas considering that buildings are a potential area for solar panel installation in urban areas. This paper is a preliminary study, undertaking the work of estimating the potential roofs of buildings and residential houses. From this study will have the knowledge and description of the potential of roofs of buildings and houses for the production of electrical energy from the sun.

## 2. Methodology

### 2.1. The Study Area

The case study in this paper is Semarang City, Indonesia, located between the 6°50' – 7°10' South and 109°35' – 110°50' East (Figure 1). The area of Semarang City is 373.70 km<sup>2</sup>, it has a tropical climate condition and the type of climate according to the Koppen classification is Tropical Monsoonal. The annual average temperature is 26.7°C. The population is made by 1,815,729 people and the population density is 4,253 per km<sup>2</sup> with a population growth of 1.6% per year [28]. Administratively, the City of Semarang is divided into 16 sub-districts [29]. Data from the Central Statistics Agency of Semarang City (2018) shows that electricity consumption in 2017 is 4,704,416 MWh/year or around 23% of the total electricity consumption in Central Java Province with an area of 32,801 km<sup>2</sup>. On a national scale, the number of customers has increased by 5.6% from the previous year, while electricity consumption has increased by 1.6% [30].

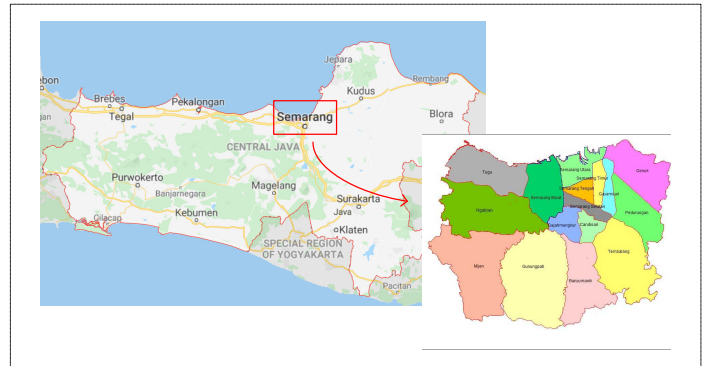


Figure 1: The Map of Semarang City, Indonesia

Topographically, the city of Semarang consists of hilly areas [31] lowland and coastal areas [32,33]. Therefore, it has an area referred to as a downtown area or lowlands and a high hills area [34]. The downtown area of the city is the center of government, trade, industry, and education activities. In contrast, the hilly areas or upper areas of the city are mostly used for residential areas [35]. In general, land use in the Semarang City area is used for roads, residential areas or housing, buildings, courtyards, industrial estates, the brackish water pond and field [36]. Geographically, the coastal area of the city of Semarang is strongly influenced by the sea wind from the north and the valley wind of Mount Ungaran from the south.

### 2.2. Evaluation of Solar Radiation

Solar radiation is a solar energy that reaches the earth's surface. The radiation consists from three components: direct beam, diffuse, and ground-reflected radiation [37]. Direct radiation is the direct emission of solar energy and intercepted by a surface without interaction with many particle in the atmosphere [38]. Diffuse radiation is scattered and intercepted in the atmosphere by aerosols and gases [39]. The radiation reflected is the one reflected from the field and surrounding surfaces [40]. The three radiations, which are direct, diffuse and reflected make up for the global radiation or total radiation reaching the earth's surface. The amount of solar radiation that reaches the earth's surface depends on its location, its atmospheric influence, and its topography. Solar radiation is affected by the



rotation of the earth's geometry and revolutions of the sun [41]. At ground level, topographic influences such as slope, elevation, and orientation affect the radiation to reach surface [42].

The Meteorology, Climatology, and Geophysics Agency (BMKG) in Semarang Climatology Station conducts periodic observations of solar radiation intensity. In order to estimate the physical potential, this study uses solar radiation measured by BMKG Climatology Station in Semarang City, Indonesia. The potential of solar radiation is estimated for electricity generation by photovoltaic on the roofs of residential areas in 16 sub-districts of Semarang City by using the data from 2014 to 2018.

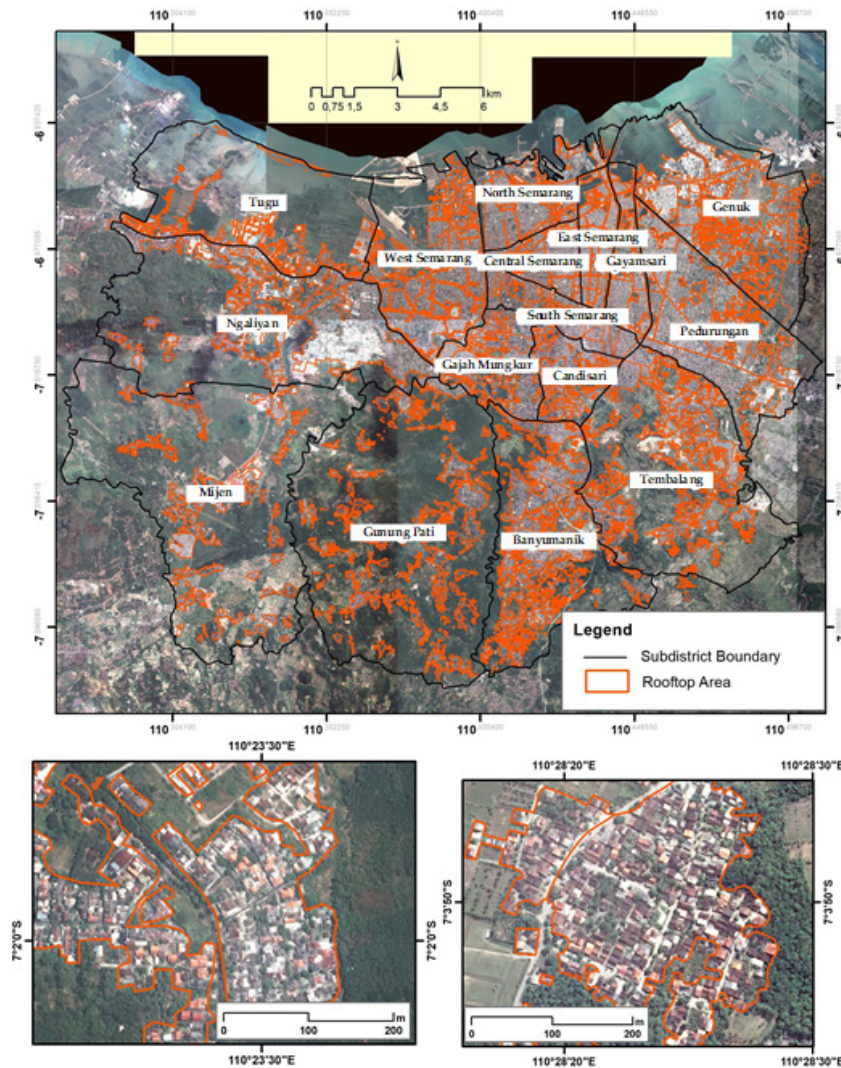
### 2.3. Evaluation of Building Roof Surface

Currently, along with the development of the population of the city of Semarang also develops residential buildings in various regions and correlates to increase the amount of electricity consumed. Some studios show that the roof of residential buildings has the potential to generate electricity from solar energy. Potential use of roofs to generate electricity using photovoltaic media conversion panels. Electric production from the roof of the

building can be intended to supply electrical energy needs so it needs attention.

The availability of the building roof area for the installation of a photovoltaic system can be determined by evaluating the geographic potential of the Semarang city. In the study, the area of the roof that has the potential to capture the energy the solar radiation is estimated through the remote sensing and Geographic Information System (GIS).

The GIS software used in this study is ArcGIS 10.5. The remote sensing data used in the form of aerial photos from the high-resolution satellite imagery with a resolution of 5m x 5m. The high-resolution satellite imagery of Semarang taken in the year 2017 was obtained from the results of the partnership with geospatial information agency. The estimation of the roof area of the settlements was conducted by the interpretation of the visual object of building manually using the Geospatial Information System (Figure 2). By using the geographic information system, the area of the roof in the whole city of Semarang spread in 16 sub-districts can be determined.



In this study for the purposes of calculating the potential of a building roof carried out several stages as follows. The roof area

for all high-rise buildings and houses is identified as a flat roof through satellite imagery analysis. Roof identification is done by



each district so that the estimated roof area is more representative of the character of the building in each district. Roof characteristics for buildings with flat and sloped roofs are given with the roof coefficient in equation (1) to determine the total roof in each district:

$$A_R = \sum_{i=1}^n A_i \times R_C \quad (1)$$

where  $A_R$  is a total rooftop area available per district in  $m^2$ ,  $n$  is the number of polygons per district,  $A_i$  = area of each roof polygon in  $m^2$  and  $R_C$  is the available roof coefficient. The coefficient value considers utilities placed on the roof such as water tanks, air conditioners system, communication towers, etc. Besides, it also considers the general construction conditions and strength of building roofs. As a general rule of thumb, type of crystalline photovoltaic module weighted 15-20 kilograms per square meter ( $kgs/m^2$ ) of dead load on the roof [43].  $R_C$  coefficient is determined around 12.5% - 15%. Roof coefficient ( $R_C$ ) considers the regulation of roofing used for PV systems has not been regulated by the Ministry of Energy and Mineral Resources of the Republic of Indonesia.

Residential area in Semarang City consist of house and high-rise buildings. The roof area according to the percentage of concrete roofs 3.95%, roof tiles 72.66%, zinc 21.12% and asbestos 2.18% [29]. Concrete-roofed buildings around 111 buildings with a minimum height of 7 floors. Residential house building refers to the number of families around 424,628 houses with roof tiles, zinc and asbestos.

#### 2.4. Evaluation of Technical Potential

Solar energy received from sunray hits the earth and is commonly called as solar radiation. Solar radiation can be utilized and converted as electricity energy with photovoltaic technology. Photovoltaic cells produce the electricity energy with absorbing photons and release electrons which can be captured in the form of electric current [44]. The cells can be grouped into modules and arrays to produce a greater amount of power or used individually to turn on small electronics. The potential for photovoltaic electricity generation on residential roofs depends on several conditions such as temporal, global, local, or spatial variables [22]. The factors that influence electricity generation are the total of solar radiation penetrating the earth's surface. The available solar radiation and the roof area are components in estimating the rooftop solar photovoltaic system potential. The technical potential of a rooftop solar photovoltaic system in Semarang City was be estimated by calculating potential capacity of the solar PV rooftop system,  $C_R$  in kilowatt-peak (kWp) using the following equation (2) [43]:

$$C_R = \left( \frac{C_M}{1.000} \right) \times \left( \frac{RCR \times A_R}{A_M} \right) \quad (2)$$

With  $A_R$  is the amount roof area available for installation the solar modules in  $m^2$ ,  $C_M$  is the individual module rated with capacity in Wp (rated at 200 peak Watts),  $A_M$  is the area of one module in  $m^2$  (sized 1,487 meters by 0,992 meter), RCR is the

roof cover ratio, which is the part of roof area that the modules will cover and value of RCR is 0.85. Energy yield (E) in kilowatt-hour (kWh) calculated using this equation (3):

$$E = C_R \times GSR \times D \quad (3)$$

GSR is the average global solar radiation over 5 years period in  $kWh/m^2/month$ . D is the derate factor value, that is converting direct current (DC) to alternating current (AC). The derate factor value ranges between 0.6 and 0.8 [43]. And typical derate factor of 0.75 used in this paper.

### 3. Result and Discussion

#### 3.1. Physical Potential

The estimation energy potential requires of solar photovoltaic an evaluation of physical potential (using solar radiation), geographic potential (roof surface availability), and technical potential (system efficiency of photovoltaic). Table 1 shows variations of average monthly global solar radiation in the Semarang City area in 12 months each year. In this study, the intensity of solar radiation was measured and recorded by BMKG, Semarang Climatology Station for 5 years. The intensity of daily global solar radiation is measured for several hours and tabulated in daily average with units of  $Cal/m^2/day$ . For analysis in this study, the unit data of global solar radiation was converted to  $kWh/m^2/month$ .

Solar radiation that reaches the surface of the earth per unit area and time is known as insolation (derived from insolation = incoming solar radiation), or sometimes referred to as global radiation. Insolation plays an important role in maintaining the continuity of life on this earth and is very dependent on place and time. Places represent the differences in latitude and the atmosphere, especially clouds. Insolation is usually expressed in units of  $Watt/m^2$ -sec which has the meaning of intensity or strength. In another form, insolation is also measured in units of hours/day, for example, the length of the sun shining on the earth in one day. One day is also referred to as the length of the day, the length of the sun on the horizon. The change in day length is not too significant in tropical regions such as in Semarang City, Indonesia, which is close to the equator. The farther away the equator is, the greater the fluctuations in the irradiation. Based on the definition issued by the *World Meteorological Organization*, solar radiation is defined as the insolation power that exceeds  $120 W/m^2$  [45].

#### 3.2. Geographical Potential

The process of land use classification is a method of roof cover detection. This method detects the roof of a building surface. The utilization of satellite image data is very useful in this case. The density of the settlement area in Semarang is very beneficial in the estimation of the area by using GIS. It enables the large – scale estimation with high accuracy. Figure 5 is the result of the analysis of the distribution rooftops area in the city of Semarang. Most researchers use GIS to estimate the cross-sectional area with a

Table 1: Global Solar Radiation in Semarang City, Indonesia

Month	2014	2015	2016	2017	2018
	kWh/m <sup>2</sup> per month				
Jan	92.37	61.94	109.47	83.91	83.91
Feb	86.41	91.97	113.53	74.79	93.75
Mar	110.37	107.67	121.36	110.70	103.39
Apr	87.37	100.64	93.05	90.64	81.58
May	90.28	102.63	119.92	101.91	86.86
Jun	80.40	87.82	109.43	75.97	79.11
Jul	93.45	93.27	121.72	96.69	78.29
Aug	120.85	101.77	134.68	108.03	116.32
Sep	120.61	111.17	139.75	108.03	119.88
Oct	129.96	143.68	144.40	97.59	106.66
Nov	109.53	131.03	142.88	85.03	117.09
Dec	88.51	94.35	83.19	88.59	88.69

Table 2: Analysis settlements rooftop area per district in Semarang City, Indonesia

District	Total Area / $A_i$ (m <sup>2</sup> )	Roof available / $A_R$ (m <sup>2</sup> )	Zone
Gunung Pati	8,719,100	1,264,269.5	High hills area
Mijen	6,580,800	954,216.0	High hills area
Tembalang	11,377,900	1,649,795.5	High hills area
Banyumanik	13,695,300	1,985,818.5	High hills area
East Semarang	3,958,200	494,775.0	Lowlands area
Candisari	5,140,700	642,587.5	Lowlands area
Genuk	13,599,200	1,699,900.0	Lowlands area
Tugu	3,303,900	412,987.5	Lowlands area
Ngaliyan	16,667,100	2,083,387.5	Lowlands area
Central Semarang	4,422,500	552,812.5	Lowlands area
North Semarang	5,648,600	706,075.0	Lowlands area
West Semarang	11,071,300	1,383,912.5	Lowlands area
Gayamsari	3,694,700	461,837.5	Lowlands area
Pedurungan	16,660,300	2,082,537.5	Lowlands area
South Semarang	6,204,600	775,575.0	Lowlands area
Gajah Mungkur	7,017,500	877,187.5	Lowlands area
Total	137,761,700	18,027,675	

Table 3: Technical potential of rooftop solar photovoltaic system in Semarang City, Indonesia

Distric	CR (kWp)	Energy Potential (MWh/month)											
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Gunung Pati	9,465	10,098	12,139	9,941	11,001	9,490	10,602	12,756	13,147	13,648	12,842	9,723	9,465
Mijen	7,144	7,622	9,162	7,503	8,303	7,163	8,002	9,628	9,922	10,301	9,693	7,338	7,144
Tembalang	12,352	13,177	15,840	12,973	14,355	12,384	13,835	16,646	17,155	17,810	16,758	12,688	12,352
Banyumanik	14,867	15,861	19,066	15,615	17,279	14,906	16,652	20,037	20,650	21,437	20,172	15,272	14,867
East Semarang	3,704	3,952	4,750	3,890	4,305	3,714	4,149	4,992	5,145	5,341	5,026	3,805	3,704
Candisari	4,811	5,133	6,170	5,053	5,591	4,824	5,389	6,484	6,682	6,937	6,527	4,942	4,811
Genuk	12,727	13,578	16,321	13,366	14,791	12,760	14,255	17,152	17,676	18,351	17,267	13,073	12,727
Tugu	3,092	3,299	3,965	3,247	3,593	3,100	3,463	4,167	4,294	4,458	4,195	3,176	3,092
Ngaliyan	15,598	16,641	20,003	16,382	18,128	15,639	17,471	21,021	21,664	22,490	21,163	16,022	15,598
Central Semarang	4,139	4,415	5,308	4,347	4,810	4,150	4,636	5,578	5,748	5,968	5,615	4,251	4,139
North Semarang	5,286	5,640	6,779	5,552	6,144	5,300	5,921	7,124	7,342	7,622	7,172	5,430	5,286
West Semarang	10,361	11,054	13,287	10,882	12,042	10,388	11,605	13,963	14,391	14,939	14,058	10,643	10,361
Gayamsari	3,458	3,689	4,434	3,631	4,019	3,467	3,873	4,660	4,802	4,986	4,691	3,552	3,458
Pedurungan	15,592	16,634	19,995	16,375	18,121	15,632	17,464	21,012	21,655	22,481	21,154	16,016	15,592
South Semarang	5,807	6,195	7,447	6,098	6,748	5,822	6,504	7,825	8,065	8,372	7,878	5,965	5,807
Gajah Mungkur	6,567	7,006	8,422	6,897	7,633	6,585	7,356	8,851	9,121	9,469	8,910	6,746	6,567
Jumlah	2,084,833	134,969	143,992	173,090	141,754	156,862	135,324	151,175	181,896	187,461	194,610	183,123	138,641

application of cities using other techniques such as LIDAR and aerial photography is highly unlikely, because of the high cost. Data are in the form of a vector. This study found the results of the calculation of the available roof area calculated using Equation 1 are shown in Table 2. The coefficient of the roof ( $R_C$ ) for the high hills area is 0,145 and the lowlands area is 0,125. The coefficient difference considers the general utility of a building roof in Semarang City.

The most potential areas of rooftops in settlements Semarang (more than 12%) are in the region of Ngaliyan and Pedurungan sub-districts. Other regions that have a big proportion of the rooftop area consist of Tembalang, Banyumanik, Genuk dan West Semarang. Those sub-districts are dense regions that have the potential of the roof distribution more than 8% of the area of the roof of the whole. Because those regions are narrow and most areas are developed with high rooftop potential. However, large areas, such as Tembalang and Banyumanik, although located in hilly areas have the potential roof area to reach 8-9% of the total area. The developed area is influenced by the level of development of the region itself. In contrast to other regions, Gunung Pati and Mijen with their wider area are less developed because these regions are the rein green belt of the city.

### 3.3. Technical Potential

Potential techniques of the photovoltaic system in rooftop can be referred to as the energy electrical generation analyzed for settlements that are distributed on the 16 sub-districts in Semarang city, Indonesia. According to the variation of GSR per month in Semarang City and using a typical for derate factor, energy potential is found and shown in Table 3.

The technical potential is indicated by the value of potential energy proportional to the area of an available rooftop. The more area of rooftop used for photovoltaic exposure leads to more intensity of solar radiation the roof of the photovoltaic. It leads to more potential energy. Table 3 shows the variation of the potential energy per month which is produced from the roof of

settlements in each sub-district. Figure 6 presents Tugu sub-district, which has the least roof area gained potential energy to 44,051 MWh/year. In contrast to Ngaliyan sub-district which has the potential of generating energy electricity 222,222 MWh/year, because the area surface of the roof provided for exposure of the photovoltaic is also most extensive than the area of the roof surface on the other sub-district.

### 3.4. Potential of CO<sub>2</sub> Mitigation

The production of electrical energy from the rooftop solar photovoltaic system model has the potential to reduce CO<sub>2</sub> emissions. The reduction in CO<sub>2</sub> emissions is proportional to the potential energy produced from the roof PV system. The CO<sub>2</sub> mitigation coefficient is used to predict CO<sub>2</sub> migration in the city of Semarang [46]. Figure 3 shows potential energy of solar photovoltaic, and prediction of CO<sub>2</sub> mitigation in Semarang City, Indonesia. Installation of solar photovoltaic rooftop in Semarang City can reduce 1,394 tonCO<sub>2</sub> in a year.

## 4. Conclusion

An overview of the potential for solar energy based on the photovoltaic roof system the potential of CO<sub>2</sub> mitigation in Semarang City has been obtained. The estimation of solar energy potential has been focused on residential areas and refers to the analysis of the use of residential roof surfaces for solar panel installation. The technique of estimating solar energy potential using GIS area analysis is very beneficial in a wide area. This technique is very useful in estimating the area precisely, with a low error rate. The land use analysis technique with GIS that uses high-resolution satellite imagery (CSRT) is more beneficial in terms of cost and time. The estimation of the rooftop area is very dependent on the type of image used. The higher accuracy of a satellite image leads to the higher precision of the area to. In addition to the use of the data with high accuracy, all the geographic data calculations are based on geographic reference data.

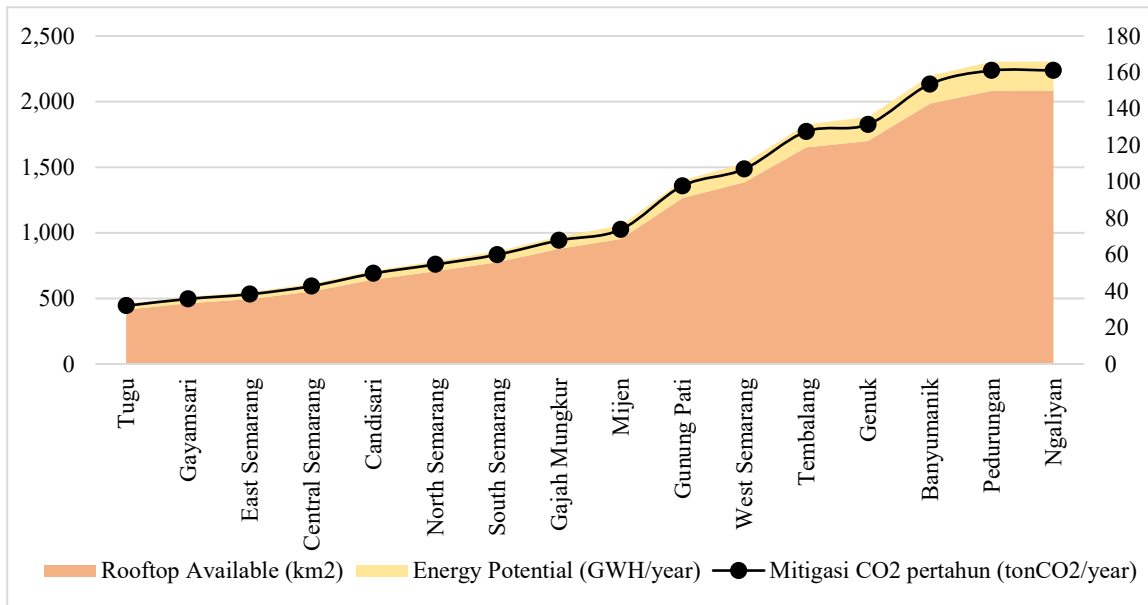


Figure 3: Potential of solar photovoltaic rooftop and prediction of mitigation CO<sub>2</sub> on the settlements in Semarang City, Indonesia

This paper shows the potential of a rooftop photovoltaic system of solar energy in the 16 districts of Semarang City is 44,051 – 222,222 MWh/year. The average monthly solar radiation intensity in the 16 district of Semarang City ranges from 61.94 - 144.40 kWh/m<sup>2</sup> per month, while the roof potential of each district varies from 412,987.50 m<sup>2</sup> to 2,083,387 m<sup>2</sup>. Based on the results of the potential for solar energy, available roof area, and daily solar radiation intensity has been described in each district, total of solar energy potential from 16 district is expected to contribute approximately 40.87% as clean or low-carbon energy to electrical energy consumption needs of Semarang City with reference electricity consumption in 2018. The value can reduce 1,394 tonCO<sub>2</sub> per year in Semarang City, Indonesia.

### Conflict of Interest

The authors declare no conflict of interest.

### References

- [1] BPSR, 67 th edition BP Statistical Review of World Wide Energy, 2018, doi:https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html.
- [2] M. Mitscher, R. Ricardo, "Economic performance and policies for grid-connected residential solar photovoltaic systems in Brazil," *Energy Policy*, **49**, 688–694, 2012, doi:10.1016/j.enpol.2012.07.009.
- [3] H.T. Nguyen, J.M. Pearce, "Estimating potential photovoltaic yield with r. sun and the open source Geographical Resources Analysis Support System," *Solar Energy*, **84**(5), 831–843, 2010, doi:10.1016/j.solener.2010.02.009.
- [4] Y. Choi, J. Rayl, C. Tammineedi, J.R.S. Brownson, "PV Analyst : Coupling ArcGIS with TRNSYS to assess distributed photovoltaic potential in urban areas," *Solar Energy*, **85**(11), 2924–2939, 2011, doi:10.1016/j.solener.2011.08.034.
- [5] L. Ko, J. Wang, C. Chen, H. Tsai, "Evaluation of the development potential of rooftop solar photovoltaic in Taiwan," *Renewable Energy*, **76**, 582–595, 2015, doi:10.1016/j.renene.2014.11.077.
- [6] A. Dedinec, V. Taseska-gjorgievska, N. Markovska, T. Obradovic, N. Duic, J. Pop-jordanov, R. Taleski, "Towards post-2020 climate change regime : Analyses of various mitigation scenarios and contributions for Macedonia," *Energy*, **94**(2016), 124–137, 2016, doi:10.1016/j.energy.2015.10.085.
- [7] S. Kulkarni, R. Banerjee, "Renewable energy mapping in Maharashtra , India using GIS," *Sustainable Cities and Regions*, 3177–3184, 2011.
- [8] N. Martín-chivelet, "Photovoltaic potential and land-use estimation methodology," *Energy*, **94**, 233–242, 2016, doi:10.1016/j.energy.2015.10.108.
- [9] E. Biyik, M. Araz, A. Hepbasli, M. Shahrestani, R. Yao, L. Shao, E. Essah, A.C. Oliveira, E. Rico, J. Luis, L. Andrade, A. Mendes, Y. Baver, "Engineering Science and Technology , an International Journal A key review of building integrated photovoltaic ( BIPV ) systems," *Engineering Science and Technology, an International Journal*, **20**(3), 833–858, 2017, doi:10.1016/j.jestch.2017.01.009.
- [10] G.E. Hassan, M.E. Youssef, Z.E. Mohamed, M.A. Ali, A.A. Hanafy, "New Temperature-based Models for Predicting Global Solar Radiation," *Applied Energy*, **179**(December 2017), 437–450, 2016, doi:10.1016/j.apenergy.2016.07.006.
- [11] B. Parida, S. Iniyar, R. Goic, "A review of solar photovoltaic technologies," *Renewable and Sustainable Energy Reviews*, **15**(3), 1625–1636, 2011, doi:10.1016/j.rser.2010.11.032.
- [12] M.C. Rodríguez-amigo, M. Díez-mediavilla, D. González-peña, A. Pérez-burgos, C. Alonso-tristán, "Mathematical interpolation methods for spatial estimation of global horizontal irradiation in Castilla-León , Spain : A case study," *Solar Energy*, **151**, 14–21, 2017, doi:10.1016/j.solener.2017.05.024.
- [13] N. Kannan, D. Vakeesan, "Solar energy for future world : - A review," *Renewable and Sustainable Energy Reviews*, **62**(62), 1092–1105, 2016, doi:10.1016/j.rser.2016.05.022.
- [14] E. Kabir, P. Kumar, S. Kumar, A.A. Adelodun, K. Kim, "Solar energy : Potential and future prospects," *Renewable and Sustainable Energy Reviews*, **82**(September 2017), 894–900, 2018, doi:10.1016/j.rser.2017.09.094.
- [15] S. Chowdhury, *Design & Estimation of Rooftop Grid-tied Solar Photovoltaic System*, Technische Universität Hamburg, 2016, doi:10.13140/RG.2.2.17599.00168.
- [16] A.K. Shukla, K. Sudhakar, P. Baredar, "Simulation and performance analysis of 110 kW p grid-connected photovoltaic system for residential building in India : A comparative analysis of various PV technology," *Energy Reports*, **2**, 82–88, 2016, doi:10.1016/j.egyr.2016.04.001.
- [17] I. Amin, N. Harun, A. Suyuti, *Indonesia Berbasis Analisis RETScreen International: The Study of Renewable Energy Potential in Eastern*, 2016.
- [18] S. Castellanos, D.A. Sunter, D.M. Kammen, "Rooftop solar photovoltaic potential in cities: how scalable are assessment approaches? To," *Environmental Research Letters*, 2017.
- [19] N. Mukisa, R. Zamora, T.T. Lie, "Feasibility assessment of grid-tied rooftop solar photovoltaic systems for industrial sector application in Uganda," *Sustainable Energy Technologies and Assessments*, **32**(February), 83–91, 2019, doi:10.1016/j.seta.2019.02.001.
- [20] D.A. Jacques, J. Gooding, J.J. Giesekam, A.S. Tomlin, R. Crook, "Methodology for the assessment of PV capacity over a city region using low-resolution LiDAR data and application to the City of Leeds ( UK )," *Applied Energy*, **124**, 28–34, 2014, doi:10.1016/j.apenergy.2014.02.076.
- [21] K. Mainzer, S. Killinger, R. Mckenna, W. Fichtner, "Assessment of rooftop photovoltaic potentials at the urban level using publicly available geodata and image recognition techniques," *Solar Energy*, **155**, 561–573, 2017, doi:10.1016/j.solener.2017.06.065.
- [22] P. Redweik, C. Catita, M.C. Brito, "3D Local Scale Solar Radiation Model Based on Urban Lidar Data," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2011, doi:10.5194/isprsarchives-XXXVIII-4-W19-265-2011.
- [23] J. Peng, L. Lu, "Investigation on the development potential of rooftop PV system in Hong Kong and its environmental benefits," *Renewable and Sustainable Energy Reviews*, **27**, 149–162, 2013, doi:10.1016/j.rser.2013.06.030.
- [24] M. Rumbayan, A. Abudureyimu, K. Nagasaka, "Mapping of solar energy potential in Indonesia using artificial neural network and geographical information system," *Renewable and Sustainable Energy Reviews*, **16**(3), 1437–1449, 2012, doi:10.1016/j.rser.2011.11.024.
- [25] S. Izquierdo, M. Rodrigues, N. Fueyo, "A method for estimating the geographical distribution of the available roof surface area for large-scale photovoltaic energy-potential evaluations," *Solar Energy*, **82**, 929–939, 2008, doi:10.1016/j.solener.2008.03.007.
- [26] S. Freitas, C. Catita, P. Redweik, M.C. Brito, "Modelling solar potential in the urban environment : State-of-the-art review," *Renewable and Sustainable Energy Reviews*, **41**, 915–931, 2015, doi:10.1016/j.rser.2014.08.060.
- [27] Indonesia Energy Council, *Buku Ketahanan Energi Indonesia Tahun 2017*, 2017.
- [28] Semarang Central Statistics Agency, *Semarang Municipality in Figures 2018*, 2018.
- [29] Semarang Central Statistics Agency, *Semarang Building Figures*, 2018.
- [30] Ministry of Energy and Mineral Resources of the Republic of Indonesia, *Performance Report of the Ministry of Energy and Mineral Resources of*



the Republic of Indonesia, 2017.

- [31] A.L. Nugraha, Hani'ah, R.D. Pratiwi, "Assessment of multi hazards in Semarang city," in AIP Conference Proceedings, 1–10, 2017, doi:10.1063/1.4987112.
- [32] M.A. Marfai, L. King, L.P. Singh, D. Mardiatno, J. Sartohadi, D.S. Hadmoko, A. Dewi, Received., "Natural hazards in Central Java Province , Indonesia: An overview," *Environ Geol*, **56**, 335–351, 2008, doi:10.1007/s00254-007-1169-9.
- [33] Husnayaen, A.B. Rimba, T. Osawa, I.N.S. Parwata, F.K. Abd. Rahman As-syakur, I.A. Astarini, "Physical assessment of coastal vulnerability under enhanced land subsidence in Semarang, Indonesia, using multi-sensor satellite data," *Advances in Space Research*, 2018, doi:10.1016/j.asr.2018.01.026.
- [34] H. Tjahjono, Suripin, Kismartini, "Spatial analysis of field vulnerability concerning landslide in Southern Semarang Environment," in IOP Conf. Series: Earth and Environmental Science, 2019, doi:10.1088/1755-1315/243/1/012012.
- [35] I. Buchori, A. Pramitasari, A. Sugiri, M. Maryono, Y. Basuki, "Adaptation to coastal flooding and inundation: Mitigations and migration pattern in Semarang City, Indonesia," *Ocean and Coastal Management*, **163**(July), 445–455, 2018, doi:10.1016/j.ocecoaman.2018.07.017.
- [36] I. Buchori, A. Sugiri, M. Mussadun, D. Wadley, Y. Liu, "A predictive model to assess spatial planning in addressing hydro- meteorological hazards : A case study of Semarang City , Indonesia," *International Journal of Disaster Risk Reduction*, (April), 0–1, 2017, doi:10.1016/j.ijdr.2017.11.003.
- [37] R. Perez, R. Seals, P. Ineichen, R. Stewartt, D. Menicucci, "A New Simpfied Version of the Perez Diffuse Irradiance Model for Tilted Sufraces," *Solar Energy*, **39**(3), 221–231, 1987.
- [38] W.A. Hetrick, P.M. Rich, S.B. Weiss, "Modeling Insolation on Complex Surfaces," in Thirteenth Annual ESRI User Conference, 1993.
- [39] L. Kumar, A.K. Skidmore, E. Knowles, "Modelling topographic variation in solar radiation in a GIS environment," *International Journal of Geographical Information Science*, **11**(5), 475–497, 1997.
- [40] Esri, ArcGIS Help 10.1. Area Solar Radiation (Spatial Analyst), 2013.
- [41] P. Fu, P. Rich, "Design and implementation of the Solar Analyst : an ArcView extension for modeling solar radiation at landscape scales," in Proceedings of the Nineteenth Annual ESRI User Conference, 1999.
- [42] M. Suri, J. Hofierka, "A New GIS-based Solar Radiation Model and Its Application to Photovoltaic Assessments," *Transactions in GIS*, **8**(2), 175–190, 2004.
- [43] Asian Development Bank, Handbook for Rooftop Solar Development in Asia, 2014.
- [44] G. Knier, How do photovoltaics work, 2011.
- [45] World Meteorological Organization, WMO Statement on the State of the Global Climate in 2018, 2008.
- [46] International Energy Agency, Global Energy and CO2 Status Report, 2018.

# Keyword Driven Image Description Generation System

Sreela Sreekumaran Pillai Remadevi Amma\*, Sumam Mary Idicula

Artificial Intelligence and Computer vision Lab, Department of Computer Science, Cochin University of Science And Technology, Kerala, 682022, India

## ARTICLE INFO

Article history:

Received: 15 May, 2020

Accepted: 07 July, 2020

Online: 09 August, 2020

Keywords:

Image captioning

Keyword Extraction

BLSTM

## ABSTRACT

Image description generation is an important area in Computer Vision and Natural Language Processing. This paper introduces a novel architecture for an image description generation system using keywords. The proposed architecture uses a high-level feature such as keywords for generating captions. The important component of caption generation is the deep Bidirectional LSTM network. The space and computational complexity of the system are smaller than that of the CNN feature-based image description generation system. The number of parameters is also small in the keyword-based image description generation system. It generates novel meaningful sentences for images. The systems performance depends on the keyword extraction system.

## 1 Introduction

Image description generation is an active research area for increasing the performance of an image search engine. The system is also useful for accessing image collections, helping visually impaired people, enhancing the education system, robotics, etc.

An image is described by using objects, attributes, actions, scenes, and spatial relationships. In this process, all the keywords of an image are mapped to a sentence called a caption. The image description is created from image features. The main image features are low level and high-level features. Here the proposed system is focused on high-level features such as keywords. The main objective of the system is to develop an image description generation system from keywords. Keyword-based image description system is a direct generation model which extracts keywords from the visual content of the image and generate sentence from keywords.

To achieve the objective, different phases are developed such as

- Keyword extraction system: The objects, attributes, actions, and scenes are extracted.
- Caption generation system: Three methods are employed for generating captions such as template-based, CFG based, and BLSTM based.

The outline of the paper is as follows. Section 2 describes the related work of the system. Section 3 explains the proposed method of the system. Section 4 discusses the experiments and results of the system. The paper is concluded in section 5.

## 2 Related works

From the literature, there are three kinds of image description generation models such as Direct generation models, Visual space model, and Multimodal space model.

Direct generation models analyse the visual content of the image and create a sentence reflecting the meaning of the visual content. Examples of these models are Google's Neural Image Caption Generator [1], BabyTalk [2], Midge [3], Karpathy's system [4] etc. The visual space model finds similar images of the query image from a visual space and transfer the description to the query image. Research [5]-[6] follows visual space model. The multimodal space model discovers similar images from visual and textual modal space, and it is treated as a retrieval problem. While, [7]-[9] systems follow Multimodal space model.

Another classification of image description generation system is Template based approach and Deep Neural Network based approach [10]-[13]. Deep Neural Network based Image captioning systems are improved using visual attention mechanisms. The first attention-based image captioning system is done by [14]. While [15]-[17] are also proposed attention-based image captioning system. While [17]-[19] follows attention-based architecture. Examples of keyword-based image captioning systems are Midge, BabyTalk, etc. Midge and BabyTalk have less accuracy because of incorrect object detection, invalid action classification, etc. So a novel system for keyword-based image captioning systems is developed using a new keyword extraction system.

\*Corresponding Author: Sreela Sreekumaran Pillai Remadevi Amma, +919497708518 & sreela148@cusat.ac.in

### 3 Proposed Method

An image is processed by Computer vision algorithms can be represented as a quadruple  $\langle O_i, Attr_i, A_i, S_i \rangle$ . Where  $O_i$  is the set of objects in the image,  $Attr_i$  is the set of attributes of each object,  $A_i$  is the set of actions of the object, and  $S_i$  is the set of scenes associated with an image. Similarly, the description of the image can be characterized as a quadruple  $\langle N_d, Adj_d, V_d, NS_{cd} \rangle$ . Where  $N_d$  is the set of nouns in the description,  $Adj_d$  is the set of adjectives in the description,  $V_d$  is the set of verbs and  $NS_{cd}$  is the set of nouns associated with a scene. Image description generation is the process of mapping image quadruple  $\langle O_i, Attr_i, A_i, S_i \rangle$  to description quadruple  $\langle N_d, Adj_d, V_d, NS_{cd} \rangle$ .

The overview of the proposed methodology is explained in figure 1. The methodology is as follows: 1) Detect Objects and scenes in the image. 2) Recognize action in an image 3) Identify attributes such as cardinality, shape, and color of the object. 4) Generate text corresponding to the image. The method is divided into two parts, such as keywords generation and caption generation.

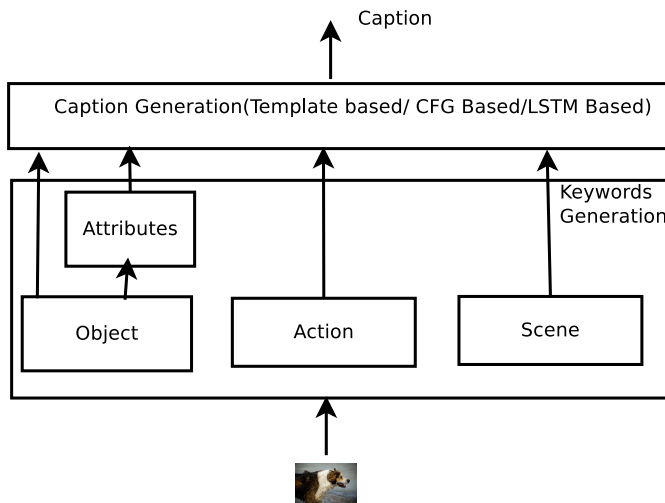


Figure 1: Proposed Architecture

#### 3.1 Keywords Generation

##### 3.1.1 Object detection

Darknet architecture is used for object detection. The images are divided into  $7 \times 7$  grid cells. For each grid cell, we apply a classifier for identifying an object. The number of layers in Darknet is 53. Object classes are obtained in this phase. The object is considered as the noun of the sentence. Top-k objects are selected for sentence construction.

##### 3.1.2 Attributes Identification

This phase identifies the color, shape, and cardinality of the object. Attributes are identified as the adjectives of the sentence.

**Color detection** The color of the object is calculated from the minimum distance between  $L^*a^*b$  value and the average of the intensity of the image. The color with the smallest distance is treated as the color of the object.

**Shape detection** The algorithm for shape detection is explained below:

```

    Compute the contour perimeter of the image;
    Approximate the contour shape to another shape with fewer
    vertices depending upon the precision using the
    Douglas-Peucker algorithm;
    if The approximated shape has three vertices then
    | it is a triangle;
    else
    | if the shape has four vertices then
    | | Compute the bounding box of contour and find the
    | | aspect ratio;
    | | if the aspect ratio is approximately equal to one then
    | | | The shape is a square;
    | | else
    | | | otherwise, it is a rectangle;
    | | end
    | else
    | | if the shape has more than four vertices then
    | | | it is treated as a circle;
    | | end
    | end
    end
    end
    
```

**Cardinality** Cardinality is obtained by the count of each object in the image.

##### 3.1.3 Action recognition

Action recognition is done in still images. Action is treated as the verb of the sentence or caption. The architecture of action recognition is explained in figure 2.

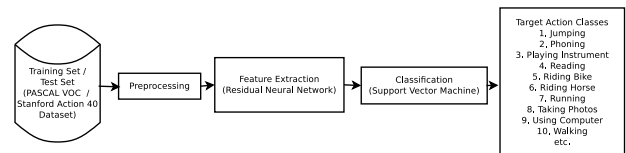


Figure 2: Proposed Architecture of action recognition

It is done using deep convolutional neural network features and Support Vector Machine. Residual Neural Network[20] is used for CNN feature extraction, and classification is done using Support Vector Machine.

##### 3.1.4 Scene Classification

Context or scene is important for caption generation. It is obtained using deep neural network-based learning. The scene is the place where an image is associated. VGG architecture is used for scene classification. Three hundred sixty-five scene categories are used for this classification.

#### 3.2 Caption generation

It is a process of generating sentences or captions from a set of keywords like objects, attributes, action, and scene. We implemented

three kinds of caption generation, such as template-based, CFG based, and Bidirectional LSTM[21] based caption generation.

### 3.2.1 Template based

Keywords are extracted as a five-tuple  $\langle object, cardinality, attribute, action, scene \rangle$ . Captions are generated from the five-tuple. Captions have a common form “There is **cardinality attribute object action** in the **scene**. The object is treated as noun, cardinality, and attributes are adjectives, the action is the verb. An example of a sentence is, “There is two blue birds fly in the sky”. The set of words in the meaning representation is fixed, and generation must make use of all given content words; and, generation may insert only gluing words (i.e., function words such as there, is, the, etc.). Advantages of this approach is that it incorporates cardinality and actions in the image captions. Disadvantages of this approach is that it generates similar kinds of captions. And also still image based action recognition generates 10 classes only.

### 3.2.2 CFG based

We design a CFG for the caption. CFG has the form represented in figure 3. Using this CFG, captions are generated.

```

S -> NP | VP
VP -> V NP | V
NP -> ADJ N | N | P N
N -> Object | Scene
V -> Action
ADJ -> Cardinality | Attribute | Cardinality Attribute
Object -> "birds"
Action -> "fly"
Cardinality -> "two"
Attribute -> "blue"
Scene -> "air"
P -> "in"
    
```

Figure 3: CFG for caption generation

### 3.2.3 Bidirectional LSTM based

It uses encoder-decoder architecture as Machine translation. Let  $(k, C)$  be a keyword and caption sentence pair. Let  $k = k_1, k_2, k_3 \dots k_m$  be a sequence of M symbols in keywords and  $C = C_1, C_2, C_3 \dots C_N$  be the sequence of N symbols in the target caption of the image. The model maximizes the following objective.

$$\theta^* = \operatorname{argmax}_{\theta} \sum_{k,C} \log p(C/k; \theta) \quad (1)$$

The encoder is simply a function of the following form

$$K_1, K_2, K_3 \dots K_M = \operatorname{encoderBLSTM}(k_1, k_2, k_3 \dots k_m) \quad (2)$$

Where  $K_1, K_2, K_3, \dots, K_M$  is a list of hidden representation of keywords of size M. *encoderBLSTM* is a bidirectional LSTM. Using chain rule, the MAP of the sequence  $p(C/k)$  can be decomposed as

$$\begin{aligned} p(C/k) &= p(C/k_1, k_2, k_3, \dots, k_m) \\ &= \prod_{i=1}^N P(C_i/C_0, C_1, C_2 \dots C_{i-1}; k_1, k_2, k_3 \dots k_m) \\ &= \sum_{i=1}^N \log P(C_i/C_0, C_1, C_2 \dots C_{i-1}; k_1, k_2, k_3 \dots k_m) \end{aligned} \quad (3)$$

where  $C_0$  is a special “beginning of sentence” symbol. During inference, we calculate the probability of next symbol given

the source encoding and the decoded target caption so far as  $p(C_i/C_0, C_1, \dots, C_{i-1}; K_1, K_2, \dots, K_M)$ .

Our decoder is organized as a combination of a fully connected neural network(FCN) layer and a softmax layer. The FCN layer produces the hidden representation of the next symbol to be predicted; which then goes through the softmax layer to generate a probability distribution over candidate vocabulary symbols.

The deep neural network architecture is drawn in Figure 4.

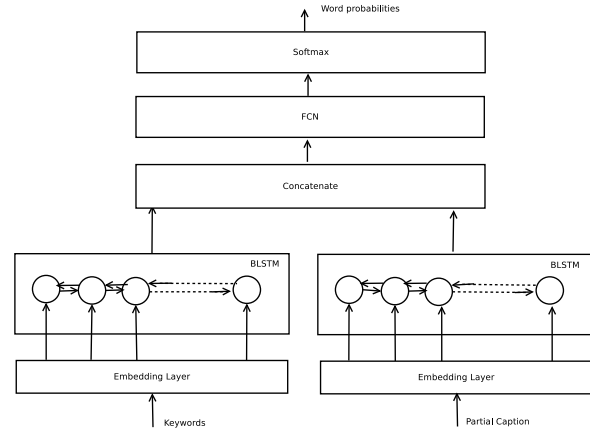


Figure 4: Bidirectional LSTM based caption generation

## 4 Results

### 4.1 Keyword Extraction

The Table 1 shows the output of keyword extraction. The table contains objects, attributes such as cardinality, shape, and color, scene, and action of images.

The Table 2 is described the various advantages and disadvantages of various phases of keyword extraction system.

### 4.2 Caption Generation

The deep neural network based caption generation model is experimented with LSTM[22] and Bidirectional LSTM. From the experiment, the accuracy graphs are plotted in figures 5 and 6. The accuracy plot explained that the accuracy of the model is more when the bidirectional LSTM is used.

The loss of the model is plotted in figures 7 and 8. From the plots, the loss of the model is less when Bidirectional LSTM is used.

Bidirectional LSTM is more suitable for designing the caption generation from keywords. The table 3 describes the captions from keywords using three methods.

The Table 4 explains the various properties of different methods used for caption generation.

### 4.3 Analysis

The computational complexity of the caption generation from keywords is small. The space complexity of this system is described by

$$s_k = |Keywords| \times \max\_cap\_Len$$



Table 1: Keywords Extraction results






Image	Objects	Attributes	Scene	Action
	Boat, Person	Boat - Color:Olive, Shape: Rectangle Person- Color: Yellow, Shape: Square	Raft, Beach, Lake	Running
	Dog	Dog - Color:Silver, Shape: Square	Sandbox, Archaeological excavation, Trench	Jumping
	Person	Person - Color:Teal, Shape: Rectangle, Cardinality: 10	Playground, Pet shop, Toy shop	Walking
	Car, Truck	Car - Color:Gray, Shape: Rectangle, Truck - Color:Gray, Shape:Rectangle	Forest path, River	-
	Person, Bicycle, Backpack	Person - Color:Teal, Shape: Rectangle, Bicycle - Color:Teal, Shape:Rectangle, Backpack- Color:Green, Shape: Rectangle	Coast, Beach, Ocean	Riding Bike
	Person, Dog, Frisbee	Person - Color:Gray, Shape: Rectangle, Dog - Color:Gray, Shape:Square, Frisbee- Color:Olive, Shape: Rectangle	Lawn, Yard, Outdoor	Walking

Table 2: The characteristics of different phases of Keywords extraction

Model	Advantages	Limitations
Object detection	Number of objects is 1000.	Vocabulary of objects is limited.
Attributes identification	Color, shape and cardinality are identified.	Only three attributes of objects are detected.
Action recognition	Number of action classes is 10.	The vocabulary of actions is limited.
Scene classification	365 scenes are classified.	Number of scenes is limited.

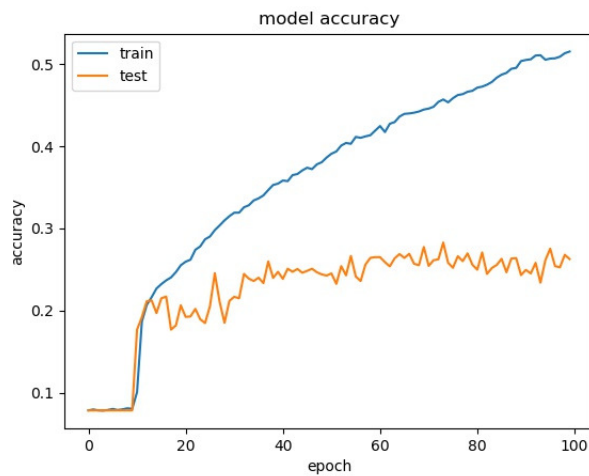


Figure 5: Accuracy plot for LSTM Based Caption generation from keywords

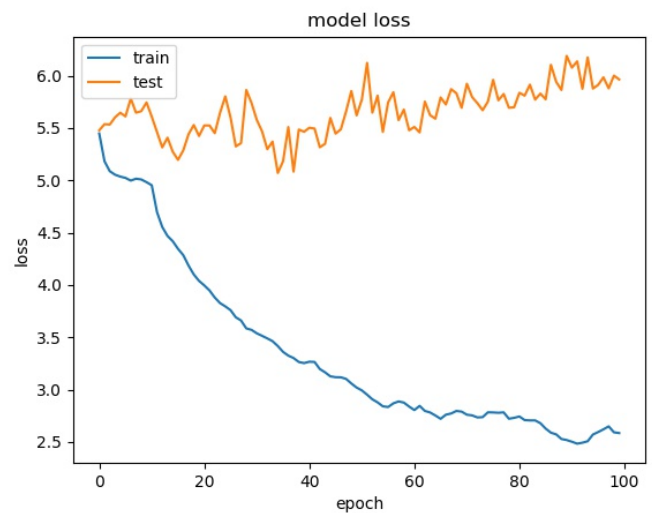


Figure 7: Loss plot for LSTM Based Caption generation from keywords

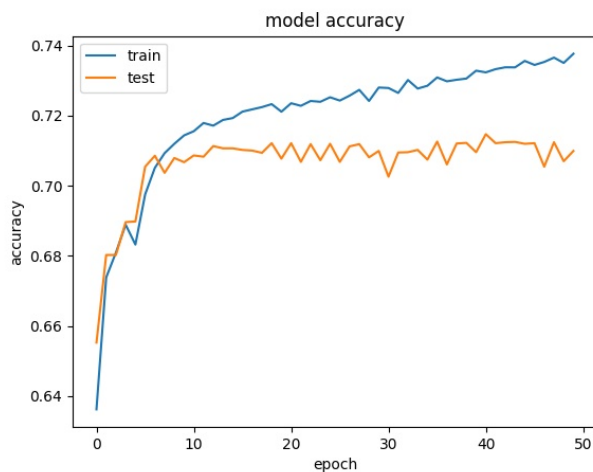


Figure 6: Accuracy plot for Bidirectional LSTM Based Caption generation from keywords

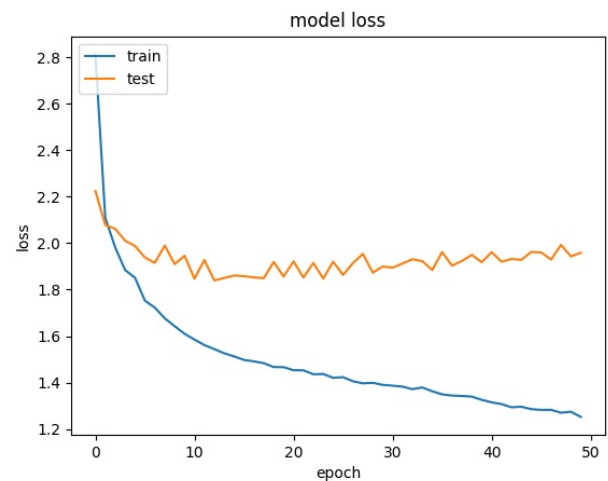


Figure 8: Loss plot for Bidirectional LSTM Based Caption generation from keywords

Table 3: Output of caption generation from keywords using different methods.

Keywords	Template based	CFG based	LSTM based
Person, Boat, Lake, Running	There is a person running in a lake.	A person running in a lake.	A Person is running in the lake with a boat.
Dog, Silver, Sandbox, Jumping	There is a silver dog jumping in sandbox.	Silver dog jumping in sandbox.	Silver dog is jumping in the sandbox.
Person, Ten, Playground, Walking	There is ten persons walking in playground.	Ten persons walking in playground.	Ten persons are walking in the playground.
Truck, Gray, Forest path	There is a gray truck in forest path.	Gray truck in forest path.	Gray truck in the forest path.
Person, Bike, Backpack, Green, Coast, Riding bike	There is a person riding a bike in bike.	Person riding a bike in a bike.	Person is riding bike.

Table 4: The characteristics of different methods of caption generation

Method	Advantages	Disadvantages
Template based	It incorporates cardinality and actions in the image captions.	It generates similar kinds of captions.
CFG based	It generates different types of captions.	It generates caption with same grammar.
BLSTM based	It creates various kinds of captions.	The vocabulary of objects, attributes, actions and scenes is limited.

where  $|Keywords|$  is the number of keywords and  $max\_cap\_len$  is the maximum caption length. The  $|Keywords|$  is less than the maximum number of keywords. Here we set the maximum number of keywords as 10.

The space complexity of the caption generation in CNN feature-based caption generation system is given by

$$s_c = |CNN\_feature| \times max\_cap\_len$$

where  $|CNN\_feature|$  is the length of the CNN feature, and it is 2208 if the CNN used is Densenet. From the analysis,  $s_k < s_c$ .

If the number of network parameters of the keyword-based caption generation system is  $p_k$  and the number of network parameters of the CNN feature-based caption generation system is  $p_c$ , then  $p_k < p_c$ . The network of keyword-based caption generation system is small.

The limitation of the system is that the performance of the system depends on the keyword extraction system. Keyword extraction is based on object detection, object identification, action recognition, and scene classification. The mean average precision (MAP) of the action recognition system is only 66%. The vocabulary of actions contains only ten words. The vocabulary size of objects

is 1000. The number of scenes is 365. So the total vocabulary contains around 1.5k words. But the vocabulary size of the flickr8k dataset is 8256. Hence the vocabulary of the keyword-based caption generation system is small.

In the future, we will concentrate on increasing the performance of the keyword extraction system. Thereby we can improve the functioning of caption generation from keywords.

## 5 Conclusions

The paper proposed a novel system for keyword-based image caption generation. The main components of the system are the keyword extraction and caption generation. It produces meaningful captions for images. The space and computational complexity of this system are small. The limitation of the system is that the performance of the system mainly depends on the keyword extraction system. So in the future, we will focus on improving the performance of the keyword extraction system.

## References

- [1] Vinyals, Oriol, A. Toshev, S. Bengio, and D. Erhan. "Show and tell: A neural image caption generator." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3156-3164. 2015. DOI: 10.1109/CVPR.2015.7298935
- [2] G. Kulkarni, V. Premraj, V. Ordonez, S. Dhar, S. Li, Y. Choi, A.C. Berg, and T.L. Berg. "Baby talk: Understanding and generating simple image descriptions.", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12), 28912903, 2013. DOI: 10.1109/TPAMI.2012.162
- [3] M. Mitchell, J. Dodge, A. Goyal, K. Yamaguchi, K. Stratos, A. Mensch, A. Berg, X. Han, T. Berg, and O. Health., "Midge: Generating Image Descriptions From Computer Vision Detections.", *Eacl*, pp. 747-756, 2012.
- [4] K., Andrej, and L. Fei-Fei. "Deep visual-semantic alignments for generating image descriptions." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3128-3137. 2015. DOI: 10.1109/CVPR.2015.7298932
- [5] V. Ordonez, G. Kulkarni, and T.L. Berg., "Im2text: Describing images using 1 million captioned photographs.", *Advances in Neural Information*, pp. 19, 2011.
- [6] P. Kuznetsova, V. Ordonez, A.C. Berg, T.L. Berg, and Y. Choi, "Collective generation of natural image descriptions.", In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers 1*. Association for Computational Linguistics, 359-368, 2012.
- [7] H., Micah, P. Young, and J. Hockenmaier, "Framing image description as a ranking task: Data, models and evaluation metrics.", *Journal of Artificial Intelligence Research* 47 853-899, 2013. <https://doi.org/10.1613/jair.3994>
- [8] Gong, Yunchao, et al. "Improving image-sentence embeddings using large weakly annotated photo collections.", *European Conference on Computer Vision*. Springer, Cham, 2014. DOI: [https://doi.org/10.1007/978-3-319-10593-2\\_35](https://doi.org/10.1007/978-3-319-10593-2_35)
- [9] R. Socher., A. Karpathy, Q. V. Le, C. D. Manning, A. Ng, "Grounded Compositional Semantics for Finding and Describing Images with Sentences.", *Transactions of the Association for Computational Linguistics* (2), 207218, 2014. DOI: 10.1162/tacl.a.00177
- [10] Y. H. Tan and C. S. Chan, "phi-LSTM: A Phrase-Based Hierarchical LSTM Model for Image Captioning." Cham: Springer International Publishing, pp. 101-117, 2017. [https://doi.org/10.1007/978-3-319-54193-8\\_7](https://doi.org/10.1007/978-3-319-54193-8_7)
- [11] Tan, Y. Hua, and C.S. Chan. "Phrase-based Image Captioning with Hierarchical LSTM Model." 2017. DOI: arXiv preprint arXiv:1711.05557
- [12] Aneja, Jyoti, Aditya Deshpande, and Alexander G. Schwing. "Convolutional image captioning." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018. DOI: 10.1109/CVPR.2018.00583
- [13] Han, Meng, W. Chen, and A.D. Moges. "Fast image captioning using LSTM." *Cluster Computing*. DOI: <https://doi.org/10.1007/s10586-018-1885-9>
- [14] Xu, Kelvin, et al., "Show, attend and tell: Neural image caption generation with visual attention.", *International Conference on Machine Learning*, 2015.
- [15] Junqi Jin, Kun Fu, Rungpeng Cui, Fei Sha, and Changshui Zhang. 2015. "Aligning where to see and what to tell: image caption with region-based attention and scene factorization." arXiv preprint arXiv:1506.06272, .
- [16] Z.Y.Y.Y. Wu and R.S.W.W. Cohen, "Encode, Review, and Decode: Reviewer Module for Caption Generation." In *30th Conference on Neural Image Processing System(NIPS)*, 2016.
- [17] Lu, Jiasen, et al. "Knowing when to look: Adaptive attention via a visual sentinel for image captioning." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 6 2017. DOI: 10.1109/CVPR.2017.345
- [18] You, Quanzeng, et al. "Image captioning with semantic attention." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016. DOI: 10.1109/CVPR.2016.503
- [19] He, Chen, and Haifeng Hu. "Image captioning with text-based visual attention." *Neural Processing Letters* 49.1: 177-185, 2019. DOI: <https://doi.org/10.1007/s11063-018-9807-7>
- [20] He, Kaiming; Zhang, Xiangyu; Ren, Shaoqing; Sun, Jian, "Deep Residual Learning for Image Recognition". *Proc. Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2016. DOI: 10.1109/CVPR.2016.90
- [21] Wang, Cheng, et al, "Image captioning with deep bidirectional LSTMs", *Proceedings of the 2016 ACM on Multimedia Conference*, ACM, 2016. DOI: <https://doi.org/10.1145/2964284.2964299>
- [22] Hochreiter, Sepp, and Jrgen Schmidhuber. "Long short-term memory.", *Neural computation* pp. 1735-1780, 1997. DOI: <https://doi.org/10.1162/neco.1997.9.8.1735>



## Tidal Propagation Based On Co-Phase Chart and Co-Range Chart in Sunda Strait, Indonesia

Denny Nugroho Sugianto<sup>1,\*</sup>, Harjo Susmoro<sup>2</sup>, Khoirol Imam Fatoni<sup>2</sup>, Virginia Stephanie Claudia<sup>1</sup>, Haris Djoko Nugroho<sup>2</sup>

<sup>1</sup>Department of Oceanography, Faculty of Fisheries and Marine Science, Center for Coastal Rehabilitation and Disaster Mitigation Studies (CoREM), Diponegoro University, Semarang, 50275, Indonesia

<sup>2</sup>Hydrography and Oceanography Center, Indonesian Navy, Indonesian Hydrography Council (IHC), Jl. Pantai Kuta V No.1, Ancol Timur, North Jakarta, DKI Jakarta, 14430, Indonesia

### ARTICLE INFO

Article history:

Received: 22 June, 2020

Accepted: 07 July, 2020

Online: 09 August, 2020

Keywords:

Co-tidal charts

Sunda Strait

Tidal propagation

Indonesia

### ABSTRACT

The characteristic of tide is different for each region. Co-tidal charts are one of the media to present an information about the character of tidal. There are two types of co-tidal chart, such as co-phase chart and co-range chart. Co-phase chart will show the distribution of tides, meanwhile the co-range chart will show the propagation direction of tides. Sunda Strait was chosen to analyze the tidal propagation pattern. The amount of the tidal station which were observed are 33 stations across Sunda Strait. This research used four harmonic constituents of tide, such as M<sub>2</sub>, S<sub>2</sub>, K<sub>1</sub> O<sub>1</sub> and the chart of Sunda Strait to establish co-tidal charts. The harmonic constituent of tide was analyzed by Admiralty Method. These data were obtained from Hydrography and Oceanography Center, Indonesian Navy (PUSHIDROSAL). The result of this research shows the tidal propagation direction of semidiurnal harmonic constituents of tide (M<sub>2</sub>, S<sub>2</sub>) that flows in south area of Sunda Strait (comes from Indian Ocean through the western part of Sumatra island). Meanwhile, in north area of Sunda Strait, the tidal propagation that comes from Indian Ocean flows through Karimata Strait. The direction of tidal propagation of diurnal harmonic constituents of tide (K<sub>1</sub>, SO<sub>1</sub>) that flows in Sunda Strait comes from Pacific Ocean through Karimata Strait. However, the tidal type in Sunda Strait is dominated by mixed tide prevailing semidiurnal and the highest amount of Z<sub>0</sub> is 100 cm, meanwhile the lowest value of Z<sub>0</sub> in Sunda Strait is 40 cm.

### 1. Introduction

The Sunda Strait is a strait connecting the island of Java and the island of Sumatra in Indonesia. On the other hand, the waters of the Sunda Strait are the waters connecting the Java Sea with the Indian Ocean [1]-[3]. Based on the Constitution of Republic of Indonesia Number 17 of 1985, the Sunda Strait is classified into ALKI I (Indonesian archipelagic sea lane) [4], where the strait connects the waters of the Indian Ocean through the Karimata Strait to the South China Sea or vice versa. Oil and gas transportation routes for energy needs in East Asia other than through the Strait of Malacca are in the Sunda Strait. The Sunda Strait is an alternative route of shipping when the traffic flow in the Malacca Strait is inhibited [5], [6]. Therefore, the Sunda Strait is the route commonly used for international shipping lanes [7],

[8] as well as the crossing path from Java Island to Sumatra Island [9]. The depth in the Sunda Strait waters ranges from 0.5 m to 1,885 m [10]-[12]. The number of ships passing through the Sunda Strait from 2007 to 2015 reached by 29,351 units. The number of vessels passing through the Sunda Strait will continue to increase each year, therefore, various regulations are developed to improve the security of traffic in the Sunda Strait [13]. Regulations aim to enhance the security of shipping in the Sunda Strait should be supported by data describing natural phenomena in these waters. One of the natural phenomena at sea that have to be understood is the change of sea level elevation or so-called tidal. Tidal in the hydrographic field is used as an object to reduce the results of bathymetry recording to obtain bathymetry value of an actual water [14]-[17].

Tidal has a dynamic character and its character is different in each region, especially the tides affect the rate of pollutants,

\*Corresponding Author: Denny Nugroho Sugianto, Email: [dennysugianto.oceanography@gmail.com](mailto:dennysugianto.oceanography@gmail.com)

sediment transport and erosion [18], [19]. Therefore, to facilitate the presentation of tidal conditions in an area is to use the chart [20]. The co-tidal chart is part of the nautical chart [21], [22]. The nautical chart serves to provide accurate data with a fast presentation system and approaches the actual situation to support maritime activities [23]. Co-tidal charts are divided into co-range chart and co-phase chart. The co-range chart provides information about the amplitude, while the co-phase chart provides information about the phase of the tide characteristics presented in contour lines [24]. More location of data measurement in the waters will facilitate the analysis of the tidal character of these waters. Study of tidal characteristic in Sunda Strait in a form of co-tidal charts can be used by the some parties to improve the management of shipping path. The location of the study is presented in Figure 1.



Figure 1 :Sunda Strait in the north is bounded by the connecting line from west to east that is from 5° 50' south latitude to 105° 47' east longitude and from 5° 53' south latitude to 105° 02' east longitude. While, in the south, the bounded line connects from west to east (from 6° 50' south latitude to 105° 15' east longitude and from 5° 56' south latitude to 105° 33' east longitude) [25].

2. Literature Study

2.1. Research Method

The methodology used in this research is descriptive quantitative method. This method aims to describe the distribution of the wave height amplitude and tidal propagation horizontally accompanied by the distance and the extent of the water area influenced by the amplitude and phase of tide.



Figure 2 : Thirty two stations of sampling points to calculate the co-tidal chart in Sunda Strait

The location of the sampling point is presented in Figure 2. The equipment was installed is tidal gauges. This device are not submerged at the highest high water level (HHWL) and are still inundated during lowest low water level (LLWL). The tidal gauges were installed perpendicularly and placed away from human activities or the flow of boats and ships. The tidal gauge should not be installed on the watershed or breakwaters in order not to cause bias and others [26].

2.2. Tidal Elevation Analysis using Admiralty Method

Tidal data were obtained using Admiralty Method by schematic and multiplier tables at data length of 15 or 29 piantans in one hour observation interval to obtain the value of tidal harmonic constituents (S0, M2, S2, N2, K1, O1, M4, MS4, P1, K2), where some of the harmonic constituents (M2, S2, K1, O1) will become the input parameters of the co-tidal chart. Equation 1 defines the numbers of Z0, meanwhile equation 2 is the equation of Formzahl which used to determine the types of tidal of a waters [27].

$$Z_0 = A(M_2) + A(S_2) + A(N_2) + A(K_2) + A(K_1) + A(O_1) + A(P_1) + A(M_4) + A(MS_4) \tag{1}$$

where:

- Z<sub>0</sub> = chart datum (m)
- A = amplitude (m);

$$F = \frac{A(K_1)+A(O_1)}{A(M_2)+A(S_2)} \tag{2}$$

where:

- F = formzahl
- A(K<sub>1</sub>) = amplitude of tidal harmonic constituent of K<sub>1</sub>
- A(O<sub>1</sub>) = amplitude of tidal harmonic constituent of O<sub>1</sub>
- A(M<sub>2</sub>) = amplitude of tidal harmonic constituent of M<sub>2</sub>
- A(S<sub>2</sub>) = amplitude of tidal harmonic constituent of S<sub>2</sub>

2.3. Construction of Co-Tidal Chart

Co-tidal chart construction requires at least three different observation stations. The datum of time used for co-tidal chart is GMT [28], ie +07.00, meanwhile references datum of tidal is low water of neap tide. The result of tidal observation from each station is amplitude height (cm), phase (deg) of each tidal harmonic constituents and the formzahl number. In addition, the period of water height in every tidal harmonic constituents can be calculated by Equation 3.

$$t = g \times \frac{1}{n} \tag{3}$$

where :

- t = time of water height (hour)
- g = phase of tidal harmonic constituent (0)
- n = angular velocity of tidal harmonic constituent (0/jam)

Co-tidal charts in offshore area can be determined based on the observation at the shore [24]. The observation results along the shore was interpolated to obtain the offshore area empirically. This research was using spline method. Spline method is a method that estimates value by using mathematical functions that minimize the total surface curvature [29], [30].

2.4. Data Analysis

The tidal analysis in this research refer to the method by Pariwono [24], which are contained of the following 3 sections :

- The aim of analyzing co-phase charts is to obtain the tidal propagation information for each tidal harmonic constituents, i.e. diurnal tidal harmonic constituents (K1, O1) and semidiurnal tidal harmonic constituents (M2, S2). The data which were observed are the source and direction of tidal wave propagation which are shown based on the high of water time contours, where tidal propagation moves from low amplitude to higher amplitudes.
- The aim of analyzing co-range charts was to obtain the information about the distribution of wave height amplitude. The data consist of the highest and lowest of wave height amplitude that shown by line contour.
- Types of tidal chart and Z0 chart were analyzed to obtain the general information about the tidal characteristic based on formzahl number and height of Z0.

3. Result

3.1. Tidal Observation

This research shows the tidal characteristic information at 2 stations of tidal observation , i.e. in Panjang Waters Lampung and Ciwandan Waters Banten. The result of tidal elevations

(shown at Fig. 3 and Fig. 4) presented the height of HHWL, LLWL, HWL, LWL and MSL. The value of HHWL, HWL, MSL, LWL and LLWL in Panjang waters, respectively are 196 cm, 179 cm, 104 cm, 42 cm and 13 cm. Meanwhile, the HHWL, HWL, MSL, LWL and LLWL in Ciwandan waters are 163 cm, 150 cm, 120 cm, 89 cm and 77 cm, respectively.

Tidal elevation in the Panjang Waters Lampung and Ciwandan Waters Banten are shown in Figure 3 and 4. High water value which is obtained in Panjang Waters is 179 cm, while in Ciwandan Waters is 150 cm. While, low water values for Panjang and Ciwandan waters are 42 cm and 89 cm, respectively.

3.2. Analysis of Tidal Harmonic Components

The results obtained are the amplitude (A) and phase (g) values of each tidal harmonic constituent, i.e. S0, M2, S2, N2, K1, O1, M4, MS4, K2 and P1 (Table 1 and Table 2). The formzahl value of Ciwandan waters, Banten is 0.436, whereas, the formzahl value of Panjang waters, Lampung is 0.482. The comparison of tidal elevation between Panjang waters, Lampung and Ciwandan waters, Banten (shown in Figure 5) is used to determine the direction of water mass movement. Figure 6 shows a plot stick value from both tidal observation stations to verify the direction of tidal wave propagation on co-tidal charts.

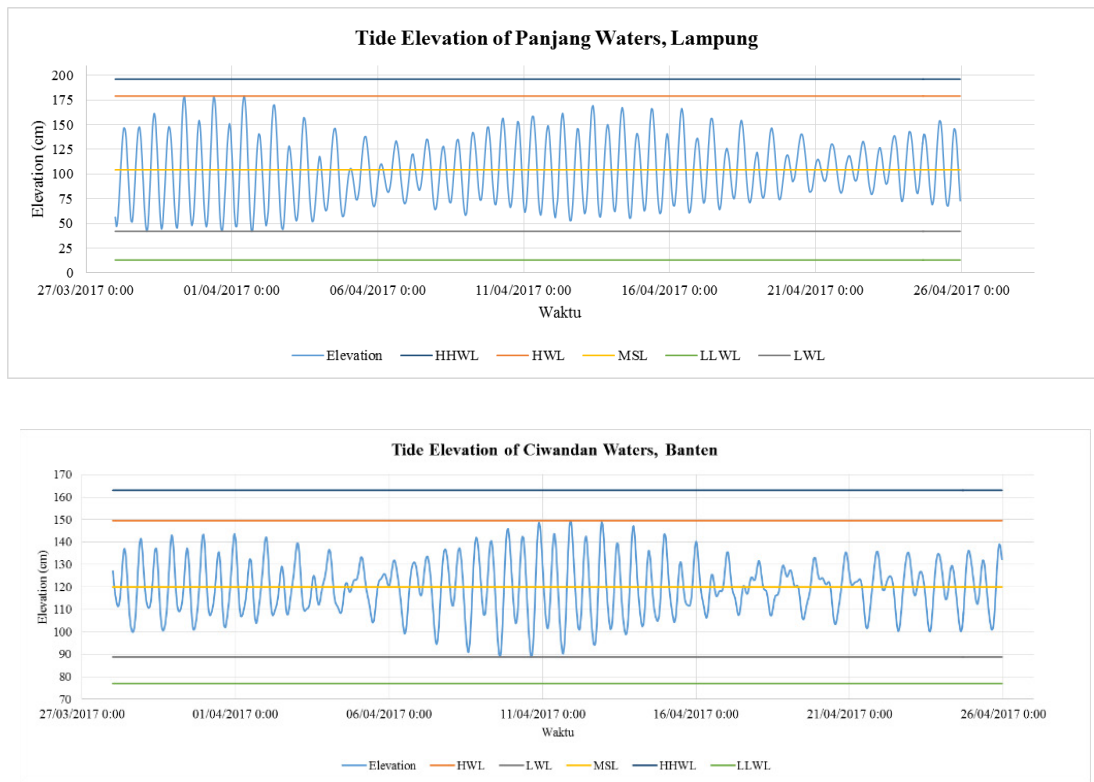


Table 1 :Tidal Harmonic Constituents at Panjang Waters, Lampung

	S0	M2	S2	N2	K1	O1	M4	MS4	K2	P1
A (cm)	104	36	14	8	16	8	0.2	0.1	4	5
g (0)		198	269	169	276	250	85	240	269	276

Table 2 : Tidal Harmonic Constituents at Ciwandan Waters, Banten

	S0	M2	S2	N2	K1	O1	M4	MS4	K2	P1
A (cm)	120	11	11	4	6	3	2	1	3	2
g (0)		214	333	216	192	76	217	56	333	192

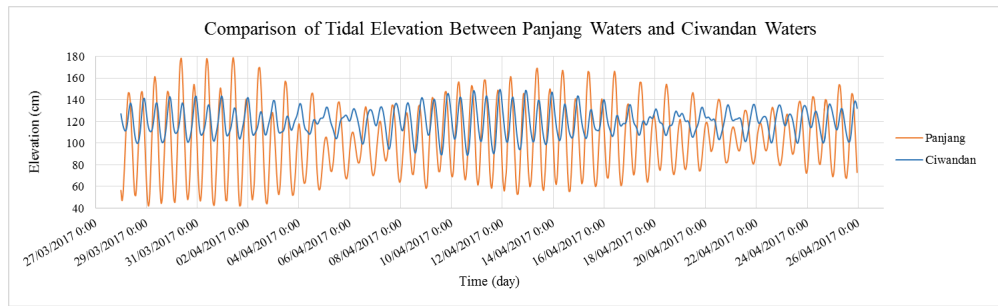


Figure 5 : Comparison of Tidal Elevation between Panjang Waters and Ciwandan Waters

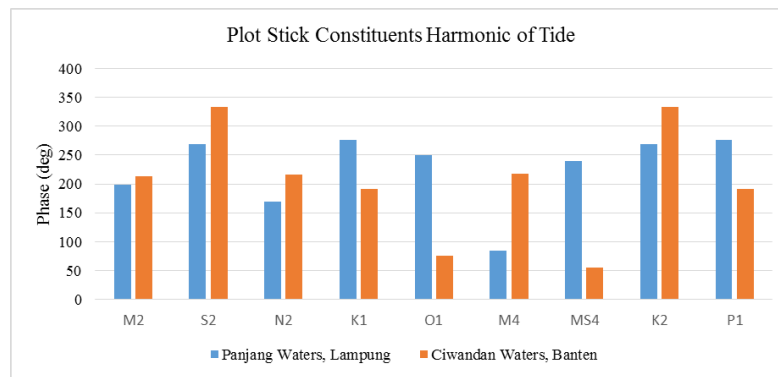


Figure 6 : Plot stick constituents harmonic of tide

Figure 5 shows the different tidal elevation of the two waters, where it can be seen that the fluctuations of sea water level length of Panjang Water are greater than the fluctuations of sea water surface of Ciwandan. Tidal steep in Panjang waters is reached up to 137 cm at neap tide, while tidal steep in Ciwandan Waters only reached 61 cm at neap tide. According to Pond and Pickard [31], the depth and width of the waters affect sea level fluctuations, therefore, the difference in tidal elevation values between Panjang and Ciwandan Waters is caused by its location, where the location of Panjang waters is in Lampung Bay and Ciwandan Waters is located in the northern part of the Sunda Strait which is directly facing to the Java Sea. Tidal types in Panjang and Ciwandan Waters are the same (mixed tide prevailing semidiurnal). Based on Figure 6, the semidiurnal harmonic constituents (including M2, S2, N2 and K2), the direction of its tidal wave propagation is from Panjang Waters to Ciwandan Waters, whereas, diurnal harmonic constituents (including O1, K1, P1), the direction of its tidal wave propagation is from Ciwandan Waters to Panjang Waters. M4 tidal constituent shows the direction of wave propagation from Panjang Waters to Ciwandan Waters, while the tidal constant MS4 shows the direction of wave propagation from Ciwandan Waters to Panjang Waters.

### 3.3. Co-tidal Chart

Figure 7a show the co-range chart which are presented by the amplitude of wave height and the tidal propagation based on time

of each tidal harmonic constituent. The contour lines define the amplitude of wave height or tidal propagation period.

Figure 7a shows co-range chart of harmonic constants M2. The distance of M2 amplitude contour has a 2 cm interval (around Bakauheni-Merak or north area of Sunda Strait) to Anyer-Kalianda waters, where the change of elevation of M2 amplitude height are from 12 cm to 30 cm. Contour density indicates the variability of the value of M2 amplitude height which is caused by tidal wave propagation influenced by depth in north area of Sunda Strait. The minimum value of M2 wave amplitude in Sunda Strait is 12 cm and the maximum height is 36 cm. The wave height of M2 amplitude in the southern area of the Sunda Strait is greater than the wave height of M2 amplitude in the north of the Sunda Strait. The M2 amplitude value in the Sunda Strait Waters is more dominant than the M2 amplitude value in the northwest part of the Java Sea, indicating that the amplitude of the M2 wave amplitude in the Sunda Strait is high. Based on the S2 co-range chart of the Sunda Strait waters which is shown in Figure 7 b, the S2 wave amplitude in the Sunda Strait Waters is quite high compared to the wave amplitudes in the western part of the Java Sea. This condition is based on contours whose value is enlarged toward the northern part of the Sunda Strait from the western part of the Java Sea, which is from 7 cm to 11 cm in Cilegon-Bakauheni area. Furthermore, the further inside of Sunda Strait area, the amplitude value of S2 become higher until it reaches the maximum height of



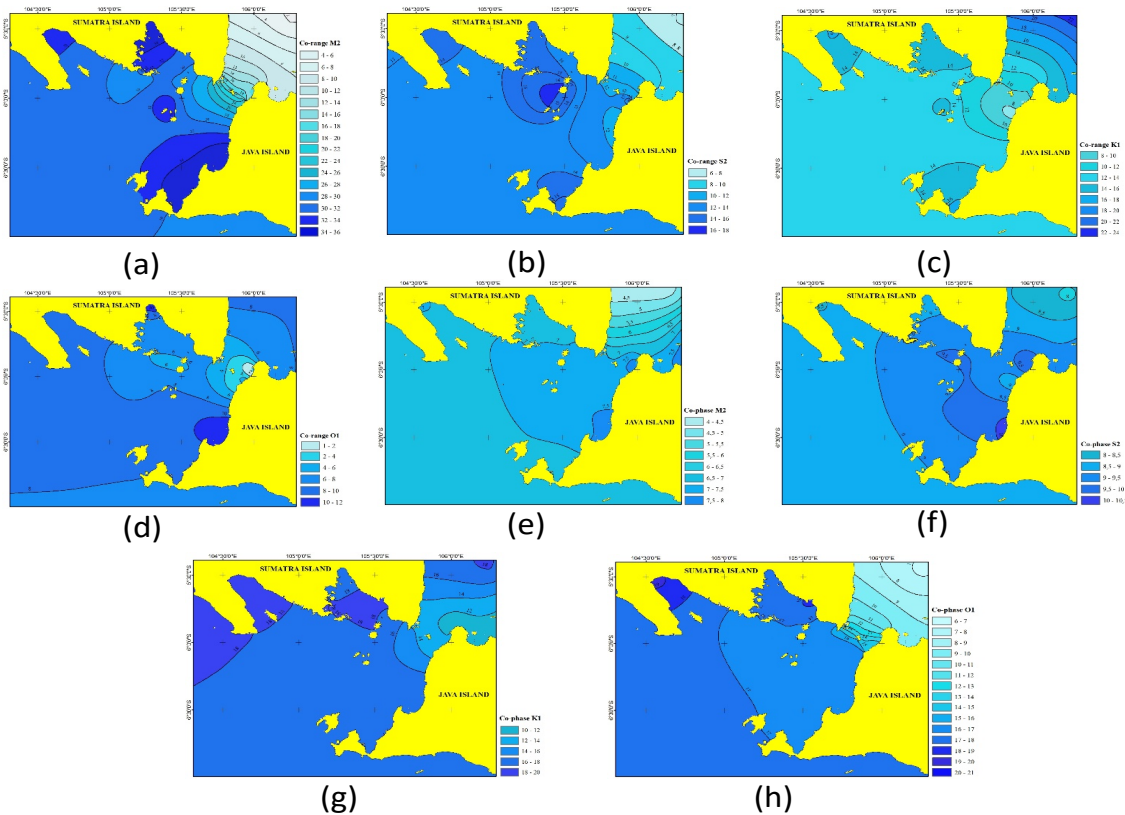


Figure 7 : (a)-(d) co-range chart (e)-(h) co-tidal chart

18 cm in Sebesi Island to Sertung Island. Several areas in the Sunda Strait have a varying degrees of amplitude of S2 based on contour density, at 14 cm contours that start from Labuhan Bangkai area, Krakatau Island to Kalianda and contour 15 cm on Legundi Island, west of Sertung island, south of Sebesi island and west Sebuku island, until the maximum S2 amplitude value on the eastern part of Sertung island and the northern part of Sebesi island. The minimum S2 wave amplitude height in the Sunda Strait waters is 10 cm and the maximum S2 wave amplitude height is 18 cm. Based on the change of wave amplitude height, the effect of the S2 wave amplitude in the Sunda Strait is higher than the effect of the S2 wave amplitude in the northwest of the Java Sea.

The co-range chart of K1 tidal harmonic constants shown in Figure 7 c that the minimum amplitude value of Sunda

Strait waters is 8 cm around Anyer Waters and maximum amplitudes, which is 18 cm in the area of Teluk Paraja and Kota Agung. According to Pariwono [24], the topography of the seabed and the shape of a waters affect the variability of the amplitude value, therefore the K1 tidal wave in the sea of the Sunda Strait tends to be the same, which is from 10 cm to 12 cm, while the amplitude value in some bay is from 14 cm to 18 cm, which is at Lampung Bay, Semangka Bay and Paraja Gulf. The northern part has a wave amplitude height which is from 8 cm to 14 cm that is influenced by the narrow gap of the two islands. Wave length amplitude are increasingly enlarged toward the western part of the Java Sea. This shows that the Sunda Strait is less influence by K1

tidal wave. Based on Figure 7 d, the distance of the O1 wave amplitude height contours in the Sunda Strait are far. In addition, the changes of wave amplitude height are not significant. The wave amplitude heights in the Ciwandan Area are from 1 cm to 2 cm and the wave amplitude height in the Merak area are from 4 cm to 6 cm. The amplitude with the height of 8 cm is found in Labuhan Waters which is connected to Krakatau Kecil Island, Sertung Island to Kiluan Waters. While in the bay area, such as the Lada Bay and Lampung Bay, the amplitude value is quite high, which is about 8 cm to 12 cm while in the south of the Sunda Strait, the amplitude values tend to be the same, which is from 8 cm to 10 cm. This condition shows the effect of O1 tide waves is small.

The tidal harmonic constant of M2 (Figure 7 e) propagates from the southeast waters of Sumatra. Then, the tidal wave propagation direction moves towards Bakauheni-Merak. The tidal wave propagation of M2 in the Sunda Strait is also influenced by the M2 tidal wave coming from the western waters of Sumatra, shown in the contour of 7 waters area of Kiluan (South Sumatera) to the West (West Java). The tidal wave propagation of M2 originating from the western waters of Sumatra and the Southeast Sumatra waters which coincides at 7.5 in Cangirin, Labuhan and Merak areas.

S2 tidal wave propagates through the waters of western Sumatra (Figure 7 f). S2 tidal wave which come from western part of Sumatra moves from 8 to Semangka Bay at 8.5 to 10 in Cilegon waters. The direction of the tidal wave propagation from the southeastern part of Sumatra and the north of the Sunda Strait that

meets or coincides in the area of Sumur, Sebesi Island, Sebuku Island to Labuhan area at 9.5 hours, S2 wave continues to reach the Lada-Labuhan Bay at 10.5. According to Fatoni [23], the double tidal wave propagating in the western waters of Sumatra were the propagation of the Indian Ocean and the tidal wave in the southeastern part of Sumatra is the tidal wave propagation from the Indian Ocean which enters through the Malacca Strait to the waters of Batam to the Karimata Strait.

The tidal wave propagation of K1 comes from the Java Sea (Figure 7 g), where the propagation of the tidal wave will partially propagate to the Karimata Strait and partly propagate to the waters of the Sunda Strait. K1 tidal wave, in Sunda Strait waters area starts at 12 in Merak-Bakauheni then its propagation direction is to the western waters of Sumatra, Sebesi Island and Bay of Semangka, which was at 18. The tidal propagation around Lampung Bay is shown at

16 to Puhawang Island at 18. Meanwhile, O1 tidal waves from the Karimata Strait and Java Sea propagate to Merak waters as shown in Figure 7 h, where the O1 tidal wave travels at 7 and will continue to propagate up to Bakauheni-Merak waters at 10. The propagation tidal waves from Bakauheni-Merak waters continue to propagate to the bay of Semangka and Kalianda at the same time, at 18 hours. O1 tide wave will continue to propagate into the waters of West Sumatra that similarly to the K1 tidal wave. The direction of O1 tidal wave propagation comes from the Pacific Ocean that travels through the Makassar Strait to the Java Sea and travels from the Natuna Sea towards the Sunda Strait. According to Fatoni [23], the K1 tidal wave originates from the Java Sea is a propagation from Pacific Ocean passing to the Sulawesi Sea through the Makassar Strait.

### 3.4. The Map of Z0 Distribution and Types of Tidal Chart

Z0 distribution and types of tidal chart are shown in Figure 8 a. The charts are shown by the consideration of formzahl and Z0 heights in The Sunda Strait. Figure 8 a is a chart of the distribution of Z0 in Sunda Strait, where the height of Z0 has a uniform amount based on the contour line. The northern part of the Sunda Strait of Z0 height has a lower tendency height, for example in Merak area Z0 height is 50 cm, while in Bakauheni area to Cilegon the contour value is 60 cm. The height of Z0 around Sunda Strait waters ranged from 50 cm to 100 cm. Tidal type in Sunda Strait waters is based on formzahl number (Figure 8 b), which showed the type of tidal in the Sunda Strait waters is affected by the amplitude height of O1, K1, M2 and S2. If the amplitude height of O1 and K1 are greater than the height of M2 and S2, then the tidal type in these waters tends to be diurnal or diurnal mixed type. Tidal types in the Sunda Strait waters are dominated by mixed tide prevailing semidiurnal, where the contours are displayed on the chart, ranging from 0.26 to 1. This indicates that the Sunda Strait waters are affected by the amplitude of M2 and S2 or semidiurnal tidal constituents.

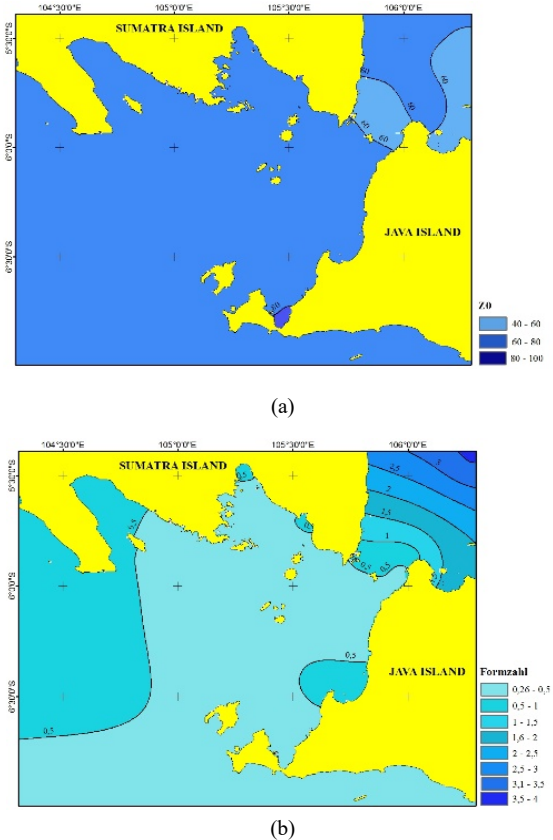


Figure 8: (a) Z0 chart (b) Type of tidal chart

## 4. Results

Co-range chart which is shown the amplitude wave in Sunda strait is affected by semidiurnal harmonic constituent (M2, S2), dominantly. The maximum height of wave amplitude in Sunda Strait is represented by M2 amplitude wave in the value of 36 cm height and the minimum height of amplitude wave is represented by O1 amplitude wave of 2 cm height. Co-phase chart shows the tidal propagation of K1 and O1 that spread from 10 to the end of 20, meanwhile the tidal propagation of semidiurnal harmonic constituent of M2 and S2 spread from 4 to 10. The tidal type in Sunda Strait is dominated by mixed tide prevailing semidiurnal and the height of Z0 is from 40 cm to 100 cm height. The tidal propagation of semidiurnal harmonic constituent (M2, S2) in Sunda Strait were sourced from Indian Ocean. The tidal wave flows through western of Sumatra Waters and Malacca Strait (flows through Batam Waters to Karimata Strait, continuously). The tidal propagation of diurnal harmonic constituent (O1, K1) in Sunda Strait are sourced from Pacific Ocean. The tidal wave flows through Java Sea and South China Sea (flows through Karimata Strait).

### Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

### Acknowledgment

The authors would like to thank to PUSHIDROSAL (Hydrography and Oceanography Center, Indonesian Navy), and

Departement of Oceanography and Coastal Rehabilitation and Disaster Mitigation Studies (CoREM) Diponegoro University for support our research.

## References

- [1] A.S. Rahmawitri, Herwi, Atmadipoera, S.S. Sukoraharjo, "Pola sirkulasi dan variabilitas arus di Perairan Selat Sunda," *Jurnal Kelautan Nasional*, **11** (3), 141-157, 2017.  
<https://doi.org/10.15578/jkn.v11i3.6115>
- [2] Y. Xu, L. Wang, X. Yin, X. Ye, D. Li, S. Liu, X. Shi, R.A. Troa, R. Zuraida, and E. Triarso, "The influence of the Sunda Strait opening on paleoenvironmental changes in the Eastern Indian Ocean," *Journal of Asian Earth Science*, **146**, 402-411, 2017.  
<https://doi.org/10.1016/j.jseae.2017.06.014>
- [3] J.E. Simanjorang, W.S. Pranowo, L. P. Sari, N. P. Purba and M. L. Syamsuddin, "Building up the database of the Level-2 Java Sea Ecoregion based on physical oceanographic parameters," *IOP Conf. Series: Earth and Environmental Science*, **176** (2018) 012009, 2018.  
<https://doi.org/10.1088/1755-1315/176/1/012009>
- [4] P.D. Anggara, T.M. Alam, D. Adrianto, and W.S. Pranowo, "The wave characteristics in Natuna Sea and its adjacent for naval operation base purposes," *IOP Conference Series: Earth and Environmental Science* **176** (1), 1 – 12, 2018.  
<https://doi.org/10.1088/1755-1315/176/1/012009>
- [5] X. Qu and Q. Meng, "Development and applications of a simulation model for vessels in the Singapore Strait," *Expert System with Application Journal*, **39**, 8430-8438, 2012.  
<https://doi.org/10.1016/j.eswa.2012.01.176>
- [6] S. Febrica, *Maritime Security and Indonesia: Cooperation, Interests and Strategies*. Maritime Security and Indonesia: Cooperation, Interests and Strategies. 2017.  
<https://doi.org/10.4324/9781315541815>
- [7] F. Azmi, C.L. Hewitt, and M.L. Campbell, "A hub and spoke network model to analyse the secondary dispersal of introduced marine species in Indonesia," *ICES Journal of Marine Science*, **72** (3), 1069-1077, 2015.  
<https://doi.org/10.1093/icesjms/fsu150>
- [8] M. Wahab, D.P. Kurniadi, T.T. Estu, D. Mahmudin, Y.Y. Maulana, and Sulistyarningsih, "Development of coastal radar network at Sunda Strait," *Telkomnika (Telecommunication Computing Electronics and Control)*, **14** (2), 507-514, 2016.  
<https://doi.org/10.12928/telkomnika.v14i2.2497>
- [9] F.J. Sluiman, "Transit vessel scheduling," *Naval Research Logistics*, **64** (3), 225-248, 2017.  
<https://doi.org/10.1002/nav.21742>
- [10] H. Lelgemann, M.A. Gutscher, J. Bialas, E. Flueh, W. Weinrebe, and C. Reichert, "Transtensional basins in the western Sunda Strait," *Geophys. Res. Lett.* **27**, 3545 – 3548, 2000.  
<https://doi.org/10.1029/2000GL011635>
- [11] C. Susilohadi, Gaedicke and Y. Djajadihardja, "Structures and sedimentary deposition in the Sunda Strait, Indonesia," *Tectonophysics Journal*, **467**, 56-71, 2009.  
<https://doi.org/10.1016/j.tecto.2008.12.015>
- [12] Dishidros, *Kepanduan Bahari Indonesia Wilayah I Edisi Tahun 2014*. Jakarta : Hydrography and Oceanography Center, Indonesian Navy, 2008.
- [13] D.P. Sobaruddin, A. Armawi, and E. Martono, "Model traffic separation scheme (TSS) di alur laut kepulauan Indonesia (ALKI) I di Selat Sunda dalam mewujudkan ketahanan wilayah," *Jurnal Ketahanan Nasional*, **23** (1), 104-122, 2017.  
<https://doi.org/10.22146/jkn.22070>
- [14] A.K. Jolly, "Co-tidal reductions in digital hydrographic survey," *International Hydrographic Review*, **4** (1), 82-86, 2003.
- [15] W. Zhang, Y. Cao, Y. Zhu, J. Zheng, X. Ji, Y. Xu, Y. Wu, and A.J.F. Hoitink, "Unravelling the causes of tidal asymmetry in deltas," *Journal of Hydrology*, **564**, 588-604, 2018.  
<https://doi.org/10.1016/j.jhydrol.2018.07.023>
- [16] van Maanen, B., G. Coco, and K. R. Bryan, "Modelling the effects of tidal range and initial bathymetry on the morphological evolution of tidal embayments," *Geomorphology*, **191**, 23-34, 2013.  
<https://doi.org/10.1016/j.geomorph.2013.02.023>
- [17] F. Ye, Y. J. Zhang, H. V. Wang, M. A.M. Friedrichs, I. D. Irby, E. Alteljevich, A. Valle-Levinson, Z. Wang, H. Huang, J. Shen, and J. Du, "A 3D unstructured-grid model for Chesapeake Bay: Importance of bathymetry," *Ocean Modelling*, **127**, 16-39, 2018.  
<https://doi.org/10.1016/j.ocemod.2018.05.002>
- [18] J.U. Kitheka, B.O. Ohowa, B.M. Mwashote, W.S. Shimbira, J. M. Mwaluma and J.M. Kazungu, "Water circulation dynamics, water column nutrients and plankton productivity in a well-flushed tropical bay in Kenya," *Journal of Sea Research*, **35** (4), 257-268, 1996.  
[https://doi.org/10.1016/S1385-1101\(96\)90753-4](https://doi.org/10.1016/S1385-1101(96)90753-4)
- [19] S. Canhanga and J. M. Dias, "Tidal characteristics of Maputo Bay, Mozambique," *Journal of Marine System*, **58**, 83-97, 2005.  
<https://doi.org/10.1016/j.jmarsys.2005.08.001>
- [20] N. Carpmann, and K. Thomas, "Tidal resource characterization in the Folda Fjord, Norway," *International Journal of Marine Energy*, **13**, 27 – 44, 2016.  
<https://doi.org/10.1016/j.ijome.2016.01.001>
- [21] L. Rose, and P. K. Bhaskaran, "Tidal propagation and its non-linear characteristics in the Head Bay of Bengal," *Estuarine, Coastal and Shelf Science*, **188**, 181-198, 2017.  
<https://doi.org/10.1016/j.ecss.2017.02.024>
- [22] W.J. Pringle, D. Wirasat, A. Suhardjo, J. Meixner, J. J. Westerink, A. B. Kennedy, and S. Nong, "Finite-Element barotropic model for the Indian and Western Pacific Oceans: Tidal model-data comparisons and sensitivities," *Ocean Modelling*, **129**, 13-38, 2018.  
<https://doi.org/10.1016/j.ocemod.2018.07.003>
- [23] K.I. Fatoni, *Pemetaan Pasang Surut dan Pola Perambatannya di Perairan Indonesia (Tidal chart and Tidal Propagation Pattern in Indonesian Seas)*. Thesis. Bogor : Bogor Agricultural University, 2011.
- [24] J.I. Pariwono, *Kondisi Pasang Surut di Indonesia Dalam Pasang Surut*. Editor : O.S.R. Ongkosongo and Suyarso. Jakarta : Indonesian Institute of Science, 1989.
- [25] IHO, *Limits of Oceans and Seas 4th Edition*. Monaco. IHO Publication S-23. 2002.
- [26] I. Mahmudin, Suyatna, Adnan, "Prediksi pasang surut menggunakan proses neural nets (backpropagation) di Pantai Indah Muara Badak Kabupaten Kutai Kartanegara Kalimantan Timur," *Jurnal Ilmu Perikanan Tropis*, **22** (1), 010-019, 2016.
- [27] Fadilah, Suripin dan D. P. Sasongko, "Menentukan Tipe Pasang Surut dan Muka Air Rencana Perairan Laut Kabupaten Bengkulu Tengah Menggunakan Metode Admiralty," *Maspari Journal*, **6** (1), 1-12, 2013.
- [28] X. Wang, Y. Chao, C. Dong, J. Farrara, L. Zhijin L, J. C. Mc Williams, J. D. Paduan, and L. K. Rosenfield, "Modeling tides in Monterey Bay, California," *Deep-Sea Research II*, **56**, 219-231, 2009.  
<https://doi.org/10.1016/j.dsr2.2008.08.012>
- [29] J.M. Pasaribu and S.H. Nanik, "Comparison of DEM SRTM interpolation techniques using inverse distance weighted (IDW), natural neighbor and spline method," *Jurnal Penginderaan Jauh*, **9** (2), 126-139, 2012.
- [30] K. Segeth, "Some splines produced by smooth interpolation," *Applied Mathematics and Computation*, **319**, 387-394, 2017.  
<https://doi.org/10.1016/j.amc.2017.04.022>
- [31] S. Pond and G.L. Pickard, *Introductory Dynamical Oceanography*. 2nd Eds. Pergamon press, 329 pp, 1983.  
<https://doi.org/10.1016/B978-0-08-057054-9.50007-9>

## Fuzzy-logical Control Models of Nonlinear Dynamic Objects

Siddikov Isamiddin Xakimovich, Umurzakova Dilnoza Maxamadjonovna\*

Department of Information processing systems and management, Tashkent State Technical University, Tashkent 100097, Uzbekistan

### ARTICLE INFO

Article history:

Received: 31 May, 2020

Accepted: 16 July, 2020

Online: 09 August, 2020

Keywords:

Control Systems

Fuzzy Logic Controller

PID Controller

Genetic Algorithm

Fuzzy Variables

Adaptation

Controller Synthesis

### ABSTRACT

The article considers the task of developing a fuzzy-logical PID-type controller for a nonlinear dynamic system. A feature of the structure is presented, which consists in simplifying its controller by decomposition. In the simplest version, three fuzzy controllers are used with one input and one output and separate rule bases. Parameters of fuzzy controllers are optimized using a genetic algorithm. A two-step controller tuning scheme for a nonlinear dynamic object is proposed. At the first step, the genetic algorithm is used to tune the linear PID controller; it is shown that the obtained coefficients are used at the output of each channel of the fuzzy PID controller. At the second step, using a genetic algorithm, a nonlinear transforming function is formed for each channel, implemented on the basis of an artificial neural network. The control algorithm is debugged and tested using the MatLab system. The results show a significant improvement in the characteristics of the transient process compared to traditional controllers.

## 1. Introduction

The most common type of industrial controller currently is the PID controller. About 90% of the controllers in commercial operation use the PID algorithm. The reason for such a high popularity is the simplicity of construction and industrial use, clarity of operation, suitability for solving most practical problems and low cost. However, the existing methods for calculating the parameters of PID controllers are oriented to linear systems, since the controller itself is a linear dynamic link. If the control object is essentially non-linear, then it is difficult to achieve a good quality of management.

A fuzzy logical controller (FLC) is a controller that contains in its structure a block of fuzzy logic inference. Usually FLC are included in series with the object controls, like traditional controllers [1], [2].

The classical theory of automatic control is focused mainly on the synthesis of linear controllers based on linearized models, however, all real objects are nonlinear. The nonlinearity of the mathematical model is expressed in the presence of static and dynamic nonlinear blocks, such as “saturation”, “dry friction”, “hysteresis”, etc. FLC, which are nonlinear in nature, can control linear objects are better than classic controls, and also manage

substantially non-linear objects for which linear controllers cannot provide acceptable quality.

The main problem of using FLC is the need to formalize the control law in the form of fuzzy rules that use linguistic variables for descriptions of inputs and outputs of the controller. First FLC used the experience of an expert to describe the law of control [3], [4], but this method suitable only for a limited range of tasks. Standard options for describing FLC rules rely on the analysis of the phase plane of the object management [5]. FLC step-by-step tuning methods similar to the Ziegler-Nichols method for proportional-integral-differential (PID) controllers [6]. But in general, the task FLC settings is an optimization task, to solve which is enough accurate computer model of the object and powerful global search algorithm [7], [8]. The task is to find a suboptimal solution, satisfying user. In [9], an *RBF* network is used, which serves to change the coefficients PID controller. Setup is done using genetic algorithm (GA).

Search engine optimization algorithms are subject-independent, the success of their application for setting up FLC depends on the choice of the optimality criterion and the method of describing the controller parameters. This work is devoted to the consideration of solutions to these problems.

## 2. Solution methods

Experience in designing industrial control systems shows that the behavior of many real dynamic systems can be approximated

\*Corresponding Author: Umurzakova Dilnoza Maxamadjonovna, Department of Information processing systems and management, Tashkent State Technical University, 100097, Uzbekistan, Email: [umurzakovadilnoz@gmail.com](mailto:umurzakovadilnoz@gmail.com)



using the transfer functions of the first or second order (possibly with delay). This feature has led to the widespread adoption of PID controllers as a simple and reliable means of automation. The PID controller equation has the form

$$u(t) = k_p e(t) + k_i \int_0^t e(\tau) d\tau + k_d \frac{de(t)}{dt} = u_p + u_i + u_d. \quad (1)$$

Often, options are used only from two terms (1) - proportional differential (PD) and proportional-integral (PI) controllers. In this case, a clear geometric interpretation of the control law is possible, since here the control surface is a plane (figure 1).

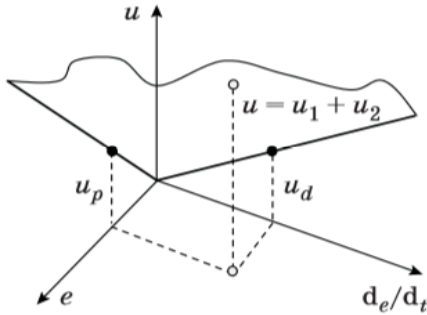


Figure 1. Scheme of the control surface of the PD controller

A fuzzy PID-type controller receives the same input signals as a linear PID-controller, but the control law here is described not by a hyperplane, but by some hypersurface. The classical approach to building FLC leads to the use of control rules with three premises (figure 2, where N and DN are normalization and denormalization coefficients).

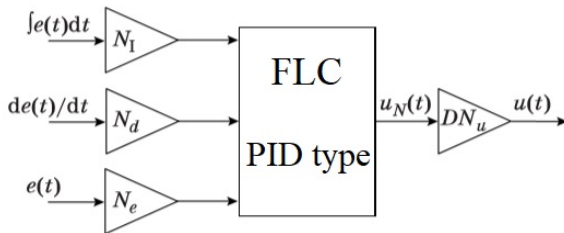


Figure 2. The scheme of the fuzzy PID-type controller with three inputs

The above path is inefficient, because when using  $n$  terms to describe each input,  $n^3$  control rules are obtained. Using fuzzy controllers, PD- and PI-type two signals are received at the input. It can be shown that if we consider at the output of the PI-type FLC not the value of the output signal  $u$ , but its increment  $du$ , then the PI-type FLC control law describes the same rules that the PD-type FLC uses.

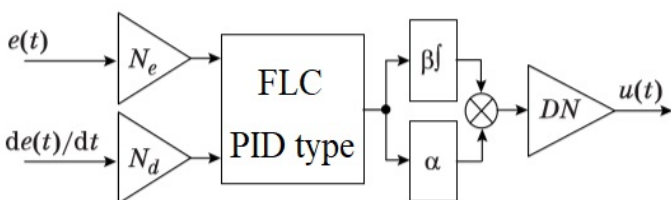


Figure 3. Scheme of a simplified description of PID-type FLC

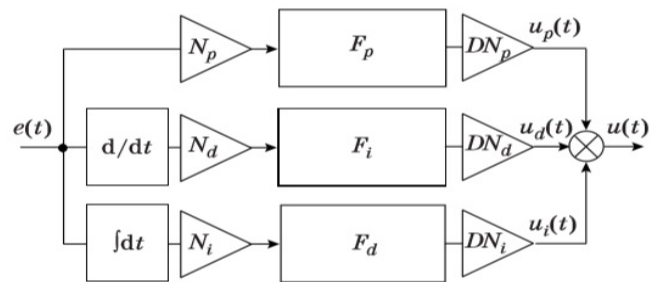
This allows you to use the structure shown in figure 3 to implement PID-type FLC (where  $\alpha$  and  $\beta$  are unknown coefficients).

This representation is often used in practice, the number of fuzzy rules here is reduced to  $n^2$ . Further simplification of the PID-type FLC description is possible when writing the fuzzy control law in a form similar to (1):

$$u(t) = F_p(e(t)) + F_i\left(\int_0^t e(\tau) d\tau\right) + F_d\left(\frac{de}{dt}\right), \quad (2)$$

where  $F_p$ ,  $F_d$ , and  $F_i$  - some nonlinear functions.

When using  $n$  terms to describe each input, only  $3n$  control rules are required here (figure 4).



Normalization coefficients are selected based on a priori data about the control system. Considering the problem of PID-type FLC synthesis as a task of improving the quality of the PID controller, the following algorithm can be used to select denormalization coefficients [10], [11]:

- A linear PID controller is being synthesized, the parameters of which  $k_p$ ,  $k_i$ ,  $k_d$  will play the role of denormalization coefficients.
- Non-linear functional dependencies are described that describe the fuzzy control law for each of the input variables.

Thus, in the first step, the basic gain factors are obtained, and in the second step, additional gain factors that are nonlinearly dependent on the input signal.

### 3. Evolutionary synthesis of nonlinear control law

The search for nonlinear dependencies in (2) can be solved in various ways, however, the most effective here is the use of population metaheuristic methods like genetic algorithm or particle swarm method [12].

The use of a genetic algorithm involves encoding the parameters of a problem using chromosomes whose constituent parts (genes) correspond to individual parameters. A set of chromosomes form a population that evolves over time. The goal of evolution is to improve suitability of chromosomes describing quality solutions to the problem.

The particle swarm method considers individual task parameters as search space coordinates. For each point, the value

of the objective function is calculated. A swarm of particles moves in the search space in the direction extremum.

When using both algorithms, the complexity of the task is determined by the number of tunable parameters and the type of objective function.

To describe nonlinear functions  $F_p$ ,  $F_d$  and  $F_i$ , you can use different methods, in particular, neural RBF networks.

Neural RBF network is a two-layer, it contains a layer of radial basis neurons and a linear output layer [13].

As the radial basis function  $\varphi$  commonly used gaussian function

$$\varphi_i(\|x - c_i\|) = \exp\left(-\frac{\|x - c_i\|^2}{2\sigma_i^2}\right),$$

where  $\sigma$  - width of the “window” of the activation function;  $c_i$  - the center of the activation RBF function of the  $i$ -th neuron;  $y$  - is the input signal.

The output of the RBF network is described by the expression

$$F(t) = \sum_{i=1}^N w_i \varphi_i(t),$$

where  $w_i$  - the weight of the neuron of the output layer.

Nonlinearity  $F_p$ ,  $F_d$  and  $F_i$  are positive, i.e. the product of any input signal and corresponding output positively. Therefore, each neuron of the RBF layer has a paired neuron in which the center has the same module, but a different sign. This allows to reduce the number of customizable options.

Thus, each neuron has two parameters, and the third parameter is the output weight of the neuron.

The optimization problem can be simplified if one pre-distributes RBF neurons to the base scale and select a fixed width of activation functions. Obviously, this operation corresponds to the linguistic description of the input variable using a set of terms (figure 5).

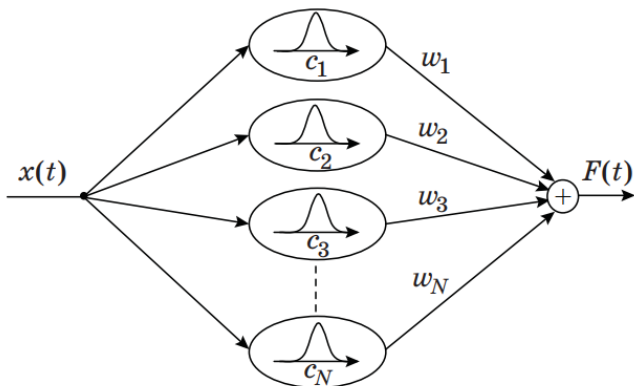


Figure 5. Approximation of static nonlinearity using an RBF network

You can also consider piecewise linear approximation of a nonlinear function. This option can be represented in the form of an RBF network in which radial basis neurons have rectangular activation functions, and the weights of the output layer correspond to the gain of linear sections.

In fact, this means replacing the linear controllers to many linear controllers, each of which is responsible for its own area input space.

Denoting the gain of each linear section as  $w_i$ , we get a vector of custom parameters  $W = \{w_1, w_2, \dots, w_N\}$ .

Description of the objective function is convenient to perform using a reference model that describes the specified transient requirements. Objective function should evaluate proximity object outputs and a reference model, for example:

$$q(W) = \sum_{i=1}^T |y_i^* - y_i|, \tag{3}$$

where  $T$  - number of time points during transition process;  $y$  and  $y^*$  - are the real and desired output value of the object.

General evolutionary optimization scheme PID-type FLC is shown in figure 6.

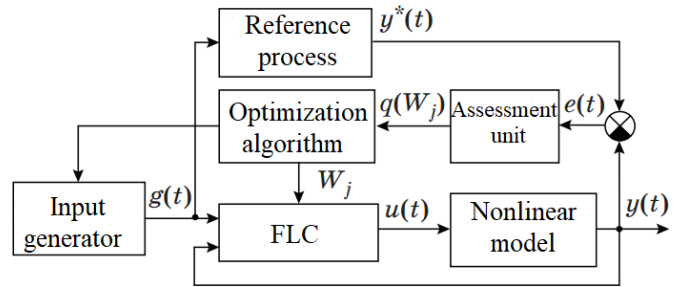


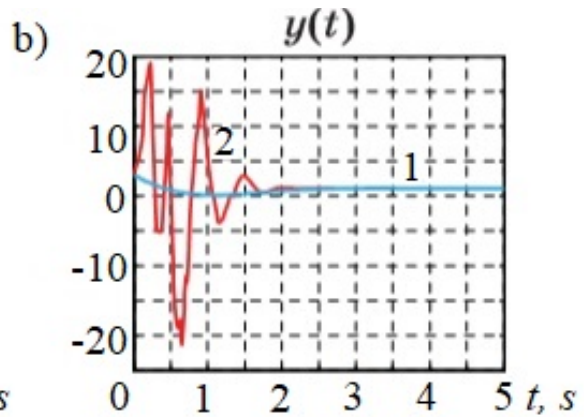
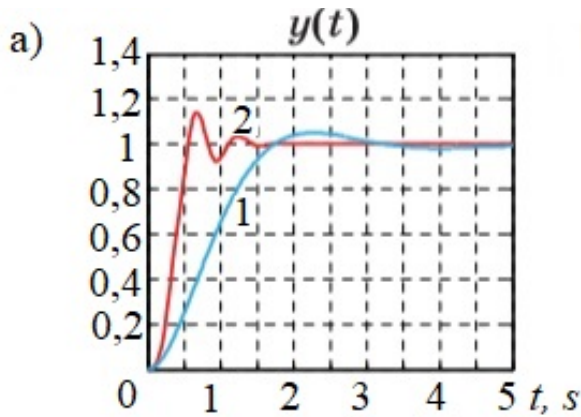
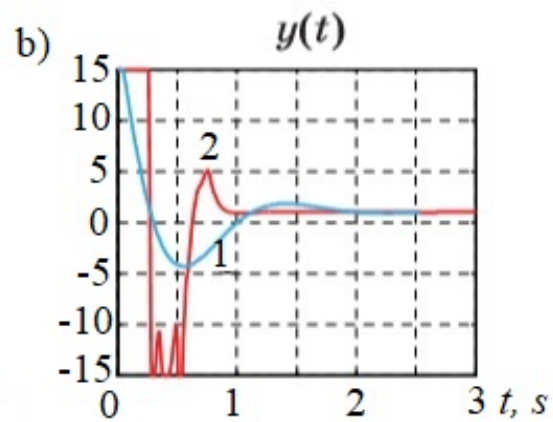
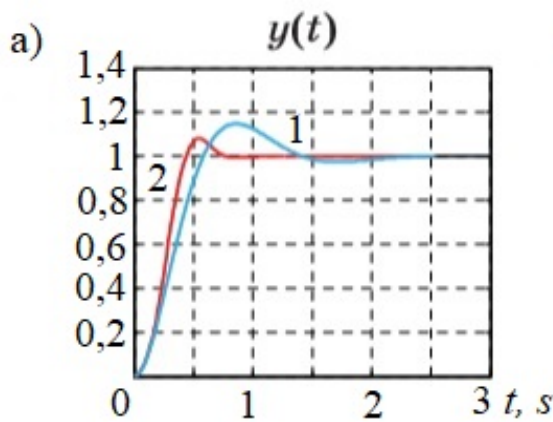
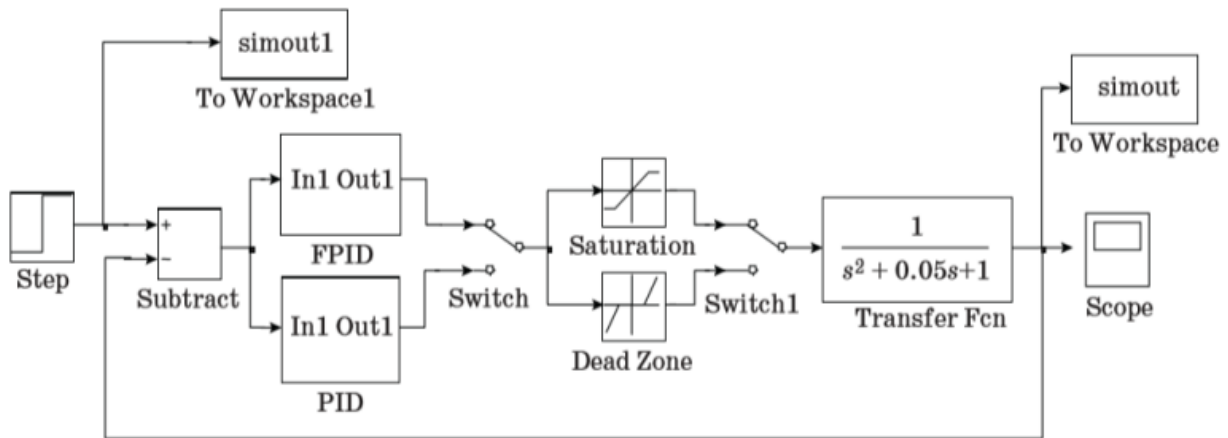
Figure 6. Evolutionary optimization of a fuzzy controller

The optimization algorithm cyclically launches an input generator that produces a test exposure to  $g(t)$ . Each time the controller is started, it gets the parameters  $W_j$  that matches chromosome genetic algorithm or particle coordinates (particle swarm method). At the end transient parameter set gets suitability rating  $q(W_i)$ . Then the population (swarm) is converted, and new testing happens. The criterion for ending the process is usually prolonged lack of improvements or exhaustion of the number of given iterations.

A convenient tool for implementing the described approach is the MatLab package with the extensions Simulink and GAtool [14].

#### 4. Implementation examples

The mathematical description of many industrial facilities (electrical, electromechanical, hydraulic, etc.) with one input and one output can be presented in the form of models containing a static non-linear link and a dynamic linear part in series [15], [16].



Nonlinearity of the “saturation” type is introduced into the model to take into account the limitations of variable levels when studying the behavior of control systems in large deviations from the equilibrium position, and also to describe the maximum levels of the control signal.

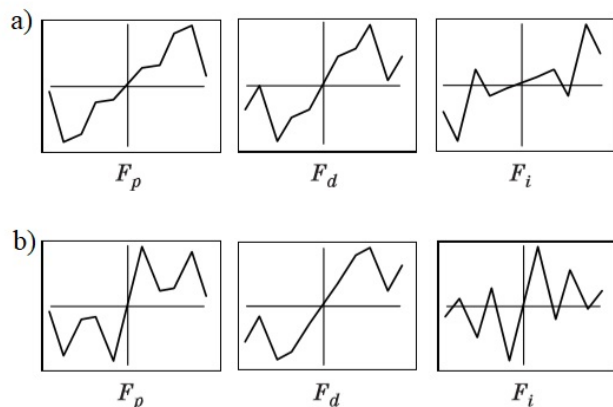
A non-linear element of the type “dead zone” takes into account the real properties of sensors of actuators and other devices with small input signals. The computational experiment diagram is shown in figure 7.

The control object is an oscillating link, to the input of which nonlinearities and a controller of the selected type can be connected. When setting up the PID controller and PID type FLC, a genetic algorithm was used. The simout blocks were used to compute (3).

The simulation results for an object with non-linearity “saturation” are presented in figure 8 a and b, and for an object with non-linearity “dead zone” - in figure 9 a and b. As follows from figure 8 a and 9 a, the transition process time was reduced by

about half, although at the same time the energy costs of management increase (see figure 8 b and 9 b).

Nonlinear functions describing the fuzzy control law for each PID-type FLC channel obtained as a result of genetic training are shown in figure 10. The resulting functions  $F_p$ ,  $F_d$  and  $F_i$  turned out to be significantly different for objects with different nonlinearities. This result is predictable, since the controller can be considered as an inverse model of the object.



## 5. Conclusion

Thus, the article considers a two-step algorithm for the evolutionary synthesis of PID-type FLC, which, in the first step, optimizes the gain of the linear PID controller, and in the second step, additional gain, nonlinearly dependent on the input signal. A comparative analysis of the simulation results shows that the use of PID-type FLC can significantly improve the parameters of the transition process. The proposed technique is simple and can be recommended for use in numerous technical applications to improve the operation of linear PID controllers.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

## References

- [1] K. Passino, S. Yurkovich, "Fuzzy Control", Addison-Wesley, NJ, 522 p. 1998. <https://www2.ece.ohio-state.edu/~passino/FCbook.pdf>.
- [2] D. Rutkovskaya, M. Pilinskiy, L. Rutkovskiy, Neyronnie seti, geneticheskii algoritmi i nechetkie sistemi- M.: Goryachaya liniya-Telekom, 452 p. 2006. <https://www.twirpx.com/file/1934241/>.
- [3] E.H. Mamdani, S. Assilian, "An Experiment in Linguistic Synthesis with Fuzzy Logic Controller", Int. J. Man-Machine Studies, 7(1), 1–13. 1975. [https://doi.org/10.1016/S0020-7373\(75\)80002-2](https://doi.org/10.1016/S0020-7373(75)80002-2).
- [4] L.P. Holmblad, J.J. Osregard, "Control of Cement Kiln by Fuzzy Logic. In: Approximate Reasoning in Decision Analysis", Eds. Amsterdam, New York, Oxford, 389–400, 1982.
- [5] P.J. Macvicar-Whelan, Fuzzy Sets for Man-Machine Interaction. Int. J. Man-Mach. Studies, 8, 687– 697, 1976.
- [6] I.X. Siddikov, D.M. Umurzakova, D.B. Yadgarova, Structural-Parametric Synthesis of an Adaptive Fuzzy-Logical System // Universal Journal of Electrical and Electronic Engineering, 7(2), 94-102. 2020. <https://DOI:10.13189/ujeee.2020.070204>.
- [7] V.A. Demchenko, Automation and modeling of technological processes at nuclear power plants and thermal power plants / V. A. Demchenko. - Odessa: Astroprint, - 308 p. 2001. <https://www.twirpx.com/file/249124/>.
- [8] O. Cordon, F. Herrera, F. Hoffmann, L. Magdalena, Genetic Fuzzy Systems: Evolutionary Tuning and learning of Fuzzy Knowledge Bases. Singapore, New Jersey, London, Hong Kong, World Scientific Publishing, 462 p. 2001. <https://DOI:10.1142/4177>.
- [9] S. Zeng, H. Hu, L. Xu, G. Li, Nonlinear Adaptive PID Control for Greenhouse Environment Based on RBF Network // Sensors 12, 5328-5348, 2012. <https://DOI:10.3390/s120505328>.
- [10] K.J. Astrom, "Revisiting the Ziegler-Nichols step response method for PID control", Journal of Process Control, 4, 635-650, 2004. <https://doi:10.1016/j.jprocont.2004.01.002>.
- [11] C. Lu, C. Hsu, C. Juang, "Coordinated control of flexible AC transmission system devices using an evolutionary fuzzy lead-lag controller with advanced continuous ant colony optimization", IEEE Transactions on Power Systems. 28(1), 385–392, 2013. <https://DOI:10.1109/TPWRS.2012.2206410>.
- [12] X.S. Yang, Engineering Optimization: An Introduction with Metaheuristic Applications. — Hoboken: John Wiley & Sons, — 347 p. [https://www.academia.edu/457294/Engineering\\_Optimization\\_An\\_Introduction\\_with\\_Metaheuristic\\_Applications](https://www.academia.edu/457294/Engineering_Optimization_An_Introduction_with_Metaheuristic_Applications), 2010.
- [13] D. Pelusi, R. Mascella, "Optimal control algorithms for second order systems", Journal of Computer Science. 9(2), 183–197, 2013. <https://doi.org/10.3844/jcssp.2013.183.197>.
- [14] I.X. Siddikov, D.M. Umurzakova, "Mathematical Modeling of Transient Processes of the Automatic Control System of Water Level in the Steam Generator", Universal Journal of Mechanical Engineering, 7(4): 139-146, 2019. <https://doi:10.13189/ujme.2019.070401>.
- [15] I.X. Siddikov, D.M. Umurzakova, H.A. Bakhrieva, Adaptive system of fuzzy-logical regulation by temperature mode of a drum boiler // IJUE Engineering Journal, 21(1), 185-192, 2020. <https://doi.org/10.31436/ijuej.v21i1.1220>.
- [16] I.X. Siddikov, D.M. Umurzakova, "Neuro-fuzzy Adaptive Control system for Discrete Dynamic Objects. International conference on information science and communications technologies applications, trends and opportunities", (ICISCT 2019). Tashkent University of information technologies named after Muhammad al-Khwarizmi. -Tashkent. 4-6 November, 2019. <https://DOI:10.1109/ICISCT47635.2019.9012027>.



## **A Framework of E-Procurement Technology for Sustainable Procurement in ISO 14001 Certified Firms in Malaysia**

Pratik Kumar Singh\*, Fadillah Binti Ismail, Chan Shiau Wei, Muhammad Imran, Syed Ashfaq Ahmed

*Faculty of Technology Management and Business, Universiti Tun Hussein Onn Malaysia (UTHM), Johor 86400, Malaysia*

---

### **ARTICLE INFO**

*Article history:*

*Received: 26 June, 2020*

*Accepted: 13 July, 2020*

*Online: 09 August, 2020*

---

*Keywords:*

*E-procurement*

*Sustainable procurement*

*Supply chain performance*

*ISO 14001*

*ICT*

*Sustainability*

*Technology acceptance model*

---

---

### **ABSTRACT**

*With the current emerging development pattern in Malaysia, E-Government has been unveiled by the Malaysian Government to be one of the multimedia super corridor flagship applications to implement digital technology to improve government operations. E-procurement was originally utilized by businesses to minimize turnaround times and prices, but recently it was often used as a platform for sustainable procurement. However, there have been few attempts to investigate E-procurement technology in Malaysia concerning sustainable procurement. Also, there is no consistent and reliable framework of E-procurement technology for sustainable procurement practices. The focal area of this research is ISO 14001 accredited firms in Malaysia, the population of these organizations is 55, sample size (N)=162 respondents including procurement managers, executives and staff which are involved directly or indirectly in the process of procuring products and services. This study is conceptual and aims to explore how the adoption of E-procurement technologies influence sustainable procurement practices within ISO 14001 accredited firms in Malaysia.*

---

### **1. Introduction**

Adopting recent digital technologies is revolutionizing the procurement function. The diffusion and implementation of modern and evolving technologies provide tremendous potential to improve production and efficiency while simultaneously protecting natural resources and reducing the cumulative environmental effects that are central to accomplishing sustainable development goals (SDG). The advent of the internet has revolutionized means of information sharing, processing and using within organizations to improve their strategic advantage [1]. Procurement plays a crucial part in achieving sustainable results through the growth, production, management, and preservation of ties between organizations and their supply chains [2]. Organizations progressively perceive that implementing information technology to their procurement procedures can gain substantial competitive benefits. The rising significance of the internet has provided possibilities for the enactment of E-procurement (EP), which is defined as using the internet or digital systems to conduct the procurement procedure for organizations to boost productivity [3]. The Boston Consulting Company's latest research finds that 9 of the top 20 Fortune 500

businesses have indicated that emerging technology is essential to their procurement operations in their annual reports [4]. Because of the environmental, economic and social impact of procurement processes and regulations, the call for sustainable practices in ISO 14001 companies has become globally important [5]. Besides, in [6], [7] authors have claimed that the designing and manufacturing sector absorbs huge quantities of non-renewable assets and products by its procedures and operations, like processing and distribution, which consume significant amounts of carbon generating high greenhouse gas (GHG) emissions. Moreover, the procurement process in the ISO 14001 and manufacturing organizations has been increasingly challenged to make a significant impact on sustainable development goals through the incorporation of social and environmental parameters in the procurement practices [8]. And this is why sustainable procurement has been instituted to promote sustainability in ISO 14001 organizations procurement. Sustainable procurement is elaborated as infusing procurement requirements and parameters that are constant with environmental protection, economic growth and societal progress through enhanced product and service quality, resource efficiency and resource management on a life-long basis [3], [9]. Numerous elements, like sustainable procurement practices (SPP), recycling, renewable activity have been discussed in the literature [10], [11], but approaches and initiatives to guarantee that

---

\*Corresponding Author: Pratik Kumar Singh, Faculty of Technology Management & Business, Universiti Tun Hussein Onn Malaysia, +60166001263 & [singhpratik.uthm@gmail.com](mailto:singhpratik.uthm@gmail.com)

SPP is widespread and rapidly implemented within the ISO certified organization are limited in the literature. According to certain figures, EP could minimize production costs by 5 to 10 percent, boost efficiency by 30-50 percent and dramatically enhance productivity, consistency and risk management [4]. Also, a new survey from the McKinsey Consultancy Company shows that chief buying officers anticipate an average savings rise of 40 percent, 30-50 percent less period spending on the transactional acquisition and a 50 percent decrease in value leakage owing to automated procurement systems [12].

This study is supported by the associated principle of the theory Technology Acceptance Model (TAM). The perceived usefulness and ease of use determine the attitude of an individual concerning their intent to use innovation and technology to act as an intermediary for the actual use of the procurement system [13]. TAM is the most cited and accepted model among other models and theories in studies of IT literature to examine the determining factors of technology reception and its application in e-commerce, online banking and supply chain [14], [15]. To ensure sustainable procurement practices, organizations need to conduct an evaluation among employees concerning the execution of E-procurement. The Malaysian Government embraced the effort to use e-Government for some of its activities. The e-Government Multimedia Super Corridor (MSC) introduced in 1997 was planned to fulfill the Vision 2020 objectives. The introduction of online procurement was the first step to resolve government procurement problems [16]. Certification of ISO 14001 proved to be an important factor in industrial practice for technical modernization. ISO 14001 is now the prevailing universal standard for the evaluation of environmental management processes (EMS). This directs emphasis on a variety of important fields, including manufacturing methods and technologies. Organizations with formal environmental protection programs (ISO 14001) have strong positive relations with several parameters of sustainable procurement also the organizational performance [17].

### 1.1. Problem Statement

In 2002, a system of E-procurement was established in Malaysia. Nevertheless, the Ministry of Finance has 30,000 public service providers enrolled, which does not effectively seem to have an existing E-procurement infrastructure [18]. Handfield et al., [19] recommend that ISO 14001 certification may be a guiding force for the adoption of sustainable environmental management standards, and is a crucial topic open to discussion and, ultimately for research in organizations. For many factors, the E-procurement analysis in ISO 14001 organizations is significant. Firstly, ISO 14001 certified organizations account for a large portion of today's economies and provide advantages to the firm. Secondly, the capacity of E-procurement is very important in contributing to the overall competitiveness of such companies [20]. The slow pace of adoption and the gradual nature of E-procurement approval is another substantial issue in implementing EP in ISO certified organizations [17]. Gunasekaran et al., [21] further stated that SME does not consider EP benefits to be sufficient. While evidential analysis has demonstrated that EP has a significant effect on sustainability efficiency in the scope of big enterprises [22]-[25], no studies have so far inspected the relationship between EPT and SPP in the context of ISO 14001 accredited organizations [3], [26]. In other terms, less attention is paid to E-procurement innovations [27],

[28]. Organizations can reduce the difficulties normally posed by planning and implementing a sustainable procurement strategy by the use of standard practices in E-procurement [29]. Sustainable procurement is the purchasing ecofriendly goods and services, and also the establishment of environmental standards in agreement, taking into account natural, social and economic considerations [30, 31]. The research concerning sustainable procurement practices is lacking particularly in Malaysia [32], [33]. Many of the research conducted on both existing procurement paradigms, i.e. EPT and SPP, found implementing the two paradigms independently [34], [35]. This strategy has culminated in a slow rate of adoption of both practices [7], [36] which hinders the ability to promote E-procurement and sustainable procurement effectively and efficiently. Limited studies have addressed possible relationships and compatibility between them [6]. However, the mechanism by which E-procurement processes influence the management of the sustainable supply chain is still an unexplained concern [37]. The theoretical fusion of the two paradigms attracted minimal attention from the researchers such as, [3] and [38] suggested the contributions of E-procurement to sustainable procurement initiatives in the procurement of projects. Nevertheless, the scarcity of comprehensive research and amalgamation frameworks for interventions to support the research paradigm prompts to find approaches to be combined and commonly used to support the two paradigms in the ISO 14001 industry. In the sense of sustainable procurement, it implies that E-procurement techniques are scarcely examined [3]. This study is aimed at establishing an integrated framework of approaches in accredited ISO 14001 corporations for E-procurement technologies for sustainable procurement activities. This research proposes a framework to analyze the relationship between EPT and SPP.

### 1.2. Conceptual Framework

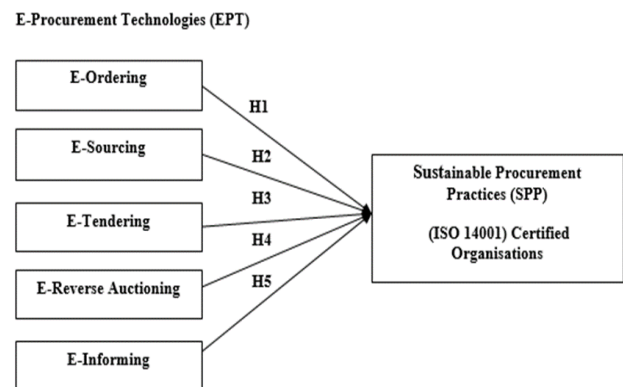


Figure 1: Theoretical Framework to Enhance Sustainable Procurement Practices

## 2. Literature Review

### 2.1. Sustainable Procurement Practices (SPP)

Sustainable growth is generally conceptualized as 3Ps representing the parameters of the Triple Bottom Line (TBL) [39]—that is, environmental (the planet), economic (profit) and social (people). Potential buyers must extend the conventional factors so that they incorporate 3P (Planet/Profit/People), whom they purchase, who / from which they purchase, their contractual terms of buying and their manufacturing processes in their purchasing sectors. They ought to include meaningful effects on this context of their decision-making criteria. Sustainable procurement is part

of a broader debate on greening the supply chain of an organization [37]. Scholars have concluded that the organizational tendency to implement new technologies and strategies linked positively to the degree of its decentralization [40]. When adopting new techniques that assume a certain degree of risk, the members of the organization must have both the authority and the confirmation that such behavior is permissible. Sustainable acquisition and supply chain management approaches have been consistently correlated with various strategic advantages such as advantageous environmental effects, positive changes in social wellbeing, enhanced perceptions of the public, cost savings, increased profitability of businesses, boosts procurement operations and stronger organizational performance [8], [41], [42]. Several states, organizations and agencies at different levels around the globe have been encouraged to implement acquiring strategies and policies that promote environmentally sustainable products and services under the Environmental Management System (EMS) [9], [43]. Successful institutionalization often calls for a strong degree of trust and coordination between the institution and its supplier network. Because of these indicators, few firms may be unable to engage in the transformation of the organizational culture to meet the demands of operational sustainable procurement. Likewise, provided that organizational values influence procurement actions on an individual basis [44]. Sustainable procurement needs manufacturers to purchase goods using renewable materials such as non-toxic, reusable and biodegradable; it promotes the elimination of waste and the reduction of hazardous materials [45]. Sustainable procurement has 12 parameters, namely environment, product differentiation, health, philanthropy, civil rights, purchasing from small local industries, policies and regulations, effective leadership, specific tendering and evaluation criteria, performance monitoring and reporting, competitive lead and technical competence. The ecosystem is one of the programs involved in the life-cycle review (LCA) by-product development; manufacturer demands comprising recycling, waste decrease and re-use; procurement of sustainable packing and buying of recycled packing. Programs to promote the integration of vendors included the purchase by firms operated by women or racial minorities [46]. On the other side, security initiatives include the smooth flow of goods onto business premises and the protected functioning of manufacturer infrastructure and also a thorough review of building management acquisition strategies. Human rights issues involve giving employees a minimum wage, tracking working practices and complying with labour legislation and regulations [46]. Also, it is essential for companies, through purchasing small and local vendors, to help local communities and economies and to obey procurement rules [33]. Good leadership often plays a significant role in a company's growth. It is seen as a vital factor as it ensures the strategic action strategy of the company. Besides, Mpeera Ntayi and Mugume [47] claimed that distributors should be assessed, recorded and controlled permanently and distributors should provide constructive feedback. As per Metcalf and Benn [48], one of the most significant facets of a corporation's success is its internal competitive edge. Meehan and Bryde [11] have said that technical skills are essential for understanding, gaining and

facilitating the change of attitudes among acquisition workers [37]. Several measures to promote sustainable procurement activities have been taken by the Government of Malaysia [49]. For the preparation of a corporate sustainability mechanism in Malaysia, the Ministry of Energy Green Technology and Water (MEGTW), and the Malaysian Finance Department are responsible. The government has the authority to elect to buy greener products and services and also to inspire businesses that take part in ecologically sound operations. There is a range of environmental initiatives, including sustainable technologies and development, emission reduction and implementation of an EMS and ISO 14001 registration in Malaysia. As mentioned by the Malaysia Department of Statistics (2019), the number of eligible accredited organizations of ISO 14001 in Malaysia until Q2- 2019 was 1586 and seen immense growth. This is a positive benefit since the EMS is a collective contribution to sustainable development by retailers, rather than a product-focused strategy. Furthermore, Mansi [50] stated that the current literature on sustainable procurement studies in emerging or developing countries including the Asian region is limited.

## *2.2. E-procurement technologies (EPT)*

E-procurement is an electronically operated functionality [51], [52] that integrates the use of ICT and buying assets to process buying orders, exchange information to contractors and to promote buying decisions. Additionally, rapid purchase cycle times offer greater versatility and even more precise order placement details [18], [53]. Based on [3], 26 driving elements display EPT as a tool for S Such drivers have been classified into six key dimensions, namely digital management, unified procurement authority, threat, integration, data analysis, and an enabling supplier partnership, that can support in handling the sustainable procurement systems through the use of ICT.

E-procurement is defined as implementing technology arrangements in the procurement phase, including purchasing, buying, tendering, auctioning and bargaining, within the framework of supply chain management [54]. The EPT is divided into five main applications: E-informing, E-reverse auctioning, E-ordering, E-tendering, and E-sourcing. Technology Acceptance Model (TAM) is an extension of the Theory of Reasoned Action (TRA) introduced by Fred Davis in 1989. TRA theory is mainly based on the idea of technology acceptance; TAM replaced TRA with two standard technological features perceived usefulness and perceived ease of use [55]. The technology acceptance model was introduced to identify the scale of usage of information systems and justified by Davis (1993). Perceived ease of use and perceived usefulness are the fundamental constructs of the Technology Acceptance Model (TAM) which can better predict with an individual's attitude and their intention towards EPT adoption and actual usage in procurement practices [56]. This theory brings an understanding that acknowledgment and use of innovation is a function of the user's emotions about the system and its perceived benefits. An individual's positive or negative attitude predicts the individual's actual behavior towards the usage of the technology. Quesada et al., [57] determined that EPT adoption has a positive effect on employee perception of both sustainable procurement procedures and procurement performance. The behavioral objectives of EPT are largely described by the specific attitude of



the individual and have a further impact on perceived performance and subjective quality [27]. Moreover, [58] concluded that E-procurement technology contributes to supply chain performance positively and smoothen the procurement process. The research on E-procurement has gained substantial coverage in the literature since utmost studies examine electronic procurement from the perspective of major companies e.g. [54]-[60], electronic procurement in SME [21], [29], [61].

### *2.3. Influence of E-informing on SPP in ISO 14001 certified organization*

E-informing is an internet-enabled data knowledge retrieval tool for interested parties to collect and share knowledge. The extent to which specific and exclusive data is distributed to a supply chain participant, therefore, allows the online knowledge exchange to include more competitive sourcing processes and better supply chain efficiency [38]. Information sharing doesn't just impart data to partners but additionally gives adequate, timely and precise data. E-informing also includes aspects such as credibility and accuracy of information shared. Moreover, e-informing provides a mechanism for organizations to control, facilitate, and conserve on exchange costs, as it improves information streams and lessens uncertainty. The e-informing application helps the organization in providing adequate information from time to time and reduces the occurrence of corruption activity ultimately leads to better procurement practices [62]-[64].

H1. E-informing influences SPP in ISO 14001 certified organizations positively.

### *2.4. Influence of E-ordering on SPP in ISO 14001 accredited organization*

The E-ordering application empowers the business to decrease costs, increase efficiency and improve client support in this way improved sustainability in procurement practices. Moreover, Gunasekaran et al., [21] declare that e-ordering is an e-commerce function where an organization permits clients to demand goods or services through web-portal. Since Internet technology is booming, having a web-based ordering system can help increase sales as it facilitates clients to put in a request for the organization's services [62]. Kim and Shunk [65] contend that e-ordering is the way toward making and approving buying requisition, submitting buying orders, receiving products and services requested, by utilizing a web-enabled software structure that improves the procurement process and supply chain performance [38]. Interested parties are provided with a web-enabled portal with adequate knowledge on procurement, such as tender advertisements, evaluations and outcomes of placement. Such permeability and receptiveness upgrade responsibility in managing the sustainable procurement and barred potential complicity and advantage by government authorities and purchasers [63].

H2. E-ordering influences SPP in ISO 14001 accredited organizations positively.

### *2.5. Influence of E-sourcing on SPP in ISO 14001 accredited organization*

E-sourcing is the process of distinguishing and selecting new suppliers to deliver products and services in a predefined category through electronic means. It is a web-based application that empowers a collaborative innovation in the period of the procurement process between supplier and buyer. The organizations are employing e-sourcing which is one of the systematic e-purchasing practices to reduce costs. The e-sourcing provides a cooperative environment for suppliers and buyers by giving a centralized platform where they can share data adequately. A single line platform also offers proposals from the different vendors for the e-sourcing program [66]. In this way, the organization can choose the supplier strategically to achieve a competitive advantage. New technology has offered to ascend to better approaches for sourcing relevant suppliers, directing business and improving procurement as well as operational performance. The utilization of the technology has prompted expanded efficiency, services and procurement practices [67].

H3. E-sourcing influences SPP in ISO 14001 accredited organizations positively.

### *2.6. Influence of E-tendering on SPP in ISO 14001 accredited organization*

E-Tendering is an application where the organization advertises through electronic requests or e-tender notices where the invoices and purchase requests, offers from suppliers, receive bids and informing suppliers about the award of contracts employing the use of web-based information exchange. E-Tendering is the way toward sending RFI (Request for Information) as well as RFP (Request for the Proposal) to suppliers and accepting responses using web technologies [67]. Industries, such as manufacturing and construction, can gain numerous advantages by implementing e-tendering, such as time and cost cutbacks in the tendering and bidding process, ensuring secure access to tendering work, improving efficiency, etc., and maintaining sustainability in the procurement procedure [68]. E-tendering usage in the purchasing process has advantages like screening and selection of qualified suppliers automated reduce the price, enhance flexibility in requests process, and quality. This leads to the greening of procurement practices and significant for supply chain performance [38]. E-tendering makes the automation of many operational tasks and procurement processes [69]. It is a protected platform that enables purchasers and suppliers to trade virtual. E-tendering structures extend the tendering requirement process up to the completion of contracts when the details and documentation are exchanged digitally

H4. E-tendering influence SPP in ISO 14001 accredited organizations positively.

### *2.7. Influence of E-reverse auctioning on SPP in ISO 14001 accredited organization*

E-reverse auctioning (ERA), also called "downward price auctions" referred to a web-based bidding mechanism that happens based on real-time procurement contracts by the eligible vendor



database [57]. The ERA has advantages to both the seller and buyer. ERA application in procurement reduces the cost under a reduction in sustainable procurement cycle time [70]. While [37] that the use of EPT like e-reverse auctioning has led to increased effective performance and efficiency in the procurement procedure. For the buyer, ERA has advantageous aspects such as reduced costs, reduced cycle time, discounted costs, create standards, increase geographical range, and it promotes accountability and transparency of the bidding procedure.

H5. E-reverse auctioning influence SPP in ISO 14001 certified organizations positively.

The literature contains a variety of articles explaining the problems of e-procurement and sustainable procurement. Whereas some empirical research [33], [71] concentrated on application expansion for sustainable procurement, fewer [53], [55] looked primarily at the deployment and cost-effectiveness of e-procurement uptake. Also, part of the research investigating SPP discussed the objectives and barriers/challenges of sustainable procurement [72], [73] while [74] looked at the views of corporate buyers on the effect of e-procurement systems on knowledge management and logistic fulfillment of quality. Review findings in both sustainable procurement and E-procurement [31], [75], [76] reviewed scientific developments trends and issues in the literature for E-procurement and sustainable procurement separately. Other studies explored the classification and assessment requirements for SPP [77], [78] and the development of collective procurement systems to lead to sustainability through a green economy for the construction and destruction of waste. There is scarce research focused on a detailed analysis of how these two fields could be integrated, which poses certain challenges in the pursuit of productivity and efficiency within procurement. Therefore, this study addresses the gap with the development of an incorporation system that provides a forum to promote research participation in supporting both measures in the future. Moreover, the research by [38] identified the relationships between e-procurement and sustainable procurement. Walker and Brammer [38] outlined the minimal research of EPT and SPP in literature, and progressively disclosed that connections between some social dimensions of SPP and EPT tend to be less established, while support structures may be developed to improve EP. This shows the need to develop how well these two approaches can be best integrated to improve sustainability and profitability in the procurement process.

### 3. Methodology

The study proposes the quantitative method. This study has utilized a positive approach and methodology because the collected data will be objective and quantitative. Deductive reasoning was often used to place hypotheses and construct comprehension until the findings were tested through scientific data. The questionnaire survey is the main instrument used to assess the proposed framework. The sampling methodology that was being picked is probability sampling. Probability sampling is led by randomly choosing a sample where every one of its subjects in the sample possesses all the qualities of the research

[www.astesj.com](http://www.astesj.com)

population. For this research, simple random sampling is used because the sampling frame was prepared and the sample represents the population. The most significant explanation behind using probability sampling in this study is that it makes samples that are profoundly representative of the population. It is likewise an unbiased random selection and a representative sample is significant in drawing inferences from the outcomes of the research. Consequently, this method is completed towards employees of the procurement department of ISO 14001 certified organization to accomplish the objective of this research. The population composed of employees of procurement departments from 55 organizations with ISO 14001 certifications of Johor, Malaysia. Hence, according to Krejcie and Morgan (1970) table for determining the sample size of the finite population, the sample size of this study comprises of 162 individuals from ISO 14001 accreditation organizations. Such respondents include procurement executives, procurement supervisors, and procurement personnel who are actively or indirectly engaged in the procurement of products or services. This research proposes techniques for demographic analysis SPSS and SEM modeling for data analysis and reflective models for estimation. One of the most significant and regular questions concerning if there is a statistical relationship between a reaction variable or dependent variables (Y) and illustrative or independent variables (Xi). An alternative to answer this question is to utilize regression analysis to model its relationship. The independent variable in this study is E-procurement technologies (E-Ordering, E-Sourcing, E-Tendering, E-Reverse auctioning, E-Informing) while the dependent variable is the Sustainable procurement practices. There are different kinds of regression analysis. The type of the regression model relies upon the sort of the distribution of Y; if it is continuous and approximate normal we use linear regression model.

In the multivariate regression, Y has a normal distribution with mean

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + \sigma(Y), \text{sd}(Y) = \sigma \text{ (independent of } X\text{'s)}$$

In this study, the equation of multivariate regression is

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + e \text{ where;}$$

Y = Sustainable procurement practices

$\alpha$  = Autonomous factors

X1 = E-Ordering

X2 = E-Sourcing

X3 = E-Tendering

X4 = E-Reverse auctioning

X5 = E-Informing

e = error term

Bs = Beta coefficients of independent variables

PLS-SMART is the preferred software tool for this simulation. The model of E-Procurement for Sustainable procurement is formulated by using Partial Least Squares Structural Equation Modeling (PLS-SEM) techniques. SEM is primarily used to study structural relationships between variables for the purpose to establish predictive and theoretical linkages [79]. SEM is employed to create a measurement model that best fits data at hand and a structural model to test the causal relationship among the constructs. It allows the researcher to support the developed theories and to select the best model by extending the standard

multivariate analysis methods including regression, factor analysis, correlation and analysis of variance (ANOVA).

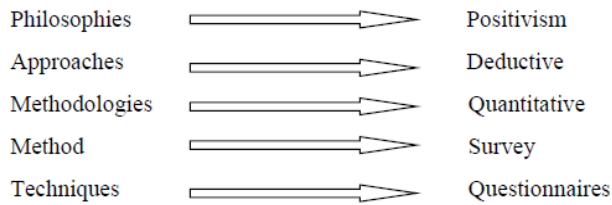


Figure 2. Summary of the selected Research Philosophy Approach

#### 4. Conclusion

The primary aim of this research is the implementation of EPT for S A new framework of E-procurement technologies for sustainable procurement will be articulated. The framework helps to improve the ecological and environmental change recognized from National Priority Area (NPA). Procurement of eco-friendly products as well as services using E-procurement inventions that cause less pollution, make economic and efficient use of resources and energy and increase interest in sustainable products and services. This research provides a roadmap of proactive planning and decision-making for industry professionals, decision-makers, business contractors and providers to facilitate sustainable e-commerce growth and sustainable procurement in ISO 14001 certified sustainable development organizations and to meet government requirements. Also, legislatures, decision-makers and regulatory committees became more productive and structured to develop plans and regulations for sustainable procurement in Malaysia. Throughout the integration context, the elements are specified to direct the execution of the future research plan to strengthen procurement processes. The implementation of sustainable procurement in Malaysia by E-procurement technologies would benefit the civil, ecological and economic aspects and thus improve citizens' quality of life. This work is therefore intended to fill a significant gap in the literature and to increase our awareness of E-procurement in context with sustainable procurement practices in Malaysia ISO 14001 certified organization. By addressing this gap, the significance of this study is justified.

#### Conflict of Interest

The authors declare no conflict of interest.

#### Acknowledgment

This research is entirely encouraged by “Research Fund E15501, Research Management Center, Universiti Tun Hussein Onn Malaysia (RMC-UTHM)”.

#### References

[1] N.N. Wimalasena, S. Gunatilake, “The readiness of construction contractors and consultants to adopt e-tendering: The case of Sri Lanka. *Construction Innovation*, **18**(3), 350-370, 2018.

[2] E. Hassini, C. Surti, C. Searcy, “A literature review and a case study of sustainable supply chains with a focus on metrics. *International Journal of*

*Production Economics*, **140**(1), 69–82, 2012. DOI:10.1016/j.ijpe.2012.01.042

[3] M. Ramkumar, M. Jenamani, “Sustainability in Supply Chain Through E-Procurement—An Assessment Framework Based on DANP and Liberatore Score. *IEEE Systems Journal*, **9**(4), 1554–1564, 2015. DOI:10.1109/jsyst.2014.2336291

[4] M. Högel, W. Schnellbacher, R. Tevelson, D. Weise, “Delivering on Digital Procurement’s Promise, Boston Consulting Group, Boston, 2018, available at: [www.bcg.com/publications/2018/delivering-digital-procurement-promise.aspx](http://www.bcg.com/publications/2018/delivering-digital-procurement-promise.aspx) (accessed June 1).

[5] A.V. Roman, “Institutionalizing sustainability: A structural equation model of sustainable procurement in US public agencies. *Journal of cleaner production*, **143**, 1048–1059, 2017. DOI: 10.1016/j.jclepro.2016.12.014

[6] A.A.M. Bohari, A. A. M., Skitmore, M., Xia, B., and Teo, M. “Green oriented procurement for building projects: Preliminary findings from Malaysia. *Journal of Cleaner Production*, **148**, 690–700, 2017. DOI:10.1016/j.jclepro.2017.01.141

[7] J.K.W. Wong, Facilitating effective green procurement in construction projects: An empirical study of the enablers. *Journal of Cleaner Production*, **135**, 859–871, 2016. DOI:10.1016/j.jclepro.2016.07.001

[8] S. Brammer, H. Walker, H. “Sustainable procurement in the public sector: an international comparative study. *International Journal of Operations and Production Management*, **31**(4), 452–476, 2011. DOI:10.1108/01443571111119551

[9] M.M. Islam, “Aspects of sustainable procurement practices by public and private organisations in Saudi Arabia: an empirical study. *International Journal of Sustainable Development and World Ecology*, **24**(4), 289–303, 2016. DOI:10.1080/13504509.2016.1209794

[10] M.H. Eriksen et al., “Strengthening requirement specification in sustainable procurement—An investigation of challenges. *Journal of Green Building*, **12**(1), 107–122, 2017. DOI: <https://doi.org/10.3992/1552-6100.12.1.107>

[11] M.J., D.J., “A field-level examination of the adoption of sustainable procurement in the social housing sector. *International Journal of Operations and Production Management*, **35**(7), 982–1004, 2015. DOI:10.1108/ijopm-07-2014-0359

[12] P. De-la-Boulaye, P. Riedstra, P. Spiller, “Driving superior value through digital procurement”, McKinsey and Company | Global Management Consulting, April, New York, 2017, available at: [www.mckinsey.com](http://www.mckinsey.com) (accessed October 2, 2019)

[13] Mohammadi, H. “Investigating users’ perspectives on e-learning: An integration of TAM and IS success model. *Computers in Human Behavior*, **45**, 359–374, 2015. DOI:10.1016/j.chb.2014.07.044

[14] P. Tobbin, “Towards a model of adoption in mobile banking by the unbanked: a qualitative study. *Emerald Group Publishing Limited*, **14**(5), 74–88, 2012. <https://doi.org/10.1108/14636691211256313>

[15] C. Tam, T. Oliveira, “Literature review of mobile banking and individual performance. *International Journal of Bank Marketing*, **35**(7), 1043–1065, 2017. <https://doi.org/10.1108/IJBM-09-2015-0143>

[16] K. Azmi “e-Procurement: A Tool to Mitigate Public Procurement Fraud in Malaysia?” *The Electronic Journal of e-Government* **13**(2), 150-160, 2015, available online at [www.ejeg.com](http://www.ejeg.com)

[17] Y. Nee, G. Abdul-Wahid, The Impact of ISO 14001 Environmental Management System (EMS) Implementation on SMEs Performance: An Empirical Study in Malaysia. *Journal of Sustainable Development*, **3**(2), 2010. doi:10.5539/jsd.v3n2p215

[18] M. Nawi, M. Nasrun, R. Deraman, J.A. Bamgbade, F., Zulhumadi, R.M. and S. Riazi, S “E-procurement in Malaysian construction industry: benefits and challenges in implementation”, *International Journal of Supply Chain Management*, **6**(1), 209-213, 2017. ISSN 2050-7399

[19] R.B. Handfield, et al., “Applying environmental criteria to supplier assessment: a study in the application of the analytical hierarchy process, *European Journal of Operational Research*, **141**(1), 70-87, 2002.

[20] H.C. Preez, H.C. and Folinas, D. “Procurement’s contribution to the strategic alignment of an organisation: findings from an empirical research study”, *Supply Chain Forum: An International Journal*, **20**(3), 1-10, 2019.

[21] A. Gunasekaran, et al., “E-procurement adoption in the Southcoast SMEs”, *International Journal of Production Economics*, **122**, 161-175, 2009. <https://doi.org/10.1016/j.ijpe.2009.05.013>

[22] A.N. Mishra, S. Devaraj, “Capability hierarchy in electronic procurement and procurement process performance: an empirical analysis”, *Journal of Operations Management*, **31**(6), 376-390, 2013.

[23] N. Ross, P.A. Bowen, D. Lincoln, “Sustainable housing for low-income communities: lessons for South Africa in local and other developing world cases. *Construction Management and Economics*, **28**(5), 433–449, 2010. DOI:10.1080/01446190903450079

[24] J. Upstill-Goddard, the role of absorptive capacity. *Engineering, Construction and Architectural Management*, **23**(4), 407-427, 2016.

- [25] Z. Wu, D.A. Ross, "Antecedents and outcomes of e-procurement adoption: an integrative model", *IEEE Transactions on Engineering Management*, **54**(3), 576-587, 2007.
- [26] D. Vries, D.K. Bayramoglu, and van der Wiele, T. "Business and environmental impact of ISO 14001. *International Journal of Quality and Reliability Management*, **29**(4), 425-435, 2012. DOI:10.1108/02656711211224866
- [27] M.G. Aboelmaged, "Predicting e-procurement adoption in a developing country: an empirical integration of technology acceptance model and theory of planned behaviour", *Industrial Management and Data Systems*, **110**(3), 392-414, 2010.
- [28] A. Ahimbisibwe, T. Wilson, T. Ronald, "Adoption of E-procurement In Uganda: Migration from the Manual Procurement Systems to the Internet. *Journal of Supply Chain Management*, **3**(1), 1-14, 2016.
- [29] A. Altayyar, J. Beaumont-Kerridge, "External Factors Affecting the Adoption of E-procurement in Saudi Arabian's SMEs. *Procedia - Social and Behavioral Sciences*, **229**, 363-375, 2016. DOI:10.1016/j.sbspro.2016.07.147
- [30] H. Srinivas, "Sustainable Development: Concepts Note series E-008. Retrieved on 20 July 2019, from <https://www.gdrc.org/sustdev/concepts.html>
- [31] H. Walker, et al., Sustainable procurement: Past, 804 present and future. *Journal of Purchasing and Supply Management*, **18**(4), 201-206. 805, 2012. <https://doi.org/10.1016/j.pursup.2012.11.003>
- [32] A.J. McMurray et al., Sustainable procurement in Malaysian organizations: Practices, barriers and opportunities. *Journal of Purchasing and Supply Management*, **20**(3), 195-207, 2014. DOI:10.1016/j.pursup.2014.02.005
- [33] H. Walker, S. Brammer, S. "Sustainable procurement in the United Kingdom public sector. *Supply Chain Management: An International Journal*, **14**(2), 128-137, 2009. <https://doi.org/10.1108/13598540910941993798>
- [34] M.A. Yahya, M. Bridge, A., Nepal, M., and Cattell, D. (2018). e-Tendering readiness in construction: the posterior model. *Construction Innovation*, **18**(2), 183-205, 2018.
- [35] B.T.H. Lim, W. Zhang, "Sustainable Procurement in Australia: Quantity Surveyors' Perception on Life Cycle Costing. *International Journal of Integrated Engineering*, **10**(2), 2018. Retrieved from <https://publisher.uthm.edu.my/ojs/index.php/ijie/article/view/2617>
- [36] A.A. Aibinu, "Using PLS-SEM technique to model construction organizations' willingness to participate in e-bidding. *Automation in construction*, **19**(6), 714-724, 2010.
- [37] J. Sarkis, Q. Zhu, "Environmental sustainability and production: taking the road less travelled. *International Journal of Production Research*, **56**(1-2), 743-759, 2017. DOI:10.1080/00207543.2017.1365182
- [38] H. Walker, S. Brammer, "The relationship between sustainable procurement and e-procurement in the public sector. *International Journal of Production Economics*, **140**(1), 256-268, 2012.
- [39] J. Elkington, "Partnerships from Cannibals with Forks: The Triple Bottom Line Of the 21st-Century Business. *Environmental Quality Management*, **8**(1), 37-51, 1998. <https://doi.org/10.1002/tqem.3310080106>
- [40] M.H. Li et al., "Adoption of electronic technologies in local US governments distinguishing between e-services and communication technologies. *The American Review of Public Administration*, **44**(1), 75-91, 2014. DOI:10.1177/0275074012460910
- [41] K.W. Deininger, D. Byerlee, "Rising global interest in farmland: Can it yield sustainable and equitable benefits? Washington, DC: World Bank Publications, 2011.
- [42] E. Prier et al., "Implementation of sustainable public procurement practices and policies: A sorting framework. *Journal of Public Procurement*, **16**(3), 312-346, 2016.
- [43] A. Varnäs, et al., "Environmental consideration in procurement of construction contracts: current practice, problems and opportunities in green procurement in the Swedish construction industry. *Journal of Cleaner Production*, **17**(13), 1214-1222, 2009.
- [44] A.V. Roman, "Public procurement specialists: They are not who we thought they were. *Journal of Public Procurement*, **15**(1), 38-65, 2015.
- [45] C.C. Hsu et al., "Supply chain drivers that foster the development of green initiatives in an emerging economy. *International Journal of Operations and Production Management*, **33** (6), 656-688, 2013.
- [46] C. Carter et al., "A framework of sustainable supply chain management: Moving toward new theory. *International Journal of Physical Distribution and Logistics Management*, **38**(5), 360-387, 2008. <https://doi.org/10.1108/09600030810882816>
- [47] J.M. Ntayi, E. Mugume, E. (2014). A taxonomy of strategic sourcing for defense forces in sub-Saharan Africa. *World Journal of Entrepreneurship, Management and Sustainable Development*, **10**(1), 13-32, 2014. doi:10.1108/wjemsd-02-2013-0019
- [48] L. Metcalf, S. Benn, "Leadership for Sustainability: An Evolution of Leadership Ability. *Journal of Business Ethics*, **112**(3), 369-384, 2012. <https://doi.org/10.1007/s10551-012-1278-6>
- [49] N. Musa et al., "Key Indicators Towards the Implementation of Green Government Procurement in Malaysia. *World Applied Sciences Journal*, **28**(13), 127-135, 2013. <https://doi.org/10.5829/idosi.wasj.2013.28.efmo.27020>
- [50] M. Mansi, "Sustainable procurement disclosure practices in central public sector enterprises: Evidence from India. *Journal of Purchasing and Supply Management*, **21**(2), 125-137, 2015. <https://doi.org/10.1016/j.pursup.2014.12.002>
- [51] S. Devaraj et al., "Impact of eBusiness technologies on operational performance: the role of production information integration in the supply chain", *Journal of Operations Management*, **25**(6), 1199-1216, 2007.
- [52] S. Mutangili, "Influence of E-Procurement Practices on Supply Chain Performance: A Case Study of Kenya Airways. *Journal of Procurement and Supply Chain*, **3**(2), 1 - 16, 2019.
- [53] A. Davila et al., "Moving Procurement Systems to the Internet: The Adoption And Use Of E-Procurement Technology Models", *European Management Journal*, **21**(1), 11-23, 2003. [https://doi.org/10.1016/S0263-2373\(02\)00155-X](https://doi.org/10.1016/S0263-2373(02)00155-X)
- [54] S. Ronchi, A. Brun, R. Golini, X. Fan, "What is the value of an IT e-procurement system?", *Journal of Purchasing and Supply Management*, **16**(2), 131-140, 2010.
- [55] R. Altounjy et al., "Moving from Bricks to Clicks: Merchants' Acceptance of the Mobile Payment in Malaysia. *International Journal of eBusiness and eGovernment Studies*, **12**(2), 136-150, 2020. DOI: 10.34111/ijebe.202012204
- [56] Renny, S. Guritno, H. Siringoringo, "Perceived Usefulness, Ease of Use, and Attitude Towards Online Shopping Usefulness Towards Online Airlines Ticket Purchase. *Procedia - Social and Behavioral Sciences*, **81**, 212-216, 2013. <https://doi.org/10.1016/j.sbspro.2013.06.415>
- [57] G. Quesada et al., "Impact of e-procurement on procurement practices and performance", *Benchmarking: An International Journal*, **17**(4), 516-538, 2010.
- [58] H.H. Chang, Y. Tsai, C. Hsu, "E-procurement and supply chain performance", *Supply Chain Management*, **18**(1), 34-51, 2013.
- [59] E.O. Ibem et al., "Survey data on e-procurement adoption in the Nigerian building industry", *Data in Brief*, **18**, 823-826, 2018.
- [60] Y.M. Tai et al., "The performance impact of implementing Web-based e-procurement systems", *International Journal of Production Research*, **48** (18), 5397-5414, 2010.
- [61] H. Hassan et al., "Factors affecting the breadth and depth of e-procurement use in small and medium enterprises", *Journal of Organizational Computing and Electronic Commerce*, **27**(4), 304-324, 2017.
- [62] J.K. Chepkwony, "E-Ordering and E-Informing on Supply Chain Performance in Retail Marketing outlets in Kenya. *Journal of Marketing and Consumer Research, An International Peer-reviewed Journal* **20**, 2016. [www.iiste.org/ISSN/2422-8451](http://www.iiste.org/ISSN/2422-8451)
- [63] A. Neupane et al., "An empirical evaluation of the potential of public e-procurement to reduce corruption. *Australas. Journal Information System* **18**(2), 21-44, 2014.
- [64] R.K. Shakya, "Procurement governance framework: Success to e-government procurement (E-GP) system implementation. In 5th International Public Procurement Conference (IPPC5), Seattle, United States ( 17-19), 2012.
- [65] J.I. Kim, D.L. Shunk, "Matching indirect procurement process with different B2B e-procurement systems. *Computers in Industry*, **53**(2), 153-164, 2004. DOI:10.1016/j.compind.2003.07.002
- [66] N.I. Shalle et al., "Effects of buyer/supplier collaboration on e-procurement performance in state corporations in Kenya. *European Journal of Management Sciences and Economics*, **1**(4), 170-185, 2014.
- [67] O. Bayazit, "Investigating the Adoption of E-Procurement Systems. *International Journal of Information Systems and Supply Chain Management*, **7**(2), 47-66, 2014. <https://doi.org/10.4018/ijisscm.2014040103>
- [68] R.L. Chilipunde, "Electronic tendering in the Malawian construction industry: the dilemmas and benefits", *Journal of Modern Education Review*, **3**, 791-800, 2013.
- [69] A.A. Costa and A. Grilo, "BIM-based E-procurement: an innovative approach to construction E-procurement", *Scientific World Journal*, 2015, 1-15.
- [70] M.H. Charki, E. Jossierand and N.B Charki, Toward an Ethical Understanding of the Controversial Technology of Online Reverse Auctions. *Journal of Business Ethics*. **98**(1), 17-37, 2010. <https://doi.org/10.1007/s10551-010-0532-z>
- [71] B.K. AlNuaimi and M. Khan, Public-sector green procurement in the United Arab Emirates: Innovation capability and commitment to change. *Journal of Cleaner Production*, **233**, 482-489, 2019. <https://doi.org/10.1016/j.jclepro.2019.06.090>
- [72] W.L. Filho, A. Skouloudis, L.L. Brandli, A.L. Salvia, L.V. Avila, and L. Rayman-Bacchus, Sustainability and procurement practices in higher education institutions: Barriers and drivers. *Journal of Cleaner Production*, **231**, 1267-1280, 2019. <https://doi.org/10.1016/j.jclepro.2019.05.202>
- [73] O. A. Ogunsanya, C.O. Aigbavboa, D.W. Thwala, and D.J. Edwards, Barriers to sustainable procurement in the Nigerian construction industry: an

- exploratory factor analysis. *International Journal of Construction Management*, 2019. <https://doi.org/10.1080/15623599.2019.1658697>
- [74] M. Ramkumar, T. Schoenherr, S.M. Wagner, and M. Jenamani, Q-TAM: A Quality Technology Acceptance Model for Predicting Organizational Buyers' Continuance Intentions for E-Procurement Services. *International Journal of Production Economics*, **216**, 333-348, 2019. <https://doi.org/10.1016/j.ijpe.2019.06.003>
- [75] S. Tiwari, C.S. Wei, and M.F. Mubarak, Sustainable procurement: a critical analysis of the research trend in supply chain management journals. *International Journal of Business Performance and Supply Chain Modelling*, 10(3), 266-282, 2019. <https://doi.org/10.1504/IJBPSM.2019.100855>
- [76] S.K. Yevu, A. Yu, The ecosystem of drivers for electronic procurement adoption for construction project procurement. *Engineering, Construction and Architectural Management*, 2019. <https://doi.org/10.1108/ECAM-03-2019-0135>
- [77] T.Y. Lam, A sustainable procurement approach for selection of construction consultants in property and facilities management. *Facilities*, 2019. <https://doi.org/10.1108/F-12-2018-0147>
- [78] T. Laosirihongthong, P. Samaranayake, S. Nagalingam, A holistic approach to supplier evaluation and order allocation towards sustainable procurement. *Benchmarking: An International Journal*, 2019. <https://doi.org/10.1108/BIJ64111-2018-0360>
- [79] J. Hair, J. Risher, M. Sarstedt, C.M. Ringle, "When to use and how to report the results of PLS-SEM", *European Business Review*, **31**(1), 2-24, 2019. <https://doi.org/10.1108/EBR-11-2018-0203>.



## Effects of Oversampling SMOTE in the Classification of Hypertensive Dataset

Nurhafifah Matondang\*, Nico Surantha

Computer Science Department, BINUS Graduate Program, Bina Nusantara University, Jakarta 11480, Indonesia

### ARTICLE INFO

Article history:

Received: 15 May, 2020

Accepted: 17 July, 2020

Online: 09 August, 2020

Keywords:

Synthetic Minority Oversampling  
Technique

Support Vector Machine

Hypertensive

### ABSTRACT

Hypertensive or high blood pressure is a medical condition that can be driven by several factors. These factors or variables are needed to build a classification model of the hypertension dataset. In the construction of classification models, class imbalance problems are often found due to oversampling. This research aims to obtain the best classification model by implementing the Support Vector Machine (SVM) method to get the optimal level of accuracy. The dataset consists of 8 features and a label with two classes: hypertensive and non-hypertensive. Overall test result performance is then compared to assess between SVM combined with SMOTE and not. The results show that SMOTE can improve the accuracy of the model for unbalanced data into 98% accuracy compared to 91% accuracy without SMOTE.

## 1. Introduction

Hypertension or high blood pressure is a disease that can possibly lead to death. Based on the report obtained from the Center of Data and Information, Indonesian Ministry of Health, hypertensive is currently still being a major health concern with a prevalence of 25,8%. On the other hand, the implementation of database technology in the health sector continues to grow rapidly. The amount of data stored in the database is increasing and requires further processing to produce valuable information and knowledge [1].

The field of science in which data can be processed into knowledge is called data mining. Data mining is a technique that includes a learning process from a machine or computer to automatically analyze and extract knowledge. Classification is one of the basic functions in data mining—a technique that can be used to predict membership of data groups. The process consists of finding a model (or function) that describes and distinguishes classes of data or concepts [2].

Support Vector Machine (SVM) is one method in classification that maps nonlinear input data to several higher dimensional spaces where data can be separated linearly, thus providing a large classification or regression performance [3].

SVM works based on the principle of Structural Risk Minimization (SRM). SRM in SVM is used to guarantee the upper

limit of generalization in the data collection by controlling the capacity (flexibility) of learning outcomes hypothesis [4]. SVM has been used extensively to classify several medical problems, such as diabetes and pre-diabetes classification [5], breast cancer [6] and a heart disease [7]. Based on previous study in liver disease dataset, SVM is known as the classifier compared to naïve bayes.

Meanwhile, problems with unbalanced data are often found due to oversampling which reduces data quality in model construction process. The imbalance of data lies in the unbalanced proportion of the number of categories between independent variables with large difference, thus the majority and minority data class are formed. This condition cause the classification model to be unequal in predicting the minority data class, even though this class still has importance as the object of modeling analysis [8]. Problems are found in the dataset used in this research, where the number of non-hypertensive classes is far greater than the number of hypertensive classes.

Unbalanced data handling needs to be done before modeling stage to develop a classification model with the highest degree of accuracy for all classes. Two techniques to tackle the issue of unbalanced data are Synthetic Minority Oversampling Technique (SMOTE) [9] and Cost Sensitive Learning (CSL) [10]. SMOTE balances the two classes by making systematic data for minority class, while CSL will take into account the impact of misclassification and provide data weighting [8]. This research will cover the performance identification of hypertensive dataset modeling that implements classification method with SMOTE and

\*Corresponding Author: Nurhafifah Matondang, Binus University, nurhafifah.matondang@binus.ac.id

without SMOTE. The main purpose of this study aims to uncover the significance of oversampling technique implementation for unbalanced data by answering the hypothesis that the combination of the SVM classification method with SMOTE can improve the accuracy of the model.

## 2. Synthetic Minority Oversampling Technique (SMOTE)

The problem of data imbalance occurs due to a large difference between the number of instances belonging to each data class. Data classes having comparatively more objects are called major classes, while others are called minor classes [11]. The use of unbalanced data in modeling affects the performance of the models obtained. Processing algorithms that ignore data imbalances will tend to be focus too much on major classes and not enough to review minor classes [9]. The Synthetic Minority Oversampling Technique (SMOTE) method is one of the solutions in handling unbalanced data with another different principle from oversampling method that has been previously proposed. Oversampling method focuses on increase random observations, while the SMOTE method increases the amount of minor class data and make it equivalent to the major class by generating new artificial instances [12].

There are many challenges in dealing with issues of data that are out of balance with the oversampling technique. These problems are related to the addition of random data which can cause overfitting [13]. The SMOTE method is one of the oversampling technique solutions which has the advantage of being successfully applied to various domains as shown in algorithm 1 [14].

---

### Algorithm 1 SMOTE Algorithm

---

1. **Function** SMOTE( $T, N, k$ )  
**Input :**  $T; N; k$  ->#minority class examples, Amount of oversampling, #nearest neighbors  
**Output:**  $(N/100) * T$  synthetic minority class samples  
**Variables:** *Sample* [] []: array for original minority class samples;  
*newindex*: keeps a count of number of synthetic samples generated, initialized to 0; *Synthetic*[][]: array for synthetic samples
  2. **if**  $N < 100$  **then**
  3.     Randomize the  $T$  minority class samples
  4.      $T = (N/100) * T$
  5.      $N = 100$
  6. **end if**
  7.  $N = (\text{int})N/100$  . -> The amount of SMOTE is assumed to be in integral multiples of 100.
  8. **for**  $i = 1$  to  $T$  **do**
  9.     Compute  $k$  nearest neighbors for  $i$ , and save the indices in the *nnarray*
  10.    POPULATE( $N, i, \text{nnarray}$ )
  11. **end for**
  12. **end function**
- 

Artificial data or synthesis is made based on k-NN algorithm (k-nearest neighbor). The number of k-nearest neighbors is determined by considering the distance between data points of all features. The process of generating artificial data for the numerical data is different from the categorical data. Numerical data are

measured by their proximity to Euclidean distance while categorical data are generated based on mode value—the value that appears most often [12]. Calculation of the distance between classes with categorical scale variables is done by the Value Difference Metric (VDM) formula, as follows:

$$\Delta(X, Y) = w_x w_y \sum_{i=1}^N \delta(x_i, y_i)^r \quad (1)$$

where

- $\Delta(X, Y)$  : the distance between observations X and Y
- $w_x w_y$  : observe weight (negligible)
- N : number of explanatory variables
- R : worth 1 (Manhattan distance) or 2 (Euclidean distance)

$\delta(x_i, y_i)^r$ : distance between categories, with the formula:

$$\delta(v_1 v_2) = \sum_{i=1}^n \left| \begin{matrix} C_{1i} & C_{2i} \\ C_1 & C_2 \end{matrix} \right| k \quad (2)$$

where

- $\delta(v_1 v_2)$  : distance between  $V_1$  and  $V_2$
- $C_{1i}$  : number of  $V_1$  that belongs to class  $i$
- $C_{2i}$  : number of  $V_2$  included in class  $i$
- I : number of classes ;  $i = 1, 2, \dots, m$
- $C_1$  : number of values of 1 occurs
- $C_2$  : number of values of 2 occurs
- N : number of categories
- K : constants (usually 1)

## 3. Proposed Research Stages

This research applies a quantitative approach for a case study of hypertensive. Overall, the steps involved consisted of three parts: (1) data pre-processing, (2) building the model and (3) evaluating model performance. The methods used are SVM, ELM, over sampling and under sampling. The performance models are compared with each other. The following are the stages of completing the methodology to be completed.

## 4. Hypertension Dataset

This research is carried out using a hypertension dataset published in the Kaggle repository sourced from National Health and Nutrition Examination Survey (NHANES) for the year 2008-2016 (<https://www.kaggle.com/felmco/nhanes-hypertensive-population-20082016/kernels>) with 24435 data rows. A total of 9 features contained in the dataset, as follows:

- a. SEQN,
- b. Gender (Male - 1 and Female - 2),
- c. Age range (from 20 years to 80 years, divided into category 1 to 6),
- d. Race (consists of 1-Mexican American, 2-Other Hispanic, 3-Non-Hispanic White, 4-Non-Hispanic Black, and 5-Other Race),
- e. BMI Range (average body mass index, starting at less than 18,5 kg defined as underweight, normal weight (between 18,5 to 24,9 kg) and overweight (more than 30 kg),

- f. Kidney health condition,
  - g. Smoking cigarette or non-smoking,
  - h. Diabetes (diabetes, non-diabetes and borderline), and
  - i. HPCLASS which has been labeled as class 1 (hypertensive) and class 0 (non-hypertensive).
- Samples of all the features in the dataset are shown in Table 1. The dataset will be pre-processed before the modeling.

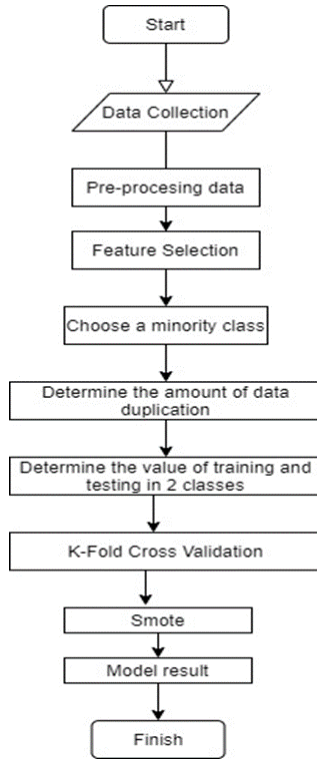


Figure 1. Research Methodology

Table 1. Sample rows from Hypertensive dataset

Seqn	Gen der	Age Range	Ra ce	Bmi Range	Kid ney	Smoke	Dia betes	Hyp class
41475	2	5	5	4	2	2	2	0
41485	2	1	2	3	2	2	2	0
41494	1	2	1	3	2	1	2	0
41516	2	6	3	4	2	1	1	0
41526	2	5	3	3	2	1	2	0
41542	2	4	4	4	2	2	1	0
41550	1	3	1	2	2	1	2	0
41576	1	5	4	4	2	1	1	0
41595	2	6	3	1	2	1	2	0
41612	2	6	3	3	2	2	2	0
41625	2	1	1	4	2	2	2	0
41639	1	2	1	4	2	2	2	0
41674	1	1	4	2	2	2	2	0
41687	1	2	1	2	2	2	2	0
41697	2	1	2	4	2	2	2	0
41711	2	6	3	3	2	2	2	0
41729	1	4	3	4	2	1	2	0

### 5. Feature Selection

In the pre-processing stage, the selection or extraction of all features of the data is carried out to get the most influential features improve the performance or accuracy of the classification model. Originally, the hypertensive dataset contains of 9 features of the hypertension dataset. The selection implemented by removing a feature "SEQN" in the first column of the dataset which has no effect and only displays the order. The label or class for this dataset is presented in "HYPCLASS" variable.

In the problem of feature selection we wish to minimize equation [15] over  $\sigma$  and  $\alpha$ . The support vector method attempts to find the function from the set  $f(x, w, b) = w \cdot \phi(x) + b$  that minimizes generalization error. We first enlarge the set of functions considered by the algorithm to  $f(x, w, b, \sigma) = w \cdot \phi(x * \sigma) + b$ . Note that the mapping  $\phi \sigma(x) = (x * \sigma)$  can be represented by choosing the kernel function K in equations. [16]:

$$K_{\sigma}(x, y) = K((x * \sigma), (y * \sigma)) = (\phi_{\sigma}(x) \cdot \phi_{\sigma}(y)) \quad (3)$$

for any K. Thus for these kernels the bounds in Theorems still hold. Hence, to minimize  $T(\sigma, \alpha)$  over  $\alpha$  and  $\sigma$  we minimize the wrapper functional  $T_{wrap}$  in equation where  $T_{alg}$  is given by the equations choosing a fixed value of  $\sigma$  implemented by the kernel. Using equation one minimizes over  $\sigma$ :

$$R^2 W^2(\sigma) = (R^2(\sigma) W^2(\sigma^0, \sigma)) \quad (4)$$

where the radius R for kernel  $K_u$  can be computed by maximizing :

$$R^2(\sigma) = \frac{\max}{\beta} \sum_i \beta_i K_{\sigma}(X_i, X_i) - \sum_{i,j} \beta_i \beta_j K_{\sigma}(X_i, X_j) \quad (5)$$

subject  $\sum_i \beta_i = 1, \beta_i \geq 0, i = 1, \dots, \ell$ , and  $W^2(\alpha^0, \sigma)$  is defined by the maximum of functional using kernel. In a similar way, one can minimize the *span* bound over  $\sigma$  instead of equation.

Finding the minimum of  $R^2 W^2$  over  $\sigma$  requires searching over all possible subsets of  $n$  features which is a combinatorial problem. To avoid this problem classical methods of search include greedily adding or removing features (forward or backward selection) and hill climbing. All of these methods are expensive to compute if  $n$  is large.

As an alternative to these approaches we suggest the following method: approximate the binary valued vector  $\sigma \in \{0,1\}^n$  with a real valued vector  $\sigma \in \mathbb{R}^n$ . Then, to find the optimum value of  $\sigma$  one can minimize  $R^2 W^2$ , or some other differentiable criterion, by gradient descent. As explained in the derivative of our criterion is:

$$\begin{aligned} \frac{\partial R^2 W^2(\sigma)}{\partial \sigma_k} &= R^2(\sigma) \frac{\partial R^2 W^2(\alpha^0, \sigma)}{\partial \sigma_k} + W^2(\alpha^0, \sigma) \frac{\partial R^2(\sigma)}{\partial \sigma_k} \\ \frac{\partial R^2(\sigma)}{\partial \sigma_k} &= \sum_i \beta_i^0 \frac{\partial K_{\sigma}(X_i, X_j)}{\partial \sigma_k} \sum_{i,j} \beta_i^0 \beta_j^0 Y_i Y_j \frac{\partial K_{\sigma}(X_i, X_j)}{\partial \sigma_k} \\ \frac{\partial W^2(\alpha^0, \sigma)}{\partial \sigma_k} &= - \sum_{i,j} \alpha_i^0 \alpha_j^0 Y_i Y_j \frac{\partial K_{\sigma}(X_i, X_j)}{\partial \sigma_k} \end{aligned} \quad (6)$$

We estimate the minimum of  $\tau(\sigma, \alpha)$  by minimizing equation in the space  $\sigma \in \mathbb{R}^n$  using the gradients with the following extra constraint which approximates integer programming.

$$R^2W^2(\sigma) + \lambda \sum_i (\sigma_i)^p \tag{7}$$

subject to  $\sum_i \sigma_i = m, \sigma_i \geq 0, i = 1, \dots, \ell$ .

For large enough  $\lambda$ , as  $p \rightarrow 0$  only  $m$  elements of  $\sigma$  will be nonzero, approximating optimization problem  $\mathcal{T}(\sigma, \alpha)$ . One can further simplify computations by considering a stepwise approximation procedure to find  $m$  features. To do this one can minimize  $R^2W^2(\sigma)$  with  $\sigma$  unconstrained. One then sets the  $q \ll n$  smallest values of  $0$  to zero, and repeats the minimization until only  $m$  nonzero elements of  $\sigma$  remain. This can mean repeatedly training a SVM just a few times, which can be fast.

### 6. Results

Hypertensive data are classified into hypertensive (1) and non-hypertensive (0) classes. 80% of available data is allocated for training set and remaining 20% for the test set. Data validation is done by the split validation method by dividing three times the test data and three times the training data. Detailed description will be explained in the following subsection.

#### 6.1 Data Visualization

Data visualization is a technique used to communicate data or information in the form of visual objects. In this section, visualization data done by using Python will be displayed in the form of bar graphs. Figure 2 shows the original results of two classifications in hypertension dataset—consists of 7399 hypertensive classes and 17035 non-hypertensive classes. Based on the graph, it can be concluded that the hypertensive class is having comparatively fewer objects than non-hypertensive class.

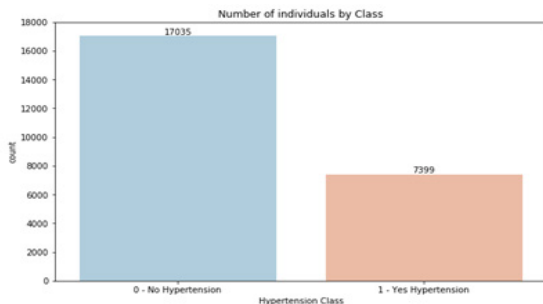


Figure.2. Proportion between Hypertensive and Non-Hypertensive

The following graph in Figure 3 shows that the sample of hypertensive patients has greater number in male gender, otherwise non-hypertensive data are dominated with female.

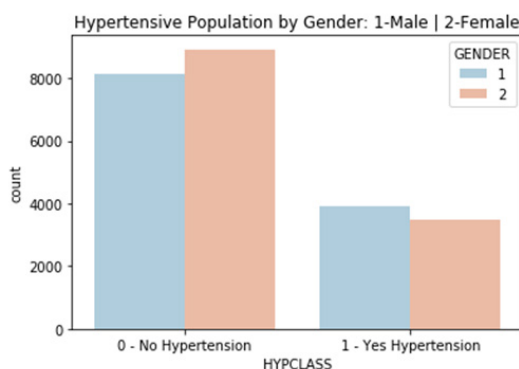


Figure 3: Population by gender hypertensive Gender

below shows that the samples of hypertensive and non-hypertensive patients are mostly dominated by the race number 3 (Non-Hispanic White).

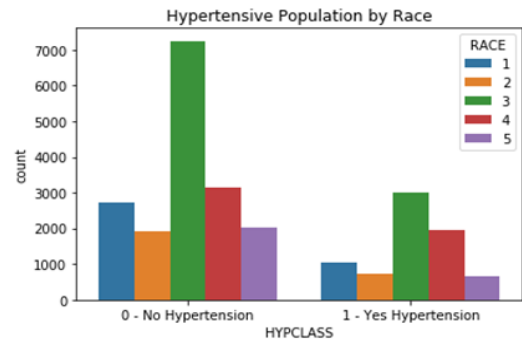


Figure 4. Population by Race

#### SVM Model and K Fold Cross Validation

The pre-processed data is then used in building the model with the Support Vector Machine (SVM) method. Validation on the SVM classifier model uses the model that has been built with 2 classes of hypertensive and non-hypertensive, in which the value of  $k = 10$  for the K-Fold Cross Validation method. The validation results correspond to the optimal accuracy based on the K-Fold method. Cross Validation rises slightly from initial experiment, where the SVM classifier model has the highest accuracy in the 5th iteration with the highest accuracy value of 95% (average value = 90.2%). The assessment method using the SVM classification method uses the 10-Fold Cross Validation method as presented in Table 1.

Table 2. K Fold Cross Validation with SVM

K-Fold											Iteration to	Accuracy Value %
	1	2	3	4	5	6	7	8	9	10		
1	1	2	3	4	5	6	7	8	9	10	1	93
2	1	2	3	4	5	6	7	8	9	10	2	90
3	1	2	3	4	5	6	7	8	9	10	3	89
4	1	2	3	4	5	6	7	8	9	10	4	89
5	1	2	3	4	5	6	7	8	9	10	5	95
6	1	2	3	4	5	6	7	8	9	10	6	91
7	1	2	3	4	5	6	7	8	9	10	7	93
8	1	2	3	4	5	6	7	8	9	10	8	87
9	1	2	3	4	5	6	7	8	9	10	9	85
10	1	2	3	4	5	6	7	8	9	10	10	90
Average											90,2	

Based on the Figure 3 in the previous sub-section, imbalance data are found in hypertensive class (minor class) which is having comparatively fewer objects than non-hypertensive class (major class). In this experiment, oversampling will be carried out on the minority with the Synthetic Minority Over-sampling Technic (SMOTE) method which is a popular method applied in order to deal with class imbalance. [9]

This technique synthesizes new samples generated from minor class to balance the dataset. New instances of the minor class



obtained by forming convex combinations from neighboring instances. Through the number of  $n\_sample = 12000$ ,  $n\_features = 2$ ,  $n\_split = 7$  and  $n\_repeats = 4$ , the accuracy obtained from SVM model is increased to 98%. Figure 5 shows the comparison the value of class 0 (non-hypertensive) and class 1 (hypertensive) after implementing SMOTE method.

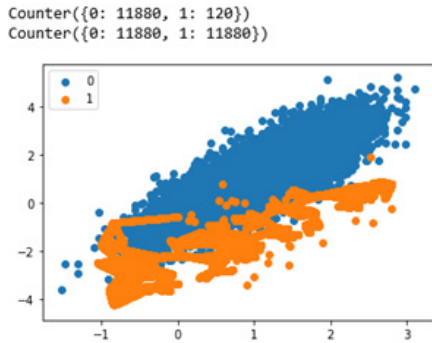


Figure.5. Hypertensive dataset experiment with SMOTE

The following is how the SMOTE algorithm works

**Step 1:** Setting the minority class set A, for each  $\chi \in A$ , the k-nearest neighbors of x are obtained by calculating the Euclidean distance between x and every other sample in set A.

**Step 2:** The sample rate N is set according to the imbalanced proportion. For each  $\chi \in A$ , N examples (i.e  $x_1, x_2, \dots, x_n$ ) are randomly selected from its k-nearest neighbors, and they construct the set  $A_1$ .

**Step 3:** For each example  $\chi_k \in A_1$  ( $k=1,2,3,\dots,N$ ), the following formula is used to generate a new example:

$$\chi'' = \chi + rand(0,1) * |\chi - X_k$$

6.4 Performance of SMOTE and Non-SMOTE Classification Results

Confusion Matrix is used to measure the performance of SVM with SMOTE and SVM performance without SMOTE to classify hypertensive. Then the calculation of accuracy, precision, and recall values is done by calculating the average value of accuracy, precision and recall in each class as shown in Table 3.

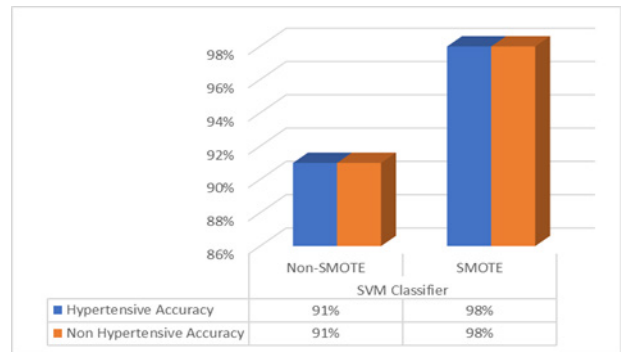
Table 3. Results of Comparison of Precision and Recall Values

Class	Non-SMOTE		SMOTE	
	SVM Classifier		SVM Classifier	
	Precision	Recall	Precision	Recall
Hypertensive	0.88	0.88	0.93	0.85
Non Hypertensive	0.94	0.94	0.88	0.94

Table 4. Average Accuracy Value

Class	Matrix	SVM Classifier	
		Non-SMOTE	SMOTE
Hypertensive	Accuracy	91%	98%
Non Hypertensive	Accuracy	91%	98%

From these results, it can be analyzed the effect of SMOTE on the performance of the SVM classification algorithm. The graphic representation of classification model performance result for hypertensive data using SVM classifier with 10-Fold Cross Validation and SMOTE and thus without SMOTE is presented in Figure 6.



The results using the combination of SVM and SMOTE, outperformed the SVM classification without SMOTE. The average accuracy based on SVM classifier with SMOTE is higher (98%) compared to SVM classifier without SMOTE (91%).

7. Conclusion

The SVM classification method with a K-Fold Cross Validation resulted on the average of 90.2% of accuracy. SVM is known as a classification method with good prediction results. Based in the experiment conducted in this research, the resulting model is increasing to optimal after running imbalanced dataset using SMOTE with the average of 98% accuracy results.

Acknowledgment

The research publication is fully supported by Bina Nusantara University.

References

- [1] M. A. Muslim, S. H. Rukmana, E. Sugiharti, B. Prasetyo, and S. Alimah, "Optimization of C4.5 algorithm-based particle swarm optimization for breast cancer diagnosis," *J. Phys. Conf. Ser.*, Figure: 983, (Snik, p. 012063, Mar. 2016. doi :10.1088/1742-6596/983/1/012063.
- [2] J. P. Jiawei Han, Micheline Kamber, *Data mining: Data mining concepts and techniques*, 3rd ed. 2012.
- [3] S. Vijayarani, S. Dhayanand, A. Professor, and M. P. Research Scholar, "Kidney Disease Prediction Using Svm and Ann Algorithms," *Int. J. Comput. Bus. Res.* 6(2), 2229–6166, 2015.
- [4] R. D. F. Adhitia, "Penilaian Esai Jawaban Bahasa Indonesia Menggunakan Metode SVM - LSA Dengan Fitur Generik," 5(1) 2009. doi:10.21609/jsi.v5i1.260.
- [5] W. Yu, T. Liu, R. Valdez, M. Gwinn, and M. J. Khoury, "Application of support vector machine modeling for prediction of common diseases: The case of diabetes and pre-diabetes," *BMC Med. Inform. Decis. Mak.*, 2010. DOI: 10.1186/1472-6947-10-16.
- [6] I. Maglogiannis, E. Zafiroopoulos, and I. Anagnostopoulos, "An intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers," *Appl. Intell.*, 1, 24–36, 2009. DOI: 10.1007/s10489-007-0073-z.
- [7] M. T., D. Mukherji, N. Padalia, and A. Naidu, "A Heart Disease Prediction Model using SVM-Decision Trees-Logistic Regression (SDL)," *Int. J. Comput. Appl.*, 16, 11–15, 2013. doi: 10.5120/11662-7250.
- [8] D. Adiangga, "Perbandingan Multivariate Adaptive Regression Spline (MARS) dan Pohon Klasifikasi C5.0 pada Data Tidak Seimbang (Studi Kasus: Pekerja Anak di Jakarta)," 2015.
- [9] N. V. Chawla, K. W. Bowyer, and L. O. Hall, "SMOTE: Synthetic Minority

Over-sampling Technique,” *Ecol. Appl.*, **30**(2), 321–357, 2002, DOI: <https://doi.org/10.1613/jair.953>.

- [10] C. Elkan, “The Foundations of Cost-Sensitive Learning,” 2001.
- [11] I. D. Rosi Azmatul, “Penerapan Synthetic Minority Oversampling Technique (Smote) Terhadap Data Tidak Seimbang Pada Pembuatan Model Komposisi Jamu,” *Xplore J. Stat.*, 1, 2013, doi:10.29244/xplore.v1i1.12424.
- [12] S. Cost and S. Salzberg, “A weighted nearest neighbor algorithm for learning with symbolic features,” *Mach. Learn.*, **10**(1), 57–78, 1993, doi:10.1023/A:1022664626993.
- [13] M. Galar, A. Fern, E. Barrenechea, and H. Bustince, “Hybrid-Based Approaches,” 2012. DOI: 10.1109/TSMCC.2011.2161285.
- [14] A. Fernandez, S. Garcia, F. Herrera, and N.V. Chawla, “SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary,” *J. Artif. Intell. Res.*, **61**, 863–905, 2018.
- [15] U. Alon *et al.*, “Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays,” *Proc. Natl. Acad. Sci. U. S. A.*, **96**(12), 6745–6750, 1999, DOI: 10.1073/pnas.96.12.6745.
- [16] J. Weston, S. Mukherjee, O. Chapelle, M. Pontil, T. Poggio, and V. Vapnik, “Feature selection for SVMs,” *Adv. Neural Inf. Process. Syst.*, 2001.

## A CNN-based Differential Image Processing Approach for Rainfall Classification

Roberta Avanzato, Francesco Beritelli\*

University of Catania, Department of Electrical, Electronic and Computer Engineering, Catania, 95125, Italy

---

### ARTICLE INFO

Article history:

Received: 15 May, 2020

Accepted: 17 July, 2020

Online: 09 August, 2020

---

Keywords:

Smart sensors

Video/Image classification

Rainfall estimation

Image processing

Convolutional Neural Networks  
(CNN)

---

### ABSTRACT

*With the aim of preventing hydro-geological risks and overcoming the problems of current rain gauges, this paper proposes a low-complexity and cost-effective video rain gauge. In particular, in this paper the authors propose a new approach to rainfall classification based on image processing and video matching process employing convolutional neural networks (CNN). The system consists of a plastic shaker, a video camera and a low cost, low power signal processing unit. The use of differential images allows for greater robustness, guaranteeing full background subtraction. As regards precision, speed and ability to adapt to variations in precipitation intensity, the proposed method achieves good performance. In particular, the results obtained from seven classes, ranging from "No rain" to "Cloudburst", applying the Discrete Cosine Transform (DCT) to the differential images on 16x16 sub-blocks show an average accuracy of 75% considering, also, the adjacent misclassification. Furthermore, the analysis of precision and sensitivity parameters yields excellent results.*

*The proposed method is very innovative, in fact, the few studies found in the state of the art use only two classification classes (No rain and Rain), while our method contains seven classification classes and overall delivers very good accuracy performance.*

---

### 1. Introduction

In recent years, in many areas of the world, countless calamitous events of hydro-geological origin have occurred. Since these events are difficult to prevent, it is important to be able to accurately estimate the level of rainfall during the occurrence of a meteorological phenomenon.

In the past, the only method for measuring rainfall levels was through the use of traditional rain gauges (such as tilt rain gauges). People have realized, however, that these instruments do not perform well in terms of temporal precision, classification accuracy and territorial coverage. Over the years, different approaches have been studied to tackle this issue, such as the use of radars and satellites [1]-[3].

Weather radars have the advantage of being able to monitor a larger area, but they are very expensive, whereas the satellite has the advantage of better spatial resolution but the disadvantage of providing a less accurate estimate [4]. Recently, due to these limitations, some researchers have studied alternative methods for estimating precipitation levels [5]-[16]. Lately, with the spread of artificial intelligence techniques, new approaches based on advanced neural networks have been introduced in the literature.

In [6]-[7], the authors investigated the use of the radio frequency signal, used in the latest generation cellular systems, as a tool for classifying rainfall intensity levels using pattern recognition method. However, in [17] an acoustic rain gauge has been proposed which is able to classify the rainfall levels in 5 classes through rainfall timbre and deep learning techniques, in particular by applying convolutional neural networks [18], [19].

In order to ensure good performance even in conditions of high environmental noise, in this paper, the authors propose a new type of rain gauge based on the analysis of images obtained from a camera that captures the rain.

In this study, the images of the rain were recorded by inserting a camera pointed towards a transparent lid of an upward-facing plastic shaker. The video signal was also analysed by conducting statistical analysis (mean, variance and first autocorrelation coefficient) of some features, while the Discrete Cosine Transform (DCT) was used to characterize the frequency content of the various frames extracted from the video.

Thanks to the low complexity of the hardware and software, the proposed system allows to generate alerts in very short times with low errors in the estimate of the level of precipitation intensity. Unlike traditional tipping rain gauges, the proposed system does

---

\*Corresponding Author: Francesco Beritelli, francesco.beritelli@unict.it

not use mechanical components subject to maintenance, but low cost and consumption electronics, with the possibility of being powered by a photovoltaic panel.

The paper is organized as follows: section 2 briefly summarizes the main studies on the classification of precipitation, section 3 illustrates the characteristics of the database; section 4 describes the acquisition system; section 5 shows the statistical and spectral analysis of the frame of images relating to the different levels of rain; section 6 describes the convolutional neural network used in this study and the dataset in the input; section 7 describes the test bed and the various tests and performance results; section 8 is devoted to conclusions.

## 2. Proposed Work

This section provides a general overview of new rainfall level estimation methods by analyzing videos recorded over the rain. The intensity of rain, in fact, can undergo tremendous variations even for short distances, less than one kilometer. It is therefore necessary to implement a prevention system especially in areas with high hydrological risk. To obtain measurements with large spatial resolution, it is necessary to distribute a substantial number of instruments such as rain gauges or weather stations, which may, nonetheless, have a major adverse impact in economic terms.

A number of studies on the classification of rainfall levels have been conducted by researchers employing various alternative methods and signals [5], [16].

These different approaches could represent a valid alternative to the typical solutions used in current meteorology. Furthermore, these innovative systems lead to improved performance in terms of time, flexibility and robustness.

In particular, there are numerous papers in which the possibility of classifying the level of rainfall through the attenuation of the electromagnetic signal [5], [6] and many others using the audio signal [6]-[11] is studied.

In this paper, the authors will focus on the use of the video signal as a tool for classifying rainfall levels and therefore this section deal with previously conducted studies in this regard.

In [12] the Electronics and Telecommunications Department of the University of Mumbai developed an algorithm that allows determining the intensity of rainfall based on a single photo only. Data processing takes place implementing a technique denominated ERAUIP, using a high-resolution slow-motion camera, which acquires the image placed on a black background with a minimum speed of 1200 FPS. The camera acquires a color image which, once filtered by noise employing a clustering technique, is converted to gray scale. Subsequently, the image is divided into rows and columns, and therefore further divided into even and odd content. After which the conversion from the gray scale image to the binary image takes place and, finally, the size of the raindrops are measured thanks to the use of a morphological filter.

This technique allows to reach high precision levels, even up to 90%. The performance will depend on the characteristics of the camera used; a full-bodied database of excellent quality images and videos is required in order to achieve better results.

A method for evaluating the raindrops present in images has been proposed in [13]. In this case image processing is performed using the 2D dual density Discrete Wavelet Transform, with the help of 2 cameras with advanced features, so that the 2 captured images have the same central point. This allows obtaining such information as the maximum and minimum size of the raindrop and the number of raindrops from the given image.

In the automotive sector, more and more sophisticated driver assistance systems are being developed to improve safety conditions behind the wheel, especially during adverse rain conditions. An approach used to detect raindrops present on a transparent screen has been developed in the context of research concerning the ITS intelligent transport algorithm, as reported in [14]. This technique detects the fuzzy raindrops that collect on the windshield, differentiating from the previously illustrated techniques as it is not necessary to adjust the focus on the windshield on which the raindrops fall.

Classification of rain levels through the acquisition of images can also be achieved by exploiting the learning abilities of a neural network [15]. In this case, a neural network was trained in order to obtain the detection of the glomerulus in the segments of the renal tissue present in digitized slides, and a classification considering such parameters as heterogeneity, staining, composition within the kidney composition etc. This is done by applying convolutional neural networks (CNN) in order to provide faster and more reliable diagnoses. Nevertheless, enormous amount of computational power is needed to train the network in an appropriate time frame.

The previously mentioned studies outline significant growth in interest concerning the replacement of the classic measurement methods with new instruments which can deliver better performance, guaranteeing a set of information that can cover a large area of the territory.

In [16], the authors propose a method for remote detection of rain via fixed cameras, using CNN. In particular, in the pre-processing phase the Sobel algorithm is used for detecting the edges and subsequently DCT is applied to the obtained images for further experiments in order to improve the classification scores. The classification results using the Sobel algorithm and the DCT are respectively 91.97% and 79.89%. There are only 2 classes considered: "Rain" and "No Rain".

Unlike all the studies present in the state of the art, our method consists in detecting and classifying, as accurately as possible, seven different levels of rainfall intensity. The proposed technique, therefore, allows to set up a rainfall measurement system with the advantage of not having mechanical parts subject to breakdown, good accuracy and high temporal resolution.

## 3. Database Description

The proposed system consists of a camera connected to a processing unit. The labelling algorithm is located in the processing unit. It is responsible for assigning the video sequences to the corresponding intensity classes thanks to the labelling performed by a traditional tipping rain gauge. After labelling, the video sequences are inserted into the database. Figure 1 shows the block scheme of the video acquisition system.



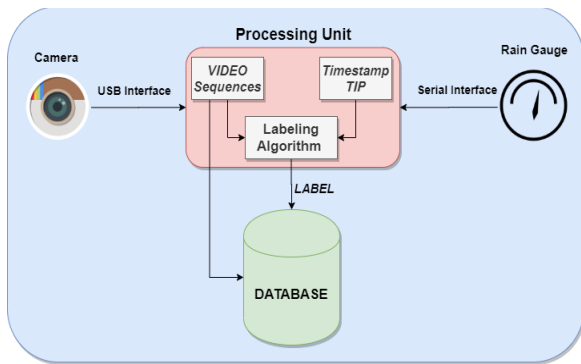


Figure 1: Block Scheme of Video Acquisition System

Table 1: Classification and Rainfall Intensity Range

Classification	Acronym	Rainfall Intensity [mm/h]
No rain	nr	< 0.5
Weak	w	[0.5 ÷ 2]
Moderate	m	[2 ÷ 6]
Heavy	h	[6 ÷ 10]
Very heavy	vh	[10 ÷ 18]
Shower	s	[18 ÷ 30]
Cloudburst	c	> 30

Video sequences are sampled with Frames Per Second (FPS) equal to 30 frames of 640×480 pixels. Table 1 lists the intensity classes that make up the dataset.

The database was created by recording natural rain characterized by all seven rainfall levels indicated in Table 1. Furthermore, during the rain recording phase, continuous checks were carried out on the tipping rain gauge used for labeling, so as to preserve it from dirt.

An image processing phase, before inserting the input data to the neural network, was applied to the video recordings: extraction of the frames with frame-rate equal to 30 FPS and offset equal to 1 frame; subtraction of the resulting frames from each other. These frames will be referred to as "differential images".

Differential images are subjected to DCT on 16×16 sub-blocks. The data obtained from the application of the DCT are standardized and input data to the neural network, described in section 6.2.

The tests were conducted using a camera inserted in a rigid plastic shaker with a transparent lid. The differential frames, relating to the various categories of precipitation, are used as input variables of a deep learning system based on a CNN classifier.

#### 4. Acquisition System Setup

The acquisition system is characterized by the following components, see Figure 2.

- Camera (a);
- Plastic shaker with a transparent cover (b);
- Tipping bucket rain gauge (c);
- Raspberry Pi (d);
- 4G dongle (e).

[www.astesj.com](http://www.astesj.com)

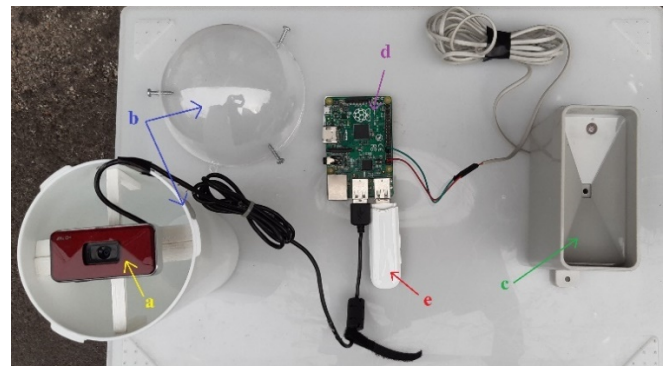


Figure 2: Hardware Components of the Acquisition System

The webcam is connected via USB cable to the Raspberry Pi, where the processing phase of the collected data is performed.

The operation of the labeling algorithm implemented within the Raspberry Pi, the connections between the hardware devices in use and the categories of recording intensity are explained in detail in [17]. The 4G modem key allows real-time data transmission [20]. Figure 3 shows the conceptual scheme of the proposed system.

In general, the processing unit implements the labeling algorithm that allows the generation of 30-second video sequences, divided by each class of rainfall intensity.

The video sequences recorded and classified using the labeling algorithm are stored in the database and subsequently provided as input to the neural network for the training phase.

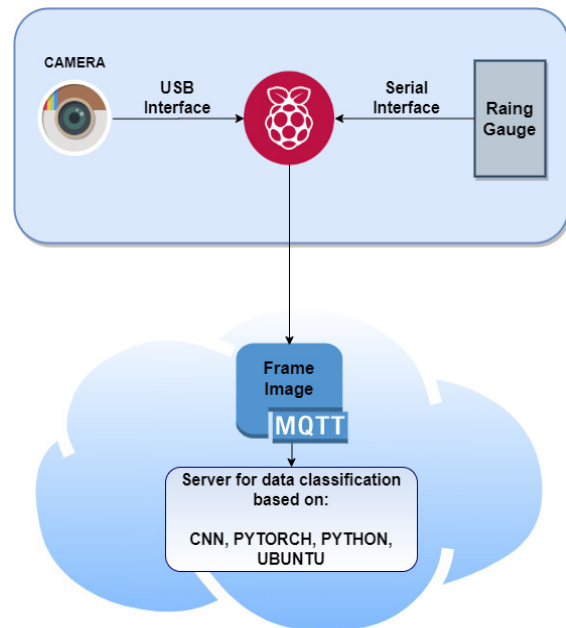


Figure 3: Conceptual Scheme of the Proposed System

#### 5. Analysis of Video Sequences

The present section deals with visual and statistical analysis of the obtained video sequences in order to further discriminate between different precipitation intensity levels.

The video signal analysis tool, DiffImg, and the calculation tool, Matlab, are used to externally analyse the differential images extracted from the videos.

In particular, DiffImg is a simple image comparison tool which takes two RGB images with the same size as the input. Some statistics are computed, and the positions at which pixels differ.

With this tool, we are able to make a comparison between different images, highlighting the differences. By analyzing the “differential images”, the authors can also view some interesting statistics, including:

- Average error: indicates the difference between the two images in terms of RGB scale value;
- Standard deviation: indicates how much the second image varies compared to the first;
- Total of error pixels: indicates the total amount of different pixels between the first image and the second image (from which the “differential image” is obtained).

After a series of evaluations, it was observed that the most interesting feature to achieve our goal is the standard deviation. To better evaluate the values assumed by this parameter, it was decided to take into consideration 7 frames extracted from the videos, for each level of precipitation, with an interval equal to 1 second.

Subsequently, the value assumed by the standard deviation of the differential image obtained from the differences between consecutive frames was evaluated, so as to obtain a total of six “differential images”. By evaluating the trend of the parameter just mentioned, in particular the average of the standard deviation values, it was observed that this value increases with the increase in the rain level, as we can see in Table 2.

As can be seen from Table 2, there is a slight correlation between the standard deviation and the intensity of precipitation. This is immediately clear if you think that as the rate of rain increases, the number of raindrops and their size increases, so the difference between two frames belonging to higher levels will be more marked than the one between the images belonging to lower intensity levels.

To illustrate the concept discussed so far, a differential frame (taken from 9 “differential image”) for each rainfall level is represented in Figure 4.

Table 2: Standard Deviation for Each Precipitation Intensity

Rain Classification	Standard Deviation							AVG
	2.0	2.0	2.0	2.1	2.1	2.0	2.0	
No rain	2.0	2.0	2.0	2.1	2.1	2.0	2.0	
Weak	3.6	3.3	2.8	2.4	2.7	2.5	3.0	
Moderate	3.3	3.3	4.0	5.7	4.5	5.3	4.4	
Heavy	3.4	8.6	5.2	5.3	7.8	5.7	5.8	
Very heavy	6.2	5.1	4.1	8.4	6	5.4	6.3	
Shower	4.0	8.3	9.6	10.0	8.2	9.3	8.1	
Cloudburst	17.3	7.4	10.7	15.4	8.5	11.4	11.9	

From Figure 4 it is possible to note that there is a visible difference between the adjacent precipitation levels from the point of view of the raindrops represented in the frame. For example, it

is clear that if in “No rain” class the image is white, in “Weak” the presence of a few raindrops is observed. The same concept may be applied to subsequent levels.

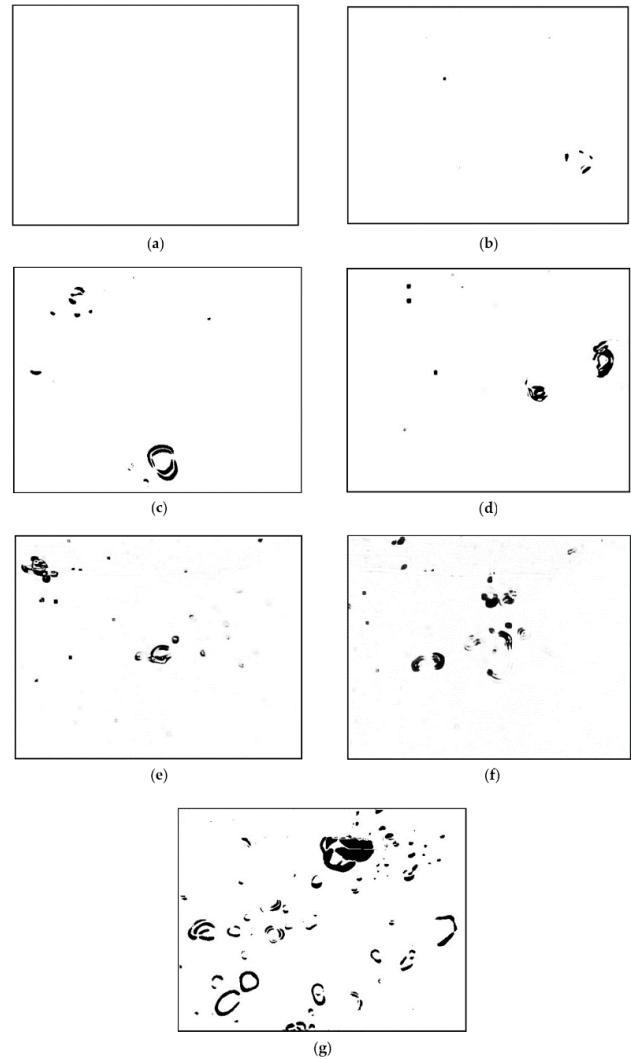


Figure 4: Examples of “Differential Image” for Each Level of Intensity

Despite some positive results obtained with the methodologies described above, some limits remain which do not allow the implementation of an autonomous system capable of achieving a reliable classification.

The different problems encountered during the experimental phase could be solved by creating a system based on machine learning by carrying out the training of a neural network through the use of more comprehensive and differentiated databases.

## 6. Convolutional Neural Network for Rainfall Classification

The paper proposes an alternative to traditional rain gauges which is based on the extraction of frames from video sequences. Rainfall levels are classified by applying a convolutional neural network.

### 6.1. CNN Architecture: SqueezeNet

In this study, the authors use a Convolutional Neural Network (CNN) called SqueezeNet [21]. It is a completely convolutional

network with reduced complexity and with layers of dropouts. This network allows to obtain good accuracy and increase performance.

To create a CNN network with reduced complexity, three main strategies are necessary:

1. Replace 3×3 filters with 1x1 filters;
2. Reduce the number of input channels on 3×3 filters;
3. Delay downsample on the network so that convolution levels have large activation maps.

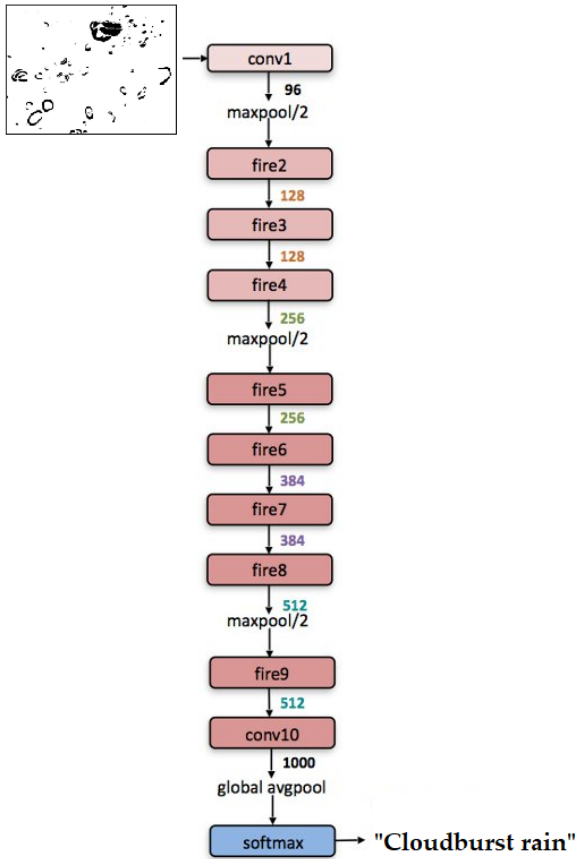


Figure 5: SqueezeNet Architecture

Figure 5 shows the SqueezeNet architecture used. The first layer and the last layer are convolutional layers, in the middle there are eight fire modules where the number of filters per fire module gradually increases. For more details see [21-22].

### 6.2. Video Learning Dataset

As described in section 3, the learning dataset was created recording the video of the rain. Figure 6 shows the general "Best practices" for inserting standardized matrices into the neural networks. The first block represents the original video, on which the extraction of the frames takes place at a frame-rate of 30 FPS.

Then, the extracted frames are brought in grayscale and subsequently, differentiated by moving between one frame and the next with an offset of 1 frame.

The difference between two images is very simple to obtain, as we know that the images are nothing more than matrices

containing numerical values in the RGB (or BW) scale of size  $M \times N$ . In this regard, it is possible to obtain what we define "differential images" by applying a subtraction between two matrices of the same size.

Therefore, since  $F_1 = M \times N$  is the first extracted frame and  $F_2 = M \times N$  the second extracted frame, the first differential image is given by (1).

$$F_1^d = F_2 - F_1. \tag{1}$$

From the matrix point of view, let  $f_{ij}^1$  be the elements of the matrix  $F_1$ ,  $f_{ij}^2$  the elements of the matrix  $F_2$  and  $f_{ij}^d$  the elements of the matrix  $F_1^d$  with  $i \in \{1, \dots, M\}$  and  $j \in \{1, \dots, N\}$ .  $f_{ij}^d$  is obtained by (2).

$$\begin{aligned} f_{ij}^d &= f_{ij}^2 - f_{ij}^1 = \\ &= \begin{pmatrix} f_{00}^2 & \dots & f_{0N}^2 \\ \vdots & \ddots & \vdots \\ f_{M0}^2 & \dots & f_{MN}^2 \end{pmatrix} - \begin{pmatrix} f_{00}^1 & \dots & 0 & f_{0N}^1 \\ \vdots & \ddots & \vdots & \vdots \\ f_{M0}^1 & \dots & f_{MN}^1 \end{pmatrix} = \\ &= \begin{pmatrix} f_{00}^d & \dots & f_{0N}^d \\ \vdots & \ddots & \vdots \\ f_{M0}^d & \dots & f_{MN}^d \end{pmatrix}. \end{aligned} \tag{2}$$

By generalizing it to all frames, the differential images will be obtained from (3).

$$F_k^d = F_k - F_{k-1}, \text{ with } k = 1, \dots, K \tag{3}$$

where  $K$  is the number of differential frames obtainable for each level of intensity and is related to the length of the recorded videos, the amount of FPS extracted from the video and the offset used.

In our method, in particular, we have adopted a function in Python language, called "absdiff", which makes up the difference between two images also reducing the noise that is obtained from a subtraction. In fact, a threshold is set at an arbitrary value between 0 and 255 (threshold 127 was used in this study); all BW pixels greater than 127 become 255; conversely, all BW pixels less than or equal to 127 become 0.

Once the differential images have been transformed into DCT matrices, a division is performed before inserting all the data within the neural network: 70% of these matrices are inserted in the training set and 30% in the test set. For each input, the probability percentage corresponding to each individual class will appear in the output.

### 6.3. Video Validation Dataset

The validation dataset was created as the dataset used in training and testing the neural network. In particular, the related videos are those that had not been included in the dataset for training the neural network.

The same procedures described in sections 3 and 6.2 of the datasets used for training and testing the network are applied to the component frames of this dataset. The following chapter will analyze the performance of this network, when this validation dataset is applied to the network input.

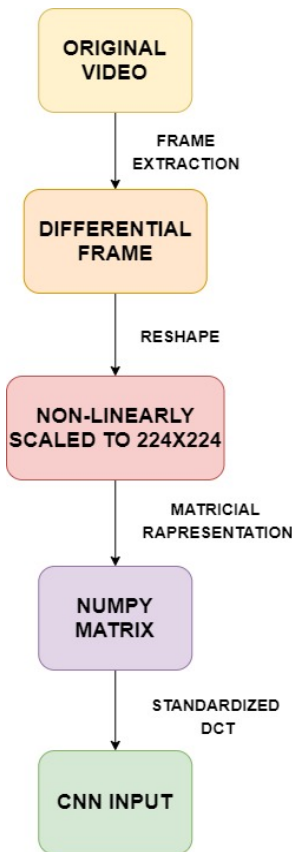


Figure 6: Image Processing Flowchart

This section will show the results of the training and validation of the convolutional neural network when the “differential images” are used as an input.

The application of the DCT to 16×16 sub-blocks show an improvement in the classification performance.

In fact, Figure 7 and Figure 8 respectively show the progress of training and test losses and training and test accuracy.

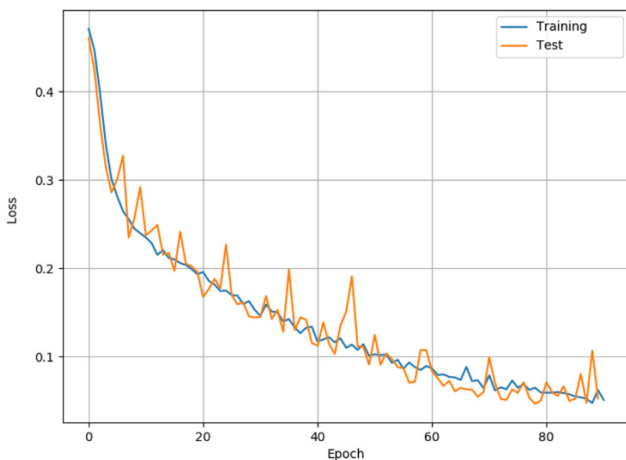


Figure 7: Training and Test Losses with Sub-block 16×16

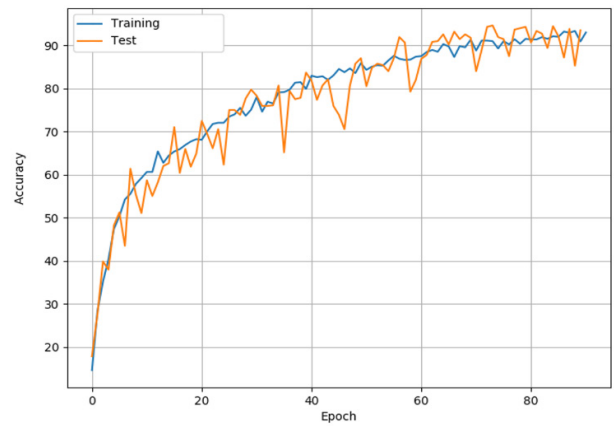


Figure 8: Training and Test Accuracy with Sub-block 16×16

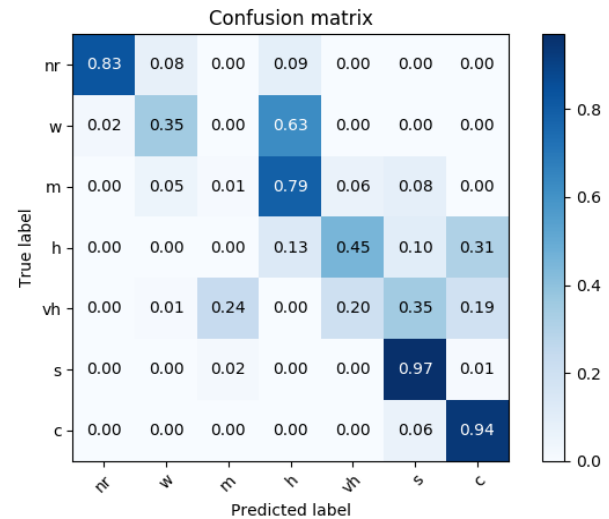


Figure 9: Confusion Matrix with Sub-block 16×16

Once the performance has been defined, it is possible to compare our results with those present in the state of the art, in particular with the study proposed in [16]. The main differences from our study lie in the number of classes used, the methods of image pre-processing and the image capture mode. In fact, in [16] the Sobel algorithm is applied for the elimination of the edges; moreover, the images of the rain are captured in an orthogonal direction to the rain itself. From the performance point of view our study seems to deliver poorer performance, but it is important to note that our method is based on a classification using seven classes, that is, considering seven levels of intensity of precipitation which are fundamental in the context of monitoring the hydrogeological risks.

The use of seven classes to classify the intensity of precipitation and the high temporal resolution of the level of intensity of precipitation allow for a more effective management of alerting mechanisms and therefore prevention and risk management in case of natural disasters-related hydro-geological risk.



## 8. Conclusions

The paper proposes an alternative to traditional rain gauges which is based on the extraction of differential frames from video sequences. Rainfall levels are classified by applying a convolutional neural network. In particular, the system is very simple, based on a plastic shaker, a video camera and a low cost/low power signal processing unit. Performance is very good in terms of precision and ability to adapt to sudden changes in rainfall intensity. The percentage of accuracy of the average classification obtained by applying the DCT for to "16×16" sub-blocks is approximately 49 %, which can reach 75% if the adjacent miss-classifications are not considered. The new video rain gauge exceeds the limits of traditional ones, as it requires no mechanical parts, specific and periodic maintenance.

## Conflict of Interest

The authors declare no conflict of interest.

## References

- [1] J.M. Trabal, D.J. McLaughlin, "Rainfall estimation and rain gauge comparison for x-band polarimetric CASA radars" in 2007 IEEE Int. Geosci. Remote Se, Barcelona, Spain, 2007. 10.1109/IGARSS.2007.4423406.
- [2] D. Nagel, "Detection of Rain Areas with Airborne Radar" in 2017 Int. Radar Symp. Proc., Prague, Czech Republic, 2017. 10.23919/IRS.2017.8008094.
- [3] A. K. Shukla, C.S.P. Ojha, R. D. Garg, "Comparative study of TRMM satellite predicted rainfall data with rain gauge data over Himalayan basin" in 2018 IEEE Int. Geosci. Remote Se, Valencia, Spain, 2018. 10.1109/IGARSS.2018.8651413.
- [4] A. K. Varma, "Measurement of Precipitation from Satellite Radiometers (Visible, Infrared, and Microwave): Physical Basis, Methods, and Limitations" in 2018 Remote Sensing of Aerosols, Clouds, and Precipitation, 2018. <https://doi.org/10.1016/B978-0-12-810437-8.00011-6>.
- [5] S.H. Fang, Y.H. S. Yang, "The impact of weather condition on radio-based distance estimation: A case study in GSM networks with mobile measurements" IEEE Trans. Veh. Technol., **65**(8), 6444–6453, 2016. 10.1109/TVT.2015.2479591.
- [6] F. Beritelli, G. Capizzi, G. Lo Sciuto, C. Napoli, F. Scaglione, "Rainfall Estimation Based on the Intensity of the Received Signal in a LTE/4G Mobile Terminal by Using a Probabilistic Neural Network" IEEE Access, **6**, 30865-30873, 2018. 10.1109/ACCESS.2018.2839699.
- [7] E.M. Trono, M.L. Guico, N.J.C Libatique, G.L. Tangonan, D.N.B. Baluyot, T.K.R. Cordero, F.A.P. Geronimo and A.P.F. Parrenas, "Rainfall Monitoring Using Acoustic Sensors" in 2012 TENCON IEEE Region 10 Conference, 2012. 10.1109/TENCON.2012.6412284.
- [8] R. Nakazato, H. Funakoshi, T. Ishikawa, Y. Kameda, I. Matsuda, S. Itoh, "Rainfall Intensity Estimation from Sound for Generating CG of Rainfall Scenes" in Inter. Workshop on Advanced Image Technology (IWAIT), Japan, 2018. 10.1109/IWAIT.2018.8369692.
- [9] N. J. C. Libatique, G. L. Tangonan, R. Gustilo, W. K. G. Seah, C. Pineda, M. L. Guico, G. Abrajano, R. Ching, J. L. Zamora, A. Espinosa, A. C. Valera, R. Lamac, H. Dy, J. Pusta, E. M. Trono, A. Gimpaya, J. R. San Luis, S. J. Gonzales, A. T. Lotho, "Design of a Tropical Rain – Disaster Alarm System, A New Approach based on Wireless Sensor Networks and Acoustic Rain Rate Measurements" in 2009 IEEE Instr. and Meas. Tech. Conference, Singapore, 2009. 10.1109/IMTC.2009.5168663.
- [10] M. Ferroundj, "Detection of Rain in Acoustic Recordings of the Environment Using Machine Learning Techniques," Master by Research Thesis, Queensland University of Technology, 2015.
- [11] T. Heittola, E. Cakir, T. Virtanen, "The Machine Learning Approach for Analysis of Sound Scenes and Events" Computational Analysis of Sound Scenes and Events, 2017. [https://doi.org/10.1007/978-3-319-63450-0\\_2](https://doi.org/10.1007/978-3-319-63450-0_2).
- [12] S. Sawant, P. A. Ghonge, "Estimation of Rain Drop Analysis Using Image Processing" International Journal of Science and Research, **4**(1), 1981 – 1986, 2015. [https://www.ijsr.net/search\\_index\\_results\\_paperid.php?id=SUB15661](https://www.ijsr.net/search_index_results_paperid.php?id=SUB15661).
- [13] S. Tawade, P. Ghonge, K. Tuckley, "Drop-size analysis using 2D Double Density Dual Tree Discrete Wavelet Transform" in 2016 International Conference on Communication and Signal Processing, 2016. 10.2991/iccasp-16.2017.62
- [14] F. Nashashibi, R. De Charette, A. Lia, "Detection of Unfocused Raindrops on a Windscreen using Low Level Image Processing" in 2010 11th International Conference on Control Automation Robotics and Vision, Singapore, 2010. 10.1109/ICARCV.2010.5707398.
- [15] J. Gallego, A. Pedraza, S. Lopez, G. Steiner, L. Gonzalez, A. Laurinavicius, G. Bueno, "Glomerulus Classification and Detection Based on Convolutional Neural Networks" MDPI Imaging, **4**(1), 20, 2018. <https://doi.org/10.3390/jimaging4010020>.
- [16] J.A. Godoy-Rosario, A.G. Ravelo-García, P.J. Quintana-Morales, J.L. Navarro-Mesa, "An Approach to Rain Detection Using Sobel Image Pre-processing and Convolutional Neuronal Networks" in 2019 Advances in Computational Intelligence, IWANN, Lecture Notes in Computer Science, vol. 11506, pp. 27 – 38. [https://doi.org/10.1007/978-3-030-20521-8\\_3](https://doi.org/10.1007/978-3-030-20521-8_3)
- [17] R. Avanzato, F. Beritelli, "A Rainfall Classification Technique based on the Acoustic Timbre of Rain and Convolutional Neural Networks" MDPI Information, **11**(4), 183, 2020. <https://doi.org/10.3390/info11040183>.
- [18] F. Beritelli, A. Spadaccini, "The Role of Voice Activity Detection in Forensic Speaker Verification" in 2011 IEEE International Conference on Digital Signal Processing (DSP), Corfu, Greece, 2011. 10.1109/ICDSP.2011.6004980.
- [19] F. Beritelli, A. Spadaccini, "A Statistical Approach to Biometric Identity Verification based on Heart Sounds" in 2010 Fourth International Conference on Emerging Security Information, Systems and Technologies, Venice, Italy, 2010. 10.1109/SECURWARE.2010.23.
- [20] F. Beritelli, A. Gallotta, C. Rametta "A Dual Streaming Approach for Speech Quality Enhancement of VoIP Service Over 3G Networks" in 2013 IEEE International Conference on Digital Signal Processing (DSP), Fira, Greece, 2013. 10.1109/ICDSP.2013.6622816.
- [21] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, K. Keutzer, "Squeezenet: Alexnet-Level Accuracy With 50x Fewer Parameters And <0.5mb Model Size" in 2016 Computer Vision and Pattern Recognition (cs.CV); Artificial Intelligence (cs.AI), 2016. arXiv:1602.07360.
- [22] A. Gholami, K. Kwon, B. Wu, Z. Tai, X. Yue, P. Jin, S. Zhao, K. Keutzer, "SqueezeNet: Hardware-Aware Neural Network Design" in Proc. IEEE CVPR, 2018. arXiv:1803.10615.

## Artificial Intelligence Approach for Target Classification: A State of the Art

Maroua Abdellaoui <sup>\*1</sup>, Dounia Daghouj <sup>1</sup>, Mohammed Fattah <sup>2</sup>, Younes Balboul <sup>1</sup>, Said Mazer <sup>1</sup>, Moulhime El Bekkali <sup>1</sup>

<sup>1</sup>LIASSE-ENSA, Sidi Mohamed Ben Abdellah University, Fes, 30050, Morocco

<sup>2</sup>EST, Moulay Ismail University, Meknes, 50050, Morocco

### ARTICLE INFO

*Article history:*

*Received: 15 May, 2020*

*Accepted: 17 July, 2020*

*Online: 09 August, 2020*

*Keywords:*

*Artificial Intelligence*

*Machine Learning*

*Deep Learning*

*Feature Extraction*

*SVD, SVM, RPCA, Sparse Coding*

### ABSTRACT

*The classification of static or mobile objects, from a signal or an image containing information as to their structure or their form, constitutes a constant concern of specialists in the electronic field. The remarkable progress made in past years, particularly in the development of neural networks and artificial intelligence systems, has further accentuated this trend. The fields of application and potential uses of Artificial Intelligence are increasingly diverse: understanding of natural language, visual recognition, robotics, autonomous system, Machine learning, etc.*

*This paper is a state of the art on the classification of radar signals. It focuses on the contribution of artificial intelligence to the latter without forgetting target tracking. This by evoking the different feature extractors, classifiers and the existing identification deep learning algorithms. We detail also the process allowing carrying out this classification.*

## 1. Introduction

The desire to organize in order to simplify has progressively evolved towards the ambition of classifying to understand and, why not, to predict. This development has led to the release of several new techniques capable of satisfying this need. This is how artificial intelligence and the study of its different techniques has become a trend that attracts the interest of researchers in different fields.

Neural networks consist of artificial neurons or nodes that are analogous to biological neurons [1]-[3]. They are the result of an attempt to design a very simplified mathematical model of the human brain based on the way we learn and correct our mistakes. Machine learning allows us to obtain computers capable of self-improvement through experience [4].

The latter is today one of the most developed technical fields, bringing together computer science and statistics, leading to artificial intelligence and data science development [5]. Due to the explosion on the amount of information available online and at low cost, machine learning has experienced tremendous progress resulting in the nonstop development of new algorithms [6]. The use of data-intensive machine learning methods is in all fields of technology and science [1], [5], [6].

Many companies and researches today claim to use artificial intelligence, when in fact the term does not apply to the technologies they use. In the same vein, there is more or less

\* Corresponding Author: Maroua Abdellaoui, Av Moulay Abdellah, Km 5, Route Imouzzar, Ensa, Fes, Morocco, (+212) 674661948. Email: Abdellaoui.marwa@gmail.com

confusion between artificial intelligence and the concept of Machine Learning, without even mentioning Deep Learning. This paper will shed light on these different concepts by detailing each of them and then focuses on the different algorithms employed to extract features and classification of detected objects.

This work complements the various state of the art studies already carried out in AI domain [1]-[3] and is mainly interested in its contribution to road safety through radar target classification (pedestrians, cyclists...).

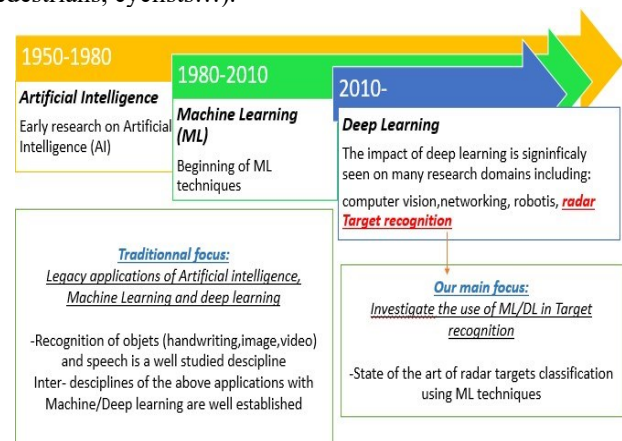


Figure 1: The evolution of artificial intelligence and the value of our work.

For this, section 2 will represent a study on the classification of radar signals and targets without forgetting the tracking. Then on section 3 we will focus on artificial intelligence by detailing its different concepts in order to eliminate any confusion. We will also study the feature extractors and the algorithms of

classification generally used in the overall classification process. A summary bringing together results of the research work in literature in this field will be carried out before concluding, thus opening the way too many perspectives

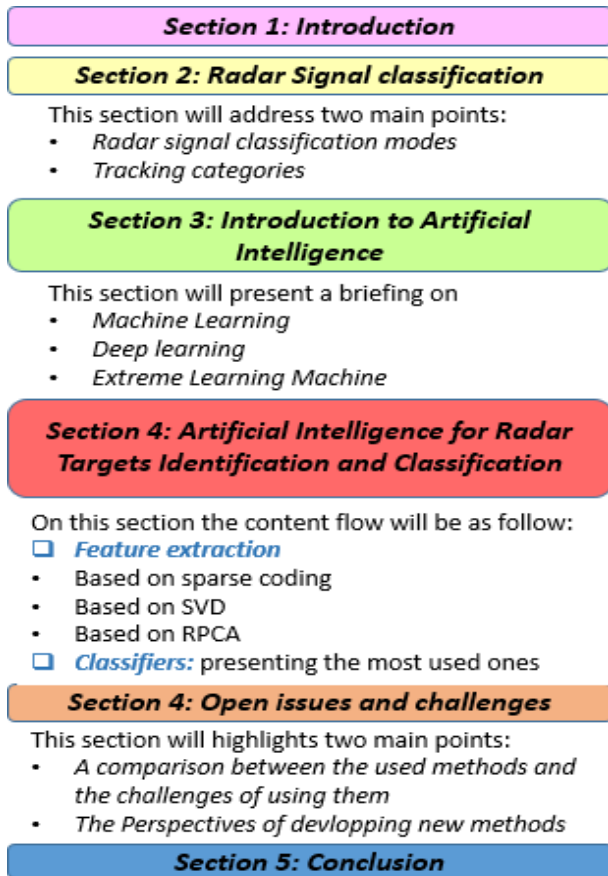


Figure 2: Content flow diagram

## 2. Radar Signal Classification

Classification can be defined as the search for a distribution of a set of elements into several categories [7]. Each category, called a class, groups together individuals who share similar characteristics. The objective is to obtain the most homogeneous and distinct classes possible. Identifying categories requires careful definition of a space in which the classification problem must be resolved [7, 8]. Such a space is often represented by vectors of parameters, as shown in Figure3, extracted from the elements to be classified and the classification is carried out by adopting a probabilistic, discriminative, neuronal or even stochastic approach.

The method of classification of radar signals is presented in Figure 4 that contains the neural network procedures. Their continuous development and improvement have made it possible to clearly understand the potential and the limits of this technique in several fields. Among these, remote sensing, signal processing, identification and characterization of targets [1, 9].

Within these modes of classification, we find; the cluster method that seeks to construct a partition of a data set so that the data from the same group exhibit common properties or characteristics. It distinguishes them from the data contained in the other groups [10, 11]. As such, clustering (or regrouping) is a

subject of research in learning stemming from a more general problem, namely classification. A distinction is made between supervised and unsupervised classification. In the first case, it is a question of learning to classify a new individual among a set of predefined classes, from training data (couples (individual, class)). Derived from statistics, and more specifically from Data Analysis (ADD), unsupervised classification, as its name suggests, consists of learning without a supervisor. From a population, it is a question of extracting classes or groups of individuals with common characteristics, the number and definition of classes not being given a priori. Clustering methods are used in a lot of domains of application ranging from biology (classification of proteins or genome sequences), to document analysis (texts, images, videos) or in the context of analysis of traces of use. Many unsupervised classification methods have been published in the literature and it is therefore hard to give an exhaustive list, despite the numerous articles published attempting to structure this very rich and constantly evolving field for more than 40 years [12-18].

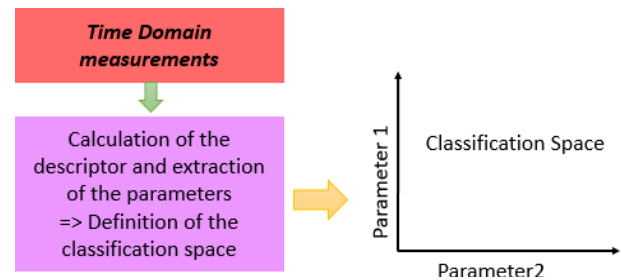


Figure 3: Illustration of the classification approach if two parameters are used to define the space in which the classification problem must be solved.

Entropy of similarity is a method from Shannon's entropy. The latter named "entropy" his definition of the amount of information. We will therefore talk indifferently in Shannon about the quantity of data generated from a message or the entropy of this message source. The classification method based on entropy will then make it possible to answer several questions which were encountered in all sociological or scientific surveys, namely: the measurement of the correlation between the characters and their selectivity, the homogeneity of the groups as well as problems of formation of new homogeneous classes [19].

Entropy methods are dedicated to the analysis of irregular and complex signals [20]. The support vector machine [7, 8], or large-margin separators, represents a set of supervised learning techniques intended to deal with the issue of discrimination and regression. Support vector machines represents an extension of linear classifiers. They were used because they are capable to work with big amount of information, the low number of hyperparameters, their performance and fiability. SVM have been used in many fields [21-23].

According to the data, the performance of support vector machines is of the same order, or even better, then that of a neural network or of a Gaussian mixture model. Also, the timescale characteristics [24], the modulation domain [25], basic function neural networks [26], Rihaczek distribution and Hough transform [27], which is a pattern recognition technique invented in 1959 by Paul Hough, subject to a patent, and used in the processing of digital images. The simplest application can detect lines present in an image, but modifications can be made to this technique to detect

other geometric shapes; it is the generalized Hough transform developed by Richard Duda and Peter Hart in 1972 [28-30], frequency estimation [28], pulse repetition interval [31], two-dimensional bispectrum [32], etc.

These methods of classification represent research in several disciplines [22, 30, 33-35]. To allow proper operation in complex signal environments with many radar transmitters, signal classification should be able to handle not determined, corrupt, and equivocal measurements reliably.

For radar classification of vehicle type and determination of speed profitably by calculation, Cho and Tseng [35] created an improved algorithm that will help smart transport applications in real time containing eight types of mode setting categorization of radar signals.

In general, during a transmission the scattered waves depend on the distance to the target, so to measure these waves we need to measure this distance. In the literature, the scattered waves received are processed by algorithms to detect the presence, the distance and the type of the target [36-40]. Among the target classification methods are the AALF, AALP and ABP method [24].

Tracking is a very important element in this process. There are different tracking methods summarized in Figure 5 and Figure 6.

### 3. Introduction to Artificial Intelligence

Artificial intelligence (AI) has come to the fore in recent years. It is used in several applications for various disciplines [41-54]. Artificial Intelligence (AI) as we know it is weak Artificial Intelligence, as opposed to strong AI, which does not exist yet. Today, machines are capable of reproducing human behavior, but without conscience. Later, their capacities could grow to the point of turning into machines endowed with consciousness and sensitivity.

AI has evolved a lot thanks to the emergence of Cloud Computing and Big Data, which is an inexpensive computing power that gives accessibility to a large amount of data. Thus, the machines are no longer programmed; they learn instead [53, 54].

The following subsections aim to highlight machine learning, deep learning and extreme machine learning respectively, in order to eliminate any confusion between these concepts.

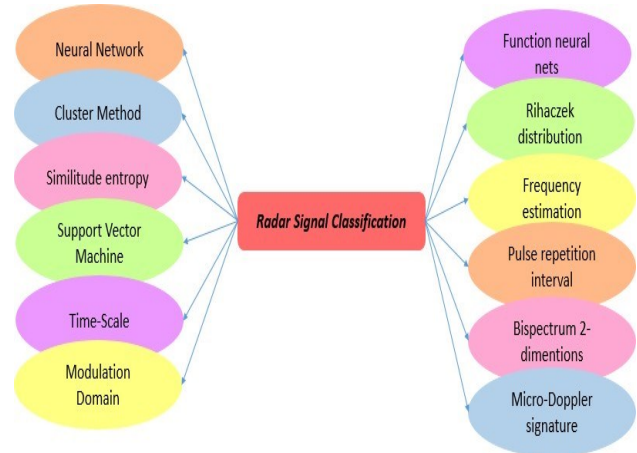


Figure 4: Radar signal classification mode

#### 3.1. Machine Learning

Machine Learning is a sub-branch of artificial intelligence, which consists of creating algorithms capable of improving automatically with experience [45-50]. We also speak in this case of self-learning systems. Machine Learning, or automatic learning, is capable of reproducing a behavior thanks to algorithms, themselves fed by a lot of information. In front of a lot of circumstances, the algorithm learns which behavior to follow and decision to take creating a model. The machine makes the tasks automate depending on the situation [54-58].

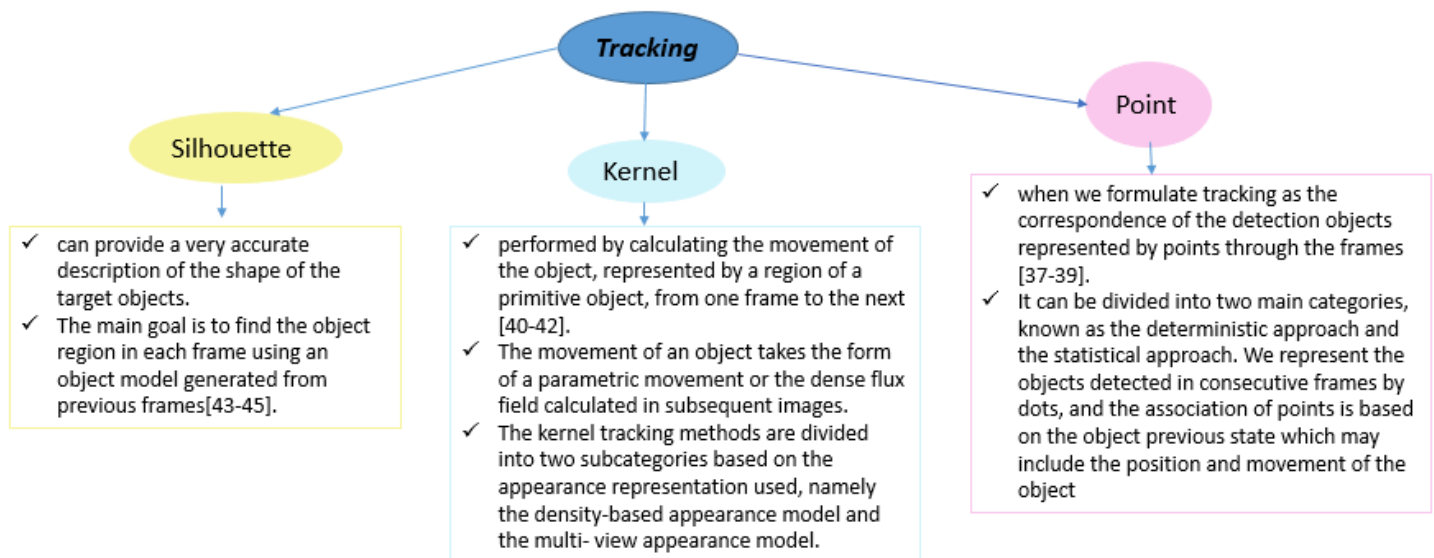


Figure 5: Tracking methods types and definition



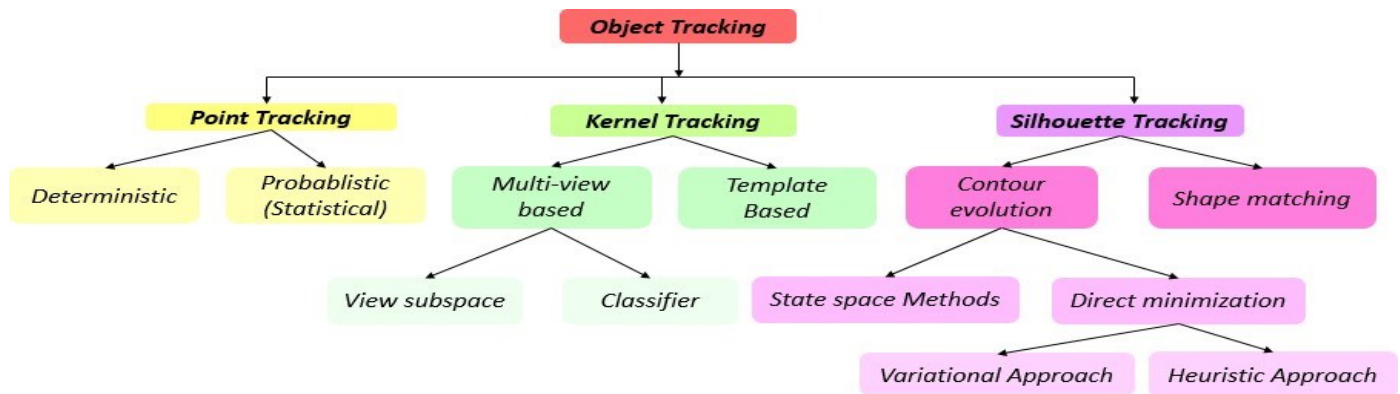


Figure 6: Representation of the different tracking categories

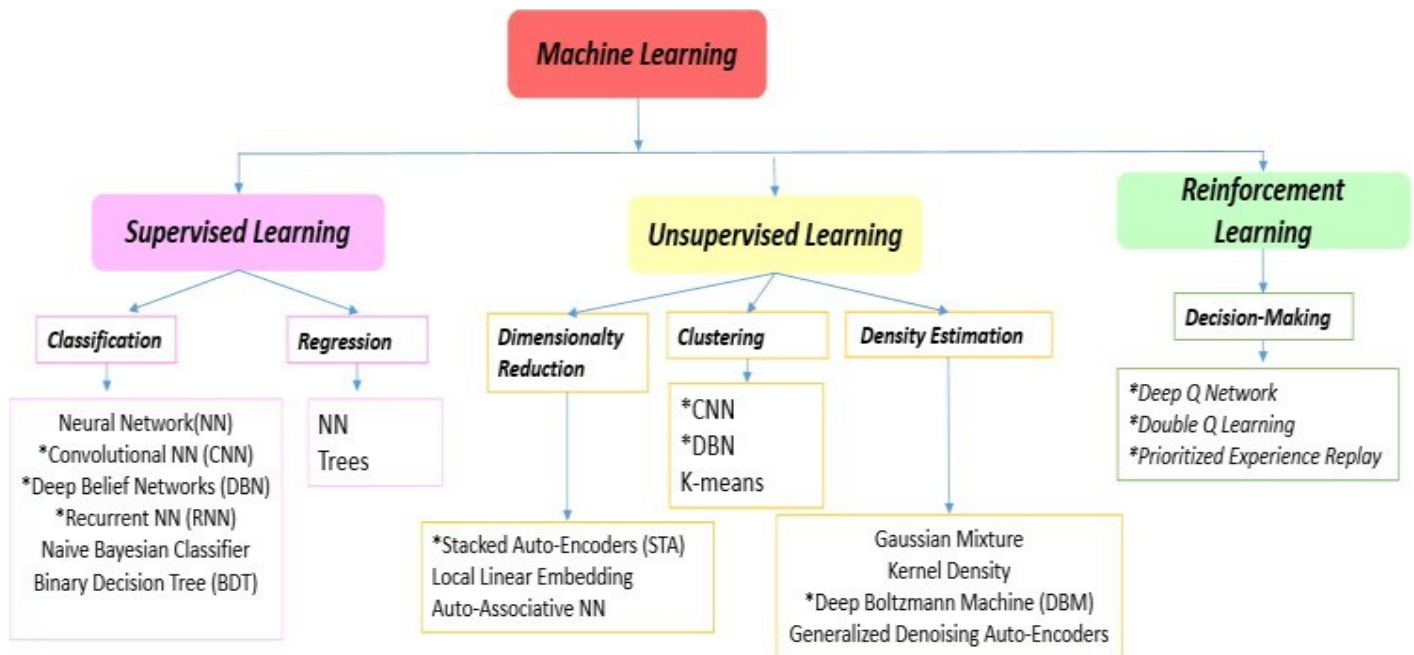


Figure 7: Representation of ML types

There are three main types of Machine Learning [57]-[62] represented in figure 7: In supervised learning, the algorithms are based on already categorized datasets, in order to understand the criteria used for classification and reproducing them [63]-[66]. In unsupervised learning, algorithms are trained from raw data, from which they try to extract patterns [67]-[70]. Finally, in reinforcement learning, the algorithm functions as an autonomous agent, which observes its environment and learns as it interacts with it [67], [69].

Machine learning is a broad field, which includes many algorithms. Among the most famous are regressions (linear, multivariate, polynomial, regularized, logistic, etc.); these are curves that approximate the data. Naïve Bayes' algorithm; which gives the probability of the prediction, in knowledge of previous events [71-76]. Clustering is always using mathematics; we will group the data into packets so that in each packet the data is as close as possible to each other [77-79].

There are also more sophisticated algorithms based on several statistical techniques such as the Random Forest (a forest of voting decision trees), Gradient Boosting, Support Vector Machine [21], [23] The learning techniques showed in Figure7

with (\*) have emerged recently with their use mainly limited to object recognition, including the classification of radar targets in urban areas.

Machine learning is a very important approach for classification. For instance, the classification of a single target e.g. a pedestrian or a cyclist is relatively simple because the micro-Doppler signatures of the pedestrian and the cyclist are different, but the problem arises when classifying overlapping targets e.g. pedestrians and cyclists. The classification here is much more difficult, which requires the intervention of deep learning techniques to deal with this issue.

### 3.2. Deep Learning

Deep Learning aim to understand concepts in a more précised way. In a neural network, successive layers of data are combined to learn the concepts. The simplest networks have only two layers: an input and an output, knowing that each one can have several hundreds, thousands, even millions of neurons. Among the most used deep learning algorithms, we have:

- Artificial neural networks (ANN): these are the simplest and are often used in addition because they sort information well

- Convolutional neural networks (CNN): Applies filters to the information collected in order to have new data (for example, bringing out the contours in an image can help to find where is the face)
- Recurrent neural networks (RNN): the best known are LSTM, which have the ability to retain information and reuse it soon after. They are used for text analysis (NLP), since each word depends on the previous few words (so that the grammar is correct)

As well as more advanced versions, such as auto-encoders [80], Boltzmann machines, self-organizing maps (SOM), etc. Figure 8 shows the key algorithms of deep learning and the research fields that are interested in it.

### 3.3. Extreme learning machine

Extreme Learning Machine (ELM) is usually used for pattern classification. It can be considered as an algorithm for direct overshadow single layer neural networks. The latter mastery the slow training speed and over-adjustment complications compared to the conventional neural network learning algorithm. ELM is based on the empirical theory of risk inferiorization. The learning process of it requires only one iteration. The multiple iterations and local minimization are avoided by the algorithm. We can find ELM useful in multiple fields and applications thanks to its robustness, controllability, good generalization capacity and its fast learning rate.

The researchers proposed modifications to the algorithms, to improve ELM [80-85] proposes the fully complex ELM (C-ELM). The latter extends the ELM algorithm from the real domain to the complex domain. Given the significant time consumed by the update procedure using the old data with the new information received, an online sequential ELM (OS-ELM) is proposed in [81-86], which can learn the training data one by one or block by block and discard the data for which the training has already been carried out. A new adaptive set model of ELM (Ada-ELM) is proposed in [82, 87] and allows better prediction performances to be obtained. It can automatically adjust the overall weights.

ELM performance is affected by hidden layer nodes. These are difficult to determine, the incremental ELM (I-ELM) [84, 89], the pruned ELM (P-ELM) [84, 89] and the self-adaptive ELM (SaELM) [85, 90] have been proposed in other works ELM achieves good results and shortens training times that takes several days in deep learning to a few minutes by ELM.

It is difficult to achieve such performance by conventional learning techniques. Example of datasets are showed in Table 1.

Artificial intelligence is widely used for the classification of radar targets. The following section will focus on the classification procedure and the different feature extractors and classifiers used in machine learning. The most famous and most used algorithms in our field (road safety) will be mentioned as well.

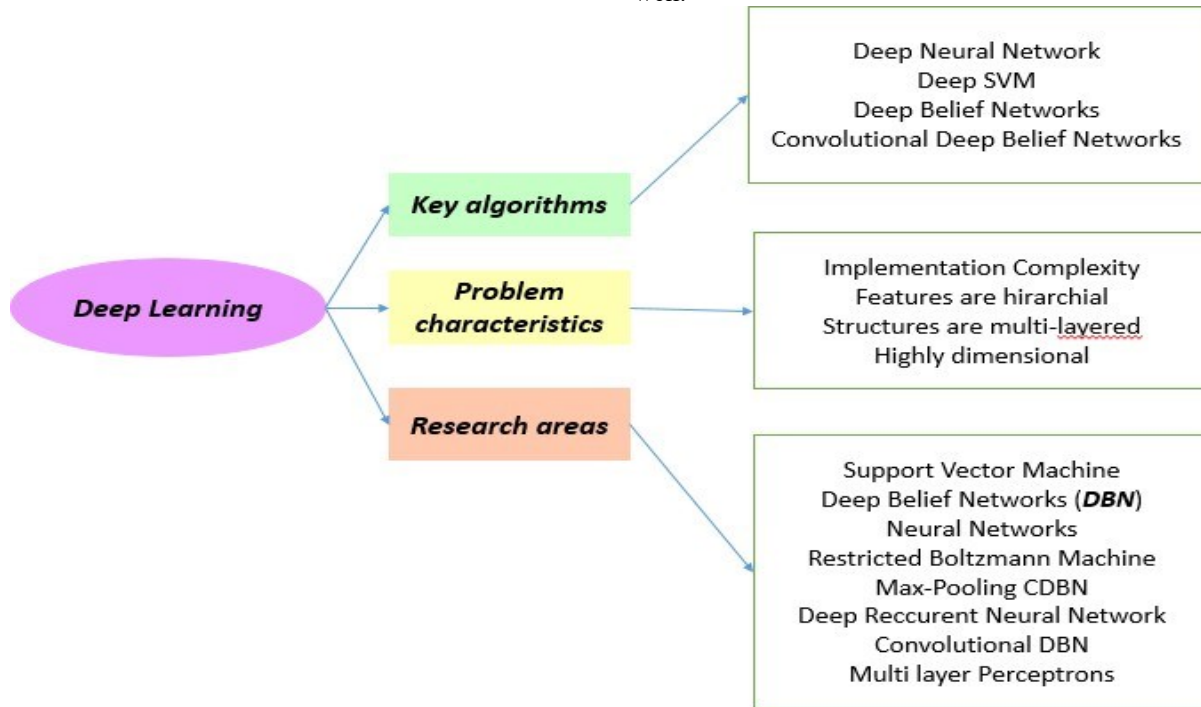


Figure 8: Algorithms and characteristic problems linked to deep learning and the research fields, which are interested in it.

Table 1: Examples of Datasets and their learning methods and training time

Datasets	Learning Methods	Testing Accuracy	Training time
MINIST OCR	▪ELM (multi hidden layers, unpublished)	99.6%	<ul style="list-style-type: none"> <li>▪ Several minutes</li> <li>▪ 281.37s (CPU)</li> <li>▪ 5.7 hours (GPU)</li> <li>▪ 19 hours (GPU)</li> <li>▪ &gt;17 hours (GPU)</li> <li>▪ &gt;17 hours (GPU)</li> </ul>
	▪ELM (multi hidden layers, ELM auto encoder)	99.14%	
	▪Deep Belief Networks (DBN)	98.87%	
	▪Deep Boltzman Machines (DBM)	99.05%	
	▪Stacked Auto Encoders (SAE)	98.6%	
	▪Stacked Denoising Auto Encoders (SDAE)	98.72%	
3D Shape Classification	▪ELM (multi hidden layers, local receptive fields)	81.39%	<ul style="list-style-type: none"> <li>▪ 306.4s (CPU)</li> <li>▪ &gt;2 days (GPU)</li> </ul>
	▪Convolutional Deep Belief Network (CBDN)	77.32%	
Traffic sign recognition (GT SRB Dataset)	▪HOG+ELM	99.56%	<ul style="list-style-type: none"> <li>▪209s (CPU)</li> <li>▪5 hours (CPU)</li> <li>▪&gt;37 hours (GPU)</li> </ul>
	▪CNN+ELM (Convolutional neural networks (as feature extractors)+ ELM (as classifiers)	99.48%	
	▪Multi-column deep neural network (MCDNN)	99.46%	

#### 4. Artificial Intelligence for Radar Target’s Identification and Classification

There is a lot of characteristics for target identification using micro Doppler signature and ML algorithms that have been studied in other research and interesting results have been presented [47, 48]. Many public datasets for target classification are introduced in [93-100].

It lays the groundwork for disentangling data into independent components [85]. PCA ignores the less important components [85]. We can use SVD in order to find PCA by truncating the less important base vectors in the original SVD matrix.

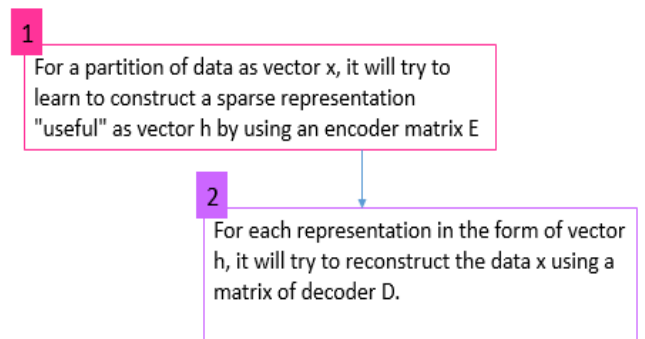
Most of the research done is on supervised learning, but very little of it uses unsupervised machine learning. The latter turns out to be one of the latest trends and added value in recent work [8, 49] and this based on sparse coding.

##### 4.1. Feature Extraction

##### 4.1.1. Feature Extraction based on sparse coding (sparse)

Sparse coding is a technique based on the study of algorithms aimed at learning a sparse / sparse useful representation of all data [99-106]. The next step consists in encoding so that each data will be in the form of sparse code. The algorithm uses information from the input to learn the sparse representation. This can be applied clearly to any type of information. We call this unsupervised learning. It will find the representation without losing any part or aspect of the data [106]. To do this, two main constraints try to be satisfied using sparse coding algorithms. Figure 9 describes them:

In reality, we just give more than the number of dimensions in which the original data is encoded or sometimes the same amount. Figure10 describes the target identification system based on sparse coding. Radar data in the time domain must be processed before sparse coding. This is done using the short-term Fourier transform (STFT), which is a method of time-frequency analysis used for micro-Doppler signatures by other researchers.



\*We note that D and E are learnable parameters.

Figure 9: Procedure using sparse coding

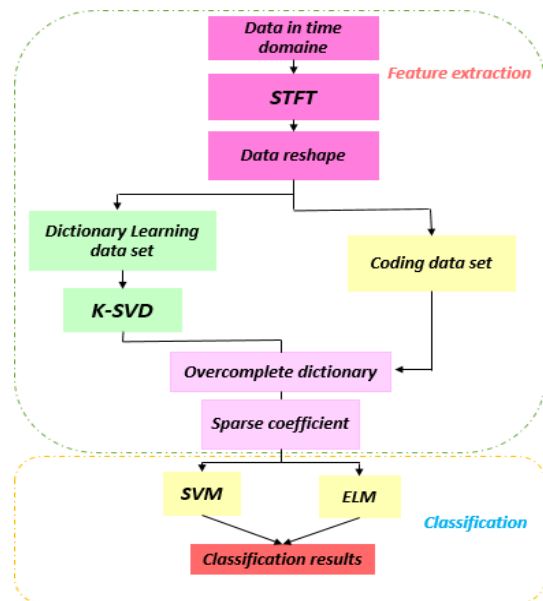


Figure 10: Functional diagram of the proposed functionality extractor based on sparse coding

We use the Hamming window to extract micro-Doppler signatures from targets. After the STFT, comes the step of the complete dictionary construction, for this each spectrogram must

be reshaped into a matrix. The micro Doppler signature is converted into a matrix with the concatenation of the two parts (real and imaginary), whose dimension is  $2U \times M$ . We generate the dataset for training using  $N$  given samples. We then gather the data received from the targets from all the angles. The dimension of the data set for each angle  $X^k$  is  $2U \times MN$ .

The resulting data set that we have trained is made of signals from the types of targets combined (pedestrians and cyclists for example).  $D$  and  $W$  (matrix coefficient) are deduced from the information of training in the optimization equation (1)

$$(\hat{D}, \hat{W}) = \arg \min_{D,W} \frac{1}{2}(DW - X) + \delta \sum_{m=1}^{2KJ} (W_m) \quad (1)$$

Reading equation (1) from the left to the right, the last term represents the error of reconstruction between the original data and its representation based on the dictionary  $D$ . A better approximation compared to the original data can be made using the minimization of this term. The work here is to play with  $D$  and  $W$  both at the same time, adjusting them to solve the equation.

Among the methods adopted for learning the complete dictionary, we have the K-SVD method. We will first search for  $W$  without touching  $D$  then on the second iterations we will search for both  $D$  and  $W$  while keeping the non-zero elements in  $W$  intact and fixed.

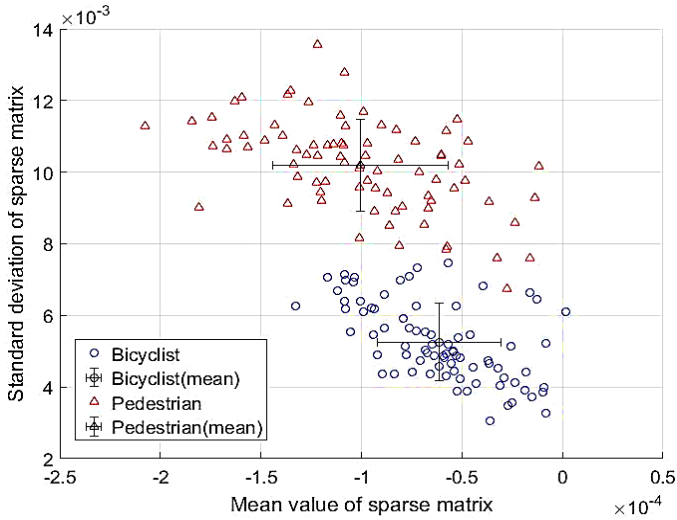


Figure 11: Reduction of sparse coding characteristics (for pedestrians and cyclists' case) [8]

SVD ensures the normalization of atoms in the dictionary to each other. K-SVD takes over between the information coding with the existing dictionary and the regular updating of the dictionary in order to obtain a better fit.

The functionality extraction based on a sparse representation can be obtained once the  $D$  dictionary has been created. At this stage, the set of sparse matrices can be directly used as classification characteristics. It should be noted that the functionality dimensions are still very important and requires reduction. For this, some numerical characteristics, such as the mean value and the standard deviation, can be calculated from the sparse matrix.

We illustrate in Figure 11 an example of the sparse coding functionalities reduced with an angle of  $30^\circ$  (this is the case of a

pedestrian and a cyclist).

Numerical characteristics can be used in the classification, even if some information are rejected by the calculation. The sparse matrix and its reduced characteristics are used in [8] to carry out the classification. Five numerical characteristics of sparse matrices are used for the classification, namely: mean values, standard deviations, maximum values, minimum values, and difference between the maximum and minimum values.

#### 4.1.2. SVD-based feature extraction

SVD makes it possible to build an empirical model, without an underlying theory, all the more precise when terms are injected into it [92]. The effectiveness of the method depends in particular on the way in which the information is presented to it. We can describe the SVD decomposition with the equation (2).

$$F = USV^T \quad (2)$$

$S$  is a diagonal matrix of singular values. We note that the components of  $S$  represent only scaling factors and therefore don't have any information of the spectrogram. The matrices  $U$  and  $V$  contains singular vectors of  $F$  in the two directions (left and right). It represents the information of both the time and the Doppler domain of an MD signature. We can use the singular vector as a characteristic for categorization.

#### 4.1.3. Robust principal component analysis RPCA

We can calculate the average frequency profile as the average value of the MD signatures along the time axis of the absolute value in the frequency domain. It is demonstrated in (3)

$$MFP(v) = \frac{1}{M} \sum_{m=1}^M STFT(v, m) \quad (3)$$

A betterment of the MFP using PCA and minimum covariance determinant estimator is discussed and presented in several works and the method is described with the equation (4)

$$F = PCA_Q(MFP(v)) \quad (4)$$

#### 4.2. Classifiers

The objective of [50] is to study the changes in the categorization performance with the parameters of signal processing and the procedures of extracting characteristics applicable to backscattered signals from the UWB radar [51-54]. In the literature, many algorithms have been used for classification. In [41], the classifiers were used: MDC [107-122], NB [111-115], k-NN [113] and SVM [21, 23, 116].

The comparison of recognition rates of the frequency FM using many discriminant analyzes (LDA) and support vector machine (SVM) are presented in Figure 12 that suggests that the support vector machine approach is a method efficient classification of radar signals with an elevated recognition percentage. SVM has the maximum failure rate ( $\leq 97\%$ ) and it is lower for LDA ( $\leq 94\%$ ) and tends to change.

Those categorization procedures are used to given processed products based whether the cross-validation or leave-a-out method. The appraisal and estimation of the performance of every classifier is done using independent tests of the learning set, thus minimizing the generalization inaccuracy.



In [7], the author used supervise learning after classifying specific objects. Two main phases known as the training phase and the test phase are in this process. Supervised regression methods are also generally used to approximate and guess the mapping between the directions of movement and the micro-Doppler signatures of the targets, among which we find the support vector regression (SVR) and the multilayer perceptron (MLP) used in [7].

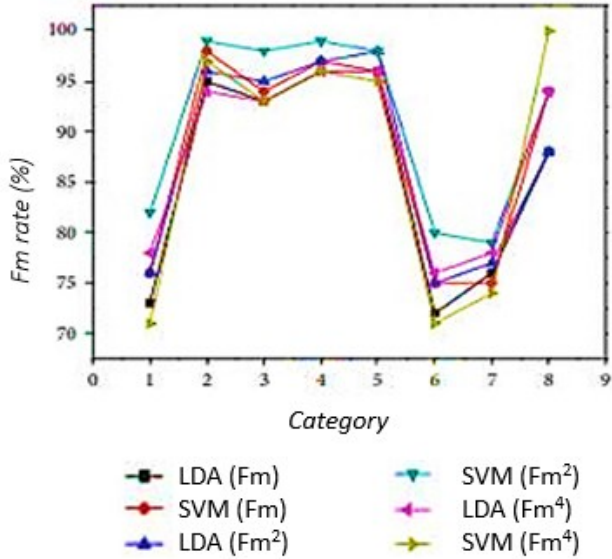


Figure 12: Recognition rate comparison adapted from [35]



Figure 13: Regression Model [7]

The regression training data set is useful to approximate the correspondence between the vectors and the direction of motion. The regression model is a function map described in figure 13. Micro-Doppler signatures of complex surface targets with moving sections are used to target the approximation of the direction of movement. Then supervised regression algorithms are applied as a solution to the problem of estimating the direction of movement [119]-[131].

The next section presents a comparison between the methods used and presents the advantages of one method compared to the [www.astesj.com](http://www.astesj.com)

others. It highlights the usefulness of the algorithms according to the desired application. The challenge here is to choose the method to be used according to the constraints presented by the studied system. The section also discusses the prospects for proposing a new procedure combining the advantages of the methods presented previously and adaptable to several uses.

### 5. Open Issues and Challenges

Learning approaches can be categorized into supervised learning and unsupervised learning. On the first one, the classifier will be designed by exploiting information already known because the training data set is previously available. This is not the case on the second one because the training information known as class labels, do not exist. We then use a group of characteristic vectors that we devise into another group of subsets called clusters. The information with similar features are portioned between subsets equally.

Table 2: Comparison between methods [8]

2sd signal	SVM Pedestrian	SVM Bicyclist	ELM Pedestrian	ELM Bicyclist
<b>RPCA</b>	Mean : 2.57 Max : 94.67 Min : 80.14	2.21 95.71 81.55	2.51 95.34 81.55	2.73 96.14 83.43
<b>SVD</b>	Mean : 2.17 Max : 96.15 Min : 85.13	2.11 97.33 89.85	2.44 96.75 86.66	2.01 97.15 896.08
<b>Sparse coding</b>	Mean : 1.85 Max : 99.97 Min : 85.13	1.79 100 91.31	1.99 100 91.55	2.26 100 91.11
<b>Reduced sparse coding</b>	Mean : 2.34 Max : 99.89 Min : 91.35	2.23 100 90.02	1.92 99.84 91.35	2.18 99.94 89.66

There is an increasing convergence towards the use of unsupervised ML using non-structured input information in several areas such as traffic engineering, definition of network anomalies, categorization of objects and optimization of road traffic and many others. Table 2 and Table 3 shows a comparison between methods. On one hand, the results in [7] [8] show that classification with characteristics based on sparse coding makes it possible to obtain the highest precision (> 96%) and that the SVD and RPCA methods are very efficient. On the other hand, the computation time for the procedure using the sparse coding functionality extractor and the support vector machine classifier is too long despite that it offers the best classification

performance.

Table 3: Method's performances comparison [8]

Extractor and classifiers	SVM	ELM
SVD	0.26s	0.22s
RPCA	0.13	0.11
Sparse Coding	1.12	0.98
Reduced sparse coding	0.87	0.86

In addition, when the parameters dimensions are minor, the total time for SVM and ELM is alike. As the size of the functionality increases, identification via ELM is faster than SVM.

We can also add that a good motion direction estimate (with an error beneath 5°) can be obtained based on the SVR-based method. The approximation conduct improves for the directions of movement towards the radar and reduce for the angles of movement at right angles to the radar sight (for a radar targets detection example).

This allows us to deduce that certain methods can be effective for certain applications and not for others. The choice of which method to use will then depend heavily on the application itself. It also prompts us to question the possibility of having a high- performance method at all levels and for all possible applications, including road safety [132-136]. The current trend is converging towards the development of new methods combining the properties of old algorithms and the expectations of new applications and adaptable to different uses.

## 6. Conclusion

The main goal of classification approach is to group the information into the adequate category based on common features. The classification makes it possible to determine the data including the unknown affiliate type or group. Classification methods can be known as approaches generating non-identical outcome. Artificial intelligence and its various techniques and algorithms help solving those issues. They are the subject of several research studies, including in the field of road safety. This paper is a state of the art of targets classification and the contribution of machine learning technologies in it. The study and comparison of the different extraction methods and classification algorithms allowed us to deduce that efficient algorithms may not be as efficient for some applications. It strongly depends on the desired application and its functionalities. The different deductions open up perspectives on the development of new approaches, adaptable to any kind of application with optimal algorithms in terms of calculation time and processing effectiveness.

## Conflict of Interest

The authors declare no conflict of interest.

## References

- [1] Z. M. Fadlullah et al., "State-of-the-Art Deep Learning: Evolving Machine Intelligence Toward Tomorrow's Intelligent Network Traffic Control Systems," in *IEEE Communications Surveys & Tutorials*, **19**(4), 2432-2455, Fourthquarter 2017. DOI: 10.1109/COMST.2017.2707140
- [2] M. Usama et al., "Unsupervised Machine Learning for Networking: Techniques, Applications and Research Challenges," in *IEEE Access*, **7**, 65579-65615, 2019. DOI: 10.1109/ACCESS.2019.2916648
- [3] Y. Xin et al., "Machine Learning and Deep Learning Methods for Cybersecurity," in *IEEE Access*, **6**, 35365-35381, 2018. DOI: 10.1109/ACCESS.2018.2836950
- [4] P. Louridas and C. Ebert, "Machine Learning," in *IEEE Software*, **33**(5), 110-115, Sept.-Oct. 2016.
- [5] M. I. Jordan, T. M. Mitchell, "Machine learning : Trends perspectives and prospects", *Science*, **349**(6245), 255-260, 2015. DOI: 10.1126/science.aaa8415
- [6] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning", *Nature*, **521**, 436- 444, May 2015. DOI:10.1038/nature14539
- [7] P. Khomchuk, I. Stainvas, I. Bilik. Pedestrian motion direction estimation using simulated automotive MIMO radar. *IEEE Trans. Aerosp. Electron.Syst.* 2016. DOI: 10.1109/TAES.2016.140682
- [8] Pedestrian and Bicyclist Identification through Micro Doppler Signature with Different Approaching Aspect Angles Rui Du, Yangyu Fan, Jianshu Wang. 10.1109/JSEN.2018.2816594, *IEEE Sensors Journal* 2018. DOI: 10.1109/JSEN.2018.2816594
- [9] J. A. Anderson, M. T. Gately, P. A. Penz, and D. R. Collins, "Radar signal categorization using a neural network," *Proceedings of the IEEE*, **78**(10), 1646-1657, 1991. DOI: 10.1109/5.58358
- [10] Hosni Ghedira, utilisation des réseaux de neurones pour la cartographie des milieux humides à partir d'une série temporelle d'images radarsat-1. 2002
- [11] Vittorio cappecchi, une methode de classification fondé sur l'entropie. *Revue française de sociologie Année 1964*. DOI: 10.2307/3319578
- [12] James Bezdek, James Keller, Raghu Krishnapuram, and Nikhil R. Pal. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. The Handbook of Fuzzy Sets Series, didier dubois and henri prade edition, 1999*. DOI 10.1007/978-0-387-24579-9
- [13] J.A. Hartigan. *Clustering algorithms*. John Wiley & Sons Inc., 1975.
- [14] A.K. Jain and R.C. Dubes. *Algorithms for clustering Data*. Prentice Hall Advanced Reference series, 1988.
- [15] A.K. Jain, M.N. Murty, and P.J. Flynn. *Data clustering: A review*. *ACM Computing Surveys*, **31**(3) :264-323, 1999. doi.org/10.1145/331499.331504
- [16] Anil K. Jain. *Data clustering: 50 years beyond k-means*. *Pattern Recognition Letters*, **31**(8) :651-666, 2010. DOI:10.1016/j.patrec.2009.09.011
- [17] Rui Xu and Donald Wunsch II. *Survey of clustering algorithms*. *IEEE Transactions on Neural Networks*, **16**(3) :645-678, May 2005. DOI: 10.1109/TNN.2005.845141
- [18] O.R. Zaiane, C.H. Lee A. Foss, and W. Wang. *On data clustering analysis : Scalability, constraints and validation*. In *Proc. of the Sixth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*. Lecture Notes in Artificial Intelligence 2336, *Advances in Knowledge Discovery and Data Mining*, Springer-Verlag, 2002. DOI: 10.1007/3-540-47887-6\_4
- [19] D. Feng et al., "Deep Multi-Modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges," in *IEEE Transactions on Intelligent Transportation Systems*, 2020. DOI: 10.1109/TITS.2020.2972974.
- [20] J. Li, A. Sun, J. Han and C. Li, "A Survey on Deep Learning for Named Entity Recognition," in *IEEE Transactions on Knowledge and Data Engineering*, 2020. DOI: 10.1109/TKDE.2020.2981314.
- [21] S. A. Osia et al., "A Hybrid Deep Learning Architecture for Privacy-Preserving Mobile Analytics," in *IEEE Internet of Things Journal*, **7**(5), 4505-4518, May 2020, DOI: 10.1109/JIOT.2020.2967734.
- [22] F. Karim, S. Majumdar and H. Darabi, "Adversarial Attacks on Time Series," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020. DOI: 10.1109/TPAMI.2020.2986319.
- [23] W. Y. B. Lim et al., "Federated Learning in Mobile Edge Networks: A Comprehensive Survey," in *IEEE Communications Surveys & Tutorials*, 2020. DOI: 10.1109/COMST.2020.2986024.
- [24] S. Mei, R. Jiang, X. Li and Q. Du, "Spatial and Spectral Joint Super-Resolution Using Convolutional Neural Network," in *IEEE Transactions on Geoscience and Remote Sensing*, 2020. DOI: 10.1109/TGRS.2020.2964288.
- [25] B. Gholami, P. Sahu, O. Rudovic, K. Bousmalis and V. Pavlovic, "Unsupervised Multi-Target Domain Adaptation: An Information Theoretic Approach," in *IEEE Transactions on Image Processing*, **29**, 3993- 4002, 2020. DOI: 10.1109/TIP.2019.2963389
- [26] T. McConaghy, H. Leung, E. Bosse, and V. Varadan, "Classification of audio radar signals using radial basis function neural networks," *IEEE Transactions on Instrumentation & Measurement*, **52**(6), 1771-1779, 2003. DOI: 10.1109/TIM.2003.820450
- [27] D. Zeng, X. Zeng, H. Cheng, and B. Tang, "Automatic modulation

- classification of radar signals using the Rihaczek distribution and Hough transform," *IET Radar Sonar & Navigation*, **6**(5), 322–331, 2012. DOI: 10.1049/iet-rsn.2011.0338
- [28] F. Gini, M. Montanari, and L. Verrazzani, "Maximum likelihood, ESPRIT, and periodogram frequency estimation of radar signals in K-distributed clutter," *Signal Processing*, **80**(6), 1115–1126, 2000. ISSN: 0165- 1684
- [29] Jean-Christophe CEXUS, THT et Transformation de Hough pour la détection de modulations linéaires de fréquence.2012.
- [30] D. Zeng, Automatic modulation classification of radar signals using the Rihaczek distribution and Hough transform. June 2012, *IET Radar Sonar ? Navigation*, 2012. DOI: 10.1049/iet-rsn.2017.0265
- [31] J. P. Kauppi, K. Martikainen, and U. Ruotsalainen, "Hierarchical classification of dynamically varying radar pulse repetition interval modulation patterns," *Neural Networks*, **23**(10), 1226–1237, 2010.
- [32] J. Han, M.-H. He, Y.-Q. Zhu, and J. Wang, "Sorting radar signal based on the resemblance coefficient of bispectrum two dimensions characteristic," *Chinese Journal of Radio Science*, **24**(5), 286–294, 2009.
- [33] H. Zhu, N. Lin, H. Leung, R. Leung and S. Theodosis, "Target Classification From SAR Imagery Based on the Pixel Grayscale Decline by Graph Convolutional Neural Network," in *IEEE Sensors Letters*, **4**(6), 1- 4, June 2020, DOI: 10.1109/LESENS.2020.2995060
- [34] W. J. Zhang, F. H. Fan, and Y. Tan, "Application of cluster method to radar signal sorting," *Radar Science and Technology*, **4**, 219–223, 2004. <https://doi.org/10.1007/s11460-007-0062-3>
- [35] "IEEE Colloquium on 'The Application of Artificial Intelligence Techniques to Signal Processing' (Digest No.42)," *IEEE Colloquium on Application of Artificial Intelligence Techniques to Signal Processing*, London, UK, 1989.
- [36] A.S. Ahmad, "Brain inspired cognitive artificial intelligence for knowledge extraction and intelligent instrumentation system," 2017 International Symposium on Electronics and Smart Devices (ISESD), Yogyakarta, 2017. DOI: 10.5772/intechopen.72764
- [37] Veenman, C. J., Hendriks, E. A., & Reinders, M. J. T. (n.d.). A fast and robust point tracking algorithm. *Proceedings 1998 International Conference on Image Processing. ICIP98*, 2010. DOI:10.1109/icip.1998.999051
- [38] Bhagya Hettige, Hansika Hewamalage, Chathuranga Rajapaksha, Nuwan Wijirasena, Akila Pemasiri and Indika Perera. Evaluation of feature-based object identification for augmented reality applications on mobile devices , Conference: 2015 IEEE 10th International Conference on Industrial and Information Systems (ICIIS), 2015. DOI: 10.1109/ICIINFS.2015.7399005
- [39] Nan Luo, Quansen Sun , Qiang Chen, Zexuan Ji, Deshen Xia. A Novel Tracking Algorithm via Feature Points Matching, 2016. <https://doi.org/10.1371/journal.pone.0116315>
- [40] JiangxiongFang, Efficient and robust fragments-based multiple kernels tracking. *AEU - International Journal of Electronics and Communications*. 2011. <https://doi.org/10.1016/j.aue.2011.02.013>
- [41] D. Comaniciu, V. Ramesh and P. Meer, "Kernel-based object tracking," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **25**(5), 564-577, May 2003, DOI: 10.1109/TPAMI.2003.1195991.
- [42] Haihong Zhang, Zhiyong Huang, Weimin Huang and Liyuan Li, "Kernel-based method for tracking objects with rotation and translation," *Proceedings of the 17th International Conference on Pattern Recognition*, 2004. *ICPR 2004.*, Cambridge, 2004. DOI: 10.1109/ICPR.2004.1334362.
- [43] A. Mondal, S. Ghosh and A. Ghosh, "Efficient silhouette based contour tracking," 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Mysore, 2013. DOI: 10.1109/ICACCI.2013.6637451.
- [44] N. R. Howe, "Silhouette Lookup for Automatic Pose Tracking," 2004 Conference on Computer Vision and Pattern Recognition Workshop, Washington, DC, USA, 2004. DOI: 10.1109/CVPR.2004.438.
- [45] G. Boudoukh, I. Leichter and E. Rivlin, "Visual tracking of object silhouettes," 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, 2009. DOI: 10.1109/ICIP.2009.5414280.
- [46] A. V. Makarenko, Deep learning algorithms for signal recognition in long perimeter monitoring distributed fiber optic sensors, 26th International Workshop on Machine Learning for Signal Processing (MLSP). *IEEE* 2016. DOI: 10.1109/MLSP.2016.7738863
- [47] F. Fioranelli, M. Ritchie, H. Griffiths. Classification unarmed/armed personnel using the NetRAD multistatic radar for micro-Doppler and singular value decomposition features. *IEEE Geosci. Remote Sens. Lett.* 2015. DOI:10.1109/LGRS.2018.2806940
- [48] J. Zabalza, C. Clemente, G. Di Caterina, J. Ren, J. J. Soraghan, S. Marshall. Robust PCA micro-Doppler classification using SVM on embedded systems. *IEEE Trans. Aerosp. Electron. Syst.* 2014. <https://doi.org/10.1109/TAES.2014.130082>.
- [49] S. Ruhela, "Thematic Correlation of Human Cognition and Artificial Intelligence," 2019 Amity International Conference on Artificial Intelligence (AICAI), Dubai, United Arab Emirates, 2019.
- [50] Doris Hooi-Ten Wong and S. Manickam, "Intelligent Expertise Classification approach: An innovative artificial intelligence approach to accelerate network data visualization," 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), Chengdu, 2010. DOI : 10.1109/ICACTE.2010.5579790
- [51] J. Mestoui et al. "Performance analysis of CE-OFDM-CPM Modulation using MIMO system over wireless channels. *J Ambient Intell Human Comput* 2019. <https://doi.org/10.1007/s12652-019-01628-0>
- [52] A. Maroua and F. Mohammed. Characterization of Ultra Wide Band indoor propagation. 2019 7th Mediterranean Congress of Telecommunications (CMT), Fès, Morocco, 2019. DOI: 10.1109 / CMT.2019.8931367
- [53] D. DAGHOUJ, et al. "UWB Coherent Receiver Performance in a Vehicular Channel", *International Journal of Advanced Trends in Computer Science and Engineering*, **9**, No 2, 2020.
- [54] M. Abdellaoui et al. "Study and design of a see-through wall imaging radar system. In *Embedded Systems and Artificial Intelligence. Advances in Intelligent Systems and Computing*, 1076. Springer, Singapour, 2020. [https://doi.org/10.1007/978-981-15-0947-6\\_15](https://doi.org/10.1007/978-981-15-0947-6_15)
- [55] C. Geng, S. Huang and S. Chen, "Recent Advances in Open Set Recognition: A Survey," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020. DOI: 10.1109/TPAMI.2020.2981604.
- [56] Cruciani, F., Vafeiadis, A., Nugent, C. et al. Feature learning for Human Activity Recognition using Convolutional Neural Networks. *CCF Trans. Pervasive Comp. Interact.* **2**, 18–32 (2020). <https://doi.org/10.1007/s42486-020-00026-2>
- [57] J. Lin, W. Kuo, Y. Huang, T. Jong, A. Hsu and W. Hsu, "Using Convolutional Neural Networks to Measure the Physiological Age of *Caenorhabditis elegans*," in *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2020. DOI: 10.1109/TCBB.2020.2971992.
- [58] F. Hussain, R. Hussain, S. A. Hassan and E. Hossain, "Machine Learning in IoT Security: Current Solutions and Future Challenges," in *IEEE Communications Surveys & Tutorials*, 2020. DOI: 10.1109/COMST.2020.2986444
- [59] A. Samat, P. Du, S. Liu, J. Li and L. Cheng, "\$\{\rm E\}^{\{2\}}\{\rm LMs\}\$ : Ensemble Extreme Learning Machines for Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **7**(4), 1060-1069, April 2014. DOI : 10.1109/JSTARS.2014.2301775
- [60] L. Bruzzone and M. Marconcini, "Domain Adaptation Problems: A DASVM Classification Technique and a Circular Validation Strategy," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **32**(5), 770-787, May 2010. DOI: 10.1109/TPAMI.2009.57
- [61] ES-SAQY et al., "Very Low Phase Noise Voltage Controlled Oscillator for 5G mm-wave Communication Systems," 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, 2020. DOI: 10.1109/IRASET48871.2020.9092005
- [62] S. Ubik and P. Žejdl, "Evaluating Application-Layer Classification Using a Machine Learning Technique over Different High Speed Networks," 2010 Fifth International Conference on Systems and Networks Communications, Nice, 2010. DOI: 10.1109/ICSN.2010.66
- [63] H. I. Bulbul and Ö. Unsal, "Comparison of Classification Techniques used in Machine Learning as Applied on Vocational Guidance Data," 2011 10th International Conference on Machine Learning and Applications and Workshops, Honolulu, HI, 2011. DOI: 10.1109/ICMLA.2011.49
- [64] L. Bruzzone and C. Persello, "Recent trends in classification of remote sensing data: active and semisupervised machine learning paradigms," 201 IEEE International Geoscience and Remote Sensing Symposium, Honolulu, HI, 2010. DOI: 10.1109/IGARSS.2010.5651236
- [65] F. Miao, P. Zhang, L. Jin and H. Wu, "Chinese News Text Classification Based on Machine Learning Algorithm," 2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Hangzhou, 2018. DOI: 10.1109/IHMSC.2018.10117
- [66] M. Ross, C. A. Graves, J. W. Campbell and J. H. Kim, "Using Support Vector Machines to Classify Student Attentiveness for the Development of Personalized Learning Systems," 2013 12th International Conference on Machine Learning and Applications, Miami, FL, 2013 .DOI: 10.1109/ICMLA.2013.66
- [67] H. Guo, W. Wang and C. Men, "A novel learning model-Kernel Granular Support Vector Machine," 2009 International Conference on Machine Learning and Cybernetics, Hebei, 2009. DOI: 10.1109/ICMLC.2009.5212413
- [68] Z. Kun, T. Ying-jie and D. Nai-yang, "Unsupervised and Semi-Supervised Two-class Support Vector Machines," Sixth IEEE International Conference on Data Mining - Workshops (ICDMW'06), Hong Kong, 2006, 813-817. DOI: 10.1109/ICDMW.2006.164
- [69] M. Dharani and M. Sivachitra, "Motor imagery signal classification using semi supervised and unsupervised extreme learning machines," 2017 International Conference on Innovations in Information, Embedded and



- Communication Systems (ICIECS), Coimbatore, 2017. DOI: 10.1109/ICIECS.2017.8276131
- [70] K. Zhao, Y. Liu and N. Deng, "Unsupervised and Semi-supervised Lagrangian Support Vector Machines with Polyhedral Perturbations," 2009 Third International Symposium on Intelligent Information Technology Application, Shanghai, 2009. DOI: 10.1109/IITA.2009.200
- [71] T. D. Bui, D. K. Nguyen and T. D. Ngo, "Supervising an Unsupervised Neural Network," 2009 First Asian Conference on Intelligent Information and Database Systems, Dong Hoi, 2009. DOI: 10.1109/ACIDS.2009.92
- [72] S. Hussein, P. Kandel, C. W. Bolan, M. B. Wallace and U. Bagci, "Lung and Pancreatic Tumor Characterization in the Deep Learning Era: Novel Supervised and Unsupervised Learning Approaches," in IEEE Transactions on Medical Imaging, **38**(8), 1777-1787, Aug. 2019. DOI: 10.1109/TMI.2019.2894349
- [73] C. Lee and H. Yang, "Implementation of Unsupervised and Supervised Learning Systems for Multilingual Text Categorization," Fourth International Conference on Information Technology (ITNG'07), Las Vegas, NV, 2007. DOI: 10.1109/ITNG.2007.107
- [74] L. Kong, G. Huang, K. Wu, Q. Tang and S. Ye, "Comparison of Internet Traffic Identification on Machine Learning Methods," 2018 International Conference on Big Data and Artificial Intelligence (BDAl), Beijing, 2018. DOI: 10.1109/BDAl.2018.8546682
- [75] R. Nijhawan, I. Srivastava and P. Shukla, "Land cover classification using super-vised and unsupervised learning techniques," 2017 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, 2017. DOI: 10.1109/ICCIDS.2017.8272630
- [76] Y. Ji, S. Yu and Y. Zhang, "A novel Naive Bayes model: Packaged Hidden Naive Bayes," 2011 6th IEEE Joint International Information Technology and Artificial Intelligence Conference, Chongqing, 2011. DOI: 10.1109/ITAIC.2011.6030379
- [77] A. H. Jahromi and M. Taheri, "A non-parametric mixture of Gaussian naive Bayes classifiers based on local independent features," 2017 Artificial Intelligence and Signal Processing Conference (AISP), Shiraz, 2017. DOI: 10.1109/AISP.2017.8324083
- [78] J. Chen, Z. Dai, J. Duan, H. Matzinger and I. Popescu, "Naive Bayes with Correlation Factor for Text Classification Problem," 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), Boca Raton, FL, USA, 2019. DOI: 10.1109/ICMLA.2019.00177
- [79] A. Tripathi, S. Yadav and R. Rajan, "Naive Bayes Classification Model for the Student Performance Prediction," 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT), Kannur, Kerala, India, 2019. DOI: 10.1109/ICICT46008.2019.8993237
- [80] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang and P. S. Yu, "A Comprehensive Survey on Graph Neural Networks," in IEEE Transactions on Neural Networks and Learning Systems, 2020. DOI: 10.1109/TNNLS.2020.2978386.
- [81] Z. Zhang, P. Cui and W. Zhu, "Deep Learning on Graphs: A Survey," in IEEE Transactions on Knowledge and Data Engineering, 2020. DOI: 10.1109/TKDE.2020.2981333.
- [82] A. Khamparia, G. Saini, D. Gupta, "Seasonal Crops Disease Prediction and Classification Using Deep Convolutional Encoder Network. Circuits Syst Signal Process **39**, 818–836 (2020). <https://doi.org/10.1007/s00034-019-01041-0>
- [83] S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey," in IEEE Transactions on Affective Computing, 2020. DOI: 10.1109/TAFFC.2020.2981446
- [84] D. Goularas and S. Kamis, "Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data," 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), Istanbul, Turkey, 2019. DOI: 10.1109/Deep-ML.2019.00011
- [85] M.-B. Li, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "Fully complex extreme learning machine," Neurocomputing, **68**(1-4), 306–314, 2005. View at: Publisher Site | Google Scholar. 2005.
- [86] N. Liang, G. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," IEEE Transactions on Neural Networks and Learning Systems, **17**(6), 1411–1423, 2006. View at: Publisher Site | Google Scholar. 2006
- [87] H. Wang, W. Fan, F. Sun, and X. Qian, "An adaptive ensemble model of extreme learning machine for time series prediction," in Proceedings of the 12th International Computer Conference on Wavelet Active Media Technology and Information Processing, ICCWAMTIP), 80–85, Chengdu, China, December 2015. View at: Publisher Site | Google Scholar. 2015
- [88] G. Huang, L. Chen, and C. Siew, "Universal approximation using incremental constructive feedforward networks with random hidden nodes," IEEE Transactions on Neural Networks and Learning Systems, **17**(4), 879–892, 2006. View at: Publisher Site | Google Scholar. 2006.
- [89] H.-J. Rong, Y.-S. Ong, A.-H. Tan, and Z. Zhu, "A fast pruned-extreme learning machine for classification problem," Neurocomputing, **72**(1–3), 359–366, 2008. View at: Publisher Site | Google Scholar. 2008.
- [90] G. G. Wang, M. Lu, Y. Q. Dong, and X. J. Zhao, "Self-adaptive extreme learning machine," Neural Comput Applic, **27**, 291–303, 2016. View at: Publisher Site | Google Scholar. 2016.
- [91] Cityscape website. [Online]. Available: <https://www.cityscapes-dataset.com> [Accessed: 08.06.2020]
- [92] ImageNet website. [Online]. Available: <http://image-net.org> [Accessed: 08.06.2020]
- [93] CIFAR-10 website. [Online]. Available: <https://www.cs.toronto.edu/~kriz/cifar.html> [Accessed: 08.06.2020]
- [94] Caltech101 website. [Online]. Available: [http://www.vision.caltech.edu/Image\\_Datasets/Caltech101](http://www.vision.caltech.edu/Image_Datasets/Caltech101) [Accessed: 08.06.2020]
- [95] website: <https://sites.google.com/view/ihsen-alouani/datasets?authuser=0> [Accessed: 08.06.2020]
- [96] website: <https://catalog.data.gov/dataset/traffic-and-pedestrian-signals> [Accessed: 06.06.2020]
- [97] website: <http://coding-guru.com/popular-pedestrian-detection-datasets/>
- [98] website: <https://www.mathworks.com/supportfiles/SPT/data/PedBicCarDat.zip>
- [99] D. Aksu and M. Ali Aydin, "Detecting Port Scan Attempts with Comparative Analysis of Deep Learning and Support Vector Machine Algorithms," 2018 International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT), ANKARA, Turkey, 2018. DOI: 10.1109/IBIGDELFT.2018.8625370
- [100] Z. Wang, J. Chen and S. C. H. Hoi, "Deep Learning for Image Super-resolution: A Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020. DOI: 10.1109/TPAMI.2020.2982166.
- [101] G. Zhu, D. Liu, Y. Du, C. You, J. Zhang and K. Huang, "Toward an Intelligent Edge: Wireless Communication Meets Machine Learning," in IEEE Communications Magazine, **58**(1), 19-25, January 2020, DOI: 10.1109/MCOM.001.1900103
- [102] A. Rehman, S. Naz, M.I. Razzak, "A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning", Circuits Syst Signal Process **39**, 757–775 (2020). <https://doi.org/10.1007/s00034-019-01246-3>
- [103] M.S.Z. Rizvi, "Learn Image Classification on 3 Datasets using Convolutional Neural Networks (CNN). FEBRUARY 18, 2020
- [104] S. M. A. Bhuiyan and J. F. Khan, "A Simple Deep Learning Network for Target Classification," 2019 SoutheastCon, Huntsville, AL, USA, 2019. DOI: 10.1109/SoutheastCon42311.2019.9020517
- [105] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang and P. S. Yu, "A Comprehensive Survey on Graph Neural Networks," in IEEE Transactions on Neural Networks and Learning Systems, 2020. DOI: 10.1109/TNNLS.2020.2978386.
- [106] Elly C. Knight, Sergio Poo Hernandez, Erin M. Bayne, Vadim Bulitko & Benjamin V. Tucker (2020) Pre-processing spectrogram parameters improve the accuracy of bioacoustic classification using convolutional neural networks, Bioacoustics, 29:3, 337-355, 2020. DOI: 10.1080/09524622.2019.1606734
- [107] R. Kountchev and R. Kountcheva, "Truncated Hierarchical SVD for image sequences, represented as third order tensor," 2017 8th International Conference on Information Technology (ICIT), Amman, 2017. DOI: 10.1109/ICITECH.2017.8079995
- [108] B. Lei, I. Y. Soon and E. Tan, "Robust SVD-Based Audio Watermarking Scheme With Differential Evolution Optimization," in IEEE Transactions on Audio, Speech, and Language Processing, **21**(11), 2368-2378, Nov. 2013. DOI: 10.1109/TASL.2013.2277929
- [109] R. Ibrahim, J. Alirezaie and P. Babyn, "Pixel level jointed sparse representation with RPCA image fusion algorithm," 2015 38th International Conference on Telecommunications and Signal Processing (TSP), Prague, 2015. DOI: 10.1109/TSP.2015.7296332
- [110] Dong Yang, Guisheng Liao, Shengqi Zhu and Xi Yang, "RPCA based moving target detection in strong clutter background," 2015 IEEE Radar Conference (RadarCon), Arlington, VA, 2015. DOI: 10.1109/RADAR.2015.7131231
- [111] Y. Guo, G. Liao, J. Li and T. Gu, "An Improved Moving Target Detection Method Based on RPCA for SAR Systems," IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019. DOI: 10.1109/IGARSS.2019.8900206
- [112] H. Sehat and P. Pahlevani, "An Analytical Model for Rank Distribution in Sparse Network Coding," in IEEE Communications Letters, **23**(4), 556-559, April 2019. DOI: 10.1109/LCOMM.2019.2896626
- [113] L. Zhang and C. Ma, "Low-rank, sparse matrix decomposition and group sparse coding for image classification," 2012 19th IEEE International Conference on Image Processing, Orlando, FL, 2012. DOI:



- 10.1109/ICIP.2012.6466948.
- [114] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," in *IEEE Transactions on Information Theory*, **45**(2), 399-431, March 1999. DOI: 10.1109/18.748992
- [115] X. Wang and W. Wang, "MRI Brain Image Classification Based on Improved Topographic Sparse Coding," 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 2018. DOI: 10.1109/ICSESS.2018.8663765
- [116] N. Zhong, T. Yan, W. Yang and G. Xia, "A supervised classification approach for PolSAR images based on covariance matrix sparse coding," 2016 IEEE 13th International Conference on Signal Processing (ICSP), Chengdu, 2016. DOI: 10.1109/ICSP.2016.7877826
- [117] M. S. Packianather and B. Kapoor, "A wrapper-based feature selection approach using Bees Algorithm for a wood defect classification system," 2015 10th System of Systems Engineering Conference (SoSE), San Antonio, TX, 2015), 498-503. DOI: 10.1109/SYSE.2015.7151902
- [118] H. Chen, R.A. Goubran, T. Mussivand, "Multi-Dimension Combining (MDC) in abstract Level and Hierarchical MDC (HMDC) to Improve the Classification Accuracy of Enoses," 2005 IEEE Instrumentation and Measurement Technology Conference Proceedings, Ottawa, Ont., 2005. DOI: 10.1109/IMTC.2005.1604204
- [119] H. Chen, R.A. Goubran, T. Mussivand, "Improving the classification accuracy in electronic noses using multi-dimensional combining (MDC)," *SENSORS*, 2004 IEEE, Vienna, 2004 .DOI: 10.1109/ICSENS.2004.1426233
- [120] K. Wang, Y. Wan and S. Shen, "Classifications of remote sensing images using fuzzy multi-classifiers," 2009 IEEE International Conference on Intelligent Computing and Intelligent Systems, Shanghai, 2009. DOI: 10.1109/ICICISYS.2009.5357646
- [121] A. Yadav and A. Swetapadma, "Combined DWT and Naive Bayes based fault classifier for protection of double circuit transmission line," International Conference on Recent Advances and Innovations in Engineering (ICRAIE- 2014), Jaipur, 2014. DOI: 10.1109/ICRAIE.2014.6909179
- [122] A. Sharmila and P. Geethanjali, "DWT Based Detection of Epileptic Seizure From EEG Signals Using Naive Bayes and k-NN Classifiers," in *IEEE Access*, **4**), 7716-7727, 2016. DOI: 10.1109/ACCESS.2016.2585661
- [123] L. Li, Y. Zhang, W. Chen, S. K. Bose, M. Zukerman and G. Shen, "Naive Bayes Classifier-Assisted Least Loaded Routing for Circuit-Switched Networks," in *IEEE Access*, **7**, 11854-11867, 2019. DOI: 10.1109/ACCESS.2019.2892063
- [124] B. Harender and R. K. Sharma, "DWT based epileptic seizure detection from EEG signal using k-NN classifier," 2017 International Conference on Trends in Electronics and Informatics (ICEI), Tirunelveli, 2017. DOI: 10.1109/ICOEI.2017.8300806
- [125] M. Sangeetha and B. N. Keshavamurthy, "Conditional Mutual Information based Attribute Weighting on General Bayesian Network Classifier," 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida (UP), India, 2018. DOI: 10.1109/ICACCCN.2018.8748481
- [126] B. H. Bei Hui, Y. W. Yue Wu, L. J. Lin Ji and J. C. Jia Chen, "A NB-based approach to anti-spam application: DLB Classification Model," 2006 Semantics, Knowledge and Grid, Second International Conference on, Guilin, 2006. DOI: 10.1109/SKG.2006.10
- [127] Z. Wu, B. Zhang, Y. Zhu, W. Zhao and Y. Zhou, "Transformer Fault Portfolio Diagnosis Based on the Combination of the Multiple Baye SVM," 2009 International Conference on Electronic Computer Technology, Macau, 2009. DOI: 10.1109/ICECT.2009.103
- [128] M. S. Acharya, A. Armaan and A. S. Antony, "A Comparison of Regression Models for Prediction of Graduate Admissions," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019. DOI: 10.1109/ICCIDS.2019.8862140
- [129] M. Yang, K. Wu and J. Hsieh, "Mountain C-Regressions in Comparing Fuzzy C-Regressions," 2007 IEEE International Fuzzy Systems Conference, London, 2007. DOI: 10.1109/FUZZY.2007.4295366
- [130] J. Y. Goulermas, P. Liatsis, X. Zeng and P. Cook, "Density-Driven Generalized Regression Neural Networks (DD-GRNN) for Function Approximation," in *IEEE Transactions on Neural Networks*, **18**(6), 1683-1696, Nov. 2007. DOI: 10.1109/TNN.2007.902730
- [131] S. Kavitha, S. Varuna, R. Ramya, "A comparative analysis on linear regression and support vector regression," 2016 Online International Conference on Green Engineering and Technologies (IC-GET), Coimbatore, 2016. DOI: 10.1109/GET.2016.7916627
- [132] T. Qunzhu, Z. Rui, Y. Yufei, Z. Chengyao and L. Zhijun, "Improvement of Random Forest Cascade Regression Algorithm and Its Application in Fatigue Detection," 2019 IEEE 2nd International Conference on Electronics Technology (ICET), Chengdu, China, 2019. DOI: 10.1109/ELTECH.2019.8839317
- [133] M. Khodjet-Kesba, "Automatic target classification based on radar backscattered ultra wide band signals. Other. Université Blaise Pascal - Clermont-Ferrand II, 2014.
- [134] M. Fattah, et al. "Multi Band OFDM Alliance Power Line Communication System", *Procedia Computer Science*, **151**, 2019, <https://doi.org/10.1016/j.procs.2019.04.146>.
- [135] V. Kumar X. Wu and coll. Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14:137, 2007.
- [136] G. Li, P. Niu, Y. Ma, H. Wang, W. Zhang, Tuning extreme learning machine by an improved artificial bee colony to model and optimize the boiler efficiency, *Knowl.-Based Syst.* **67** , 2014.

## Assessing the Operator's Readiness to Perform Tasks of Controlling by the Unmanned Aerial Platforms

Dmytro Kucherov<sup>1\*</sup>, Olha Sushchenko<sup>2</sup>, Andrii Kozub<sup>3</sup>, Volodymyr Nakonechnyi<sup>4</sup>

<sup>1</sup>Department of Computerized Control Systems, National Aviation University, Kyiv, 03058, Ukraine

<sup>2</sup>Department of Aerospace Control Systems, Institute, National Aviation University, Kyiv, 03058, Ukraine

<sup>3</sup>National Space Facilities Control and Test Center, State Space Agency, Kyiv, 01010, Ukraine

<sup>4</sup>Department of Cyber Security and Information Protection, Taras Shevchenko National University, Kyiv, 01033, Ukraine

---

### ARTICLE INFO

Article history:

Received: 09 July, 2020

Accepted: 19 July, 2020

Online: 09 August, 2020

---

Keywords:

Operator of unmanned aerial platform

Assessment criteria

Quality of operator activity

Method of hierarchical analysis

---

---

### ABSTRACT

Together with the intensity of development in the field of technology of unmanned platforms and their effective use for solving various tasks of peacetime and war, the requirements for the training of the operator managing the platform also increase. This fully applies to personnel providing the flight of manned means. Nevertheless, there are significant differences in the requirement of operator preparedness for an unmanned platform. This is, first of all, control in conditions of remoteness from an unmanned platform, orientation on the display of the control panel, delays in the passage of information, and possibly a complete loss of communication. In such conditions, the requirements for the reaction time of the operator increase, he must also have the ability to anticipate the development of the situation, be able to work with available equipment for a long time. These and several other criteria determine the general criterion of operator productivity, which is introduced in the work. The productivity criterion is a linear convolution function of particular criteria with some weighting factors, the exact values of which are unknown. A detailed analysis of particular criteria made it possible to establish their inconsistency and heterogeneity. The article proposes an approach that allows us to eliminate the inconsistency of local criteria by separately calculating weight coefficients for each part based on the hierarchy analysis method. The basic properties of the proposed approach are also given; modeling confirms the correctness of the solutions. This approach can be useful in the certification of operators of various fields of activity.

---

### 1. Introduction

Modern successes in the field of microelectronics and electromechanics have made it possible to create unmanned aerial vehicles that are effectively used to solve various problems [1]. The control of an unmanned vehicle can be performed manually, automatically and semi-automatically, where the role of the human as the machine operator is very important both in preparing and conducting the flight. The operator's remoteness from the control means creates several difficulties in flight control, primarily related to control delay, inadequate presentation of information about the environment, and the dynamics of the control process itself.

A typical flight control system includes an operator, on-board receiving and control equipment, and a ground-based transmitting station, usually a tablet computer with a software interface, and the aircraft itself. The control of a military unmanned vehicle usually involves two human operators, one of them is responsible for controlling the movement of the vehicle (aircraft operator), and the second ensures the fulfillment of the target task (operator of the target task load). Therefore, the management of an unmanned aerial platform is not a simple task but requires training the operator in skills that will help him cope with the control difficulties in conditions of natural and artificial interference, loss of a communication channel, loss of line of sight, delays and distortions of transmitted information, control of many devices, etc.

---

\*Corresponding Author: Dmytro Kucherov, Dept. of Computerized Control Systems, Faculty of Cybersecurity, Computer and Software Engineering, National Aviation University, Kyiv, Ukraine, [d\\_kucherov@ukr.net](mailto:d_kucherov@ukr.net)

Completing the target at the terminal point requires significantly more time to prepare the platform for flight in automatic than in manual or semi-automatic modes. These conditions imply the fulfillment of many requirements for the practical skills of operators that can ensure the implementation of the target in real flight conditions.

Our contribution consists of developing the versatile method's assessment of the productivity of operators based on measurement of their indicators such as reaction time, the accuracy of the operation, stress resistance, the fitness of mind and hands, preparedness, and foresight. The proposed method is suitable not only for assessing the productivity of UAV operators but operators of some other industries, where high requirements imposed on the organization of labor.

The rest paper is organized as follows: a literature review is given in the next section, the control system structure and typical operator actions are presented in Section 3, the problem statement is given in Section 4, the solution method is presented in Section 5, numerical examples and discussion of input data are given in Section 6, and conclusions are discussed in Section 7.

## **2. Review of Previous Researches**

This work is an extended version of the report [1] presented at the 5th International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), where a method for evaluating the operator's activity taking into account the study of his psycho-physical reactions is proposed.

An analytical review of the influence of the human factor on the efficiency of using unmanned aerial vehicles (UAVs) in the current and immediate future is presented in [2]. The study emphasizes the need for special training by the UAV human operator, which, in principle, differs from that to which the pilots are involved. The review also notes the need for additional scientific research regarding the skills, knowledge, and abilities of UAV operators and ways to evaluate them.

In [3], the importance of the influence of a human operator on the efficiency and quality of control of unmanned aerial vehicles is noted. The paper notes that about 70% of all UAV-related accidents are caused by erroneous human activities. Despite the high level of automation, a person is not excluded from the operation of the control system, but rather, on the contrary, takes a leading place in the preparation, conduct, and implementation of tasks, eliminating the risks associated with the death of the pilot. The presence of several operators involves the coordination of their actions and interaction to achieve the ultimate goal of using one or a group of devices.

In [4-8], factors are considered that affect the preparation and quality of exploitation of unmanned platforms by operators. So, the effect of spatial capabilities obtained by the operator through the use of UAVs was evaluated in [4]. A sense of space gives confidence to the operator when acting in whichever situation at any time of the day and for any type of purpose.

In [5], a simulator description is proposed for a UAV operator with wide simulation capabilities, namely: flight control, wireless communications, and ground equipment. However, the article does

not provide the parameters and methods for evaluating the operator.

In [6], the UAV operator errors were presented in the form of a chi-square curve of the Pearson distribution, obtained by constructing a histogram of the probability density of measurements of errors in operator activity.

The authors of [7] propose evaluating the capabilities of the UAV operator according to the profile generated by special software. As estimates, metrics are proposed that are based on a combination of fuzzy logic methods and cluster analysis. The proposed techniques allow us to evaluate the operator not only of a single device but when managing a group.

Errors of a technical nature were analyzed in [8]. It indicates the technical problems that the operator encounters, namely, the limited field of view, delays in the transfer of information, poor operator understanding of the situation in which the unmanned vehicle is located, and errors on-board equipment.

The methodical complex for accelerated training of UAV operators is presented in [9], the components of which are control situations with corresponding difficulties and the intensity of their implementation for a certain time. It is assumed that during the preparation of the operator by this complex, the generalized indicator of operator readiness assumes some steady-state value.

In the study the authors of [10] the temporal norms of operator training for the formation of the required psycho-physiological characteristics of personnel for unmanned aviation systems. The study relies on training programs for the air force. An assessment of the quality of preparedness of operators and methods for their certification in the study is not considered.

The usefulness of practical developments aimed at training operators who are ready to work in conditions when the equipment is on the verge of failure is proposed in [11].

In [12], a system for modeling dynamics is given, which allows one to analyze factors affecting the training of the professional skills of coal mine operators. The work explores operator training over a long time. Also, priority is determined, an important place among which is the ability to work in difficult conditions.

For some industries, it is important to maintain operator productivity in conditions when the order of actions changes [13, 14]. In this case, there is a need to redistribute the procedure and change the training schedule, and then it is useful a loss training procedure is proposed. The authors also introduced an indicator of operator experience [13]. The skill indicator in the study is represented by a multi-purpose function that takes into account the skill indicator and working time. In research [14] proposes an additional indicator, the standardization point.

The Italian concern Terna [15] has developed three main programs for training and certification of operators in real-time: the initial program, sliding (continuous) training, and the general program for other interested parties. The training is configured for the specific role of each operator in real-time and covers all its operational activities. Besides, an operating training simulator (OTS) has been developed by the company.

The authors of [16] evaluated the mental workload of operators whose main professional activity is air traffic control and students by comparing continuous exposure close to the infrared range and a mobile touch panel of the same range worn on the operator's head. The study revealed the possibility of using a flexible touch panel to assess the mental workload of operators involved in air traffic control based on the repetition of measurements and the ANOVA technique.

The work [17] presents a methodology for the development of psychomotor skills through the use of teaching means. The results of the study showed a positive influence on the development of psychomotor skills in people of different categories.

The article [18] considers the assessment of human hearing by methods by a system of clinical audiometry. A biotechnological system for assessment and stimulating the human auditory analyzer under the influence of various factors was developed and tested, which allowed us to study the effect of sound frequencies on the patient.

The method of training a group of agents is presented in [19]. Training is based on the introduction of a Q-function, the result of which depends on the joint actions of a pair of agents. Unfortunately, the result is characterized by the limited elements involved in the learning process, and simple scenarios are considered.

In [20], [21], the UAV control system is studied, which can become the basis for the construction of a simulator designed for training, preparing, and certification of operators.

Based on the papers review, we can conclude that there is strong usefulness in developing and implement appropriate methods that will make the operator certification process more accurate.

### 3. Features of the Controlling of Unmanned Platform

The section briefly discusses the generalized structural composition of a typical unmanned platform, the main problems of flight control, and the requirements for operator actions.

#### 3.1. System structure

The structure of an automated UAV control system with a human operator in the control loop has the form shown in Figure 1. The operator's workplace is a remote control that provides several controls that are available on the display screen.

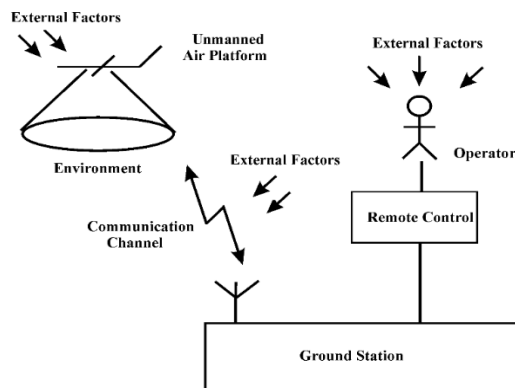


Figure 1: The structure of an automated control system of UAV

#### 3.2. Flight Control Problems

A vulnerable place to control the UAV platform is the flight control channel. The problem here is the limited bandwidth of the radio control channel since the narrow-band path is considered susceptible to interfering signals that occur in the radio channel. In-flight, communication losses are also possible; in places with strong reflections, erroneous or false signals and radio control commands can be received. Besides, the complexity of radio command control lies in the low data rate.

#### 3.3. Operator Actions

The operator controlling the flight of the UAV must know the modes and parameters of the control channel that ensure a stable flight of the platform; be able to make adjustments to flight control; know the conditions for switching manual and semi-automatic modes to autopilot; be prepared to exit autopilot with appropriate authorization.

The operator responsible for carrying out the useful task should be well aware of the capabilities of the equipment used, be able to control existing cameras, be able to extract useful information from the received data, and also closely cooperate with the flight control operator.

### 4. Problem Statement

Long-term work with the console puts forward some specific requirements for the operator. The situation displayed on the remote control screen is similar to a computer game, and therefore the operator's actions are similar to the reactions of an experienced gamer. The main ones among them are the reaction to the event, its accuracy, perception, and understanding of the displayed situation with a certain delay, fitness of mind and hands, characterized by resistance to long work; stress resistance, determined by the degree of resistance to negative influences and unexpected situations; foresight is the possibility of seeing or feeling probable future events that may occur.

The following performance indicators of the operator are offered:

- the reaction time  $t$  is considered the main parameter characterizing the behavior of the operator; it is determined by the time the irritant occurs before the reaction to the event;
- the accuracy of the operation  $\varepsilon$ , which is the degree to which the data read by the operator corresponds to the values read from the screen;
- exposure  $\tau$  is the time the operator understands the displayed situation;
- fitness of mind and hands  $v$ , characterized by the frequency of errors that occur in the work of the operator;
- stress resistance  $k$  shows the degree of resistance to negative impacts of pressure factors;
- preparedness  $\Sigma$  is a set of knowledge, skills, and conditions of mental and physiological functions that establish its ability to perform specified actions with a certain quality;
- foresight  $\lambda$  is the ability to see or feel the possible future that will take place soon.



The analysis of the selected indicators shows their heterogeneity. There are time indicators estimated by absolute values, this is the reaction time  $t$  and the exposure time  $\tau$ , measured in seconds. The frequency of errors  $\nu$ , we will measure in relative units, while  $\nu \leq \nu_{max} \leq 1$ , the value  $\nu_{max}$  corresponds to an unacceptable level of errors. In the interest of generality, we write  $\nu_{min} \leq \nu \leq \nu_{max}$ , where the desired value is  $\nu_{min} \approx 0$ . Stress tolerance  $k$ , preparedness  $\Sigma$ , and foresight  $\lambda$  should be evaluated in scores. The accuracy of the operation  $\varepsilon$  when using, for example, a stylus or other manipulator, we will evaluate the distance between the true value and the exact coordinates, in the scale of the display screen this value is measured in millimeters.

Assuming that the real indicators of the operator can be in some intervals allowed by various standards, we will suggest that the measured parameter is in the intervals

$$\underline{\xi} \leq \xi_i \leq \bar{\xi}, \quad (1)$$

where  $\underline{\xi}$ ,  $\bar{\xi}$  are the minimum and maximum values of the measured parameters, and  $i=1..N$ ,  $N$  is the number of indicators.

The heterogeneity of the given parameters is also manifested in the inconsistency of the values of the indicators. For example, the values of reaction time  $t$ , accuracy  $\varepsilon$ , the fitness of mind and hands  $\nu$  should be small, while exposure  $\tau$ , stress resistance  $k$ , preparedness  $\Sigma$ , and foresight  $\lambda$  for a good operator should correspond to higher values in their ranges.

It is proposed to evaluate the generalized indicator of operator suitability by the relation

$$J(\xi_i | i=1..N) \geq J_{min}, \quad (2)$$

where  $J_{min}$  is a certain threshold value of the indicator by which a decision is made on the operator's admission to the activity. We will call the  $J$  as the performance of the operator.

The task is to find a method for determining the value  $J_{min}$  of expression (2), in the presence of measured parameters  $\xi_i$  satisfying inequality (1), which will allow us to establish inequality (2), which will give an objective assessment of the operator's readiness to carry out the task of controlling an unmanned platform taking into account the influence heterogeneities  $\xi_i$ .

### 5. Technique of Solving

The section discusses the approach to calculating the criterion and the methodology for calculating operator productivity.

#### 5.1. Operator Assessment Criteria

The operator's productivity estimation based on the introduced set of indicators will be calculated in the form of a linear function that written as follow

$$J = \sum_{i=1}^N a_i \sigma_i, \quad (3)$$

where  $a_i$  is the  $i$ th weight coefficient,  $\sigma_i$  is the  $i$ th normalized parameter, i.e.

$$\sigma_i = \frac{\bar{\xi} - \xi_i}{\xi - \bar{\xi}}, \quad (4)$$

where  $\xi_i$  is the measured value of the parameter,  $i = 1, \dots, N$ . The restriction is imposed on the weighting coefficients

$$\sum_{i=1}^N a_i = 1. \quad (5)$$

Now the task of assessing the quality of operator training is reduced to determining the value of  $J_{min}$  with unknown weights  $a_i$ .

Since the introduced indicators have different effects on the value of  $J_{min}$ , and for the most part are contradictory, for example, indicators such as reaction time  $t$ , accuracy (error)  $\varepsilon$  and fitness  $\nu$  must be small and the others should be as large as possible, it makes sense to separate these factors into two opposite groups, and solve the maximization problem taking into consideration the equal contribution of each component, i.e. define a threshold value as

$$J_{min} = \max_{a_i} \left( \frac{1}{m_1} \sum_{i=1}^{m_1} a_i \sigma_i^+ + \frac{1}{m_2} \sum_{i=m_1+1}^{m_2} a_i \sigma_i^- \right), \quad (6)$$

where  $m_1 + m_2 = N$ , in the case the considered parameters  $N=7$ .

Expressions (3), (4), (6), and taking into account (1) - (5) as well, provide finding the threshold  $J_{min}$ .

#### 5.2. Calculation Approach

The computational procedure involves this sequence of steps

1. Measured are the values of the parameter vector  $\xi$  satisfying inequality (1).
2. We calculate the values of the coefficients  $a_i$  by the method of hierarchical analysis, including the preparation of the matrix  $A$  of pairwise comparisons of size  $p \times p$ , where  $p$  is the number of criteria. We set the diagonal elements of the matrix  $A$  equal to 1, calculate the values of the elements  $a_{ij}$ , taking into account the importance of the measured parameter so that the parameter takes a value in the range, for example, from 1 to 9. In case, when we deal with the most important parameter it takes the value 9, otherwise the least important will be 1. The element  $a_{ji} = 1 / a_{ij}$ .
3. Next, the fulfillment of the consistency condition for the matrix of pairwise comparisons is checked by calculating the consistency coefficient CR

$$CC = \frac{p(p_{max} - p)}{1.98(p-1)(p-2)}. \quad (7)$$

If  $CC \leq 0.1$ , the level of consistency of matrix  $A$  is considered acceptable; otherwise, matrix  $A$  is checked.

4. Determine the value of  $J_{min}$  (6).

5. We measure the parameters of the vector  $\xi$  for a particular human operator, apply (2), (1) and (4).

Corollary 1. The quantity  $J$  satisfies the condition  $0 < J \leq J_{max}$ .

Corollary 2. The value  $J_{max} = 1$ .

**6. Simulation**

An important point in modeling is the selection of suitable parameters and the evaluation of the results. These issues are addressed in this section.

*6.1. Some information about parameter values*

The reaction time of the operator is determined by the delay times of the signal  $t_i$  in the  $i$ th control device and the delay of the signal from the moment the operator recognizes the situation to the corresponding reaction  $t_{del}$ , or

$$t = \sum_{i=1}^n t_i + t_{del}, \tag{8}$$

where  $n$  is the total number of controls. The delay entered by the operator far exceeds the control delay of the computer. This value is in the range of 100 ... 500 ms, which corresponds to the reaction of the gamer. The best indicator is the smaller value.

The exposure time  $\tau$  of the operator’s eye is also determined by two factors: the ability to notice a new luminous object and the time of awareness of this object. A trained eye is capable of detecting flashes of light lasting 5-10 ms, however, this parameter depends on the surrounding light. American scientist B. Libet found that conscious human activity manifests itself after 300-500 ms of the fact that he decided to move his hand [22].

The accuracy  $\varepsilon$  of the operator’s actions is determined by the thickness of the user device and the size of the spot on the screen, which the operator must reliably overlap by the device used. The operator’s action will be performed in case of a 50% capture of the screen spot with the stylus tip. For example, if the diameter of the stylus tip is 1.9 - 2 mm, then approximately 1 mm is the accuracy of the operator [23].

The level of operator errors  $v$  depends on the degree of preparedness, workload, and fatigue due to long work. According to experimental data obtained during the processing of visual signals [24], the following results were obtained: for beginners who do not have the experience, the error rate is 0.27, for operators who have completed 4 or more training, this indicator is 0.018. Since there are, as a rule, no untrained operators, the maximum value of the indicator can be taken as 0.15.

Indicators evaluated on a scale of points, which include stress resistance  $k$ , preparedness  $\Sigma$ , foresight  $\lambda$ , can also be assessed on a categorical scale in terms: excellent, good, satisfactory, poor. Having chosen this scale as a basis, it is possible to make such a correspondence. The category “excellent” is estimated at 8-9 scores, “good” - 5-7 scores, “satisfactory” - 3-4 scores, “poor” - 1-2 scores.

*6.2. Numerical examples*

The initial data accepted as the norm are given in Table I. Following the procedure described in the solution section and formulas (3), (4), (6) for  $N = 7$ , we found weight coefficients that satisfy the consistency condition for the initial matrices of pairwise comparisons (Table 1).

Table 1: Normal Data

para meter	$t$	$\varepsilon$	$\tau$	$v$	$k$	$\Sigma$	$\lambda$
value	.25 s	.9 mm	.4 s	.07	3	3	2
$\sigma_i$	.625	.5	.5	.6	.25	.25	.25

$$a_t = 0.1017, a_\varepsilon = 0.1835, a_\tau = 0.2602,$$

$$a_v = 0.4545, a_k = 0.5972, a_\Sigma = 0.2847, a_\lambda = 0.1181$$

Using the hierarchy analysis method [25], weights were determined that satisfy the consistency condition for the initial matrices of pairwise comparisons  $A$ . Then, following (6), we obtained the value  $J_{min} = 0.2788$ . It should be noted that condition (5) is only satisfied with the components of expression (6), enclosed in brackets.

Now we will check the weights for other, obviously worse values of the considered parameters, for example, for the data in Table 2.

Table 2: The Second Set of Data

para meter	$t$	$\varepsilon$	$\tau$	$v$	$k$	$\Sigma$	$\lambda$
Value	.6 s	1.5 mm	.55 s	.2	2	2	2
$\sigma_i$	-.25	-2.5	-.25	-.38	.88	.88	.88

We obtain the value  $J = 0.1536$ , which does not satisfy condition (2).

Now consider the best indicators than in Table 2 and evaluated these results, Table 3.

Table 3: The Third Set of Data

para meter	$t$	$\varepsilon$	$\tau$	$v$	$k$	$\Sigma$	$\lambda$
Value	.2	.85	.35	.05	8	8	8
$\sigma_i$	.75	.75	.75	.76	.125	.125	.125

We obtain the value  $J = 0.8783$ , which fully satisfies condition (11). The results obtained completely coincide with the proposals put forward, and also confirm the study conducted in [26]. Also, we expect that the developed methodology turns out to be enough effective in practical application with any set of measurement indicators.

**7. Conclusions**

A study of the expansion of research on unmanned platforms confirms the need for special training of operators for maintenance, operation and the performance of tasks as intended.

The paper proposes an approach based on the entering of a generalized indicator of the operator's productivity, which is a linear function of partial heterogeneous quality criteria with unknown weight coefficients. The described method is a modification of the method for analyzing process hierarchies, based on the identification of indicators of the same type, the effectiveness of which is confirmed by numerical examples. The considered approach can be used in the certification of the operator after passing the training process. Our further research we will plan to be directed to the development and implementation of software for the simulator complex with assessing the training of UAV operators following the developed methodology.

### Conflict of Interest

The authors declare no conflict of interest.

### Acknowledgment

The authors thank both the authorities of the State Space Agency of Ukraine, Taras Shevchenko National University, and National Aviation University the especially leadership of the Faculty of Cybersecurity, Computer and Software Engineering for their support during the preparation of this paper. The authors also thank the anonymous reviewers, whose comments significantly improved the content of the paper.

### References

- [1] D. Kucherov, O. Sushchenko, A. Kozub, "Operator Training for Unmanned Aerial Vehicles Control" in 2019 IEEE 5th International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), 22-24 Oct., Kiev, Ukraine, 1-4, 2019. <https://doi.org/10.1109/APUAVD47061.2019.8943918>
- [2] J. S. McCarley, C.D. Wickens, "Human Factors Implications of UAVs in the National Airspace", 2019. [https://www.researchgate.net/publication/228358350\\_Human\\_factors\\_implications\\_of\\_UAVs\\_in\\_the\\_national\\_airspace](https://www.researchgate.net/publication/228358350_Human_factors_implications_of_UAVs_in_the_national_airspace), last accessed 2019/03/13.
- [3] R. M. Taylor, "Human Automation Integration for Supervisory Control of UAVs" in Virtual Media for Military Applications, RTO-MP-HFM-136, 12-1-12-10, 2006.
- [4] J. Y. C. Chen, "Effects of operator spatial ability on UAV-guided ground navigation" in 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2-5 March 2010, Osaka, Japan, Proceedings, 139 – 140, 2010. <https://doi.org/10.1109/HRI.2010.5453227>
- [5] J. Wu, W. Wang, J. Zhang, B. Wang, "Research of a kind of new UAV training simulator based on equipment simulation" in 2011 International Conference on Electronic & Mechanical Engineering and Information Technology, 12-14 Aug., Harbin, China, Proceedings, 4812 – 4815, 2011. <https://doi.org/10.1109/EMEIT.2011.6024116>
- [6] O. V. Kozhokhina, V. M. Gribov, L. V. Blahaia, "Processing statistical data about UAV operator errors", in IEEE 3rd International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), 13-15 October, Kyiv, Ukraine, Proceedings, 124 – 127, 2015. <https://doi.org/10.1109/APUAVD.2015.7346578>
- [7] V. Rodriguez-Fernandez, H.D. Menendez, D. Camacho, "Automatic profile generation for UAV operators using a simulation-based training environment", Lect. Notes Artif. Int., 5 (1), 37 – 46, 2016. <https://doi.org/10.1007/s13748-015-0072-y>
- [8] K. W. Williams, "Human Factors Implications of Unmanned Aircraft Accidents: Flight-Control Problems", Federal Aviation Administration, April, 1 – 6, 2006. DOT/FAA/AM-06/8. [https://doi.org/10.1016/S1479-3601\(05\)07008-6](https://doi.org/10.1016/S1479-3601(05)07008-6).
- [9] Y. A. Gunchenko, S. A. Shvoroov, N. D. Rudnichenko, V. D. Boyko, "Methodical complex of accelerated training for operators of unmanned aerial vehicles" in 2016 4th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), 18-20 Oct., Kiev, Ukraine, 130-133, 2016. <https://doi.org/10.1109/MSNMC.2016.7783124>
- [10] S. M. Zlepko, L. G. Koval, D. Kh. Shtofel, V. O. Homolynskiy, M. I. Palamarchuk, "Peculiarities of Psycho-Physiological Selection of Operators for Unmanned Aviation Systems" in 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), 25-29 Feb., Lviv-Slavskve, Ukraine, Ukraine, 628-631, 2020. <https://doi.org/10.1109/TCSET49122.2020.235508>
- [11] J. Dagle, "'Partial panel' operator training: advanced simulator training to enhance situational awareness in off-normal situations" in 2006 IEEE Power Engineering Society General Meeting, 18-22 June, Montreal, Que., Canada, 1-2, 2006. <https://doi.org/10.1109/PES.2006.1709229>
- [12] D. Hong-bin, G. Tao-li, D. Xiao-li, "A System Dynamics simulation analysis of the Influencing Factors of coalmine special operators' competency by training" in 2nd IEEE International Conference on Information Management and Engineering, 16-18 April, Chengdu, China, 367-370, 2010. <https://doi.org/10.1109/ICIME.2010.5477580>
- [13] H. Haraguchi, T. Kaihara, N. Fujii, D. Kokuryo, "A study of designing for the operator training under an order change in cell manufacturing system" in 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), 20-23 Sept., Tsukuba, Japan, 802-806, 2016. <https://doi.org/10.1109/SICE.2016.7749216>
- [14] H. Haraguchi, "A Study on Operator Allocation Method Considering the Productivity and the Training Effect in Labor-Intensive Manufacturing System" in 2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 15-18 Dec., Macao, Macao, Macao, 1236 – 1239, 2019. <https://doi.org/10.1109/IEEM44572.2019.8978613>
- [15] C. Coluzzi, M. Contu, G. Pecoraro, S. Orlandi, F. Allella, A. Pascucci, S. Moroni, "Real time operators training: Tema experience in implementing the Commission Regulation (EU) 2017/1485" AEIT International Annual Conference (AEIT), 18-20 Sept. 2019, Florence, Italy, Italy, p. 1 – 5, 2019. <https://doi.org/10.23919/AEIT.2019.8893413>
- [16] H. Ayaz, P. A. Shewokis, S. Bunce, K. Izzetoglu, B. Willems, B. Onaral, "Optical brain monitoring for operator training and mental workload assessment", NeuroImage 59, 36-47, 2012. <https://doi.org/10.1016/j.neuroimage.2011.06.023>
- [17] K.G. Selivanova, O.V. Ignashchuk, L.G.Koval et al. "Computer-aided system for interactive psychomotor testing," Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments, Proc. SPIE, vol. 10445, 2017. <https://doi.org/10.1117/12.2280815>.
- [18] M.V. Bachynskiy, O.Yu. Azarkhov, D.Kh. Shtofel et al. "System and algorithm for evaluation of human auditory analyzer state," Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments, Proc. SPIE, vol. 10445, 2017. <https://doi.org/10.1117/12.2280738>.
- [19] Wiem Zemzem, Ines Hosni, "A New Distributed Reinforcement Learning Approach for Multiagent Cooperation Using Team-mate Modeling and Joint Action Generalization", Adv. Sci. Technol. Eng. Syst. J. 5(2), 1-12, 2020. <https://doi.org/10.25046/aj050201>
- [20] D. P. Kucherov, A. M. Kozub, "Mathematical model of UAV as agent of multi-agent system" in Dependable Systems, Services, and Technologies (DESSERT-2018), May 24-27, Kyiv, Ukraine, Proceedings, 358 – 362, 2018. <https://doi.org/10.1109/DESSERT.2018.8409156>
- [21] D. Kucherov, O. Sushchenko, A. Rasstrygin, S. Zhdanov, A. Kozub, "Synthesis of the switching control law for a quadrotor autopilot", Int. J. of Eng. and Tech. (UAE), 7(4), 3065-3069, 2018. <https://doi.org/10.14419/ijet.v7i4.16368>
- [22] B. Libet, Mind time: The temporal factor in consciousness, Perspectives in Cognitive Neuroscience, Harvard University Press, 2004.
- [23] Dektak stylus capabilities. [https://www.bruker.com/fileadmin/user\\_upload/8-PDF-Docs/SurfaceAnalysis/StylusProfilometry/ApplicationNotes/AN526-Dektak\\_Stylus\\_Capabilities\\_How\\_to\\_Choose\\_the\\_Corre.pdf](https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/SurfaceAnalysis/StylusProfilometry/ApplicationNotes/AN526-Dektak_Stylus_Capabilities_How_to_Choose_the_Corre.pdf)
- [24] A methodological approach to determining the influence of the human factor on the performance of information systems. (in Russian) <https://habr.com/ru/company/cognitive/blog/209266/>
- [25] T. L. Saaty, "Relative Measurement and its Generalization in Decision Making Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors the Analytic Hierarchy/Network Process" Rev. R. Acad. Cien. Serie A. Mat., 102(2), 251 – 318, 2008. <http://www.rac.es/ficheros/doc/00576.PDF>
- [26] A. A. Avetistova, "Psychological features of players in computer games", Psychology. J of the Higher School of Econ., 8 (4), 35 – 58, 2011.

## Short-Term Dynamic Exchange Rate Model: IFEER Concept Development

Anton Kuzmin\*

Department of Data Analysis, Decision-Making and Financial Technologies, the Financial University, Ramsey 125993, Russia

---

### ARTICLE INFO

Article history:

Received: 24 April, 2020

Accepted: 20 July, 2020

Online: 09 August, 2020

---

Keywords:

Exchange Rate

Model

IFEER

Nonlinear Dynamics

Short-Term Imbalances

Russian Ruble

US Dollar

---

---

### ABSTRACT

The new model of the short-term exchange rate dynamics was constructed. First of all, the most interesting were the reasons of the deviation from the medium-term equilibrium. The author's IFEER-concept (International Flows Equilibrium Exchange Rate) was used as a base and it was developed. In this study, due to the short-term modeling period the differential approach was applied. The result was an integrated version of the exchange rate dynamics model. The main result of mathematical modeling is a nonlinear multi-factor functional dependence of the exchange rate. The result dynamic functional dependence differs from the previous medium-term dependencies by the type of the internal dynamic function. Economically, this function in the short-term period is responsible for explosive changes in the exchange rate dynamics. The basis for mathematical modeling was the system of fundamental economic factors that affect the dynamics of the exchange rate. The influence of crisis events on the Russian financial market in the short term was studied. The conducted research allowed us to analyze and evaluate the quantitative impact of the short-term effects of the dynamics of the exchange rate of the Russian ruble to the US dollar.

---

## 1. Introduction

This paper is an extension of work originally presented in Twelfth International Conference "Management of large-scale system development" (MLSD'2019) [1].

The theory of exchange rate determination is one of the most important components of modern economic theory. Mathematical models of the medium-term dynamics of the exchange rates are widely presented in the works of classics in [2]-[7], and the monetarist school pay considerable attention to modeling the medium-term and the long-term dynamics of exchange rates. The same can be said about DSGE-models in [8].

Medium-term modeling of the exchange rates including the Russian ruble, based on the IFEER-concept (International Flows Equilibrium Exchange Rate concept), was carried out by [9]. This concept is based on mathematical modeling of the foreign trade operations and the capital flows as the most important macroeconomic fundamental determinants of the exchange rate movements. Several studies (for example, in [10]) consider sources of exchange rate fluctuations in other cases.

Against this background, only a few studies are devoted to the short-term exchange rate dynamics. And first of all, it is necessary to note here the works explained in [11-13].

Mathematical modeling of short-term dynamics will reveal the mechanisms of explosive changes in the exchange rate in crisis periods.

This study will build a new model of the short-term exchange rate dynamics. First of all, we will be interested in the reasons for the deviation from the medium-term equilibrium, and these aspects will be formalized mathematically. But in any case, the basis for mathematical modeling will be the fundamental economic factors that affect the dynamics of the exchange rate. Thus we will use the author's IFEER-concept and develop it.

## 2. Conceptual Framework of Exchange Rate Modelling

In the basic version, discrete flows are considered in the previous author's works [9]. For exchange rates that have market pricing, in a certain period we defined in the  $i$ -th market transaction:

$e_i$  - the nominal exchange rate,

$D_i$  - the volume of the market transaction in the foreign currency,

$R_i$  - the volume of the market transaction in the national currency.

---

\* Corresponding Author: Anton Kuzmin, Moscow, Russian Federation, [a\\_kuzmin@rambler.ru](mailto:a_kuzmin@rambler.ru)



From a financial point of view these variables are linked as follows:

$$e_i D_i = R_i. \tag{1}$$

As a result, in the *i*-th transaction:

$$e_i = R_i / D_i. \tag{2}$$

However, in this study due to the short-term modeling period author will also apply a differential approach.

Let's define  $R(t)$  as a streaming continuously differentiated function of funds in the national currency in the foreign exchange market. Accordingly,  $D(t)$  is a streaming continuously differentiated function of funds in the foreign currency and  $e(t)$  is the nominal bilateral exchange rate at time  $t$ .

By analogy with (1) in the differential form at  $t$ :

$$e(t) \frac{\partial D(t)}{\partial t} = \frac{\partial R(t)}{\partial t}. \tag{3}$$

As a result:

$$\frac{\frac{\partial R(t)}{\partial t}}{\frac{\partial D(t)}{\partial t}} = e(t). \tag{4}$$

By analogy with the previously used discrete IFEER-concept [9] we introduce the weight function  $w(t)$  for the exchange rate. In the discrete version, the weight function  $w(i)$  showed the contribution of the particular *i*-th transaction, depending on the volume in the foreign currency:

$$w(i) = \frac{D_i}{\sum_{j=1}^T D_j}.$$

In the integral version, the weight function  $w(t)$  is also based on the volume of funds in the foreign currency in a certain period  $T$ :

$$w(t) = \left[ \frac{\frac{\partial D(t)}{\partial t}}{\int_T \frac{\partial D(t)}{\partial t} dt} \right]. \tag{5}$$

Thus we will define the exchange rate as an integral in the certain period  $T$ :

$$e(T) = \int_T w(t)e(t)dt. \tag{6}$$

Substitute the weight function (5) in the formula (6):

$$e(T) = \int_T w(t)e(t)dt = \int_T \left[ \frac{\frac{\partial D(t)}{\partial t}}{\int_T \frac{\partial D(t)}{\partial t} dt} \right] e(t)dt =$$

$$= \int_T \left[ \frac{\frac{\partial D(t)}{\partial t} e(t)}{\int_T \frac{\partial D(t)}{\partial t} dt} \right] dt.$$

Let's use properties (3) and (4) and take the denominator beyond the integral:

$$e(T) = \int_T \left[ \frac{\frac{\partial D(t)}{\partial t} e(t)}{\int_T \frac{\partial D(t)}{\partial t} dt} \right] dt = \int_T \left[ \frac{\frac{\partial R(t)}{\partial t}}{\int_T \frac{\partial D(t)}{\partial t} dt} \right] dt =$$

$$= \frac{\int_T \frac{\partial R(t)}{\partial t} dt}{\int_T \frac{\partial D(t)}{\partial t} dt}.$$

Now we can disaggregate the streaming functions according to the balance of payments accounts:

$$e(T) = \frac{\int_T \frac{\partial R(t)}{\partial t} dt}{\int_T \frac{\partial D(t)}{\partial t} dt} = \frac{\int_T \frac{\partial R^{CA}(t)}{\partial t} dt + \int_T \frac{\partial R^K(t)}{\partial t} dt}{\int_T \frac{\partial D^{CA}(t)}{\partial t} dt + \int_T \frac{\partial D^K(t)}{\partial t} dt}.$$

Here funds with the upper index  $CA$  belong to the current balance, funds with the upper index  $K$  belong to the capital flows balance.

This formula is the main intermediate result of the IFEER-concept, allowing us to identify foreign trade operations and capital flows as the main factors of the exchange rate movements.

Let's denote:

$$\int_T \frac{\partial R^{CA}(t)}{\partial t} dt = I_T,$$

$$\int_T \frac{\partial R^K(t)}{\partial t} dt = K_T^-,$$

$$\int_T \frac{\partial D^{CA}(t)}{\partial t} dt = E_T,$$

$$\int_T \frac{\partial D^K(t)}{\partial t} dt = K_T^+.$$

From an economic point of view,  $E_T$  is the volume of funds as the supply of the foreign currency from export operations,  $I_T$  is the demand in the national currency for the foreign currency from import operations.  $K_T^-$  is the capital outflow as the demand in the national currency for the foreign currency and  $K_T^+$  is the capital inflow as the supply of the foreign currency.

At this conceptual level, the exchange rate dependence has a dynamic form by  $T$ :

$$e(T) = e_T = (I_T + K_T^-) / (E_T + K_T^+) \quad (7)$$

However, for the two involved parties, the nominal exchange rate is determined symmetrically because the country's export is another country's import and the capital outflow is the inflow for the other side. Here the asterisk shows the financial and economic variables of the opposite counterparty in this two-country world:

$$e_T = (E_T^* + K_T^{*+}) / (I_T^* + K_T^{*-}) = 1/e_T^*$$

It should be noted that the dependence (7) is natural from the economic point of view, since, for example, the increase in demand for imports and an increase in capital outflow leads to an increase in the exchange rate:

$$e = e^{*(-1)} = f_e(I^+, (K^-)^+, E^-, (K^+)^-)$$

The upper sign "-" or "+" here and further on this factor shows that the function respectively strictly decreases or increases.

In terms of partial derivatives, for example:

$$\partial f_e(I^+, K^+, E^-, K^-) / \partial K^+ < 0$$

### 3. The Two-Country Model: The Level of the Current Account Balance

We will build a two-period model of the exchange rate in periods  $t$  and  $t-1$ . Since we are going to use point estimates, we will replace the large  $T$  as a period index with the small  $t$  for convenience.

Under these conditions formula (7) will take a more decent look, which has already been used by us:

$$e_t = (I_t + K_t^-) / (E_t + K_t^+) = (E_t^* + K_t^{*+}) / (I_t^* + K_t^{*-}) = 1/e_t^* \quad (8)$$

One of the most important components of the foreign currency inflow to the domestic market is export revenue. However, its level in the period  $t$  is determined by decisions of national exporters in the previous period  $t-1$ :

$$E_t = I_t^* = k_E^I P_t^* (k_E^{II} (Q_t^{*x+1} Q_{t-1}^{*x+1})^\delta) (e_{t-1}^R)^z = k_E P_t^* ((Q_t^{*x+1} Q_{t-1}^{*x+1})^\delta) (e_{t-1}^R)^z \quad (9)$$

here:  $k_E$  is a constant,

$P_t^*$  - the aggregate foreign price level,

$Q_t$  - the total domestic real output,

$e_{t-1}^R$  - the indicator of the international competitive advantages.

The exports of one side in this two-country world are imports of the other side.

The dependence (9) indicates that only the part  $k_E^I$  of export revenue comes to the domestic currency market. In general, the volume of exports is determined by the part  $k_E^{II}$  of the total domestic real output  $Q_t$  in the foreign prices  $P_t^*$ .

Here intertemporal averaging of the total real output  $Q_t$  is used for modeling. Note, that  $\frac{1}{x+1} + \frac{x}{x+1} = 1$ . In this case, the income at  $t$  is given more weight than the income at  $t-1$ . The method of intertemporal averaging of  $Q_t$  should not have a significant impact since it is the least volatile factor.

One of the most important factors of the model is the indicator of the international competitive advantages. This is the real bilateral exchange rate in the previous period  $t-1$ :

$$e_{t-1}^R = e_{t-1} P_{t-1}^* / P_{t-1} \quad (10)$$

Thus, in formula (9) we take into account the possible different degrees of the export functions responses to changes in the indicator of the international competitive advantages and the total domestic real output. For these purposes, we use various adjustable parameters in degrees:  $\delta, x, z$ . Their relationships will be shown later.

Functional dependence (9) at the same time also has a natural functional character from the economic point of view with respect to the system of the main determinants of the exchange rate:

$$E_t = f_{E_t}(P_t^*, Q_t^+, Q_{t-1}^+, e_{t-1}^{R+})$$

In accordance with the previous modeling, the demand for the foreign currency from imports is determined completely symmetrically for the other side:

$$I_t = E_t^* = k_I P_t (Q_t^{*x+1} Q_{t-1}^{*x+1})^\lambda (e_{t-1}^R)^y \quad (11)$$

Here:  $k_I$  is a constant,

$Q_t^*$  - the total foreign real output.

For the purposes of simulation, it is necessary to impose the restrictions:  $x \neq -1, x=z-y$ .

The properties of the stream functions of the current balance operations:

$$E^*=I, I^*=E.$$

Here it is important to note, that the dependencies (10) and (11) we define for the medium-term environment. But we can assume

that these dependencies retain their mathematical factor structure even for the short-term modeling. This is due to the economic inertia of exports and imports in the world economy.

**4. The Two-Country Model: the Level of the Capital Flows**

Of course, the capital flows are one of the most important factors of the exchange rates dynamics. And the developed author's IFEER-concept allows us to conduct mathematical modeling taking into account the capital flows.

We need to accept the hypotheses about mathematical forms of the basic capital flows dependencies. The capital inflow dependence is an increasing function by the aggregate foreign prices, the indicator of international competitive advantages, and the total domestic real output.

In this context, it is important to note that these functions must be modified for short-term exchange rate modeling in relation to the previously discussed functions in the author's work [9]. Here the constants  $k$  becomes functions by time. For the functional dependence of the capital outflow, for example, it is an increasing function in a crisis period. This shows a short-term increase in capital outflow compared to the medium-term equilibrium dynamics.

For the capital outflow, thus, the dependence must satisfy the following condition:

$$K_t^- = f_{K_t^-}(k_{K_t^-}^+(t), P_t^+, Q_t^{*+}, Q_{t-1}^{*+}, e_{t-1}^{R-})$$

Thus we define the functional dynamic dependence of the capital outflow:

$$K_t^- = k_{K_t^-}(t) P_t (Q_t^{*x+1} Q_{t-1}^{*x+1})^\rho (e_{t-1}^R)^y. \quad (12)$$

The capital outflow is a part of the domestic product produced in foreign prices.

In the short-term period in the currency market the crisis phenomena lead to a sharp increase in the capital outflow. Mathematically this is expressed in the strict increase function  $k_{k_t^-}(t)$  by  $t$  in the dynamic formula (12):

$$\frac{\partial k_{k_t^-}(t)}{\partial t} > 0.$$

At the same time, the functional dynamic dependence of the capital inflow:

$$K_t^+ = k_{K_t^+}(t) P_t^* (Q_t^{x+1} Q_{t-1}^{x+1})^\theta (e_{t-1}^R)^z. \quad (13)$$

Perfectly symmetrical, in the domestic currency market the crisis phenomena leads to a significant reduction of the capital inflows: a strict decrease of the function  $k_{k_t^+}(t)$  by  $t$  in the formula (13):

$$\frac{\partial k_{k_t^+}(t)}{\partial t} < 0.$$

Here we use the hypotheses about the type of the functions  $k_K(t)$ , that are not constants for the capital flows in the short-term period. This distinguishes this model from the previously developed author's models.

In accordance with the methodology of modeling dependencies are interrelated:

$$K_t^{*+} = K_t^- = k_{K_t^-}(t) P_t (Q_t^{*x+1} Q_{t-1}^{*x+1})^\rho (e_{t-1}^{*R})^{-y}.$$

The properties of the capital stream functions:

$$K^{*-} = K^+, \\ K^{*+} = K^-$$

Let's substitute the functional dependencies (9)-(13) into the basic formula (8):

$$e_t = e(t, k(t), Q(t), P(t), P^*(t)) = \frac{P_t (Q_t^{*x+1} Q_{t-1}^{*x+1})^\rho (e_{t-1}^R)^y (k_t (Q_t^{*x+1} Q_{t-1}^{*x+1})^{\lambda-\rho} + k_{K_t^-}(t))}{P_t^* k_E (Q_t^{x+1} Q_{t-1}^{x+1})^\theta (e_{t-1}^R)^z (k_E (Q_t^{x+1} Q_{t-1}^{x+1})^{\delta-\theta} + k_{K_t^+}(t))}. \quad (14)$$

Let's introduce the dynamic aggregate function  $k(t)$ :

$$\frac{(k_t (Q_t^{*x+1} Q_{t-1}^{*x+1})^{\lambda-\rho} + k_{K_t^-}(t))}{(k_E (Q_t^{x+1} Q_{t-1}^{x+1})^{\delta-\theta} + k_{K_t^+}(t))} = (k(t))^{x+1}. \quad (15)$$

With (15) in mind, we can rewrite (14):

$$e_t = \frac{P_t (Q_t^{*x+1} Q_{t-1}^{*x+1})^\rho (k(t))^{x+1}}{P_t^* k_E (Q_t^{x+1} Q_{t-1}^{x+1})^\theta (e_{t-1} \frac{P_{t-1}^*}{P_{t-1}})^x}.$$

Thus, move the member  $(e_{t-1})^x$  to the left side:

$$e_t (e_{t-1})^x = \left( k(t) \frac{P_t}{P_t^*} Q_t^{*\rho/x+1} Q_{t-1}^{-\theta/x+1} \right) \times \left( k(t) \frac{P_{t-1}}{P_{t-1}^*} Q_{t-1}^{*\rho/x+1} Q_{t-1}^{-\theta/x+1} \right)^x.$$

Let's make a temporary separation of the modeling variables and re-define  $\psi = \rho/x+1$  and  $\zeta = \theta/x+1$  for convenience:

$$e(t, k(t), P(t), P^*(t), Q(t)) = k(t) \frac{P_t}{P_t^*} (Q_t^{*\psi}) (Q_t^{-\zeta}). \quad (16)$$

**5. Assessment of Short-Term Dynamic Effects**

The main result of mathematical modeling is the new nonlinear multi-factor functional dependence of the exchange rate. Here, it

$$k(t)$$

Economically, this function in the short-term period is responsible for explosive changes in the exchange rate dynamics. So, this is the main fundamental difference between this model and the previous author's models [9], etc. In those works, these functions are constants in the medium-term in various modifications.

There is significant stability of  $(Q_{t-1}^{\delta/\delta+1} Q_t^{1/\delta+1})$  in the dynamic functional dependence (15) in comparison with other fundamental factors in the short-term period. So, a strict decrease of the internal function  $k_{k^+}(t)$  and a strict increase of the internal function  $k_{k^-}(t)$  guarantees us a strict increase of the aggregate function  $k(t)$  in a crisis period:  $\frac{\partial k(t)}{\partial t} > 0$ . Mathematically this reflects the reasons for the deviation from the medium-term equilibrium.

Direct economic assessment of the short term impact of the function  $k(t)$  is quite difficult. But we will try to do it using the example of the Russian ruble.

Let's use the main result of the author's work [9]. The medium-term dynamic functional dependence of the exchange rate of the Russian ruble against the US dollar is:

$$e_t = k' \frac{P_t}{P_t^*} Q_t^{-\psi} \tag{17}$$

here:  $k'$  is a constant,

$\psi$  - an estimated parameter.

For research and verification of this model we used the Russian CPI (Consumer Price Index) as the model variable  $P$  (Federal service of state statistics of Russia), the index of the real Russian GDP as the model variable  $Q$ , the prices of blend crude oil as the model variable  $P^*$  (Intercontinental Exchange, US dollars per brent-mix barrel, Bloomberg information agency, Bloomberg Terminal).

For our research the most appropriate is the crisis period 2013 – 2015. This period is associated with the currency crisis and the significant depreciation of the Russian ruble.

The ruble/dollar exchange rate (USD/RUR) is the main rate on the Russian foreign exchange market. The value of the ruble exchange rate in relation to other world currencies is determined through the cross-rate system.

Thus, as the starting point of the research period, December 2013 was selected. At that time, both the international environment and the domestic macroeconomic situation were stable. The exchange rate of the ruble at that moment was 32,73 USD/RUR according to the Central Bank of Russia.

This study suggests using the least squares method with normalization of the exchange rate values. Since the exchange rate

changes quite seriously during crisis periods, it is more important to use relative indicators of changes.

Thus, in accordance with numerical simulation, the solution of the parametric minimization problem

$$\left[ \begin{array}{l} \min_{\psi} \sum_t \left( \frac{e_t(\psi) - e(\text{nominal})_t}{e(\text{nominal})_t} \right)^2 \\ \psi \geq 0 \end{array} \right]$$

is  $\psi = 0,5$ . Here  $e(\text{nominal})$  is the nominal exchange rate USD/RUR, according to the Bank of Russia monthly calculations at the end of each period.

In Figure 1 (monthly data, author's calculations) we present the dynamics of the exchange rate of the Russian ruble against the US dollar in accordance with the nominal exchange rate  $e(\text{nominal})$  and the result dynamic functional dependence of this research  $e(\text{Theor})$  (17).

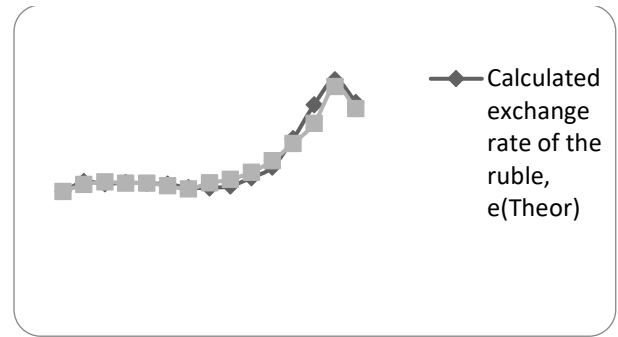


Figure 1: The calculated and nominal exchange rates of the Russian ruble to the US dollar (USD/RUR, monthly data, 2013 - 2015).

The average of absolute normalized deviations of the calculated and nominal exchange rates was 3%, and the average of normalized deviations was 0,3%. As a result, this allows us to conclude that the model quality is high.

This research has shown a high correlation between the exchange rate of the ruble, the inflation, the Russian balance of payments, and other fundamental variables (for example, we can also recommend a work [14]).

Using the methodology, presented in this paper, and the author's work [15] (in Russian) it can be shown, that short-term dynamic functional dependence of the exchange rate of the Russian ruble against the US dollar is:

$$e_t = k'(t) \frac{P_t}{P_t^*} Q_t^{-\psi} \tag{18}$$

Here, the system of main fundamental economic factors of the exchange rate dynamics in relation to the dependencies (16) and (17) is saved.

In Figure 2 (Moscow Exchange, executions TOM) in the period 2013 – 2015 solid line represents the medium-term dynamics of the exchange rate of the Russian ruble against the US dollar. We schematically overlaid data from the author's work [9] and Figure 1. Based on the presented data, we can estimate the cumulative contribution of the function  $k'(t)$  to changes in the ruble exchange rate during crisis periods in the short-term in the



amount of 20-25%. The contribution to the crisis period in 2015-2016 can also be estimated at the same amount approximately.

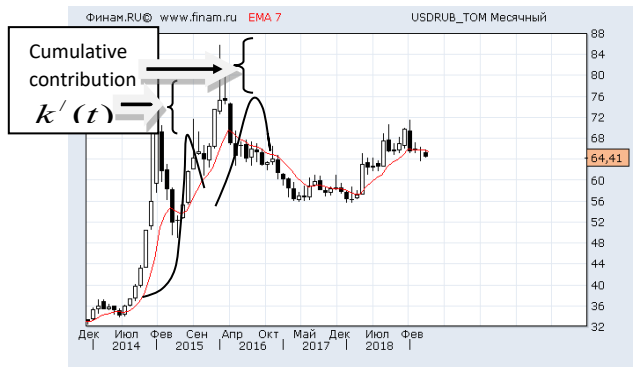


Figure 2: USD/RUR quotes (Moscow Exchange, monthly data, executions TOM, Japanese candlesticks, 2013 – 2018, Information agency \*FINAM\*)

## 6. Conclusion

In this study, the new model of short-term exchange rate dynamics was constructed. First of all, we were interested in the reasons for the deviation from the medium-term equilibrium.

The author used the IFEER-concept as a base and developed it. In this study due to the short-term modeling period, the differential approach was applied. The result was an integrated version of the exchange rate dynamics model. The functional dependencies of the export-import operations and the capital flows were mathematically formalized.

The main result of mathematical modeling is the nonlinear multi-factor dynamic functional dependence of the exchange rate that differs from the previous medium-term functional dependencies by the type of the internal dynamic function. Economically, this function in the short-term period is responsible for explosive changes in the exchange rate dynamics. But in any case, the basis for mathematical modeling was the system of the fundamental economic factors that affect the dynamics of the exchange rate.

The most important limitation of mathematical modeling is the significant difficulty in directly economic assessment of the

internal functions  $k_{k^+}(t)$  and  $k_{k^-}(t)$  that determine short-term capital flows in crisis periods.

But the influence of crisis events on the model in the short term is studied on the Russian financial market at a previous time by assessment of the short-term impact of the function  $k(t)$ . The conducted research allowed us to analyze and evaluate the quantitative impact of the short-term effects of the dynamics of the exchange rate of the Russian ruble against the US dollar. The estimated indicators allow us to conclude that the Bank of Russia needs to pursue a more active interventionist policy to maintain the stability of the national currency.

At the same time, such economic assessments are individual and specific for each currency and each crisis period.

The developed IFEER-concept allows us to model the exchange rate dynamics in other economic conditions. First of all,

this concerns the future research of the long-term dynamics of the main world currencies and the Russian ruble.

## Conflict of Interest

The author declares no conflict of interest.

## References

- [1] A. Kuzmin, "Modeling of Short-Term Exchange Rates Dynamics" in Twelfth International Conference "Management of large-scale system development" (MLSD 2019), Publisher: IEEE, 2019, <https://doi.org/10.1109/MLSD.2019.8911067>
- [2] R.A. Mundell, "Capital Mobility and Stabilization Under Fixed and Flexible Exchange Rates" *Canadian Journal of Economics and Political Science*, November, 475-485, 1963. <https://doi.org/10.2307/139336>
- [3] A.C. Stockman, "A Theory of Exchange Rate Determination" *Journal of Political Economy*, August, **88**, 673-698, 1980. <https://doi.org/10.1086/260897>
- [4] P.B. Clark, R. MacDonald, Exchange Rates and Economic Fundamentals: A Methodological Comparison of BEERs and FEERs, IMF Working Paper 98/67, Washington: International Monetary Fund, March, 1998.
- [5] P.B. Clark, R. MacDonald, Filtering the BEER a permanent and transitory decomposition, IMF Working Paper 00/144, Washington: International Monetary Fund, 2000.
- [6] M. Mussa, "The Theory of Exchange Rate Determination" in Exchange Rate Theory and Practice, J.E.O. Bilson, R.C. Marston (eds.), Chicago: University of Chicago Press, NBER, 13-78, 1984.
- [7] L. Killian, "Exchange Rates and Monetary Fundamentals: What Do We Learn from Long-Horizon Regressions?" *Journal of Applied Econometrics*, 14( 5), 491-510, 1999. [https://doi.org/10.1002/\(SICI\)1099-1255\(199909/10\)14:5<491::AID-JAE527>3.0.CO;2-D](https://doi.org/10.1002/(SICI)1099-1255(199909/10)14:5<491::AID-JAE527>3.0.CO;2-D)
- [8] M. Obstfeld, K. Rogoff, "Exchange Rate Dynamics Redux" *Journal of Political Economy*, **103** (3), 624-660, 1995. <https://doi.org/10.3386/w4693>
- [9] A. Kuzmin, "Exchange Rate of the Ruble Modeling" *Advances in Systems Science and Applications*, **19**(4), 87-93, 2019. <https://doi.org/10.25728/assa.2019.19.4.830>
- [10] I. Chowdhury, "Sources of Exchange Rate fluctuation: empirical evidence from six emerging market countries" *Applied Financial Economics*, **14**(1), 697-705, 2004. <https://doi.org/10.1080/0960310042000243538>
- [11] R. Dornbusch, "Capital Mobility, Flexible Exchange Rate and Macroeconomic Equilibrium" in Recent Developments in International Monetary Economics, E. Claassen, P. Salin (eds.), North-Holland, 261-278, 1976.
- [12] R. Dornbusch, "Expectations and Exchange Rates Dynamics" *Journal of Political Economy*, **84**, December, 1161-1176, 1976. <https://doi.org/10.1086/260506>
- [13] R. Dornbusch, "Equilibrium and Disequilibrium Exchange Rates" *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, Berlin, **102**(6), 573-599, 1982. <https://doi.org/10.3386/w0983>
- [14] L. Strizhkova, "The relationship between inflation, exchange rate and parameters of economic policy (on an example of Russia)" *Bulletin of the Institute of Economics of the Russian Academy of Sciences*, **5**, 156-176, 2017. (Article in Russian with an abstract in English)
- [15] A. Kuzmin, "Modeling short-term dynamics of the ruble exchange rate" *Money and credit*, **8**, 59-65, 2015. (Article in Russian with an abstract in English)

## The Sound of Trust: Towards Modelling Computational Trust using Voice-only Cues at Zero-Acquaintance

Deborah Ooi Yee Hui<sup>1</sup>, Syaheerah Lebai Lutfi<sup>1,\*</sup>, Syibrah Naim<sup>2</sup>, Zahid Akhtar<sup>3</sup>, Ahmad Sufri Azlan Mohamed<sup>1</sup>, Kamran Siddique<sup>4,\*</sup>

<sup>1</sup>Universiti Sains Malaysia, School of Computer Sciences, Penang, 11800, Malaysia

<sup>2</sup>Woosong University, Technology Department, Endicott College of International Studies, Daejeon, 34606, Republic of Korea

<sup>3</sup>University of Memphis, Tennessee, 38152, United States of America

<sup>4</sup>Xiamen University Malaysia, Sepang, 43900, Malaysia

### ARTICLE INFO

Article history:

Received: 16 May, 2020

Accepted: 26 June, 2020

Online: 09 August, 2020

Keywords:

True

Voice

Artificial neural network

Zero acquaintance

### ABSTRACT

Trust is essential in many interdependent human relationships. Trustworthiness is measured via the effectiveness of the relationships involving human perception. The decision to trust others is often made quickly (even at zero acquaintance). Previous research has shown the significance of voice in perceived trustworthiness. However, the listeners' characteristics were not considered. A system has yet to be produced that can quantitatively predict the degree of trustworthiness in a voice. This research aims to investigate the relationship between trustworthiness and different vocal features while considering the listener's physical characteristics, towards modelling a computational trust model. This study attempts to predict the degree of trustworthiness in voice by using an Artificial Neural Network (ANN) model. A set of 30 audio clips of white males were obtained, acoustically analyzed and then distributed to a large group of untrained Malaysian respondents who rated their degree of trust in the speakers of each audio clip on a scale of 0 to 10. The ANOVA test showed a statistically significant difference of trust ratings across different types and intensities of emotion, duration of audio clip, average fundamental frequencies, speech rates, articulation rates, average loudness, ethnicity of listener and ages of listener ( $p < .01$ ). The findings conclude that Malaysians tend to trust white males who talk faster and longer, speak louder, have an  $f_0$  between 132.03Hz & 149.52Hz, and show a neutral emotion or rather stoic (arousal < .325). Results suggest that Indians are the most trusting Malaysian ethnic group, followed by Bumiputera from East Malaysia and then followed by Malays. Chinese are the least trusting Malaysian ethnic group. The data was fed into an ANN model to be evaluated, which yielded a perfect percentage accuracy (100%) in degree of trustworthiness 39.70% of the time. Given a threshold of two-point deviation, the ANN had a prediction accuracy of 76.86%.

## 1. Introduction

Trust is a fundamental part of various commercial and interpersonal relationships, especially wherever risk, uncertainty, or interdependence exist. Human trust is influenced by perception, intuition, environment and many other factors to the human centred scheme that is generally defined in any communication systems, natural and synthetic environment in everyday human

interaction. Investigating state of the art in theoretical studies in trust and extending the finding as a base to symmetrically model and simulate trust in synthetic agents is crucial, especially in this digital era. In [1] it is stated that an abundance of information about other people can be determined through only a quick observation of their expressive behaviour. Humans can gauge trustworthiness in zero acquaintance situations [1], [2]. Zero acquaintance is a situation in which perceivers make judgments about targets they are given no opportunity to interact with [3]. Several other studies have been conducted previously with regards to zero acquaintance

\*Corresponding Authors: Syaheerah Lebai Lutfi & Kamran Siddique, Emails:

[syahherah@usm.my](mailto:syahherah@usm.my), [kamran.siddique@xmu.edu.my](mailto:kamran.siddique@xmu.edu.my)

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050456>

[4]–[7], with each experiment utilizing thin slices that last from 5 seconds to 30 seconds each. In a successful study by Mohammad and Kiritchenko [8], only short 10-second audio clips were used to detect the effects of prosodic features of voice on personality. Therefore, the 10s length is enough to capture personality impressions at zero acquaintance while lightening the load on the listeners.

In recent years, many studies have shown the significance and importance of voice in perceived trustworthiness. These studies have analysed the effect of various vocal cues (e.g., pitch, speech rates, and fundamental frequency ( $f_0$ )) on trustworthiness. However, prior research did not consider the characteristics of the listener (the rater/ trustor). It is important to include the characteristic of the listener as well, since trust is a concept that involves two parties [9].

Recently, chatbots or robots have started taking over the roles of teachers and tutors [10]. One such example of a chatbot would be Jill Watson, who was created to handle forum posts by students [11]. However, chatbots communicate primarily via text, which would be inconvenient visually impaired students or students who suffer from dyslexia. Furthermore, [12] indicated that a chatbot's suggestions are not fully trusted by humans. Since voice has a significant impact on perceived trustworthiness, people might trust the chatbots more if the chatbots had a trustworthy voice. Therefore, this research aims to determine the specific characteristics displayed in a "trustworthy" voice. When it comes to chatbots, many studies have been conducted to investigate the aptitude and efficiency of the chatbot applications, ignoring the voice characteristics. One interesting study in the potential of having chatbot imitate human ability during a conversation is by [13]. The outcome of the study showed that the "uncanny valley" effect (where people become uncomfortable to a lifelike robot to a certain point) is lowered by using psychophysiological responses in simple chatbox developments compared to a multifaceted one.

A prediction model would be useful to measure the degree of trustworthiness of the chatbot's synthetic voice before releasing it to the public. There has yet to be a system that can quantitatively predict the degree of trustworthiness in a voice.

This research aims to investigate the relationship between vocal cues and trustworthiness at zero acquaintance, and to predict the degree of trustworthiness of a voice by using an Artificial Neural Network (ANN) model. However, there are several constraints.

Firstly, the chosen audio clips only consist of white male speakers and the accents of the voices are therefore only limited white accents. Secondly, the survey does not consider the listener's emotional state at the point of answering. Thirdly, due to time and financial constraints, this survey was only open to Malaysians. Hence, all the findings of this study are only valid and applicable among Malaysians.

## 2. Background and Related Work

In recent years, the significance of voice in perceived trustworthiness has been investigated in many previous studies [14]–[18]. Many papers have also investigated the effect of individual vocal cues on trust at zero acquaintance.

In [19] it is noted that pitch has a small role in detecting cheating probability of infidelity among committed romantic partners, but it does not represent the entire picture. Hughes and Harrison [19] did not consider other acoustical measures or any combination of different acoustic measures in the research.

The authors in [20] utilized data-driven voice computational modelling to study the relationship between acoustics and perceived trustworthiness in the single word "hello". The findings demonstrate a strong acoustical basis for perceived voice personality. However, this research did not take the listener characteristics into consideration at all.

In [18] it is investigated the effect of trustworthiness impressions on vocal expressive and person characteristics and how the relationship can be explained by acoustical profiles. In [18] it is indicated that positive speaker valence but not arousal, consistently predict greater perceived trustworthiness. Furthermore, female as compared with male speakers and younger voices compared with older were perceived as more trustworthy. However, [18] did not consider the ethnicity of the listener. Ethnicity of listeners is vital as interpersonal trust is largely influenced by culture [21]–[23].

A computer based, acoustic profile of Steve Job's tone of voice was created by Niebuhr et al. [24] Findings indicate that all the melodic features of charisma could be found in his tone of voice. However, this study is purely from the author's assumption of Steve Jobs as a charismatic speaker, and the author did not take into consideration the perception of the general public. Furthermore, this study only investigated the perception of charisma in a voice and not trustworthiness.

With regards to evaluating the voice of an artificial intelligence (A.I.) instructor, [10] indicated that students who perceive themselves as being "older" trust the older A.I. voice more than a younger A.I. voice. Although this study did consider the listener's age, they did not consider other acoustic or prosody characteristics of the A.I. instructor's voice.

There also exists some degree of dissension regarding the direction of association between fundamental frequency ( $f_0$ ) and trustworthiness. On one hand, a deception study indicated that participants consistently raised their  $f_0$  when lying [25] while another study indicated the opposite [26]. Yet another further study found no differences in  $f_0$  or other acoustic characteristics when analyzing deceptive and truthful messages [27]. This situation highlights the fact that there exist inconsistent results from research regarding the direction of association between fundamental frequency ( $f_0$ ) and trustworthiness.

Research has also used prosody features to measure trustworthiness. In [28] the authors used the investment game to assess the effect of prosody characteristics on trust attributes. Although higher investments (higher trust) were associated with faster articulation rate and higher pitch, this finding might have been influenced by game behavior and game turn, both of which have the highest effect on investments (trust).

Evidently, previous research has only considered the impact of several different vocal cues on trust at zero acquaintance. This research aims to provide a more comprehensive analysis of the effect of vocal cues on trust at zero acquaintance by including more

vocal features. The specific acoustic characteristics that will be investigated in this research are (a) Mean f0 (b) Min f0 (c) Max f0 (d) Mean Intensity/Loudness (e) Min Intensity (f) Max Intensity (g) Articulation Rate (h) SD of f0 (i) Speech Rate.

Most of the researches mentioned did not consider the characteristics of the listener (the rater). When the characteristics of the listener were considered, the ethnicity of the listener was not taken into consideration. It is important to include the characteristics of the listener as well, since trust is a concept that involves two parties [9]. This study will, at the very least, take the physical characteristics of the listener into consideration e.g. age, gender, English proficiency, and ethnic groups. In addition to that, this study will also consider the arousal and valence of the emotion displayed by the speaker.

### 3. Recent Algorithms Used to Measure / Predict Trust

Very few studies have attempted to quantitatively measure the degree of trustworthiness. Korovaiko and Thomo [29] have attempted to predict trust from online review ratings by using Random Forest models and Support Vector Machines, and [30] have tried predicting trust and distrust between two strangers in social networks by using an inference algorithm. Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar.

ANN models have also been used to create a broker-assisting trust and reputation system and to investigate the potential determinants of relationship quality in relationship marketing [31, 32]. Although the research conducted by [32] is closely related to this research, it considers trust as one of the measurements of relationship quality instead of the feature to be predicted.

Lee et al. [33] used a Trust Hidden Markov Model to predict whether an individual is willing to behave cooperatively (trust) with their novel partner based on gestural cues. Although Lee et al. [33] does use artificial intelligence to predict trust at zero acquaintance, this research investigates the effect of gesture and not voice on interpersonal trust.

Evidently, there has been no prior research using artificial intelligence to predict the degree of trustworthiness in a person's voice. Therefore, this study aims to predict the degree of trustworthiness in a person's voice by using an ANN model that will take both the acoustic characteristics of the speaker's voice and the physical characteristics of the listener into consideration.

A typical artificial neural network model (ANN) was used in this research. The ANN is a feedforward neural network with an input layer of source neurons, one or more hidden layer(s) of computational neurons, and an output layer of computational neurons. The input layer receives input signals and redistributes these signals to all neurons in the hidden layer. The hidden layer detects the features and represent the features in the input patterns by manipulating the weights of the neurons. The output pattern of the entire network is established by the output layer. The ANN is also capable of supervised learning through the backpropagation learning algorithm. In other words, the ANN is able to learn from its mistakes and improve the accuracy of its predictions automatically [34].

The use of a feedforward backpropagation ANN model was due to several reasons: a) capable of supervised training, b) minimization of training errors, c) robustness to noisy data, d) high computational speed across large input datasets. The main disadvantage is the many variables and the training time which must be considered when constructing an ANN.

In this experiment, a large number of variables must be considered when predicting the degree of trustworthiness, where the exact relationship between each variable on trustworthiness still remains a mystery. As such, the use of an ANN model is the most suitable approach to represent such a complex relationship and adequately predict the degree of trustworthiness in a voice.

## 4. Materials and Methods

The methodology undertaken in this research is similar to the methodology undertaken by Hughes and Harrison [19], Imhof [35] and Smith et al. [36]. Participants were informed about the study procedure and given the option to opt out of the online survey at any point of time.

### 4.1. Participants

This study was conducted in Universiti Sains Malaysia. Malaysia has a multicultural population that consist of 3 main ethnic groups; Chinese, Malay and Indian. There also exist other minority ethnic groups from East Malaysia, such as Kadazans, Bajaus and Dayaks [37], [38]. Since these minority ethnic groups share the bumiputera (native/indigenous) status with the Malays, we will consider this minority ethnic groups as "Other Bumiputera".

In total, the participants of this study consisted of 86 Malaysian individuals (44 men and 42 women) of which 40 were Chinese, 21 were Malay, 15 were Indian and only 10 were from Other Bumiputera. The participants had a mean age of 24.337 years (SD 5.517, 17-53). 64 of them (74.42%) were students while 22 (25.58%) were non students. The participants were also asked to rate their own English proficiency. 12 (13.95%) participants considered their English to be Highly Proficient, 35 (40.70%) satisfactory, 19 (22.09%) proficient, 15 (17.44%) modest, 2 (2.33%) expert and 3 (3.49%) limited command of English.

### 4.2. Stimuli

Six labelled audio clips were sourced from the One-Minute Gradual-Emotion Recognition (OMG-Emotion) dataset [39]. The dataset comprises of 567 YouTube videos, which were divided into clips of varying length and labelled based on utterances. At least five independent subjects annotated each utterance using an arousal/valence scale and emotion tags based on the six universal emotions from Ekman (fear, disgust, happiness, sadness, surprise and anger) [40].

The 6 audio clips to be used in this research were chosen based on the following conditions: a) each audio clip must be about 10s long, b) it must be from a white male speaker and c) there must be a balanced number of audio clips representing each different emotion.

The audio clips were digitally manipulated with Praat into a higher and lower pitch and speech rate version. The degree to which it was manipulated was based from previous research



conducted by [41], [42]. This manipulation results in a total of 30 audio clips being formed (6 original and 6\*4 modified clips), which will be used for this research.

The following acoustic characteristics of each audio clip was extracted with Praat and tabulated: (a) Mean f0 (b) Min f0 (c) Max f0 (d) Mean Intensity/Loudness (e) Min Intensity (f) Max Intensity (g) Articulation Rate (h) SD of f0 (i) Speech Rate.

#### 4.3. Procedure

86 Malaysian respondents were asked via an online survey to listen to the audio clips and rate their degree of trust in the speakers of each audio clip on a scale of 0 to 10 (0 being absolutely untrustworthy and 10 being most trustworthy). Respondent details such as whether they were Malaysian or non-Malaysian, Student or nonstudent, Gender, Age, Ethnicity and English proficiency were also collected in the survey. The responses were recorded and tabulated.

Each individual response was mapped to the acoustic characteristics and the emotion annotations of the respective audio clips as obtained from the data set. After data cleaning, the total number of valid participants were reduced to 74 (N=2220).

Analysis of Variance (ANOVA) was used to evaluate the null hypothesis that there is no statistically significant difference between each vocal cue or rater characteristic on the trustworthiness ratings. Each vocal cue or rater characteristic was used as the dependent variable and the trust ratings were used as the independent variable. The Welch test was used when the vocal cues violated the assumption of homogeneity of variance.

Since arousal and valence are closely related, a Multivariate Analysis of Variance (MANOVA) test was conducted with arousal and valence as the dependent variables and trust ratings as the independent variable. However, the ANOVA and MANOVA test does not tell us exactly how the mean varies between different groups of the data. Therefore, a means plot of each vocal cue on trust rating was carried out.

The data was split into train, validation and test sets. The train set trained the neural network and the validation set verified the prediction accuracy. Finally, the model was run on the test set to determine the model performance on unfamiliar data.

#### 4.4. Artificial Neural Networks

A feedforward backpropagation ANN was used in this research. The model aims to obtain a low mean-squared-error (MSE) as well as a high percentage of low-deviation predictions. The network was built with Keras API. The model was compiled with Adam optimizer with learning rate = 0.0005,  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = \text{None}$ ,  $\text{decay} = 0.0003$  and  $\text{amsgrad} = \text{False}$ . The loss function used in this neural network is categorical cross-entropy, with mean squared error and accuracy as the metrics for checkpointing.

After many simulations, the architecture of the ANN that produced the lowest mean squared error is as follows:

1. Input layer: Dense, 28 input neurons, ReLu activation.
2. Hidden layer: Dense, 40 neurons, ReLu activation.

3. Dropout layer: Dropout rate=0.2
4. Hidden layer: Dense, 30 neurons, ReLu activation
5. Dropout layer: Dropout rate=0.4
6. Hidden layer: Dense, 16 neurons, ReLu activation  
Output layer: Dense, 11 output neurons, Softmax activation

The data underwent feature selection with the score function  $f_{\text{classif}}$  from SelectKBest prior to being used for ANN. Findings indicate that only the maximum f0, average f0 and minimum f0 had insignificant results ( $p > .05$ ) and hence these features were removed from the data set. The input layer of the ANN has 28 nodes for each of the acoustic characteristics, emotion annotations as well as listener characteristics. The output layer has 11 nodes for predicting the degree of trust on a scale of 0 to 10. Softmax activation was used in the output layer for multiclass classification.

The model was fit with the training data set (n=1506) and run over 600 epochs, with batch size of 12 and the validation data set (n=377) was simultaneously used to check for overfitting. EarlyStopping callback was used to ensure minimal overfitting by monitoring accuracy. Finally, the ANN was tested with the testing dataset.

The predictions obtained were recorded and the results were evaluated by measuring the deviation of the predicted trust ratings from the original. In this study, the threshold for acceptable deviation in predictions is arbitrarily set to 2. This method of model evaluation has not previously been used in any prior research. The use of a novel evaluation approach is due to several reasons.

Firstly, prior research that uses Artificial Intelligence to predict trust considers trust as a binary variable. However, Yulin [43] indicates that trust is a spectrum. To simplify and bin the responses into either trustworthy or untrustworthy would significantly reduce the richness of the data and the subsequent results. Hence, for this study, the author maintains the view of trust as a spectrum instead of binary variables.

By setting the threshold to 2, predictions that are less than or equal to 2 marks away from the actual rating would then also be considered as accurate. As the original purpose of the ANN model is to measure the degree of trustworthiness of a chatbot's synthetic voice, it is more realistic to provide an acceptable margin of error, rather than ensuring that the model predicts a specific number.

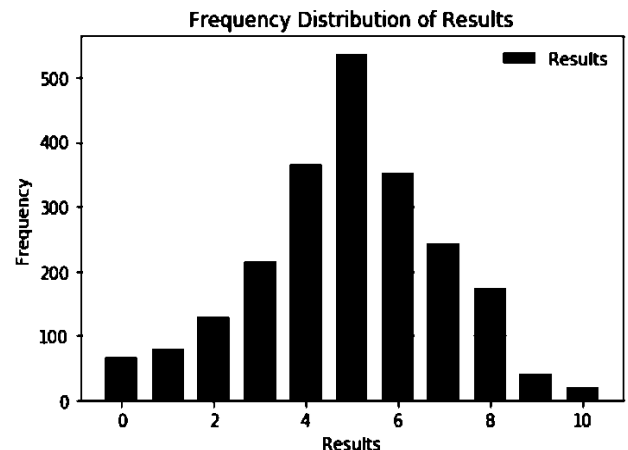


Figure 1 Frequency Distribution of Results.

### 5. Results

A graphical view of the frequency distribution of the results are shown in Figure 1, It is observed that the results are normally distributed around the mean ( $\mu = 4.90$ ,  $SD = 2.052$ ) with skewness =  $-0.228$  and kurtosis =  $-0.060$ . From Figure 1, it is observed that the majority of trust ratings ( $n = 537$ ) were 5, and the two extreme trust ratings (0 and 10) were less commonly chosen ( $n = 68$ ,  $n=19$ ) respectively.

#### 5.1. The Relationship between Vocal Cues and Trustworthiness

There was a statistically significant difference between arousal of emotion, valence

of emotion, type of emotion, duration of audio clip, speech rate, articulation rate, average loudness, ethnicity of rater and age of rater on trust ratings ( $p < .01$ ). A more detailed view of the results is shown in Table 1.

Table 1: ANOVA / Welch test results between each individual vocal cue on trust ratings

Vocal Cue	Type of Test	df1	df2	F	p	$\omega$
Duration	Welch	10	311.15	10.73	0	0.75
Speech rate	ANOVA	1	10	4.46	0	0.14
Articulation rate	ANOVA	1	10	5.95	0	0.16
Average Loudness	Welch	10	306.56	4.77	0	0.58
Arousal	ANOVA	1	10	3.36	0	0.12
Valence	Welch	10	307.34	2.41	0.009	0.4
Age of rater	Welch	10	327.13	7.88	0	0.7

There was also a statistically significant difference across trust ratings when considered jointly on arousal and valence. Pillai's Trace = .028,  $F(20, 4358) = 3.147$ ,  $p = .000$ , partial  $\eta^2 = .014$ . The average  $f_0$  was split into 4 groups using the first, second and third percentile as cutoff points. The groups were: Low  $f_0$  ( $M=107.03$ ,  $SD=6.87$ ,  $n=511$ ), LowMedium  $f_0$  ( $M=125.23$ ,  $SD=4.52$ ,  $n=584$ ), Medium-High  $f_0$  ( $M=139.03$ ,  $SD=5.57$ ,  $n=511$ ), and High  $f_0$  ( $M=191.19$ ,  $SD=29.58$ ,  $n=584$ ). The average  $f_0$  was used as the independent variable and the trust ratings were used as the dependent variable. Likewise, the emotion and ethnicity of rater were also used as the independent variable. The results are as shown in Table 2.

Table 2 ANOVA / Welch test results between trust ratings and the groups of each characteristics

Characteristic	Type of Test	df1	df2	F	p	$\omega$
Emotion	Welch	5	76.69	19.05	0	0.74
Ethnicity of rater	Welch	3	833.733	33.29	0	0.75
Average $f_0$	ANOVA	1	3	12	0	0.13

The means plot indicate that Malaysians trust more when the white male speaker a) speaks for a longer period (12s) than a shorter period (10.5s), b) speaks faster (170wpm and/or 3.6 syllables per second) than slower (153wpm and/or 3.1 syllables per second), c) has a louder voice (roughly about 56dB), d) has an

average  $f_0$  which is between 132.03Hz and 149.52Hz, e) shows a neutral emotion or less emotion ( $\text{arousal} < .3250$ ). Finally, findings among Malaysians indicate that Indians are the most trusting ethnic group with trust ratings of ( $5.5 < \mu < 5.75$ ), followed by "Other Bumiputera" ( $5.0 < \mu < 5.25$ ) and then Malay ( $\mu < 4.75$ ). Chinese are the least trusting Malaysian ethnic group. Graphical representations of the means plots are as shown in Figure 2 to Figure 9.

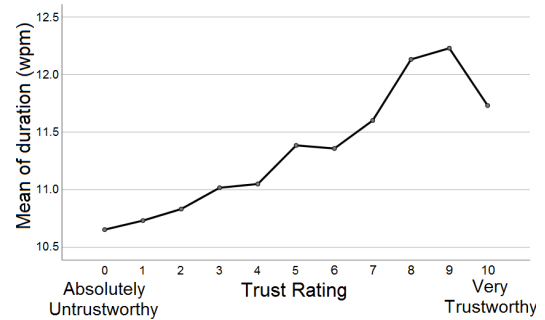


Figure 2 Means plot of duration across trust ratings.

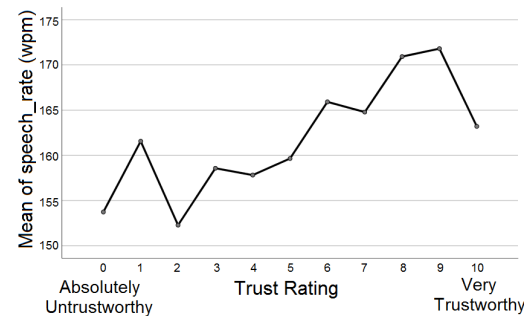


Figure 3 Means plot of speech rate across trust ratings.

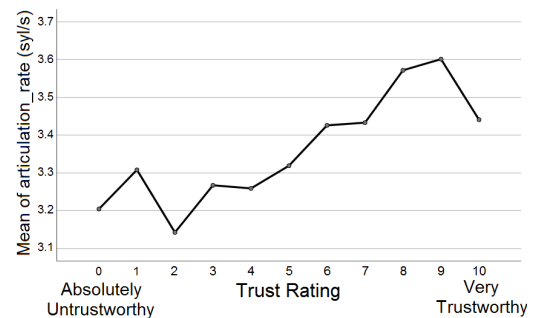


Figure 4 Means plot of articulation rate across trust ratings.

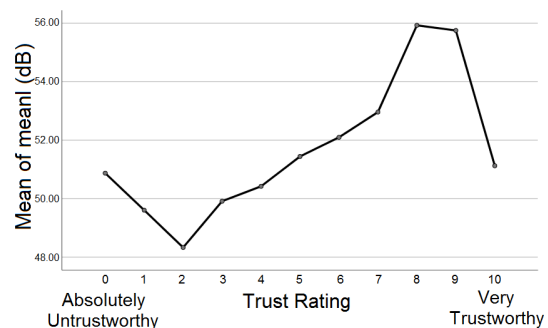


Figure 5 Means plot of loudness across trust ratings.

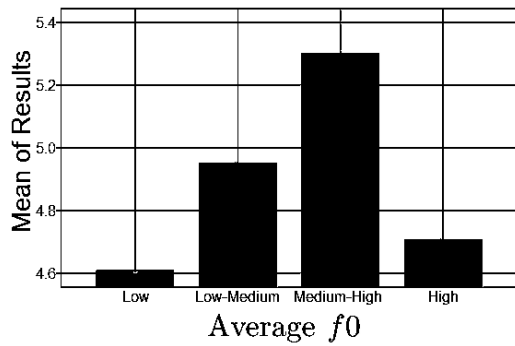


Figure 6 Means plot of f0 across trust ratings.

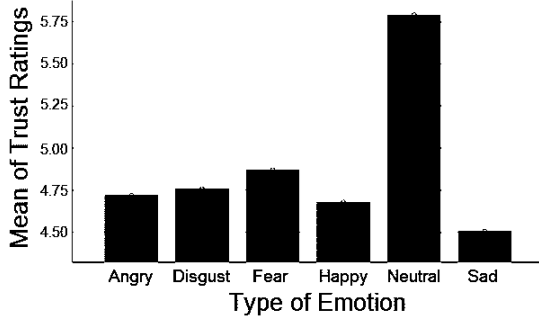


Figure 7 Means plot of emotion across trust ratings.

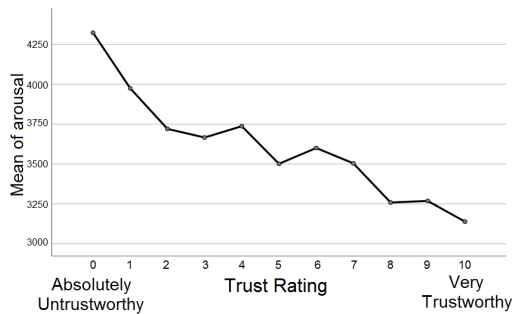


Figure 8 Means plot of arousal across trust ratings.

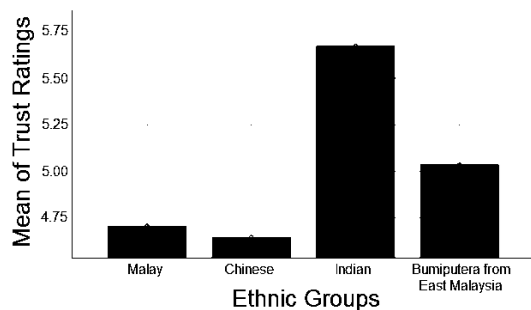


Figure 9 Means plot of ethnic groups across trust ratings.

### 5.2. Artificial Neural Network

The results of the ANN predictions are as shown in Figure 10, where "freq" represents frequency of occurrence and "difference" represents difference between the predicted and original value. It indicates how much deviation has happened between the predicted results and the original trust rating, and how often that deviation has occurred. Results indicate that the majority of

predictions (76.86%) have an acceptable difference (the difference is less than or equal to 2), and only a few predictions (23.14%) are very far off the mark (having a difference of more than 2).

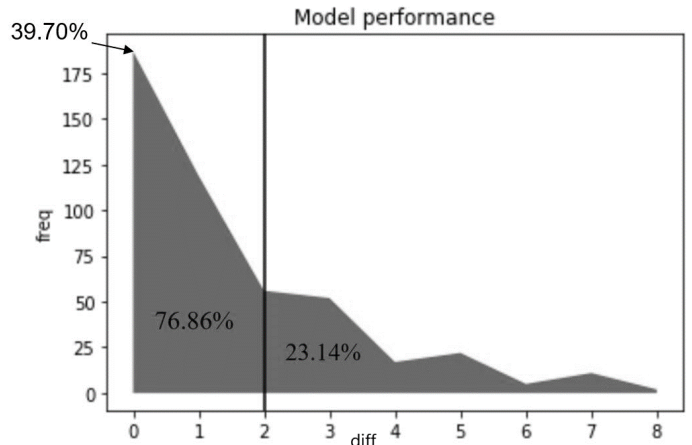


Figure 10 The cumulative percentage performance of ANN.

### 5.3. Discussion and Conclusion

The accurate predictions of the ANN clearly signify a relationship between vocal cues and trustworthiness. Contrary to previous related research, findings indicate that Malaysians tend to trust white male voices the most when the average f0 of the voices fall between 132.03Hz and 149.52Hz. This could potentially explain the reason behind the degree of dissension regarding the direction of association between different vocal features. It is not a direct association but rather, it is a parabolic curve with its peak when the average f0 is between 132.03Hz and 149.52Hz.

Findings prove that Malaysians tend to trust people who talk faster and for a longer time versus slower and shorter. This could be attributed to knowing more of the context of the audio clip and then judging the perceived trustworthiness according to the speaker's knowledge of the topic. Results also suggest that Indian Malaysians trust more than the other ethnic groups in Malaysia, while Chinese Malaysians trust the least. The underlying reasons for this relationship are still unknown, and it could potentially be investigated in future work.

Findings show that Malaysians tend to trust people with louder voices. This relationship could be due to the clarity of the audio clip - it is easier to identify the words being said when the voice is louder. Hence, clarity of voice could potentially be used as another variable when carrying out future work.

With regards to the ANN, 39.70% of the testing data set has been perfectly predicted by the neural network. With the threshold set at 2, the percentage accuracy of the neural network has risen to 76.86%. With such high trust prediction accuracy, the neural network model is potentially capable of fulfilling its original purpose.

In conclusion, there was a statistically significant difference between the arousal of emotion, valence of emotion, type of emotion, duration of audio clip, average f0, speech rate,

articulation rate, average loudness, ethnicity of rater and age of rater on trust ratings ( $p < .01$ ).

The findings from this research also conclude that Malaysians tend to trust white males who a) talk faster and longer vs slower and shorter, b) speak louder, c) have an f0 between 132.03Hz and 149.52Hz, d) show a neutral emotion or rather stoic (arousal < .325). Furthermore, results prove that Indians are the most trusting Malaysian ethnic group, followed by Bumiputera [1] from East Malaysia and followed by Malays. Chinese are the least trusting Malaysian ethnic group.

Lastly, the degree of trustworthiness has been quantitatively measured and successfully predicted by using an ANN model. The ANN model perfectly predicts the trustworthiness of a voice to an accuracy of 39.70% and given a threshold of two-point deviation, had a prediction accuracy of 76.86%. This Neural Network model will hopefully prove invaluable when used to predict the degree of trustworthiness in a chatbot's synthetic voice.

Although these conclusions have been derived and proven to be the best of the authors' abilities, there is still room for further validation. As this research is the first of its kind, especially pertaining to the use of a Neural Network model to predict the degree of trustworthiness in a voice, there is a need for more researchers to replicate or potentially improve the neural network model. Future researchers investigating this topic could possibly expand the variety of audio clips to encompass both male and female voices and people from various races. Future work could potentially include listener's emotional state and clarity of the audio clip as a factor in perceived trustworthiness. Lastly, this research would bring more value if future researchers are able to obtain responses from an international audience.

### Conflict of Interest

The authors declare no conflict of interest.

### Acknowledgment

This work was supported by the Research Management Center, Xiamen University Malaysia under the XMUM Research Program Cycle 3 (Grant XMUMRF/2019-C3/IECE/0006). The authors also thank Universiti Sains Malaysia for the partial funding of this work from the grant no. 304/PKOMP/6315137 and 1001/PKOMP/8014001.

### References

[1] N. Ambady, M. A. Krabbenhoft, and D. Hogan, "The 30-sec sale: Using thin-slice judgments to evaluate sales effectiveness," *J. Consum. Psychol.*, **16**(1), 4–13, 2006. [https://doi.org/10.1207/s15327663jcp1601\\_2](https://doi.org/10.1207/s15327663jcp1601_2).

[2] M. Stirrat and D. I. Perrett, "Valid Facial Cues to Cooperation and Trust: Male Facial Width and Trustworthiness," *Psychol. Sci.*, **21**(3), 349–354, 2010. <https://doi.org/10.1177/0956797610362647>.

[3] D. A. Kenny and T. V. West, "Zero acquaintance: Definitions, statistical model, findings, and process," *First impressions*, 129–146, New York, Guilford Press, 2008.

[4] N. Ambady and R. Rosenthal, "Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness," *J. Pers. Soc. Psychol.*, **64**(3), 431, 1993. <https://doi.org/10.1037/0022-3514.64.3.431>.

[5] N. Eisenkraft, "Accurate by way of aggregation: Should you trust your intuition-based first impressions?," *J. Exp. Soc. Psychol.*, **49**(2), 277–279, 2013. <https://doi.org/10.1016/j.jesp.2012.11.005>.

[6] M. A. Hecht and M. LaFrance, "How (Fast) Can I Help You? Tone of Voice and Telephone Operator Efficiency in Interactions 1," *J. Appl. Soc. Psychol.*, **25**(23), 2086–2098, 1995. DOI: 10.1111/j.1559-1816.1995.tb02389.x.

[7] N. Ambady, F. J. Bernieri, and J. A. Richeson, "Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream," *32*, Academic Press, 201–271, 2000. [https://doi.org/10.1016/S0065-2601\(00\)80006-4](https://doi.org/10.1016/S0065-2601(00)80006-4).

[8] S. M. Mohammad and S. Kiritchenko, "Using nuances of emotion to identify personality," *Proc. ICWSM*, 2013. <https://arxiv.org/abs/1309.6352>.

[9] H. H. Höhmann and E. Malieva, "The concept of trust: Some notes on definitions, forms and sources," *Trust Entrep.*, 7–23, 2005.

[10] C. Edwards, A. Edwards, B. Stoll, X. Lin, and N. Massey, "Evaluations of an artificial intelligence instructor's voice: Social Identity Theory in human-robot interactions," *Comput. Human Behav.*, **90**, 357–362, 2019. <https://doi.org/10.1016/j.chb.2018.08.027>.

[11] A. Goel, B. Creeden, M. Kumble, S. Salunke, A. Shetty, and B. Wiltgen, "Using watson for enhancing human-computer co-creativity," in *2015 AAAI Fall Symposium Series*, 2015.

[12] A. Følstad, C. B. Nordheim, and C. A. Bjørkli, "What makes users trust a chatbot for customer service? An exploratory interview study," in *International Conference on Internet Science*, 194–208, 2018. [https://doi.org/10.1007/978-3-030-01437-7\\_16](https://doi.org/10.1007/978-3-030-01437-7_16).

[13] L. Ciechanowski, A. Przegalinska, M. Magnuski, and P. Gloor, "In the shades of the uncanny valley: An experimental study of human-chatbot interaction," *Futur. Gener. Comput. Syst.*, **92**, 539–548, 2019. <https://doi.org/10.1016/j.future.2018.01.055>.

[14] J. Fenwick, L. Barclay, and V. Schmie, "Chatting: an important clinical tool in facilitating mothering in neonatal nurseries," *J. Adv. Nurs.*, **33**(5), 583–593, 2001. <https://doi.org/10.1046/j.1365-2648.2001.01694.x>.

[15] N. Bos, J. Olson, D. Gergle, G. Olson, and Z. Wright, "Effects of Four Computer-mediated Communications Channels on Trust Development," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2002), 135–140. <https://doi.org/10.1145/503376.503401>.

[16] L. Qiu and I. Benbasat, "Online Consumer Trust and Live Help Interfaces: The Effects of Text-to-Speech Voice and Three-Dimensional Avatars," *Int. J. Human-Computer Interact.*, **19**(1), 75–94, 2005. [https://doi.org/10.1207/s15327590ijhc1901\\_6](https://doi.org/10.1207/s15327590ijhc1901_6).

[17] S. Greenspan, D. Goldberg, D. Weimer, and A. Basso, "Interpersonal Trust and Common Ground in Electronically Mediated Communication," in *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, 2000), 251–260. <https://doi.org/10.1145/358916.358996>.

[18] A. Schirmer, Y. Feng, A. Sen, and T. B. Penney, "Angry, old, male-and trustworthy? How expressive and person voice characteristics shape listener trust," *PLoS One*, **14**(1), e0210555, 2019. <https://doi.org/10.1371/journal.pone.0210555>.

[19] S. M. Hughes and M. A. Harrison, "Your Cheatin' Voice Will Tell on You: Detection of Past Infidelity from Voice," *EPsychol.*, **15**(2), p. 1474704917711513, 2017. <https://doi.org/10.1177/1474704917711513>.

[20] P. Belin, B. Boehme, and P. McAleer, "The sound of trustworthiness: Acoustic-based modulation of perceived voice personality," *PLoS One*, **12**(10), 1–9, 2017. <https://doi.org/10.1371/journal.pone.0185651>.

[21] P. M. Doney, J. P. Cannon, and M. R. Mullen, "Understanding the influence of national culture on the development of trust," *Acad. Manag. Rev.*, **23**(3), 601–620, 1998. <https://doi.org/10.2307/259297>.

[22] D. Gefen and T. H. Heart, "On the need to include national culture as a central issue in e-commerce trust beliefs," *J. Glob. Inf. Manag.*, **14**(4), 1–30, 2006. <https://doi.org/10.4018/jgim.2006100101>.

[23] S. Zaheer and A. Zaheer, "Trust across borders," *J. Int. Bus. Stud.*, **37**(1), 21–29, 2006. <https://doi.org/10.1057/palgrave.jibs.8400180>.

[24] O. Niebuhr, J. Voße, and A. Brem, "What makes a charismatic speaker? A computer-based acoustic-prosodic analysis of Steve Jobs tone of voice," *Comput. Human Behav.*, **64**, 366–382, 2016. <https://doi.org/10.1016/j.chb.2016.06.059>.

[25] L. Anolli and R. Ciceri, "The voice of deception: Vocal strategies of naive and able liars," *J. Nonverbal Behav.*, **21**(4), 259–284, 1997.

[26] M. Zuckerman, R. S. DeFrank, J. A. Hall, D. T. Larrance, and R. Rosenthal, "Facial and vocal cues of deception and honesty," *J. Exp. Soc. Psychol.*, **15**(4), 378–396, 1979. [https://doi.org/10.1016/0022-1031\(79\)90045-3](https://doi.org/10.1016/0022-1031(79)90045-3).

[27] C. Kirchhübel and D. M. Howard, "Detecting suspicious behaviour using speech: Acoustic correlates of deceptive speech--An exploratory investigation," *Appl. Ergon.*, **44**(5), 694–702, 2013. <https://doi.org/10.1016/j.apergo.2012.04.016>.

[28] I. Torre, L. White, and J. Goslin, "Behavioural mediation of prosodic cues to implicit judgements of trustworthiness," 2016. <https://doi.org/10.21437/SpeechProsody.2016-167>.

[29] N. Korovaiko and A. Thomo, "Trust prediction from user-item ratings," *Soc. Netw. Anal. Min.*, **3**(3), 749–759, 2013. <https://doi.org/10.1007/s13278-013-0122-z>



- [30] T. DuBois, J. Golbeck, and A. Srinivasan, "Predicting trust and distrust in social networks," in 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing, 2011, 418–424. <https://doi.org/10.1109/PASSAT/SocialCom.2011.56>.
- [31] B. Zong, F. Xu, J. Jiao, and J. Lv, "A broker-assisting trust and reputation system based on artificial neural network," in 2009 IEEE International Conference on Systems, Man and Cybernetics, 2009, 4710–4715. <https://doi.org/10.1109/ICSMC.2009.5346098>.
- [32] D. Bejou, B. Wray, and T. N. Ingram, "Determinants of relationship quality: an artificial neural network analysis," *J. Bus. Res.*, **36**(2), 137–143, 1996. [https://doi.org/10.1016/0148-2963\(95\)00100-X](https://doi.org/10.1016/0148-2963(95)00100-X).
- [33] J. J. Lee, B. Knox, and C. Breazeal, "Modeling the dynamics of nonverbal behavior on interpersonal trust for human-robot interactions," in 2013 AAAI Spring Symposium Series, 2013. <http://hdl.handle.net/1721.1/69244>.
- [34] L. Ekonomou, "Greek long-term energy consumption prediction using artificial neural networks," *Energy*, **35**(2), 512–517, 2010. <https://doi.org/10.1016/j.energy.2009.10.018>.
- [35] M. Imhof, "Listening to Voices and Judging People," *Int. J. List.*, **24**(1), 19–33, 2010. <https://doi.org/10.1080/10904010903466295>.
- [36] B. L. Smith, B. L. Brown, W. J. Strong, and A. C. Rencher, "Effects of Speech Rate on Personality Perception," *Lang. Speech*, **18**(2), 145–152, 1975. <https://doi.org/10.1177/002383097501800203>.
- [37] M. N. Ismail, S. S. Chee, H. Nawawi, K. Yusoff, T. O. Lim, and W. P. T. James, "Obesity in Malaysia," *Obes. Rev.*, **3**(3), 203–208, 2002. <https://doi.org/10.1046/j.1467-789X.2002.00074.x>.
- [38] V. Swami and A. Furnham, "Self-assessed intelligence: Inter-ethnic, rural–urban, and sex differences in Malaysia," *Learn. Individ. Differ.*, **20**(1), 51–55, 2010. <https://doi.org/10.1016/j.lindif.2009.11.002>.
- [39] P. Barros, N. Churamani, E. Lakomkin, H. Siqueira, A. Sutherland, and S. Wermter, "The OMG-Emotion Behavior Dataset," in Proceedings of the International Joint Conference on Neural Networks, 2018, 2018-July. <https://doi.org/10.1109/IJCNN.2018.8489099>.
- [40] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," *J. Pers. Soc. Psychol.*, **17**(2), 124–129, 1971. <https://psycnet.apa.org/doi/10.1037/h0030377>.
- [41] J. Trouvain, S. Schmidt, M. Schröder, M. Schmitz, and W. J. Barry, "Modelling personality features by changing prosody in synthetic speech," 2006. <https://doi.org/10.22028/D291-25920>.
- [42] H. Quené, "On the just noticeable difference for tempo in speech," *J. Phon.*, **35**(3), 353–362, 2007. <https://doi.org/10.1016/j.wocn.2006.09.001>.
- [43] G. Yulin, "The Spectrum of Trust and Distrust," *Jiangsu Soc. Sci.*, **1**, 2012.

## Study of the Effect of Abnormalities in the External Ear Inducing Hearing Problems

Jihane Melloui\*, Jamila Bakkoury, Omar Bouattane

SSDIA Laboratory, ENSET of Mohammedia, Hassan II University of Casablanca, Mohammedia, 28830, Morocco

### ARTICLE INFO

Article history:

Received: 30 June, 2020

Accepted: 17 July, 2020

Online: 09 August, 2020

Keywords:

Tinnitus

Ear canal

Acoustic model

Anomaly

Resonant frequency

### ABSTRACT

Tinnitus is a phenomenon for which the patient hears sound in the absence of any external sound source. To this day, there is no cure for this phantom sound perception. But it can be masked temporarily to help relief the patient's pain. In order to allow this, a better understanding of the phenomenon is needed. A validated acoustic model of the outer ear developed by the authors is used in this study. This model allows to study the effect of the presence of an anomaly (a cavity, a swelling or a foreign fluid) in the human auditory canal. These anomalies are modeled as a modification of the section of the ear canal or as an alternation of the medium of sound propagation in the ear canal. A parametric study involving the position, width and height of the anomaly as well as the sound velocity in the ear canal is conducted. The results obtained make it possible to conclude on the effect of each parameter on the frequency response of an auditory canal with anomaly.

### 1. Introduction

Tinnitus is the perception of a sound for which there is no external stimulus; patients suffering from it perceive a sound which may take the form of ringing, buzzing, beeping, hissing or whooshing [1]. This condition can impair one or both ears. According to national health surveys, almost 10 per cent of adults endure a form of tinnitus [2]. The prevalence of tinnitus among employees subject to workplace noise is 15 per cent [3].

This condition has consequences that can be continuous or intermittent, long-term or short-term, with a sudden or incremental occurrence. This can also be associated with the cardiac or respiratory rhythm. Its magnitude also changes during the day, with a period of morning calm and an evening or nightly rise (nocturnal rhythm) [4].

In addition, tinnitus is related not only to alternation in the auditory system such as speech perception, localization of sound and auditory attention [5-8], but also to affective disorders such as depression or anxiety, insomnia, confusion and lower quality of life [9-11].

There is no official classification of tinnitus recognized by all scientists and doctors. The most common types are two: One dividing tinnitus by manifestation (objective or subjective tinnitus) and the other defining it by the potential cause of tinnitus [12].

Objective tinnitus is the rarest but also the easiest to identify. It is a perceived pulsatile sound produced by the body (vascular or

muscular cause) which the examiner may also hear. This type of tinnitus represents about 10 per cent of cases of tinnitus. The 90 per cent remaining are subjective [13].

Subjective tinnitus is only heard by the patient. The frequency perceived does not come from the body, but from one or more body system failures that may be from the cortex to the external auditory canal [14]. It can be caused by ear problems (tinnitus aurium) or cerebral problems (tinnitus cerebri) [1]. This tinnitus is a perceived sensation in the absence of any outer physical source (phantom sound) and its source is hard to determine [14]. It is characterized as persistent after continuous occurrence for 6 months [15].

The symptom of tinnitus can be induced by multiple causes and is often accompanied by loss of hearing at high frequencies linked to ageing. Different parts of the hearing system can cause tinnitus, such as the outer ear [16]. For example, an excessive ear wax can cause pressure in the ear canal and causes a change in the resonance frequency at the eardrum.

According to the second type of classification (potential cause of tinnitus) and, as in [17], the origin of tinnitus may be categorized as:

- Otological origin: The auditory system (external, middle or inner ear) or the auditory nerve. In this case, tinnitus is usually accompanied by hearing loss and sometimes dizziness.
- Central origin: The relay level of central auditory canal;

\*Corresponding Author: Jihane Melloui, [jihane.melloui@gmail.com](mailto:jihane.melloui@gmail.com)

- Non otological origin: vascular, muscular, cerebral or cervical;
- Unknown origin: if no lesion and no other symptom are detected.

Several efforts have been made in the past to suppress or treat tinnitus. Some of these interventions included the use of drugs for the treatment of tinnitus [18-19]. In some serious cases, the sensation was eliminated by taking drastic actions, like cutting a specific nerve into the ear [20]. This completely removes the ability to hear using that ear.

The improved understanding of tinnitus physiopathological processes during the last decade has opened the way for many medical therapeutic researches [21-25]. The most common treatments used for tinnitus managing are cognitive and behavioral therapy (CBT), tinnitus retraining therapy (TRT) and sound therapy (SD)[26-31].

Cognitive and behavioral therapy is based on the premise that "what matters is not what happens to us, but how we do it." [25]. In fact, this treatment does not enhance tinnitus appreciation, but allows the individual to ignore it by emotional adjustments and behavioral changes [26].

Tinnitus retraining therapy is one of the most common techniques for managing tinnitus. Its neurophysiological model indicates that the limbic mechanism plays an important role in the perception of tinnitus [27]. The TRT principle is based on the patient's ability to become accustomed to the tinnitus signal.

Sound therapy is a program that uses external sound to manage tinnitus either via acoustic or electric stimulation [24]. It offers relief from tinnitus for both the short and long term. This treatment aims to minimize patient perception, effectively reducing the signal-to-noise ratio of tinnitus to ambient or surrounding noise [29-31].

Previous tinnitus suppression attempts involved the use of electrical stimulation [32]. In practice, it does not suppress tinnitus, but rather tries to mask the current tinnitus sound with another sound produced by the electrical stimuli. The key purpose of electrical stimulation is to eliminate tinnitus without any additional sound being applied. The direct electrical currents used cause severe damage to the biological tissue and can therefore not be used in the long term [32].

This problem can be fixed by means of an approach in which the tinnitus patient is subjected to a hearing test of varying pitches to determine the frequency of the perceived tinnitus sound. A noise generator (white noise, filtered noise, pink noise, ambient sounds ...) is used to mask tinnitus [24, 33]. This generator provides signals both above and below the specified tinnitus frequency.

The corresponding frequency of tinnitus was between 9000 and 11000 Hz for 86 per cent of the 771 patients in an experiment led by Johnson in [34]. The median tinnitus frequency of the 29 patients taking part in Neff's experiment [33] was 4046 Hz with a standard deviation of 2212,25 Hz. In Reavis experiment [29] on 20 participants, the mean frequency of tinnitus was 6929 Hz with a standard deviation of 2090 Hz.

It should be mentioned that such treatments do not heal tinnitus, but rather superimpose stimulus on the original tinnitus symptom in an attempt to mask the perceived sound. Given the multidisciplinary aspect of the condition, it is better to merge approaches for better results [35-37]. A deeper understanding of the tinnitus syndrome is required to make these masking methods more effective. To this end, further investigations must be carried out into causes that stimulate the phantom frequencies and the loss of hearing at high frequencies.

In this study, an analysis is carried out with the intention of further understanding the effect of certain anomalies that induce tinnitus on the resonance frequency of the human outer ear. An acoustic model previously developed by authors is used in the current research by inserting various diseases or anomalies into it, either through introducing acoustic resonators or by acting on some of its parameters.

The anomalies investigated are those that can lead to:

- A change of section at any position along the ear canal;
- A change in the propagating medium of the sound wave in the outer ear [38].

The introduction of a resonator simulating a swelling makes it possible to model, for example, an excessive presence of earwax, a pimple, an infection or a foreign body in the ear canal (Figure 1). The addition of an acoustic resonator simulating a cavity can make it possible to model anomalies such as lesions of the ear canal. The presence of a foreign fluid in the ear canal can be modeled by acting on the wave propagation velocity parameter in the outer ear. If this fluid is water, the swimmer's ear disease can be mentioned.

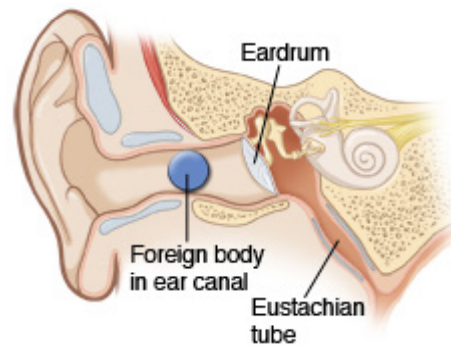


Figure 1: External auditory canal with a foreign body

## 2. Acoustic model of the outer ear

### 2.1. Outer ear

The outer ear plays a very significant function in the transmission of human sensory information from the environment to the tympanic membrane. It functions as a low frequency filter, a resonator for the middle range and direction dependent filter on high frequencies to enhance spatial awareness [39].

By providing a narrow entrance to the auditory system, the external ear protects the eardrum from mechanical damage. It provides directional amplification of the incident sound pressure level. It also transforms the incoming sound spectrum which was subjected to diffraction and amplifies acoustic waves at higher

frequencies [40]. These properties makes that the ear plays a significant role in spatial hearing.

In an attempt to comprehend these characteristics, it is better to divide the external ear system into its functional elements: the head, torso and pinna acting as diffraction bodies, the concha and the ear canal acting as acoustic resonators, and the eardrum providing an acoustic termination.

Such components must be seen as part of an integrated network, since they work together to transfer sound pressure from the free field to the middle ear entrance, depending on the frequency, direction and anatomy of the hearing system. Human ear models can provide a better understanding of how the human ear responds when it is artificially stimulated.

An acoustic model taking account of this integration is previously developed by authors [41] and is used in this paper to model anomalies present in the outer ear. This model is presented briefly in Figure 2. The purpose of this study is to examine the impact of anomalies on the external ear function. This is done by comparing the frequency response of a healthy external ear to an external ear with an anomaly.

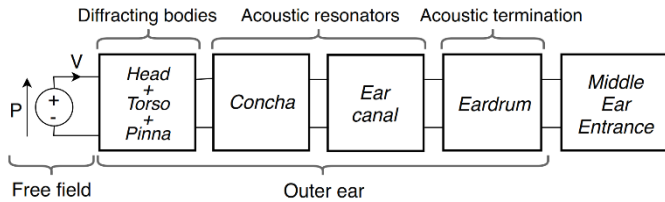


Figure 2: Acoustic Model of the External Ear

### 2.2. Outer ear acoustic model

The acoustic model presented associates with each part of the outer ear an acoustic resonator modeled by a transfer matrix. These transfer matrices can model a segment of the concha, auditory canal or even the eardrum. These resonators can have several forms (Figure 3) and are used, in this study, to model segments of the ear canal as well as abnormalities present in the outer ear. The corresponding transfer matrix of the outer ear (with or without anomaly) is obtained by multiplying the transfer matrixes of each segment.

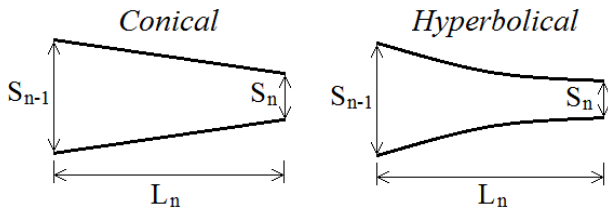


Figure 3: Acoustic resonator segment of different forms with parameters

The ear canal is divided into several segments (Figure 4), the number of segments depends on the precision sought in the calculation. A detailed study of the form as well as the optimal number of segments to approach with good precision the actual shape of the auditory canal is carried out previously by the authors in [41]. This model adopts the sound pressure  $p_i$  and the mass velocity  $v_i$  as the two state variables for linking the two ports of the  $i$ -th acoustic resonator.

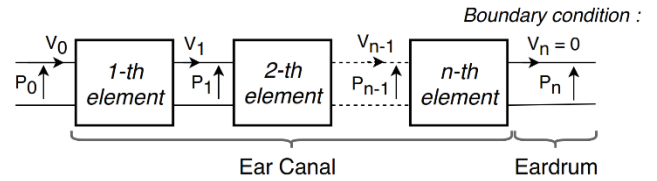


Figure 4: Segments of the ear canal acoustic model and boundary conditions

As illustrated in Figure 4,  $p_0$  and  $v_0$  represent the acoustic pressure and the mass velocity at the input of first resonator coming from the free field. The acoustic pressure and the mass velocity  $p_n$  and  $v_n$  are the outputs of the acoustic model of the outer ear which is also the entrance of middle ear. At the output of the  $n$ -th resonator,  $v_n=0$  simulates the infinite impedance of the tympanic membrane since it behaves as an obstacle. This is used as a boundary condition for calculating the frequency response of the model.

The transfer matrix of the conical acoustic resonator  $T_C$  is described in (1). The section parameters are shown in Figure 3.

$$T_C = \frac{1}{1 + \gamma L_n} \begin{bmatrix} T_{1C} & T_{2C} \\ T_{3C} & T_{4C} \end{bmatrix} \quad (1)$$

where:

$$T_{1C} = \cos(K L_n) + \frac{\gamma}{K} \sin(K L_n) ; T_{2C} = -j \frac{c}{S_{n-1}} \sin(K L_n) ;$$

$$T_{3C} = \frac{-j S_n}{c K} \left[ \frac{-\gamma^2 L_n}{1 + \gamma L_n} \cos(K L_n) + \left( K + \frac{\gamma^2}{(1 + \gamma L_n) K} \right) \sin(K L_n) \right] ;$$

$$T_{4C} = \frac{S_n}{S_{n-1}} \left( \cos(K L_n) - \frac{\gamma}{(1 + \gamma L_n) K} \sin(K L_n) \right) \quad (2)$$

$$\gamma = \frac{1}{L_n} \left( \sqrt{S_n/S_{n-1}} - 1 \right) \quad (3)$$

The wave number  $K$  (in radius/m) takes into account the attenuation of the sound in the human ear as advised in [42]. It should be mentioned that this segment can be used to model uniform acoustic resonator simply by taking  $S_n$  equal to  $S_{n-1}$ . The transfer matrix of the hyperbolic acoustic resonator  $T_H$  is described in (4).

$$T_H = e^{\delta L_n} \begin{bmatrix} T_{1H} & T_{2H} \\ T_{3H} & T_{4H} \end{bmatrix} \quad (4)$$

where:

$$T_{1H} = \cos(K_h L_n) - \frac{\delta}{K_h} \sin(K_h L_n) ; T_{2H} = -j \frac{c}{S_{n-1}} \frac{K}{K_h} \sin(K_h L_n)$$

$$T_{3H} = \frac{-j K S_n}{K_h c} \sin(K_h L_n) ; T_{4H} = \cos(K_h L_n) + \frac{\delta}{K_h} \sin(K_h L_n) \quad (5)$$

$$K_h = \sqrt{K^2 - \left( \frac{\ln(S_n/S_{n-1})}{2L_n} \right)^2} \quad (6)$$

$$\delta = \frac{\ln(S_n/S_{n-1})}{2L_n} \quad (7)$$



The acoustic model is applied to two of the 15 auditory canals published by Stinson et al. [43], referred to in "Canal 4" and "Canal 6." Canal 4 and Canal 6 are measured from the right ear of a man and the left ear of a woman, respectively. Therefore, this study takes into account both ears and the sex of the individual.

Complete anatomy of the canals studied is available in [43]. The dimensions of the acoustic resonator segments and the variation in diameter over the length of both canals are shown in Figures 5 and 6.

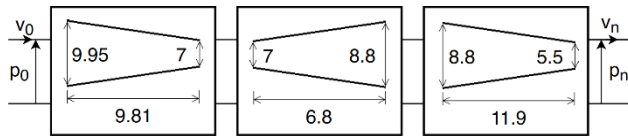


Figure 5: Parameters of the acoustic model of "Canal 4"

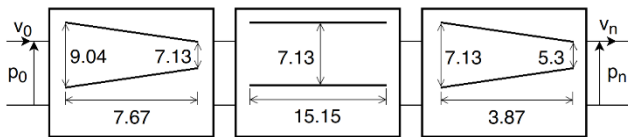


Figure 6: Parameters of the acoustic model of "Canal 6"

Canal 4 has a length of 28.51 mm and canal 6 has a length of 26.69 mm. Despite the fact that these lengths are of the same order of magnitude, there is a very different variation in the section along the ear canal. Both canals are modeled using three acoustic resonators in series. This allows to approach with good precision the true forms of these ear canals. Other segments of acoustic resonators are injected thereafter in order to model anomalies present in these healthy auditory canals.

### 3. Anomaly modeling

The normal function of transmission of the ear is compromised by alterations in its anatomy. It can be due to individual differences or technical blockages, such as foreign body parts, hearing aid implants, infections or tympanic membrane perforation.

The anomalies studied are those that can cause a difference in the section at any location along the external auditory canal or an alternation in the sound propagation medium in the outer ear. The first type of anomalies, consisting of an obstacle present in the ear canal, may be modeled by swelling or cavities. The parameters concerning this type of anomaly are presented in Figure 7. The second type can be modeled by the presence of a fluid in the auditory canal, which would cause the sound transmission medium to change from the air to the fluid contained in the auditory canal. These abnormalities can cause the resonance frequency of the ear to alter.

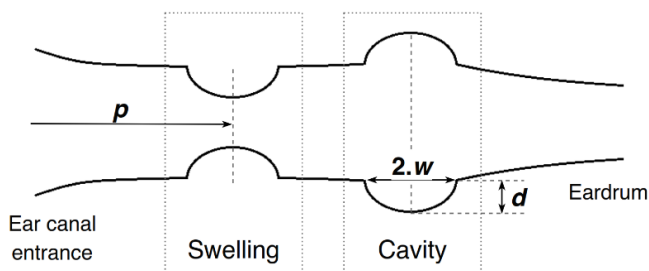


Figure 7 : Parameter definition of anomalies that change the section of the ear canal

Acoustic models of canals 4 and 6 are those of healthy outer ears into which these anomalies will be introduced (Figure 8). Each anomaly is modeled considering its shape. It is modeled by two acoustic resonators in series which can have a conical, hyperbolic or even a uniform section. Given the nature of the obstacle type anomaly studied in this paper, it is modeled using two segments of hyperbolic acoustic resonators as shown in Figure 7.

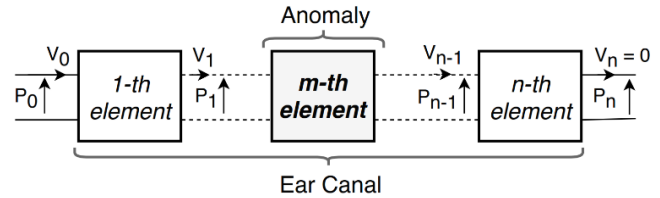


Figure 8 : Insertion of an acoustic resonator modeling anomaly in a healthy auditory canal model

A parametric analysis will be carried out using the acoustic model acquired (auditory canal with anomaly). In this analysis, the effect of position "p", width "w" and depth "d" of the obstacle type anomaly on the frequency response of the outer ear is investigated. The second type of anomaly concerns the study of the effect of the propagation speed of sound "c" in the ear canal.

Parameters "p" and "w" are expressed as a function of the total length "L" of the ear canal. The position values of the anomaly investigated in this work are 0.2L, 0.35L, 0.5L, 0.65L and 0.8L. The values considered for the anomaly width are 0.05L, 0.1L, 0.15L, 0.2L and 0.25L. The depth parameter "d" is expressed as a function of the width of the anomaly to hold the anomaly in proportion. This parameter can be set to 0.5w, 0.75w, 1w, 1.25w and 1.5w. In order to simulate the different levels of severity of the diseases that cause the presence of fluid in the ear canal, the speed of sound "c" parameter can be set to 350, 500, 650, 800 and 950m/s. This results in a test matrix of 65 cases.

Only one case is presented in Figure 9 given the large number of cases studied. This acoustic model of the "Canal 4" ear canal has a swelling of a width of  $w = 0.15L$  and a height of  $d = 1w$  at a distance of  $p = 0.65L$  from the entrance of the auditory canal. The same procedure applies in the other cases. In cases of abnormalities involving the presence of foreign fluids in the ear canal, the dimensions of the model are similar to those of the healthy model shown in Figures 5 and 6 (only the speed of sound parameter is modified).

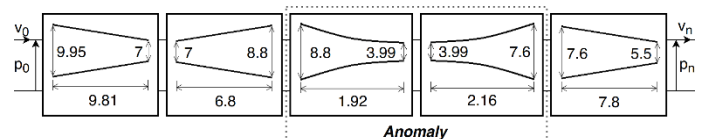


Figure 9 : Acoustic resonator modeling a swelling anomaly ( $p=0.65L$ ,  $w=0.15L$  and  $d=1w$ ) present in auditory canal « Canal 4 »

### 4. Results and discussion

The sound pressure level (SPL) presented in (8) is computed using Matlab. The results are reported in two formats: figures and tables. The figures show the variation in sound pressure level with frequency. The numerical values of the resonance frequencies as well as the gains at these frequencies are presented in the tables to improve the readability of the results.

$$SPL = 20 \log \left| \frac{p_n}{p_0} \right| \quad (8)$$

#### 4.1. Effect of the propagation medium

Figure 10 shows the effect of the medium of sound propagation, which is characterized by the speed of sound “C”, on the frequency response of the auditory canal. In this simulation, the greater the value of “C”, the greater the volume of fluid contained in the ear.

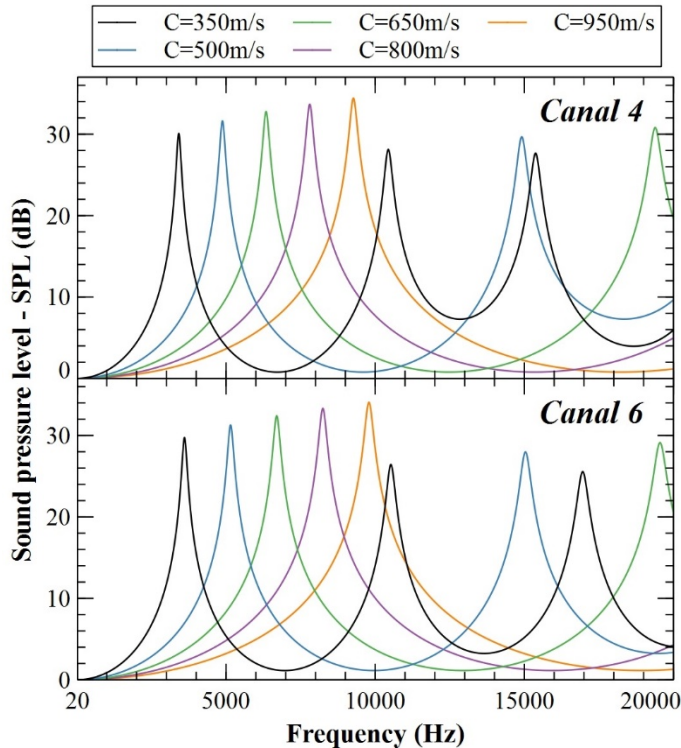


Figure 10 : Frequency response of Canal 4 and Canal 6 to different sound speed values

In the case of “Canal 4”, it can be shown from Figure 10 that the first resonant frequency occurred at 3416, 4898, 6341, 7811 and 9272 Hz at sound velocity of 350, 500, 650, 800 and 950 m/s, respectively. The first resonant frequency for “Canal 6” is 3626, 5154, 6688, 8246 and 9793 Hz, respectively, at 350, 500, 650, 800 and 950 m/s.

It can be observed that only the ear canals with a propagation medium inducing a sound speed of 350, 500 and 650 m/s have a second resonant frequency. This resonant frequency for “Canal 4” is 9413, 14935 and 19388 Hz, respectively, at 350, 500 and 650 m/s. For “Canal 6”, these values are 10529, 15037 and 19556 Hz at sound velocity of 350, 500 and 650 m/s, respectively.

For the third resonant frequency, only the canal providing a propagating medium with a sound velocity of 350 m/s is concerned. “Canal 4” resonates at 15383 Hz and “Canal 6” resonates at 16960 Hz. These ear canals are healthy canals.

These results show that the presence of fluid in the ear canal tends to reduce the number of resonant frequencies of the ear canal in the audible band. The resonance frequency values obtained for “Canal 6” are higher than those found for “Canal 4”. That is due to the disparity in the physiology of these two auditory canals. The

presence of a foreign fluid in the ear canal leads to an increase in the speed of sound propagation in this canal, which has a significant effect on the frequency of the resonance. Indeed, the frequency of resonance rises with the speed of sound propagation.

Considering the first resonance frequency, it can be reported that the gains at these frequencies for canals 4 and 6 are almost identical. This gain ranges by an order of magnitude from 30 dB at a sound transmission speed of 350 m/s to 34 dB at a speed of 950 m/s. From these findings, it can be assumed that an increase in the amount of fluid in the ear canal leads to an increase in the sound pressure level.

#### 4.2. Effect of anomaly position

In order to study the effect of the position of a cavity or swelling on the frequency response of the human auditory canal, the parameters of the anomaly width “w” shall be set to 0.15L and the depth or height “d” of the anomaly shall be set at 1w.

Figure 11 indicates the evolution of the SPL as a result of the frequency in the audible band. This concerns the case of swollen auditory canals with various anomaly positions. Figure 12 shows that of auditory canals with a cavity. Table 1 indicates the relative deviation between the resonance frequency findings for a healthy canal and canals with anomaly. The table further shows the variations in SPL observed at these frequencies. It should be noted that for the calculation of these deviations, a healthy auditory canal is used as a reference for all the canals with anomaly.

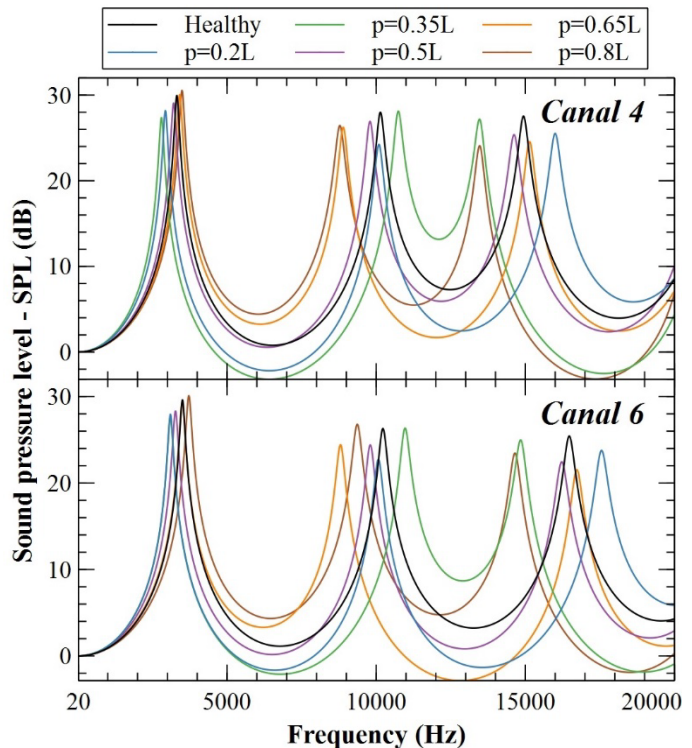


Figure 11 : Frequency response of the auditory canals with a swelling anomaly of different positions p (w= 0.15L and d=1w)

A negative SPL for the swelling located at 0.2L and 0.35L in “Canal 4”, for the frequency bands [4918-7890] Hz and [4592-8344] Hz can be seen from the reading of Figure 11. In these same bands, gains for anomalies of positions 0.65L and 0.8L are significantly higher than those of the healthy canal. Between

15500Hz and 19100Hz, the frequency response of the canals with anomalies at 0.35L and 0.8L is almost identical and presents negative gains.

The frequency response of “Canal 6”, consisting of a 0.2L and 0.35L swelling, shows negative gains in the [5200-8400] Hz band. Same observation for 0.35L and 0.8L in the frequency band [17000-20000] Hz. There are also negative gains for swelling cases in the 0.2L and 0.65L positions in the [12365-14774] Hz and [10951-14664] Hz bands, respectively. The healthy ear canal amplifies the sound of these frequencies. When such abnormalities have occurred, we find that this ear canal starts to decrease the sounds heard at these frequencies.

According to Figure 11 and taking into account the results obtained for “Canal 4”, there are maximum differences of -18.77, -22.82, -10.74, -21.6 and -24.77 dB between the SPL curve of the healthy canal and those of the auditory canals with anomalies at positions 0.2L, 0.35L, 0.5L, 0.65L and 0.8L, respectively. These differences are observed at 14924, 14947, 10175, 10156 and 14955 Hz frequencies at 0.2L, 0.35L, 0.5L, 0.65L and 0.8L respectively. Negative deviations mean that the sound perceived at these frequencies is reduced. This can lead to a slight loss of hearing.

Absolute variations of 15.39, 18.12, 9.26, 19.04 and 20.01 dB between the healthy canal SPL and the canals with anomaly positions of 0.2L, 0.35L, 0.5L, 0.65L and 0.8L, respectively. Such variations are found at 16021, 13464, 9754, 8886 and 8768 Hz frequencies for 0.2L, 0.35L, 0.5L, 0.65L and 0.8L, respectively. Similar observations may be made from “Canal 6”. Indeed, results show very similar variations but with slightly higher frequencies (of the order of 1500 Hz). These positive variations (amplification) may lead to an over-sensitivity to sounds having these frequencies in some cases and for some patients.

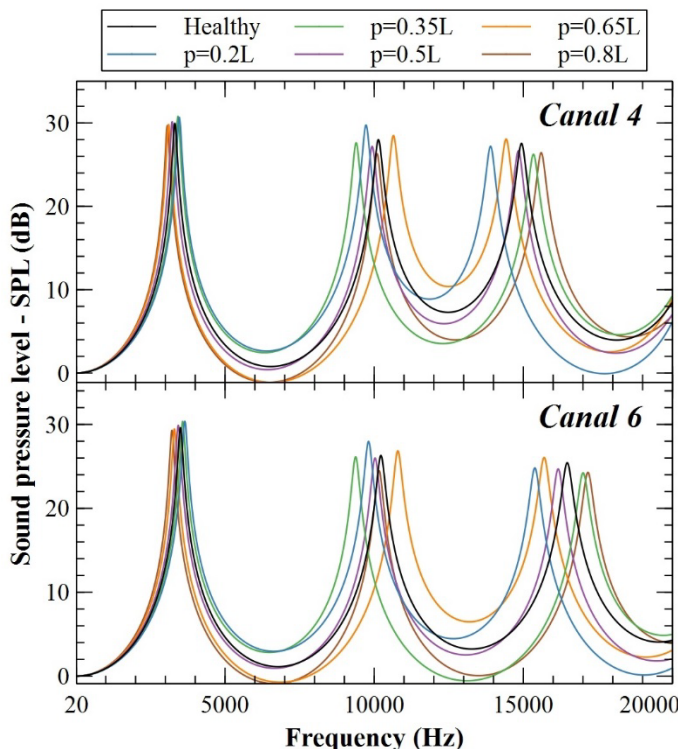


Figure 12 : Frequency response of the auditory canals with a cavity anomaly of different positions p (w= 0.15L and d=1w)

Figure 12 shows that the auditory canals 4 and 6 with a cavity type anomaly do not generally have negative SPL values. From this figure and taking into consideration the obtained results for “Canal 4”, the maximum differences are -18.88, -16.68, -7.54, -12.92 and -14.65 dB between the healthy canal SPL and the auditory canals with 0.2L, 0.35L, 0.5L, 0.65L and 0.8L location anomalies, respectively. Those very variations are found at 14962, 10162, 10197, 10113 and 14911 Hz frequencies at 0.2L, 0.35L, 0.5L, 0.65L and 0.8L respectively.

Absolute differences of 15.91, 16.02, 6.15, 12.74 and 12.53 dB between the healthy canal SPL and the canal with anomalies of 0.2L, 0.35L, 0.5L, 0.65L and 0.8L respectively. These differences are observed at frequencies 13886, 9374, 9875, 10673 and 15630 Hz for positions 0.2L, 0.35L, 0.5L, 0.65L and 0.8L respectively.

For “Canal 6”, these SPL differences are of the same order of magnitude for all the studied cases. However, these differences are observed at higher frequencies of approximately 1500 Hz at 0.2L and 0.8L positions, 6000 Hz at 0.5L and 0.65L positions and almost similar at 0.35L position.

This confirms the observation obtained from the swollen ear canals. However, the variations found for the anomaly of the cavity are smaller than those reported for the anomaly of the swelling.

Table 1: Effect of the position of swelling and cavity anomalies on frequency response for canal 4 and 6

	Resonant frequencies			SPL at resonant frequencies		
	RF-1	RF-2	RF-3	SPL (RF-1)	SPL (RF-2)	SPL (RF-3)
<b>Canal 4</b>						
<b>Healthy</b>	3317 Hz	10140 Hz	14940 Hz	29,96 dB	28,00 dB	27,54 dB
<b>Swelling</b>						
p=0.2L	-11,6%	-0,49%	7,10%	-5,94%	-13,4%	-7,33%
p=0.35L	-15,6%	5,92%	-9,84%	-8,68%	0,39%	-1,27%
p=0.5L	-3,26%	-3,41%	-2,14%	-3,00%	-3,86%	-7,88%
p=0.65L	2,89%	-12,3%	1,34%	0,30%	-6,29%	-10,8%
p=0.8L	5,19%	-13,4%	-9,84%	1,97%	-5,57%	-12,5%
<b>Cavity</b>						
p=0.2L	4,49%	-4,08%	-6,96%	2,37%	6,25%	-1,23%
p=0.35L	3,11%	-7,34%	2,68%	2,77%	-1,39%	-4,68%
p=0.5L	-2,68%	-2,05%	-0,67%	0,60%	-2,89%	-3,20%
p=0.65L	-6,33%	5,03%	-3,41%	-0,53%	1,75%	1,92%
p=0.8L	-7,54%	-0,39%	4,42%	-0,90%	-6,04%	-4,03%
<b>Canal 6</b>						
<b>Healthy</b>	3504 Hz	10230 Hz	16470 Hz	29,61 dB	26,30 dB	25,43 dB
<b>Swelling</b>						
p=0.2L	-14,0%	-1,56%	6,62%	-5,64%	-13,6%	-6,53%
p=0.35L	-11,5%	7,23%	-9,90%	-6,62%	0,19%	-1,77%
p=0.5L	-6,65%	-4,18%	-1,52%	-4,36%	-7,19%	-11,7%
p=0.65L	-0,20%	-13,9%	1,58%	-0,91%	-7,11%	-15,3%
p=0.8L	6,14%	-8,45%	-10,9%	1,69%	1,86%	-7,75%
<b>Cavity</b>						
p=0.2L	4,22%	-4,09%	-6,56%	2,60%	6,35%	-2,44%
p=0.35L	2,11%	-8,33%	3,28%	2,57%	-0,61%	-4,76%
p=0.5L	-2,11%	-1,96%	-1,82%	0,91%	-1,18%	-2,87%
p=0.65L	-5,74%	5,47%	-4,68%	-0,51%	2,05%	2,48%
p=0.8L	-8,11%	-0,59%	4,25%	-1,05%	-6,92%	-4,52%



The calculation of the relative deviations found in Table 1 uses the results reported for healthy auditory canal cases as a reference. Results presented in this table show that for the first resonance frequency (RF-1) and for the swelling case, there is a maximum deviation of -15.6% for “Canal 4” with an anomaly at position 0.35L and a maximum deviation of -14% for “Canal 6” with an anomaly at position 0.2L. In the case of a cavity, there is a maximum deviation of -7.54% for “Canal 4” with an anomaly located at 0.8L and a maximum deviation of -8.11% for “Canal 6” with an anomaly located at 0.8L.

In the case of the second resonance frequency (RF-2) and in the case of swelling, there is a maximum deviation of -13.4% for “Canal 4” with an anomaly at 0.8L and a maximum deviation of -13.9% for “Canal 6” with an anomaly at 0.65L. In the case of a cavity, there is a maximum deviation of -7.34% for “Canal 4” with an anomaly at position 0.35L and a maximum deviation of -8.33% for “Canal 6” with an anomaly at position 0.35L.

There is a maximum deviation of -9.84% in the third resonance frequency (RF-3) and for the swelling anomaly in “Canal 4” with an anomaly at 0.35L and 0.8L. For “Canal 6”, one can observe a deviation of -9.9% for an anomaly located at 0.35L and a maximum deviation of -10.9% for an anomaly at 0.8L. In the case of cavity, there is a maximum deviation of -6.96% for “Canal 4” with an anomaly located at 0.2L and a maximum deviation of -6.56% for “Canal 6” with an anomaly at 0.2L.

It can be deduced that a decrease of the first resonance frequency may be caused by a swelling located at the entry of the auditory canal or a cavity situated towards the tympanic membrane. The result is completely reversed at the second resonance frequency. A swelling effect causes a drop of this frequency at the input and output of the auditory canal for the third resonance frequency. A cavity acts in the same way at the end of the auditory canal.

The effect of anomalies on the SPL at resonance frequencies can also be observed from Table 1. For auditory canals with swelling, there is a maximum deviation of -13.4% corresponding to the SPL at RF-2 for “Canal 4” with an anomaly at 0.2L and a maximum deviation of -15.3% corresponding to the SPL at RF-3 for “Canal 6” with an anomaly at 0.65L. For the auditory canals with a cavity, a maximum deviation of 6.25% corresponding to the SPL at RF-2 for “Canal 4” with an anomaly at position 0.2L and a maximum deviation of -6.92% corresponding to the SPL at RF-2 for “Canal 6” with an anomaly at position 0.8L are observed.

These deviation values remain low, which means that, unlike the swelling anomaly, this anomaly has little effect on the amplitude of the resonance peaks of the ear canal.

#### 4.3. Effect of anomaly width

The parameter of the anomaly location “p” is set to 0.5L and the depth or height “d” is set to 1w to evaluate the influence of the width “w” of a swelling or cavity on the frequency response of the human auditory canal. Figure 13 shows the results obtained for canals 4 and 6 consisting of a swelling with values of different widths. Figure 14 displays the findings for the case of a cavity. Table 2 describes the relative deviations of the resonant frequency and the SPLs at these frequencies, taking as a reference the healthy auditory canal.

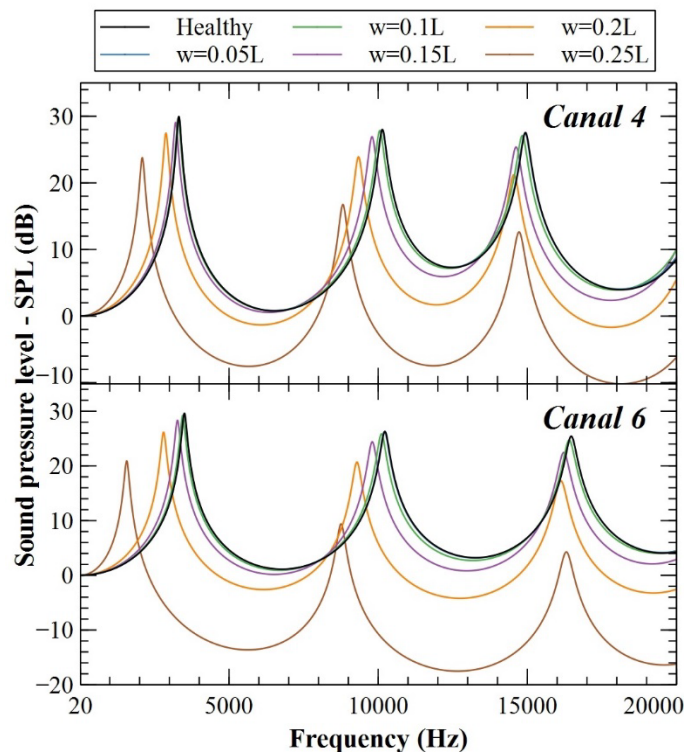


Figure 13 : Frequency response of the auditory canals with a swelling anomaly of different widths  $w$  ( $p=0.5L$  and  $d=1w$ )

Figure 13 indicates negative magnitudes relating to the reduction of SPL through the ear canal with swelling. These values are -7.55, -7.46 and -10.15 dB, observed in the case of an anomaly of width of 0.25L in “Canal 4” in the frequency bands [3197-7972] Hz, [9666-13960] Hz and [15438-20000] Hz respectively. The anomaly of 0.2L width is also affected by this effect, but at a smaller scale. Therefore, one can assume that the greater the width of the swelling, the greater the risk of hearing loss.

In the case of “Canal 6”, negative gains are noted for anomalies of width 0.2L and width 0.25L for frequency bands [4589-7607] Hz, [10776-14626] Hz and [17637-20000] Hz, and [2288-8398] Hz, [9102-16050] Hz and [16564-20000] Hz respectively. The maximum values observed for the 0.25L width anomaly in the bands mentioned above are -13.62, -17.51 and -16.38 dB. By comparing, on the one hand, the width of these frequency bands and, on the other hand, the minimum SPL values for canals 4 and 6, it can be concluded that “Canal 6” is more affected by a swelling anomaly than “Canal 4”. This is because of their different anatomies.

Maximum absolute deviations between the healthy “Canal 6” SPL curve and those of the swollen canals with anomaly widths 0.05L, 0.1L, 0.15L, 0.2L and 0.25L are -0.52 dB at 10347 Hz, -4.31 dB at 10309 Hz, -12.7 dB at 10262 Hz, -22.76 dB at 10246 Hz and -38.37 dB at 10238 Hz, respectively. Positive deviation values are also observed and are 0.42 dB at 10070 Hz, 3.35 dB at 10025 Hz, 9.21 dB at 9764 Hz, 15.88 dB at 2797 Hz and 18.52 dB at 1548 Hz for swollen “Canal 6” with anomaly widths of 0.05L, 0.1L, 0.15L, 0.2L and 0.25L, respectively.

It can be deduced from these values that the SPL reduces exponentially with an increase in the width of the swelling.



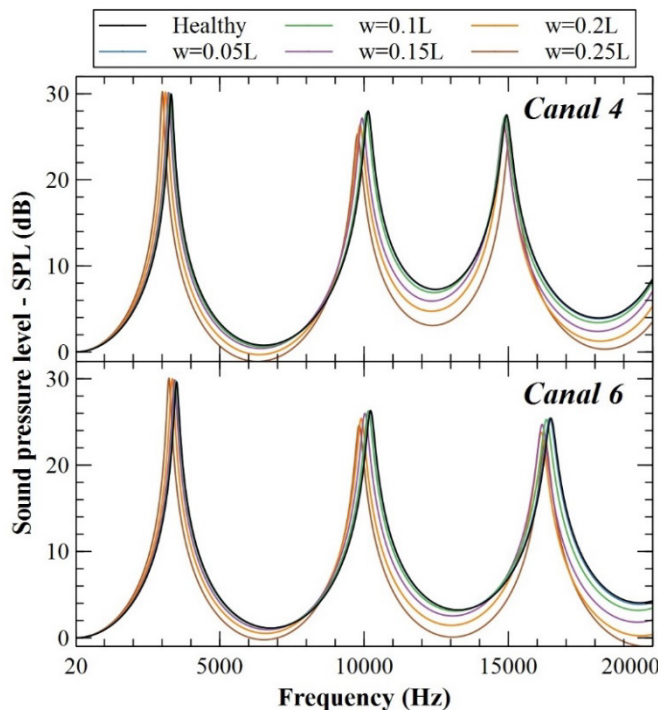


Figure 14 : Frequency response of the auditory canals with a cavity anomaly of different widths  $w$  ( $p=0.5L$  and  $d=1w$ )

Figure 14 shows that auditory canals 4 and 6 with an anomaly of the cavity type have no negative SPL values. This is in agreement with the observation made in the study of the effect of the anomaly position.

From this figure and taking into consideration the obtained results for “Canal 6”, the maximum absolute deviations are -0.98, -4.34, -8.42, -10.06 and -12.61 dB measured between the healthy canal SPL and the auditory canals with 0.05L, 0.1L, 0.15L, 0.2L and 0.25L width anomalies, respectively. Those variations occurs at 16629, 16580, 16545, 16544 and 10265 Hz frequencies for 0.05L, 0.1L, 0.15L, 0.2L and 0.25L widths respectively. As in the majority of the cases studied, it can be noted that the maximum absolute deviations occur at high frequencies.

For similar anomaly widths, the cavity effect is much less pronounced than swelling effect. This is because the swelling presents an obstacle in the ear canal, while the cavity presents a wider opening of the latter. This observation can easily be made by comparing the shape of the curves in Figures 13 and 14.

Figure 14 also shows that the difference between the SPL curves on the healthy canal and the anomaly canals increases exponentially with the “w” width. It can be deduced that this effect is much more present in the case of swelling anomalies by comparing the differences obtained.

Results presented in Table 2 show that the maximum relative deviations for resonance frequencies RF-1 and RF-2 are observed for an anomaly of a width of 0.25L. That applies to the two types of abnormalities investigated (swelling and cavity). The first resonance frequency (RF-1) for the swelling case presents a maximum deviation of -36.9% and -55.2% for “Canal 4” and “Canal 6”, respectively. In the case of a cavity, there is a maximum deviation of -9.29%, -7.56% for “Canal 4” and “Canal 6”, respectively.

Table 2: Effect of the width of swelling and cavity anomalies on frequency response for canal 4 and 6

	Resonant frequencies			SPL at resonant frequencies		
	RF-1	RF-2	RF-3	SPL (RF-1)	SPL (RF-2)	SPL (RF-3)
<b>Canal 4</b>						
<b>Healthy</b>	3317 Hz	10140 Hz	14940 Hz	29,96 dB	28,00 dB	27,54 dB
<b>Swelling</b>	w=0.05L	0,03%	0,00%	-0,07%	-0,20%	0,04%
	w=0.1L	-0,63%	-0,79%	-0,74%	-0,93%	-0,61%
	w=0.15L	-3,26%	-3,41%	-2,14%	-3,00%	-3,86%
	w=0.2L	-13,0%	-7,88%	-2,61%	-8,38%	-14,6%
	w=0.25L	-36,9%	-13,0%	-1,41%	-20,63%	-40,2%
<b>Cavity</b>	w=0.05L	-0,18%	-0,10%	-0,07%	0,07%	-0,14%
	w=0.1L	-1,15%	-0,59%	-0,40%	0,30%	-0,89%
	w=0.15L	-2,68%	-2,05%	-0,67%	0,60%	-2,89%
	w=0.2L	-5,85%	-2,74%	-0,40%	0,73%	-5,79%
	w=0.25L	-9,26%	-3,59%	0,66%	0,97%	-9,57%
<b>Canal 6</b>						
<b>Healthy</b>	3504 Hz	10230 Hz	16470 Hz	29,61 dB	26,30 dB	25,43 dB
<b>Swelling</b>	w=0.05L	-0,29%	-0,20%	0,00%	-0,24%	-0,23%
	w=0.1L	-1,66%	-1,17%	-0,36%	-1,32%	-1,71%
	w=0.15L	-6,65%	-4,18%	-1,52%	-4,36%	-7,19%
	w=0.2L	-19,9%	-9,19%	-2,06%	-11,65%	-21,4%
	w=0.25L	-55,2%	-14,4%	-0,97%	-29,35%	-64,3%
<b>Cavity</b>	w=0.05L	-0,11%	-0,20%	-0,18%	0,17%	0,08%
	w=0.1L	-0,66%	-0,78%	-0,91%	0,54%	-0,11%
	w=0.15L	-2,11%	-1,96%	-1,82%	0,91%	-1,18%
	w=0.2L	-4,42%	-3,18%	-2,00%	1,28%	-3,38%
	w=0.25L	-7,56%	-3,89%	-1,21%	1,52%	-6,46%

In the case of the second resonance frequency (RF-2) and for swelling anomaly, there is a maximum deviation of -13% and -14.4% for “Canal 4” and “Canal 6”, respectively. In the case of a cavity, there is a maximum deviation of -3.59% and -3.89% for “Canal 4” and “Canal 6”, respectively.

There is a maximum deviation of -2.61% in the third resonance frequency (RF-3) for the swelling anomaly of width 0.2L in “Canal 4”. For “Canal 6”, one can observe a deviation of -2.06% for an anomaly of the same width. In the case of cavity, there is a maximum deviation of -6.67% for “Canal 4” with an anomaly of width 0.15L and a maximum deviation of -2% for “Canal 6” with an anomaly of width 0.2L.

From these deviations, it can be deduced that the effect of the increase in the anomaly width is the maximum for RF-1. This effect is less pronounced for RF-2 and less so for RF-3. It should be noted that the effect on RF-3 is insignificant, regardless of the type of anomaly and its width. However, the influence of this anomaly has a major effect on the SPL at this frequency (RF-3).

The maximum relative deviations of the SPL at resonance frequencies is observed at the third resonance frequency (RF-3). Its value is -9.77% and -9.99% for “Canal 4” and “Canal 6”, respectively in the case of a cavity. In the case of swelling

anomaly, there is a maximum deviation of -54.8% and -83.20% for “Canal 4” and “Canal 6”, respectively. From these values, it can be noted that the drop in SPL caused by an increase in anomaly width is very significant, especially at the third resonance frequency.

4.4. Effect of anomaly height/depth

The anomaly position parameter “p” is set to 0.5L and the width “w” is set to 0.15L in order to evaluate the influence of the height/depth “d” of the swelling or cavity on the frequency response of the human auditory canal. In the case of a swelling anomaly, the parameter “d” being studied is called height, and in the case of a cavity anomaly, one can talk about the depth.

Figure 15 shows the results obtained for canals 4 and 6 consisting of a swelling with values of varying heights. Figure 16 shows the data for the cavity case. Table 3 illustrates the relative deviations of the resonant frequency and the SPLs at these frequencies, taking as a reference the healthy auditory canal.

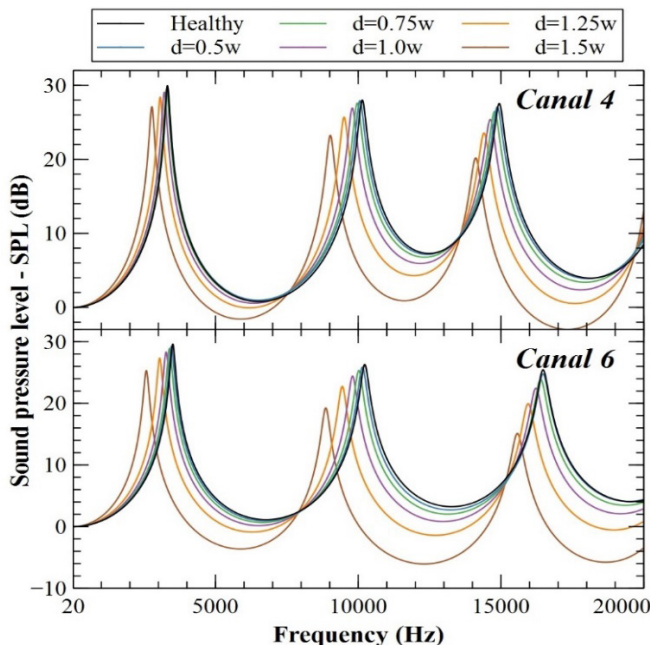


Figure 15 : Frequency response of the auditory canals with a swelling anomaly of different heights d (p = 0.5L and w = 0.15L)

Figure 15 indicates negative magnitudes relating to the reduction of SPL through the ear canal with swelling. The maximum of these values, in “Canal 4”, are -1.58 and -2.96 dB, observed in the case of an anomaly of 1.50w height in the frequency bands [4737-7026] Hz and [15776-18849] Hz respectively. Thus, the greater the swelling height, the more risk of hearing loss can be expected.

In the case of “Canal 6”, at frequency bands [5381-7195] Hz, [11601-13834] Hz and [18210-20000] Hz, negative SPL values are noted for anomalies of height 1.25w. Same finding for the anomaly of the 1.50w height in the [4127-7510] Hz, [10037-14501] Hz and [16630-20000] Hz frequency bands. The maximum values observed for the 1.25w height anomaly in the above mentioned bands are -0.865, -1.4 and -0.53 dB, respectively. The maximum values measured for the anomaly of 1.50w height are -3.62, -6.05 and -5.77 dB, respectively.

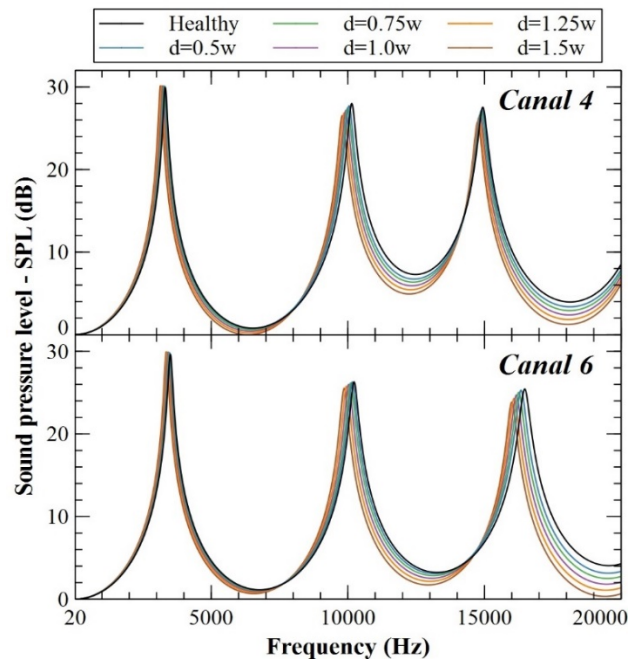


Figure 16 : Frequency response of the auditory canals with a cavity anomaly of different depths d (p = 0.5L and w = 0.15L)

Table 3: Effect of the height/depth of swelling and cavity anomalies on frequency response for canal 4 and 6

	Resonant frequencies			SPL at resonant frequencies			
	RF-1	RF-2	RF-3	SPL (RF-1)	SPL (RF-2)	SPL (RF-3)	
<b>Canal 4</b>							
Healthy	3317 Hz	10140 Hz	14940 Hz	29,96 dB	28,00 dB	27,54 dB	
Swelling	d=0.5w	0,24%	-0,79%	-0,47%	-0,87%	-0,36%	-1,71%
	d=0.75w	-0,9%	-1,71%	-1,07%	-1,67%	-1,46%	-3,92%
	d=1.00w	-3,26%	-3,41%	-2,14%	-3,0%	-3,86%	-7,88%
	d=1.25w	-7,60%	-6,26%	-3,55%	-5,24%	-8,25%	-14,49%
	d=1.5w	-16,5%	-11,0%	-5,49%	-9,55%	-17,0%	-26,7%
Cavity	d=0.5w	-0,63%	-0,89%	-0,20%	0,43%	-1,04%	-0,98%
	d=0.75w	-1,60%	-1,46%	-0,47%	0,53%	-1,89%	-1,96%
	d=1.00w	-2,68%	-2,05%	-0,67%	0,60%	-2,89%	-3,20%
	d=1.25w	-3,92%	-2,71%	-0,94%	0,63%	-4,04%	-4,65%
	d=1.5w	-5,28%	-3,40%	-1,20%	0,63%	-5,25%	-6,28%
<b>Canal 6</b>							
Healthy	3504 Hz	10230 Hz	16470 Hz	29,61 dB	26,30 dB	25,43 dB	
Swelling	d=0.5w	-1,37%	-0,78%	0,06%	-1,28%	-1,63%	-2,63%
	d=0.75w	-3,22%	-2,05%	-0,49%	-2,47%	-3,69%	-5,98%
	d=1.00w	-6,65%	-4,18%	-1,52%	-4,36%	-7,19%	-11,72%
	d=1.25w	-13,0%	-7,7%	-3,16%	-7,7%	-13,54%	-21,55%
	d=1.5w	-26,4%	-13,3%	-5,40%	-14,52%	-26,77%	-40,54%
Cavity	d=0.5w	-0,57%	-0,68%	-0,79%	0,61%	-0,11%	-0,55%
	d=0.75w	-1,26%	-1,27%	-1,28%	0,81%	-0,53%	-1,53%
	d=1.00w	-2,11%	-1,96%	-1,82%	0,91%	-1,18%	-2,87%
	d=1.25w	-3,11%	-2,74%	-2,31%	1,01%	-2,02%	-4,44%
	d=1.5w	-4,25%	-3,57%	-2,85%	1,05%	-2,97%	-6,21%

It should be mentioned that the frequency bands mentioned above in the case of anomaly of 1.25w height are far more restricted than that of anomaly of 1.50w height. It can be seen from these values that the influence of the height of the anomaly is less significant than the other effects studied, but remains quite important especially for the swelling case.

As in the study of the effect of the position and the width of the anomaly, the SPL curve presents no negative values in the case of the auditory canals in the presence of a cavity (Figure 16). Curves obtained for auditory canals with anomalies show greater deviations for "Canal 6" than for "Canal 4". This can be easily confirmed by reading Figure 16 and Table 3.

Considering the results shown in Figure 16 for "Canal 6", the maximum absolute deviations are -3.82, -6.15, -8.42, -10.51 and -12.38 dB measured between the healthy canal SPL and the auditory canals with 0.5w, 0.75w, 1w, 1.25w and 1.50w depth anomalies, respectively. Those variations occur at 16589, 16567, 16545, 16535 and 16530 Hz frequencies for 0.5w, 0.75w, 1w, 1.25w and 1.50w depths respectively. It can be noted that the maximum absolute deviations occur at high frequencies for all considered cases. These deviations are observed at an almost identical frequency (around 16550 Hz) for all studied cases. This can also be shown that these differences correlate linearly with increasing depth d parameter.

By taking into account, for example, the results of the swelling case in "Canal4" available in Table 3, it can be observed that the deviations in resonance frequency varies from -0.24% for the height of 0.1w to -16.5% for the height of 1.5w. This shows that the resonance frequency varies linearly with the height / depth of the anomaly. It can also be found that the SPL deviation at the resonance frequency varies from -0.87% at a height of 0.1w to -9.55% at a height of 1.5w. This shows that the SPL at resonance frequency varies in an exponential manner with the height/depth of the anomaly.

Reading from Table 3, maximum relative deviations in terms of frequency resonance and SPL at those frequencies are all observed for 1.5w, which is the highest depth/height value considered in the analysis. In the case of "Canal4" with a 1.5w depth swelling, the deviation is -16.5%, -11%, -5.49% at the first, the second and the third resonant frequencies respectively. The deviation in SPL is -9.55%, -17%, -26.7% at the first, the second and the third resonant frequencies respectively.

The effect of increasing the anomaly depth is most pronounced for RF-1. This effect is less significant for RF-2 and even less for RF-3. The opposite effect is observed with respect to the SPL at these frequencies. The same observation can be made from both canals (4 and 6) as well as the two anomalies (cavity and swelling).

## 5. Conclusion

An investigation is conducted in this paper with the purpose of better understanding the influence of some abnormalities which may cause tinnitus. A previously developed and validated acoustic model associates an acoustic resonator with each part of the outer ear. Such resonators can have several forms which may be used to model a human auditory canal. The model is used for adding various anomalies into it, either by incorporating acoustic

resonators or by operating on the parameters of it. The abnormalities examined are those that may induce a change in the external auditory canal section, or an alternation in the external ear's sound transmission medium.

The acoustic model is applied to two auditory canals that are found in the literature where anomalies are inserted. A parametric study is conducted taking into account the influence of an anomaly's location, width and depth simulating an obstacle on the outer ear's frequency response. This research also discusses the influence of the propagation speed of sound in the ear canal to model a form of ear condition like swimmer's ear. This is done by contrasting the frequency response of a healthy external ear to an external ear with an anomaly.

It can be concluded in the light of the results that the presence of fluid in the ear canal tends to decrease its number of resonant frequencies in the audible band. The more fluid the ear canal has, the higher the resonant frequency and the higher the sound pressure level at it.

A swelling at the entrance of the auditory canal or a cavity at the end of it may result in a significant drop in the first resonance frequency value. The finding at the second frequency of resonance is entirely inverted. A swelling effect causes this frequency to drop for the third resonance frequency at the input and output of the auditory canal. A cavity at the end of the auditory canal functions in the same way. The auditory canal SPL's response decreases exponentially with an increase in the anomaly's width. The effect of the anomaly width increase is the maximum for first resonant frequency. The drop in SPL at third resonance frequency is the most significant.

The effect of the anomaly's height is less significant than the other parameters studied, but remains very important especially for the case of swelling. For this case, the more the height of the anomaly, the more probability of hearing loss can be expected. The deviation between the resonant frequency for the healthy canal and the one with anomaly correlate linearly with increasing height/depth parameter. However, the deviation of the SPL at resonance frequency varies exponentially with the height/depth.

Across all the cases analyzed, it is observed that, when contrasting it with the swelling anomaly, the cavity has less effect on the frequency response of the auditory canal. This is due to the fact that the swelling in the ear canal presents an obstacle while the cavity presents a wider opening. The frequency response of auditory canals with the same anomaly is not identical. This amounts to the disparity of their anatomies.

All anomalies cause changes in SPL, resulting in increases in some frequency ranges and, at the same time, attenuation in other ranges. This can lead to hearing loss or over-sensitivity to sound having these frequencies. This could lead to tinnitus with severities that rely on the physiology and psychology of the individual.

## References

- [1] C. Coelho, R. Santos, Roberto, K. Campara, R. Tyler, "Classification of Tinnitus" *Otolaryngologic Clinics of North America*, **S0030-6665(20)**, 30052-9, 2020. <https://doi.org/10.1016/j.otc.2020.03.015>
- [2] J. M. Bhatt, H. W. Lin, N. Bhattacharyya, "Prevalence, severity, exposures, and treatment patterns of tinnitus in the United States" *JAMA Otolaryngol Head Neck Surg.*, **142(10)**, 959-965, 2016. <https://doi.org/10.1001/jamaoto.2016.1700>



- [3] E. A. Masterson, C. L. Themann, S. E. Luchhaupt, J. Li, G. M. Calvert, "Hearing difficulty and tinnitus among US workers and non-workers in 2007" *American Journal of Industrial Medicine*, **59**(4), 290–300, 2016. <https://doi.org/10.1002/ajim.22565>
- [4] E. Josse, *Les acouphènes. Traitement par l'hypnose, Resilience-pys*, 2006.
- [5] K. J. Trevis, N. M. McLachlan, S. J. Wilson, "Cognitive Mechanisms in Chronic Tinnitus: Psychological Markers of a Failure to Switch Attention" *Frontiers in psychology*, **7**, 2016. <https://doi.org/10.3389/fpsyg.2016.01262>
- [6] D. Ivansic, O. Guntinas-Lichius, B. Müller, G. F. Volk, G. Schneider, C. Dobel, "Impairments of Speech Comprehension in Patients with Tinnitus-A Review" *Frontiers in aging neuroscience*, **9**, 2017. <https://doi.org/10.3389/fnagi.2017.00224>
- [7] P. Hyvärinen, "Neurophysiologically-based approaches to tinnitus diagnostics and treatment," Ph.D Thesis, Aalto University, 2016.
- [8] L. Jagoda, N. Giroud, P. Neff, A. Kegel, T. Kleinjung, M. Meyer, "Speech perception in tinnitus is related to individual distress level - A neurophysiological study" *Hearing Research*, **367**, 48–58, 2018. <https://doi.org/10.1016/j.heares.2018.07.001>
- [9] B. Langguth, A. B. Elgoyhen, "Current pharmacological treatments for tinnitus" *Expert Opinion on Pharmacotherapy*, **13**(17), 2495–2509, 2012. <https://doi.org/10.1517/14656566.2012.739608>
- [10] D. A. Colucci, "Insomnia and Tinnitus: A Worrisome Cycle" *The Hearing Journal*, **70**(6), 43–45, 2017. <https://doi.org/10.1097/01.HJ.0000520670.39014.4c>
- [11] S. Weidt, A. Delsignore, M. Meyer, M. Rufer, N. Peter, N. Drabe, T. Kleinjung, "Which tinnitus-related characteristics affect current health-related quality of life and depression? A cross-sectional cohort study" *Psychiatry Research*, **237**, 114–121, 2016. <https://doi.org/10.1016/j.psychres.2016.01.065>
- [12] T. Rodebaugh, E. Lenze, "Tinnitus" *JAMA*, **323**, 2020. <https://doi.org/10.1001/jama.2020.0697>
- [13] E. Coulon, "Les acouphènes ou l'impossible silence: étiologie, physiopathologie et tentatives de traitement," Ph.D Thesis, Mixed Faculty Of Medicine And Pharmacy Of Rouen, 2002.
- [14] I. Adamchic, T. Toth, C. Hauptmann, M. Walger, B. Langguth, I. Klingmann, P. A. Tass, "Acute effects and after-effects of acoustic coordinated reset neuromodulation in patients with chronic subjective tinnitus" *NeuroImage: Clinical*, **15**, 541–558, 2017. <https://doi.org/10.1016/j.nicl.2017.05.017>
- [15] B. Mazurek, H. Olze, H. Haupt, A. J. Szczepek, "The more the worse: the grade of noise-induced hearing loss associates with the severity of tinnitus" *International Journal of Environmental Research and Public Health*, **7**(8), 3071–3079, 2010. <https://doi.org/10.3390/ijerph7083071>
- [16] O. Zagólski, "Management of Tinnitus in Patients with Presbycusis" *International Tinnitus Journal*, **12**(2), 175–178, 2006.
- [17] N. Julien, O. Sterkers, "Les acouphènes" *La vie médicale*, **16**, 661–667, 1990.
- [18] P. S. Guth, J. Risey, W. Briner, P. Blair, H. T. Reed, G. Bryant, C. Norris, G. Housley, R. Miller, "Evaluation of amino-oxyacetic acid as a palliative in tinnitus" *Annals of Otolaryngology, Rhinology, and Laryngology*, **99**(1), 74–79, 1990. <https://doi.org/10.1177/000348949009900113>
- [19] R. Salvi, E. Lobarinas, W. Sun, "Pharmacological Treatments For Tinnitus: New And Old" *Drugs of the future*, **34**(5), 381–400, 2009. <https://doi.org/10.1358/dof.2009.034.05.1362442>
- [20] E. C. Nam, R. Lewis, H. H. Nakajima, S. N. Merchant, R. A. Levine, "Head rotation evoked tinnitus due to superior semicircular canal dehiscence" *The Journal of Laryngology & Otology*, **124**(3), 333–335, 2010. <https://doi.org/10.1017/S0022215109991241>
- [21] D. M. Baguley, D. McFerran, D. Hall "Tinnitus" *The Lancet*, **382**(9904), 1600–1607, 2013. [http://doi.org/10.1016/S0140-6736\(13\)60142-7](http://doi.org/10.1016/S0140-6736(13)60142-7)
- [22] R. F. Cima, D. Kikidis, B. Mazurek, H. Haider, C. Cederroth, A. Noreña, A. Lapira, A. Bibas, D. Hoare, "Tinnitus healthcare: a survey revealing extensive variation in opinion and practices across Europe" *BMJ Open*, **10**, 2020. <http://doi.org/10.1136/bmjopen-2019-029346>
- [23] F. Zamiri, M. Moosapour, R. Hoseinabadi, "Tinnitus Treatments and Managements" *Global Journal of Otolaryngology*, **13**(5), 01–04, 2018. <https://doi.org/10.19080/GJO.2018.13.555874>
- [24] R.S Hallam, S.C Jakes, C. Chambers, R. Hinchcliffe, "A comparison of different methods for assessing the intensity of tinnitus" *Acta Otolaryngol*, **99**, 501–508, 1985. <https://doi.org/10.3109/00016488509182253>
- [25] H. Selye, *Du rêve à la découverte*, Éditions de La Presse, 66–68, 1973.
- [26] T. Fuller, R. F. Cima, B. Langguth, B. Mazurek, J. Vlaeyen, D. Hoare, "Cognitive behavioural therapy for tinnitus, Cochrane Database of Systematic Reviews, 2020. <http://doi.org/10.1002/14651858.cd012614.pub2>
- [27] P. J. Jastreboff, "25 Years of tinnitus retraining therapy" *HNO*, **63**, 307–311, 2015. <https://doi.org/10.1007/s00106-014-2979-1>
- [28] C. Formby, R. Scherer, "Rationale for the Tinnitus Retraining Therapy Trial" *Noise & health*, **15**, 134–142, 2013. <https://doi.org/10.4103/1463-1741.110299>
- [29] K. M. Reavis, V. S. Rothholtz, Q. Tang, J. A. Carroll, H. Djalilian, F. G. Zeng, "Temporary suppression of tinnitus by modulated sounds" *Journal of the Association for Research in Otolaryngology : JARO*, **13**(4), 561–571, 2012. <https://doi.org/10.1007/s10162-012-0331-6>
- [30] R. Soleimani, M. M. Jalali, T. Hasandokht, "Therapeutic impact of repetitive transcranial magnetic stimulation (rTMS) on tinnitus: a systematic review and meta-analysis" *European archives of oto-rhino-laryngology*, **273**(7), 1663–1675, 2016. <https://doi.org/10.1007/s00405-015-3642-5>
- [31] M. Pedemonte, Chapter 7 – Tinnitus treatment during sleep, *The auditory system in sleep 2nd edition*, 2018. <https://doi.org/10.1016/B978-0-12-810476-7.00007-5>
- [32] ATA Newsletter, American Tinnitus Association, **10**(1), 1–4, 1985.
- [33] P. Neff, L. Zielonka, M. Meyer, B. Langguth, M. Schecklmann, W. Schlee, "Comparison of Amplitude Modulated Sounds and Pure Tones at the Tinnitus Frequency: Residual Tinnitus Suppression and Stimulus Evaluation" *Trends in hearing*, **23**, 1–16, 2019. <https://doi.org/10.1177/2331216519833841>
- [34] R. M. Johnson, J. Fenwick, "Masking Levels (Minimum Masking Levels) and Tinnitus International Frequency" *The Journal of Laryngology & Otology*, **98**(S9), 63–66, 1984. <https://doi.org/10.1017/S1755146300090132>
- [35] M. Ralli, A. Greco, R. Turchetta, G. Altissimi, M. de Vincentiis, G. Cianfrone, "Somatosensory tinnitus: Current evidence and future perspectives" *Journal of International Medical Research*, **45**(3), 933–947, 2017. <https://doi.org/10.1177/0300060517707673>
- [36] G. Shekhawat, K. Kobayashi, G. Searchfield, "Methodology for studying the transient effects of Transcranial Direct Current Stimulation combined with Auditory Residual Inhibition on Tinnitus" *Journal of Neuroscience Methods*, **239**, 28–33, 2014. <https://doi.org/10.1016/j.jneumeth.2014.09.025>
- [37] H. Teismann, A. Wollbrink, H. Okamoto, G. Schlaug, C. Rudack, C. Pantev, "Combining transcranial direct current stimulation and tailor-made notched music training to decrease tinnitus-related distress--a pilot study" *PloS one*, **9**(2), 1–11, 2014. <https://doi.org/10.1371/journal.pone.0089904>
- [38] J. Melloui, O. Bouattane, J. Bakkoury, "Study of the effect of a cause of tinnitus on the resonant frequency of the outer ear" in *1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, Meknes, Morocco, 2020. <https://doi.org/10.1109/IRASET48871.2020.9092217>
- [39] B. B. Ballachanda, "Theoretical and applied external ear acoustic" *Journal of the American Academy of Audiology*, **8**(6), 411–420, 1998.
- [40] E. A. G. Shaw, *Acoustical Characteristics of the Outer Ear*, Encyclopedia of Acoustics, M.J. Crocker (Ed.), 1325–1335, 2007. <https://doi.org/10.1002/9780470172537.ch105>
- [41] J. Melloui, O. Bouattane, J. Bakkoury, "Acoustic model of the human outer ear" *International Journal of Engineering & Technology*, **7**(4), 3286–3293, 2018. <https://doi.org/10.14419/ijet.v7i4.15842>
- [42] H. Hudde, A. Engel, "Measuring and modeling basic properties of the human middle ear and ear canal. Part II: Ear canal, middle ear cavities, eardrum, and ossicles" *Acta Acustica united with Acustica*, **84**(5), 894–913, 1998.
- [43] M.R. Stinson, B.W. Lawton, "Specification of the geometry of the human ear canal for the prediction of sound-pressure level distribution" *The Journal of the Acoustical Society of America*, **85**(6), 2492–2503, 1989. <https://doi.org/10.1121/1.397744>



## Neural Network-Based Fault Diagnosis of Joints in High Voltage Electrical Lines

Marco Bindi<sup>\*1</sup>, Igor Aizenberg<sup>2</sup>, Riccardo Belardi<sup>1</sup>, Francesco Grasso<sup>1</sup>, Antonio Luchetta<sup>1</sup>, Stefano Manetti<sup>1</sup>, Maria Cristina Piccirilli<sup>1</sup>

<sup>1</sup>University of Florence, Department of Information Engineering, Firenze 50139, Italy

<sup>2</sup>Manhattan College, Department of Computer Science, NY, a 10471, USA

### ARTICLE INFO

Article history:

Received: 10 June, 2020

Accepted: 20 July, 2020

Online: 09 August, 2020

Keywords:

High-Voltage transmission line

Complex neural network

Smart monitoring

Fault diagnosis

Electrical joint modeling

Testability analysis

### ABSTRACT

*In this paper a classification system based on a complex-valued neural network is used to evaluate the health state of joints in high voltage overhead transmission lines. The aim of this method is to prevent breakages on the joints through the frequency response measurements obtained at the initial point of the network. The specific advantage of this kind of measure is to be non-intrusive and therefore safer than other approaches, also considering the high voltage nature of the lines. A feedforward multi-layer neural network with multi-valued neurons is used to achieve the goal. The results obtained for power lines characterized by three and four junction regions show that the system is able to identify the health state of each joint, with an accuracy level greater than 90%.*

## 1. Introduction

In the management of a long electrical transmission line, fault location represents an important aspect, because it allows to reduce recovery times and increase network availability [1]. Currently, high voltage lines are protected by devices which allow to identify the fault position and put out of service the non-functioning part of the grid. For example, the distance protection employed on overhead transmission lines uses Intelligent Electronic Devices (IEDs) to estimate the impedance between the protection point and the failure point by measuring the line voltage and current [2]. In this way it is possible to obtain the distance from the fault, exploiting the correspondence between conductor impedance and length of the line section. In the last years new protection devices have been introduced, based on traveling wave detection; they allow to measure the impulsive signals generated by failures with greater precision [3]. However, these methods allow to locate the failure only once it has occurred [4].

In this paper, an extension of a work originally presented at the International Conference on Soft Computing & Machine Intelligence (ISCMi 2019) [1], a system for monitoring the health status of the joints [1], [5] is presented. The system allows a preventive maintenance, so increasing the availability of the line. The joints (or junction regions) are the connection points between

two different parts of the same phase conductor, and they represent one of the most stressed parts of overhead electrical lines [6], [7]. The proposed method exploits the equivalent lumped circuit of conductors and junction regions [8], [9]. In particular, the model of the elementary section of the network consists in the connection of the equivalent circuits of joint and conductor. The model of the whole line is obtained by the cascade connection of several elementary sections. The health status of the joint is obtained by measuring the network frequency response. To this aim it is necessary to establish a nominal working range for each electrical parameter of the line model. Any departure from this interval can be interpreted as a symptom of a possible fault. The deviation from the nominal condition is obtained through a comparison between the theoretical frequency response and measurements taken at several frequencies. The results of this comparison are traced back to the corresponding variation of the joint parameters. This operation is executed by a neural network able to classify the severity of the joint degradation. The selected neural network is a feedforward Multi-Layer network with Multi-Valued Neurons (MLMVN) whose training phase is carried out using the theoretical frequency response simulated on *SapWin IV* and *MatLab*<sup>®</sup>. In this way it is possible to obtain a smart monitoring approach capable of classifying the measurements obtained through the PLC (Power Line Communication) systems [10]. This type of equipment is usually integrated on high and medium voltage networks to use the phase conductor as a channel for the

\* Corresponding Author: Marco Bindi, 50139 Via S. Marta 3, Firenze, Italy, [m.bindi@unifi.it](mailto:m.bindi@unifi.it)

transmission of information. Since the communication method is based on the modulation of the carrier wave, PLC systems can be used to generate signals with different frequencies obtaining measurements of the line transfer function [11].

The paper is organized as follows. Section 2 describes the model of conductors and joints. Section 3 is focused on the preliminary steps of the procedure, while section 4 is dedicated to the simulation procedure. Finally, Sections 5 and 6 report some specific application cases and conclusions, respectively.

## 2. Line Modelling

The main physical components of a line are junction regions and conductor stretches. The derivation of the equivalent circuits of these components represents the starting point for the line modelling. It is then necessary to establish the relationships between fault mechanisms and variation of electrical parameters. Only the joint degradation is considered in this paper and two different failure mechanisms are introduced: oxidation process and partial breakage of the joint structure. In this way it is possible to obtain three different intervals for the values of the electrical parameters, corresponding to the nominal situation, to the oxidation condition and to the presence of partial breaking [1].

### 2.1. Conductor modelling

In this paper an ACSR (Aluminum Conductor Steel Reinforced) with a diameter of 22.8 mm is used (Figure 1a). The equivalent model of the conductor (Figure 1b) contains four electrical elements and is called “canonical  $\pi$ -model”.

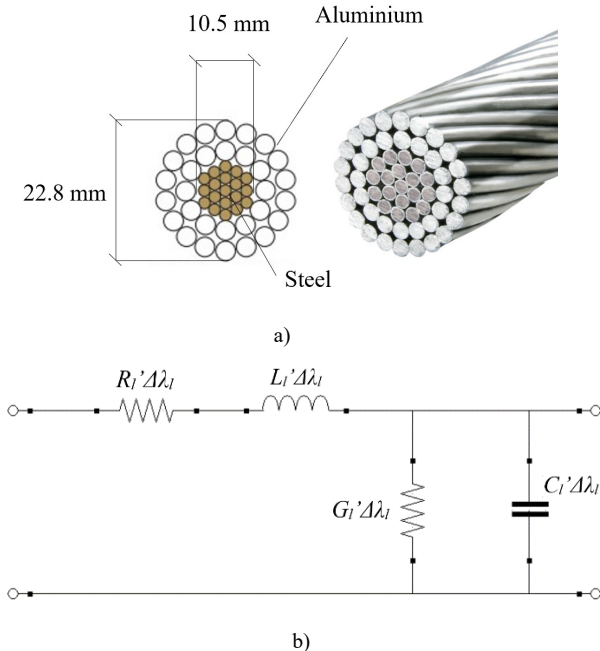


Figure 1 – a) Physical structure of the conductor; b) Lumped circuit of the conductor.

It is commonly used to study wave propagation and power flow analysis [12], [13]. The term  $\Delta\lambda_l$  represents the conductor length, while the values of the parameters  $R_l'$ ,  $L_l'$ ,  $G_l'$ ,  $C_l'$  depend on the mechanical characteristics of the phase conductor; all these parameters are quantities per unit of length [14].

### 2.2 Joint modelling

Junction regions represent the connection points between two different parts of the same phase conductor, and they must guarantee the electrical continuity along the network [15], [16]. Bolted joints are used close to the pylons due to their high mechanical strength [17]. However, no available literature exists that describes their electrical behaviour; an acceptable alternative consists in modelling them with the better-known model of the solder joint, for which some information is available about its degradation mechanisms [18], [19]. The physical structure of the joint is shown in Figure 2a.

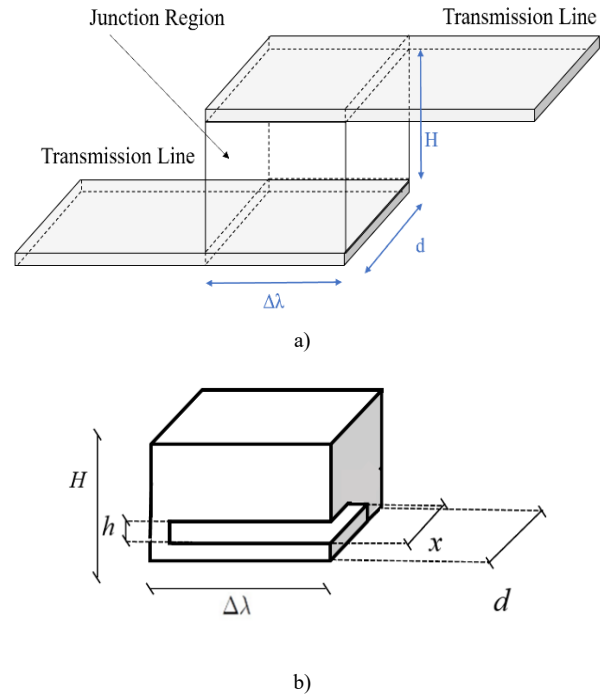


Figure 2: a) Physical structure of the junction region; b) Braking parameters.

$\Delta\lambda$  is the length of the junction region,  $d$  is its width and  $H$  its height. The characteristic parameters of the braking mechanism are shown in Figure 2b where  $x$  represents the crack width and  $h$  is its height. The physical model and the equivalent lumped circuit of the joint are shown in Figure 3a and Figure 3b, respectively. In the equivalent circuit, the value of  $R_{sj}$  and  $L_{sj}$  in nominal conditions can be calculated by  $R_{sj} = \frac{\rho H}{\Delta\lambda d}$  and  $L_{sj} = \frac{\mu_0 \mu_r H}{2\pi} \left[ \ln \frac{2H}{\Delta\lambda + d} + 0,5 \right]$  respectively, while the capacitance value is nominally zero. If the frequency value used for the measurements exceeds the network frequency, it is necessary to consider the skin effect. For this reason, the penetration depth of the current (a frequency dependent parameter, calculated by  $\delta = \frac{1}{\sqrt{\pi \sigma \mu_0 \mu_r f}}$ ) is introduced and (1) is used to evaluate the resistance  $R_{sj}$ .

$$R_{sj} = \frac{\rho H}{2\delta (\Delta\lambda + d - 2\delta)} \quad (1)$$

In oxidation conditions the resistance changes its value due to the increase of resistivity, while the partial breakage mechanism induces variations on the electrical parameters of the joint. In this work the consequentiality between oxidation mechanism and partial breakage process is assumed.

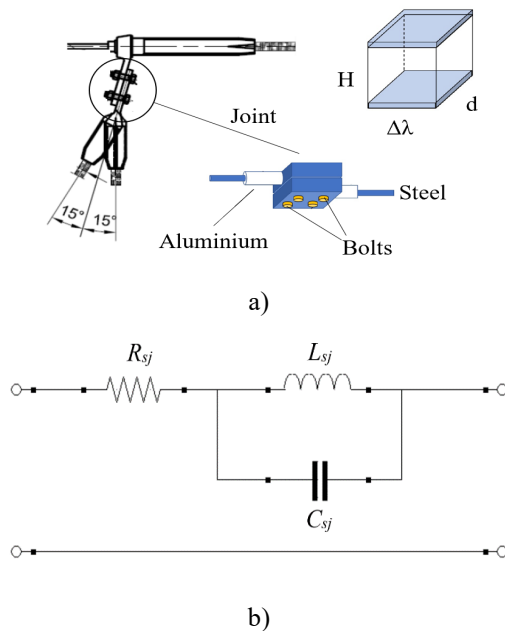


Figure 3 – a) Physical model of the joint; b) Lumped circuit of the joint.

The formulas (2), (3) and (4) are obtained by introducing the typical breaking mechanism of solder joints and describing the variations of  $R_{sj}$ ,  $L_{sj}$  and  $C_{sj}$  as functions of the crack width  $x$  and the crack height  $h$  [18]. Actually, only the  $R_{sj}$  and  $L_{sj}$  variations are considered, because the sensitivity of the  $C_{sj}$  variation is low.

$$R_{sj} = \begin{cases} \rho \left[ \frac{H-h}{2\delta(\Delta\lambda+d-2\delta)} + \frac{h}{2\delta(\Delta\lambda+d-x-2\delta)} \right] & (d-x) \geq 2\delta \\ \rho \left[ \frac{H-h}{2\delta(\Delta\lambda+d-2\delta)} + \frac{h}{\Delta\lambda(d-x)} \right] & (d-x) < 2\delta \end{cases} \quad (2)$$

$$L_{sj} = \frac{\mu_0 \mu_r H}{2\pi} \left[ \ln \frac{2(H-h)}{(\Delta\lambda+d)} + 0,5 \right] + \frac{\mu_0 \mu_r H}{2\pi} \left[ \ln \frac{2H}{(\Delta\lambda+d-x)} + 0,5 \right] \quad (3)$$

$$C_{sj} = \frac{\epsilon_0 \epsilon_r x \Delta\lambda}{h} \quad (4)$$

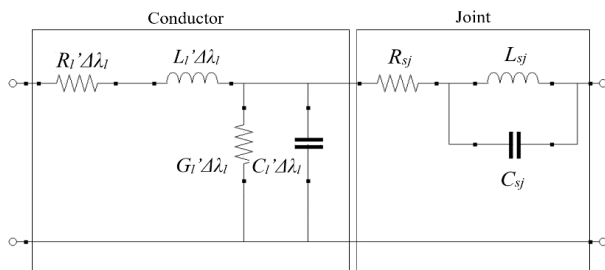


Figure 4: Elementary section of the line.

### 3. Problem outline

Once the equivalent lumped circuits of joint and conductor have been built, it is possible to describe the whole network by connecting in cascade the elementary line sections shown in Figure 4. Each line section between two pylons is represented by an elementary section and each joint model corresponds to the pylon position. A simplification is used in this case because there should be two joints and a short section of the conductor at the pylon. This is not a problem for the theoretical basis of the procedure presented

in this paper, but it should be taken into account in practical applications [4, 1].

The next step of the procedure consists in the simulation of the model. This phase is needed for testability analysis, frequencies selection and neural network training, as shown in the following.

#### 3.1. Testability analysis

It goes without saying that an a-priori knowledge of how many and which parameters are actually identifiable is quite useful in any fault diagnosis procedure, since it allows to correctly select the measurement test points. Testability analysis is exactly aimed at providing this kind of information [20, 21, 22]. It consists of two steps, testability evaluation and ambiguity groups identification. The testability gives the solvability degree of the fault diagnosis equations [23], i.e. it allows establishing the number of parameters whose value can be determined, starting from measurements taken on a given number of test points. The maximum possible value of testability is the number  $N$  of parameters present within the circuit under test. This maximum value is often not reached; this means that, in case of fault, there are parameters that are indistinguishable one to each other. In other words, it is possible to identify a set that contains the faulty component, but it is not possible to exactly identify the faulty one inside the group. The ambiguity groups are all those groups constituted by sets of indistinguishable parameters. The determination of the ambiguity groups gives information about which circuit parameters are identifiable.

In the procedure presented in this paper, testability analysis is a preliminary phase, aimed at discovering which and how many parameters of the line model are identifiable. This operation is executed starting from the network function, which acts as reference for the measurements, taken at selected frequencies. Consequently, it is necessary to simulate the line model in order to determine its network function.

#### 3.2. Optimal frequency selection

As previously stated, the fault identification is the result of a comparison between the measurements at the test point and the nominal amplitude response, evaluated at the same frequencies. As shown in [24], measurement errors and manufacturing tolerances, which affect the result in a similar manner, depend on the choice of the frequencies. A proper selection allows minimizing the effect of these errors, increasing the probability to detect the fault. A method for carrying out this selection is presented in [24] and, in this work, it is used to obtain one test frequency. However, three octaves of the selected frequency are also used in the simulation procedure in order to make the system more robust and better train the complex neural network. It is worth pointing out that in [24] the use of more than one frequency has the scope to increase the number of equations, so that their number is at least as high as that of the unknown parameters. In the procedure presented in this paper there is no parameter identification, so there are no fault diagnosis equations to solve. Nevertheless, it is necessary to determine the network function of the line model in order to apply the procedure in [24], based on the network function derivatives.

#### 3.3. Neural network training

The diagnosis method proposed in this paper uses neural networks, which do not require the derivation of fault equations.

The analysis of the circuit model is however necessary for the training of the neural network, as it will be shown in the next section. The used intelligent classifier is a complex neural network, that presents excellent performances compared to other machine-learning techniques. It is based on a feedforward multilayer neural network with multi-valued neurons (MLMVN), characterized by a derivative free learning algorithm [25], an alternative algorithm based on the linear least square (LLS) methods [26] to reduce the high computational cost of the original backpropagation procedure, a soft margin method [27, 28] in order to make the network a good classifier. Further details will be given later in the application description.

#### 4. Simulation Procedure

All the fundamental steps in the classification of the joint conditions are presented in this paragraph. First it is necessary to define the main characteristics of the network, the type of conductor and the physical size of the joints. Referring to the technical data sheets of the ACSR shown in Table 1, it is possible to obtain the value of the DC electrical resistance per unit of length  $R_l'$  at ambient temperature (293K) and, consequently, the value of the resistivity  $\rho_l = \frac{R_l' \pi (\frac{D}{2})^2}{\Delta \lambda_l}$ . The conductor characteristics are extracted from [17], while the capacitance  $C_l'$  and the inductance  $L_l'$  are calculated assuming the transposition of the phase conductors. The formulas used in this case consider the number of conductors, their relative position and their diameter:

$$L_l' = 0,46 \log\left(\frac{\Delta'}{(D/2)}\right) = 1.28 \text{ mH/km} \quad (5)$$

$$C_l' = \frac{0,024}{\ln\left(\frac{\Delta'}{(D/2)}\right)} = 3.74 \text{ nF/km} \quad (6)$$

where  $\Delta'$  is the equivalent distance between the conductors ( $\Delta' = \sqrt[3]{\Delta_{12}\Delta_{23}\Delta_{31}}$ ). The physical size of the joint and the nominal value of the DC electrical resistivity are extracted from [15] and shown in Table 1. The joint resistivity can be calculated through the equation  $\rho_{sj} = \frac{R_{sj} d \Delta \lambda}{H}$ . In nominal conditions the inductance value  $L_{sj}$  is 1.5  $\mu\text{H}$  and the capacitance  $C_{sj}$  is calculated by (4) assuming very low values for the breaking parameters ( $h = 0,001H$  and  $x = 0,001d$ ). In this way, it is possible to take into account the parasitic effects and maintain the same circuit topology for the nominal condition and the partial breaking condition. Therefore, the nominal value of the joint capacitance is 0.01 pF.

#### 4.1. Tolerance on the conductor electrical parameters

As previously stated, the conductors are assumed not faulty, hence they are set to their nominal values. Their resistance  $R_l'$  depends on temperature and is subjected to the skin effect, on its time depending on the frequency  $f_m$  used for the measurements. The penetration depth of the current  $\delta_l$  is given by the standard relation  $\delta_l = \frac{\sqrt{\rho_l}}{\sqrt{\pi \mu_l f_m}}$ , where  $\rho_l$  and  $\mu_l$  are respectively the resistivity of the conductor and the relative magnetic permeability. In order to obtain the correct value of the resistance, a parameter  $K_0 = \frac{(D/2)}{\delta_l}$  is introduced, which is calculated as the ratio between the radius of the conductor and the penetration depth of the current. The electrical resistance of the conductor at ambient temperature (293K) can be obtained by multiplying  $R_l'$  for a coefficient  $K$  so defined:

$$\begin{cases} K = \left[1 + \frac{K_0^4}{3}\right] & (K_0 < 1) \\ K = \left[0,25 + K_0 + \frac{3}{K_0}\right] & (K_0 > 1) \end{cases} \quad (7)$$

Then the final form of the electrical resistance is obtained by  $R_{l'fm} = KR_l'$ . As well known, the resistivity of a conductor changes its value with respect to the temperature. This means that there is a tolerance around the nominal resistance due to the conductor temperature. The definition of the resistance range is very important for verifying the performance of the monitoring system, since the neural network considers all the conductor components fixed at their nominal values, but the measurements of the frequency response change according to their real values. Therefore, the energy balance equation of transmission line is introduced [29, 30]:

$$C_m \nu \pi A_t \frac{dT_c}{dt} = R_l'(1 + \alpha(T_c - 293))I^2 + \alpha_s Q_s A_s - \alpha_c (T_c - T_a) A_s - 5,67 \alpha_r \left[ \left(\frac{T_c}{100}\right)^4 - \left(\frac{T_a}{100}\right)^4 \right] A_r \quad (8)$$

where:  $T_c$  is the temperature of the conductor;  $T_a$  is the air temperature surrounding the conductor;  $C_m$  is the specific heat of the conductor materials;  $\nu$  is the density of the conductor materials;  $A_t$  is the conductor section;  $A_s$  is the surface area receiving illumination with unit of length;  $A_c$  is the convection surface area;  $A_r$  is the radiation surface area per unit of length;  $\alpha_c$  is the coefficient of convective heat;  $\alpha_r$  is the radiation factor of the conductor materials;  $\alpha_s$  is the sunshine absorption rate of the

Table 1: Characteristic of the line conductor and bolted joint

ACSR 22.8	Line Voltage [kV]	Outer Diameter D [mm]	Wire diameter (Al) [mm]	Number of wires (Al)	Wire Diameter (Steel) [mm]	Number of wires (Steel)	$R_l'$ [ $\Omega/\text{Km}$ ]
		132	22.8	3.5	26	2.8	7
Bolted joint	H [m]		$\Delta \lambda$ [m]		d [m]		$R_{sj}$ [ $\mu\Omega$ ]
	0.038		0.008		0.008		60



Table 2: Characteristics of the conductor materials

$C_m$ [J/(kg K)]	$v$ [kg/m <sup>3</sup> ]	$A_t$ [m <sup>2</sup> ]	$A_s$ [m <sup>2</sup> /m]	$A_r$ [m <sup>2</sup> /m]	$A_c$ [m <sup>2</sup> /m]	$\alpha_s$	$\alpha_c$ [W/(kg/m <sup>3</sup> )]	$\alpha_r$
0.88 e-3	2.7 e3	$\pi(D/2)^2$	$\pi(D/2)$	$\pi D$	$\pi D$	0.6	5.8	0.2

conductor;  $Q_s$  is the radiation intensity of the sun and sky. Within the equation there is the term  $R_l'$  because it produces heat due to the Joule effect. It is necessary to observe that this is the value of the DC resistance considered equal to the resistance value at 50 Hz. The characteristics of the conductor material are extracted from [29] and reported in Table 2.

The expression of temperature variations as a function of time is obtained by solving the energy balance equation. In order to solve the equation, the initial conditions of the conductor, the radiation intensity  $Q_s$  and the air temperature  $T_a$  must be fixed. Typical values of  $Q_s$  and  $T_a$ , based on the measurements obtained in Italy in 2011, are used in this work. In the following,  $Q_{smax}$  represents the average of the maximum radiation levels for a specific month and  $T_{amax}$  that of the maximum air temperatures. The average of the minimum temperature levels  $T_{min}$  is considered as the mean temperature in the night period and, for this reason, it is associated with a zero-radiation level. For simplification, the average values for each season are considered in Table 3. Once the load current has been fixed, the incidence of the environmental conditions on the resistance value can be obtained.

Table 3: Average values of solar radiation and temperature for each season

	Winter	Spring	Summer	Autumn
$Q_{smax}$ [W/m <sup>2</sup> ]	392	753	750	357
$T_{amax}$ [K]	281	295	301	284
$Q_{smin}$ [W/m <sup>2</sup> ]	0	0	0	0
$T_{amin}$ [K]	275	286	292	279

For example, if in summer the conductor temperature is calculated using  $(Q_{smax}, T_{amax})$ , the maximum resistance error in the night period is about 16%. Considering the interval 0÷500 A for the load current and setting the environmental conditions, it is possible to obtain the incidence of the load on the resistance value. For example, in summer, the influence of the load current is about 7%. The worst situation for each season is presented in Table 4 and the errors are calculated assuming the maximum difference between the presumed and the real current. Similarly, the difference between the assumed and actual environmental conditions is maximized.

Table 4: Worst error for each season

Selected load current [A]	Real load current [A]	Error %			
		Winter	Spring	Summer	Autumn
500	0	17.34	25.16	24.92	16.1

In this case the maximum error is made in spring and its percentage value is about 25,16%. By calculating the temperature with intermediate starting conditions between  $(Q_{smax}, T_{amax})$  and  $(Q_{smin}, T_{amin})$  and fixing the current at 250 A, it is possible to consider a 10÷12% variation range for the conductor resistance. Where  $(Q_{smax}, T_{amax})$  and  $(Q_{smin}, T_{amin})$  are selected and real environmental condition respectively. The same variation range is extended to each electrical parameter of the conductor.

#### 4.2. Circuit designing on SapWin IV

The simulator *SapWin IV* (Symbolic Analysis Program for Windows) [31, 32] is used for executing the circuit analysis and obtaining the symbolic network function of the model. The line admittance calculated at the starting point of the line is the network function. The first step of the fault procedure is the testability analysis, carried out by the program LINFTA [20] exploiting the *SapWin IV* simulation. Despite the fact that only the joint parameters are variable, all the electrical components are taken into account in the testability analysis. The circuits characterized by 2, 3, 4 and 5 elementary sections have been tested and the corresponding testability values are always maximum. The second step consists in the determination of the optimum set of frequencies. A program associated with the package *SapWin IV* performs this operation [24]. In this phase many circuit simulations are needed, at various different frequencies. The availability of the network function in symbolic form, as that provided by *SAPWIN*, is fundamental for keeping the processing times within reasonable limits. A single simulation is in fact sufficient, because, once the symbolic network function is available, it can be directly used to derive the response at all the desired frequencies. In the third step of the fault procedure the symbolic network function extracted from *SapWin* is processed on *MatLab*<sup>®</sup> to obtain the training samples. Through a *MatLab*<sup>®</sup> script, it is possible to set the conductor parameters at the numerical values and generate the training samples by varying the electrical components of each joint in its fault classes. Each conductor parameter is generated randomly in an interval of 10% around its nominal value, so taking into account the effects of environmental conditions and load current.

#### 4.3. Fault classes and neural network setup

Three possible health states are considered for each junction region: nominal condition, structure oxidation and partial breakage. In order to obtain the correct classification of the working conditions by using the neural network, it is necessary to set the corresponding intervals for each electrical parameter. As mentioned above, the nominal value of the joint resistance is obtained from [17], while those of the inductance and capacitance have been calculated. Using the results presented in [17], it is also possible to define the resistance values in case of presence of the oxidation process (Table 5). These values represent the DC resistances of the joint in each operating condition and the

Table 5: Values of the joint resistance in the different oxidation conditions

Nominal condition	Low oxidation	Medium oxidation	High oxidation
$R_{sj} = 60 \mu\Omega \div 2.5 m\Omega$	$R_{sj} = 2.5 m\Omega \div 5 m\Omega$	$R_{sj} = 5 m\Omega \div 100 m\Omega$	$R_{sj} = 100m\Omega \div 2 \Omega$
$\rho_{sj} = 10^{-7} \div 4 \cdot 10^{-6} \Omega \cdot m$	$\rho_{sj} = 4 \cdot 10^{-6} \div 8.5 \cdot 10^{-6} \Omega \cdot m$	$\rho_{sj} = 8.5 \cdot 10^{-6} \div 1.7 \cdot 10^{-4} \Omega \cdot m$	$\rho_{sj} = 1.7 \cdot 10^{-4} \div 3.4 \cdot 10^{-3} \Omega \cdot m$

corresponding resistivities. The resistance values are given by (1) and we can note that the incidence of the skin effect decreases as the oxidation process increases. In fact, the oxidation process increases the resistivity of the material and, consequently, also increases the depth of the current penetration. Therefore, the skin effect is not relevant in the oxidation conditions, while it could introduce a false positive error in the nominal conditions.

inductance variation with respect to the size of the crack is presented in Figure 5b and Figure 6a. According to the inductance value previously calculated, a tolerance of 0.2  $\mu H$  is used to define the nominal conditions.

As shown in Figure 5b the value of  $L_{sj}$  leaves the interval of nominal conditions for  $h$  and  $x$  included between 10% and 65% of  $H$  and  $\Delta\lambda$  respectively.

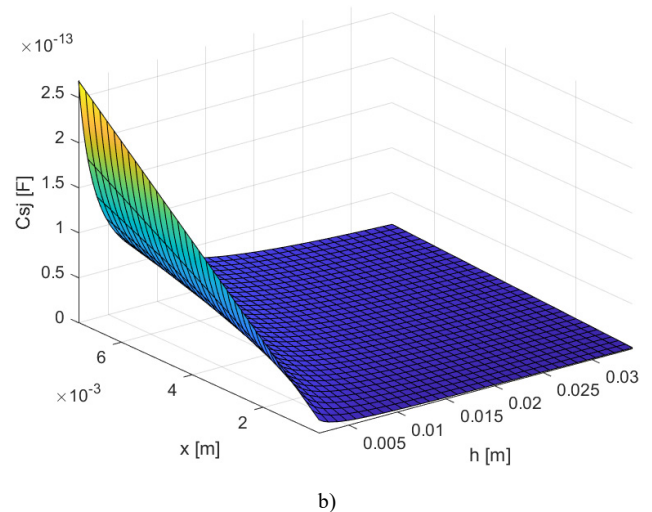
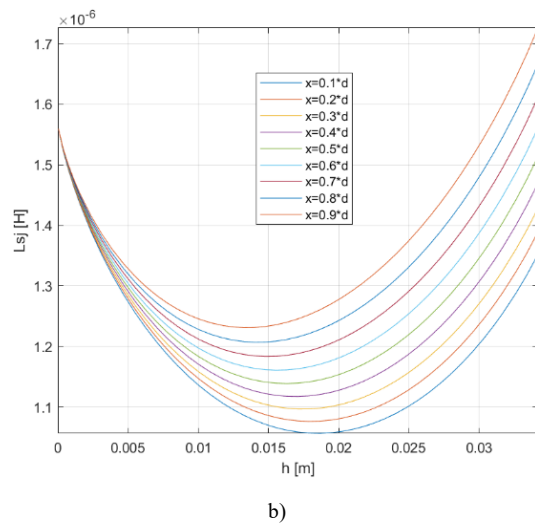
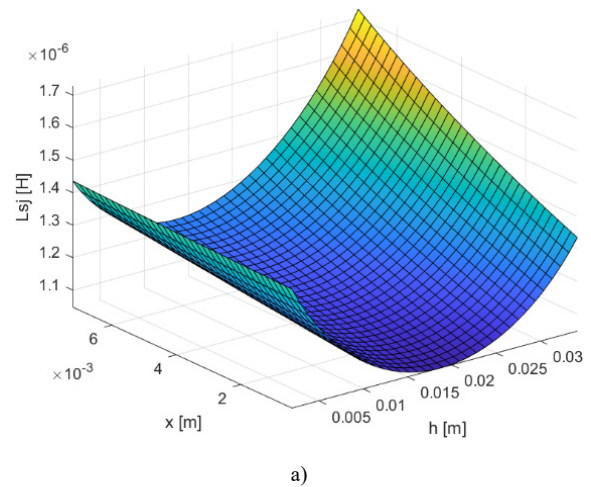
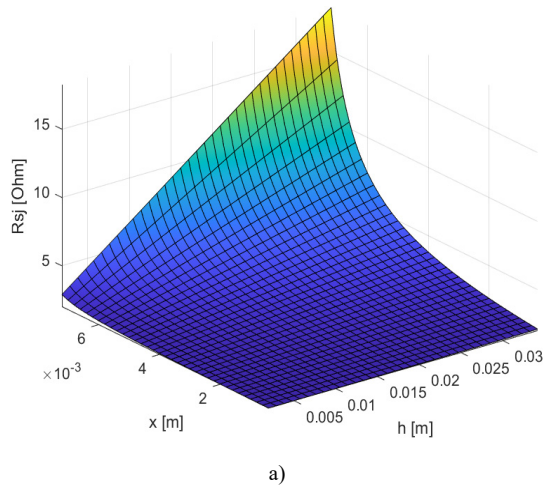


Figure 5: a) Resistance of the joint with respect to the crack height and crack width; b) Inductance of the joint with respect to the crack height and crack width

Figure 6: a) Inductance of the joint with respect to the crack height and crack width; b) Capacitance of the joint with respect to the crack height and crack width

The skin effect can be also neglected in correspondence with the partial damage of the joint structure since the consequentiality between oxidation and breaking mechanism has been assumed. For this reason, only the second relation presented in (2) is used to describe the variation of the joint resistance with respect to the size of the crack (Figure 5a). In this case the resistivity  $\rho_{sj} = 3,4 \cdot 10^{-3} \Omega m$  is used and  $\delta = d/2$ . The resistance of the joint  $R_{sj} = 2 \Omega$  represents the upper limit in the oxidation conditions. The

Since it is reasonable to consider equal percentage variations of  $x$  and  $h$ , since 65% of breakage represents a very high level of degradation, the interval chosen for the braking conditions is  $L_{sj} = 1 \div 1.3 \mu H$  neglecting the size of the crack for which the inductance returns in the nominal interval. Concerning the capacitance of the joint, its variation with respect to the size of the crack is shown in Figure 6b. Unfortunately, the sensitivity of the

Table 6: Fault classes

Joint Parameters	Nominal Condition	Oxidation Condition	Breaking Condition
$R_{sj}$	$60 \mu\Omega \div 2.5 m\Omega$	$2.5 m\Omega \div 2 \Omega$	$> 2 \Omega$
$L_{sj}$	$1.5 \pm 0.2 \mu H$	$1.5 \pm 0.2 \mu H$	$1 \div 1.3 \mu H$
$C_{sj}$	$0.01 pF$	$0.01 pF$	$0.01 pF$

line frequency response with respect to the capacitance is low and, consequently, the value of  $C_{sj}$  is fixed to the nominal value.

Starting from these results the fault classes are defined. They are shown in Table 6. Once the equivalent lumped circuit has been obtained and the value of each electrical parameter has been defined, the complex neural network must be adapted in order to classify the health state of the joints.

The inputs of the neural network are the complex values corresponding to the network function at the selected frequencies. The network is characterized by two layers of neurons and two output neurons for each junction region (Figure 7).

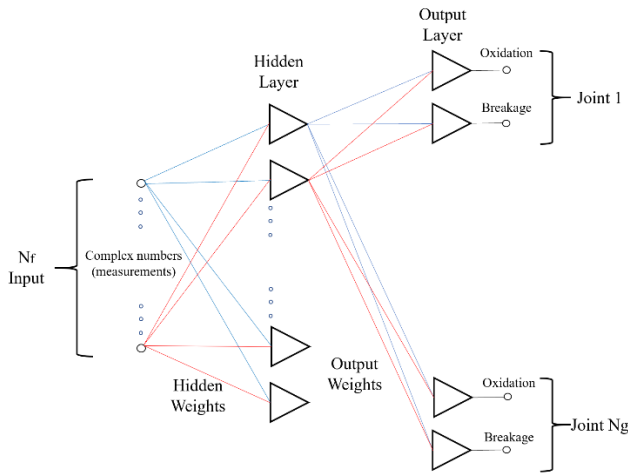


Figure 7: Structure of the neural network

Each output neuron divides the complex plane in two different sectors (Figure 8) representing the nominal conditions and the presence of the failure mechanism.

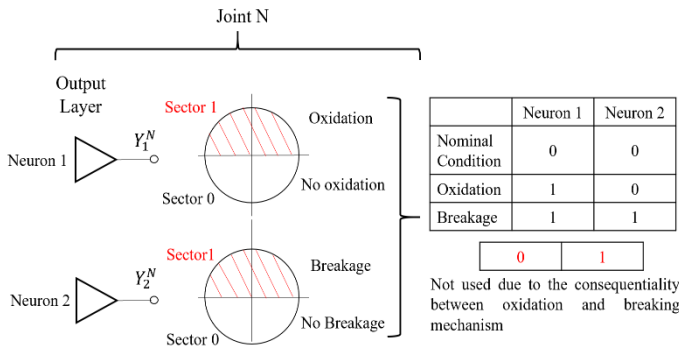


Figure 8: Output combinations for each neuron

The first output neuron of each joint takes into account the oxidation process, while the second one considers the partial breaking mechanism. When the output level of both neurons is low

(sector zero), the corresponding junction region is within the nominal conditions. If the first output is high (sector one) and the second one is low (sector zero), the combination represents the oxidation of the joint. If each output presents a high level (sector one), the corresponding junction region has partial damage. The last combination is not taken into consideration due to the hypothesis of the consequentiality between oxidation and partial breaking mechanism.

This paper uses a complex neural network implemented on MatLab® version R2019b, but the whole code can also be used in previous versions of this software. Once the number of joints  $N_g$  has been set, there are  $N_c = 3^{N_g}$  possible output combinations, since each junction region can be oxidized, broken or fully functional.

#### 4.4. Neural network training and testing

The complex neural network used in this work, like other perceptron neural networks, has a good learning speed and allows to significantly reduce the computational cost by having a derivative free learning algorithm. Furthermore, the complex nature of this network is very suitable for dealing with electrical quantities and usually it is possible to obtain a better generalization capability in comparison with other neural networks and neuro-fuzzy networks. The learning rule is shown in (9):

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \frac{\alpha_r}{(n+1)|Z_k|} (D - Y)\bar{\mathbf{X}} \quad (9)$$

where  $\mathbf{w}_{k+1}$  is the vector of the corrected weights,  $\alpha_r$  is the learning constant,  $n$  is the number of the inputs,  $Z_k$  is the weighted sum,  $D$  is the desired output,  $Y$  is the actual output and  $\bar{\mathbf{X}}$  is the conjugate complex vector of the inputs. Considering  $N_{rs}$  random samples for each combination, the matrix of the data sets presents  $N_t = N_{rs}N_c$  rows. Each row contains  $N_f$  magnitude measurements and  $N_f$  phase measurements of the same network function, with  $N_f$  number of test frequencies.

$$\begin{bmatrix} |M_1^1|_{f_1} & |\phi_1^1|_{f_1} & \dots & |M_1^1|_{f_{N_f}} & \dots & O_1^1 & O_2^1 & \dots \\ |M_2^1|_{f_1} & |\phi_2^1|_{f_1} & \dots & |M_2^1|_{f_{N_f}} & \dots & O_1^1 & O_2^1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ |M_{N_{rs}}^{N_c}|_{f_1} & |\phi_{N_{rs}}^{N_c}|_{f_1} & \dots & |M_{N_{rs}}^{N_c}|_{f_{N_f}} & \dots & O_1^1 & O_2^1 & \dots \end{bmatrix} \quad (10)$$

Equation (10) shows the structure of the matrix containing the data sets. For example,  $|M_2^1|_{f_1}$  represents the second measure of magnitude corresponding to the first combination made at the frequency  $f_1$ ,  $|\phi_2^1|_{f_1}$  represents the second measure of phase corresponding to the first combination made at the frequency  $f_1$  and  $O_2^1$  is the second output of the first joint. Initially the weights of the neural network are randomly chosen and, subsequently, for each sample the error between the desired output and the actual one is

calculated. The procedure shown in [26] is used to obtain the correction of the weights. This method uses a part of the total samples  $N_a$ , while the remaining part is used for the testing procedure. The backpropagation formulas calculate the error for each weight of the hidden layer and, then, the system shown in (11) is obtained.

$$\begin{bmatrix} 1 & X_1^1 & \dots & X_{N_f}^1 \\ 1 & X_1^2 & \dots & X_{N_f}^2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_1^{N_a} & \dots & X_{N_f}^{N_a} \end{bmatrix} \begin{bmatrix} \Delta W_0 & \Delta W_0 & \dots & \Delta W_0 \\ \Delta W_1^{11} & \Delta W_1^{21} & \dots & \Delta W_1^{N_f 1} \\ \vdots & \vdots & \dots & \vdots \\ \Delta W_{N_f}^{11} & \Delta W_{N_f}^{21} & \dots & \Delta W_{N_f}^{N_f 1} \end{bmatrix} = \begin{bmatrix} \delta_{11}^1 & \delta_{21}^1 & \dots & \delta_{N_f 1}^1 \\ \delta_{11}^2 & \delta_{21}^2 & \dots & \delta_{N_f 1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{11}^{N_a} & \delta_{21}^{N_a} & \dots & \delta_{N_f 1}^{N_a} \end{bmatrix} \quad (11)$$

For example,  $X_1^2$  is the first line input corresponding to the second sample,  $\Delta W_1^{21}$  represents the correction for the weight between the first input and the second neuron of the first layer,  $\delta_{21}^2$  is the error on the second neuron of the first layer corresponding to the second sample. The values of the corrections are calculated through the Q-R decomposition and the same procedure is used for the output layer on the system shown in (12). The weight adjustment is repeated until the error satisfies the stopping criteria indicated in [27]. Therefore, there are two different stopping criteria: the first concerns the maximum number of classification errors that can be committed during the test phase and this means that the operator can set the maximum number of outputs that can be outside the desired sector. The second stopping criterion is chosen according to the rules of the soft margin technique and consists of a tolerance range on the distance between the value of each output and the bisector of the desired sector.

$$\begin{bmatrix} 1 & Y_{11}^1 & \dots & Y_{N_j 1}^1 \\ 1 & Y_{11}^2 & \dots & Y_{N_j 1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Y_{11}^{N_a} & \dots & Y_{N_j 1}^{N_a} \end{bmatrix} \begin{bmatrix} \Delta W_0 & \Delta W_0 & \dots & \Delta W_0 \\ \Delta W_1^{12} & \Delta W_1^{22} & \dots & \Delta W_1^{N_m 2} \\ \vdots & \vdots & \dots & \vdots \\ \Delta W_{N_j}^{12} & \Delta W_{N_j}^{22} & \dots & \Delta W_{N_j}^{N_m 2} \end{bmatrix} = \begin{bmatrix} \delta_{12}^1 & \delta_{22}^1 & \dots & \delta_{N_m 2}^1 \\ \delta_{12}^2 & \delta_{22}^2 & \dots & \delta_{N_m 2}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{12}^{N_a} & \delta_{22}^{N_a} & \dots & \delta_{N_m 2}^{N_a} \end{bmatrix} \quad (12)$$

### 5. Results

All the fundamental steps in the classification of the joint conditions are presented in this paragraph to evaluate the performances of the proposed diagnostic method. The main objective of these simulations is to show that the system is able to identify the fault classes presented in Table 6, so confirming the

validity of the presented prognostic approach for the predictive maintenance of high voltage electrical lines. A branch of electrical network containing three junction regions is the first configuration considered for the simulations. In this case, since the conductor length between two consecutive joints is 300 meters, the diagnostic system analyses approximately 900 meters of network. The analytical procedure shown in the paragraph 3.2 is used to obtain the optimal test frequency ( $f_m$ ) and, starting from this value, also the next three octaves are considered to realize the dataset. This means that four signals are used to measure the line frequency response and, consequently, the network function magnitude and phase are calculated for each test frequency. The equivalent line admittance is used in this work. It corresponds to the ratio between the output current and the input voltage measured at the starting point of the network. From the theoretical point of view, the main steps to obtain the measurements are the equivalent circuit simulation on SapWin IV and the dataset generation on MatLab<sup>®</sup>. The first step is the equivalent circuit simulation, which allows to obtain the symbolic formula of the line network function. For this reason, a cascade of three elementary sections (Figure 4) is realized on SapWin IV. The joint electrical parameters are kept symbolic, while the conductor components are set to their nominal values (as already said, taking into account the tolerances). Once the equivalent line admittance has been obtained, a specific MatLab<sup>®</sup> code is used to generate the dataset. The main task of this code is to calculate the analytical formulas of magnitude and phase for each test frequency; subsequently the neural network inputs are calculated for each health state of the line by replacing the symbolic parameters with their numerical values. Since there are 3 junction regions in the first simulation ( $N_g = 3$ ), 27 different combinations of the health state of the joints are obtained ( $N_c = 3^{N_g} = 27$ ) and, for each of them, 100 random samples ( $N_{ts} = 100$ ) are generated to complete the dataset. This means that each electrical parameter of the joints is randomly selected within one of the intervals shown in Table 6. The choice of 100 random samples is the result of some tests made to obtain a good compromise between the speed of processing and the representativeness of the dataset; it is not an optimized choice but it could, using specific techniques. Finally, the MatLab<sup>®</sup> script uses the organization of the outputs shown in section 4 to associate the exact health state of the joints with each sample. Table 7 summarizes the situation for the first simulation. Finally, it is necessary to select the characteristics of the neural network, i.e. number of hidden neurons, initial values of the weights, learning rate, backpropagation procedure, computational method used for the weight adjustment. The last two are described in paragraph 4, while the fixed parameters of the neural network are selected through a heuristic approach. In particular, the number of hidden neurons is chosen through some tests to obtain good performances and a good generalization capability. A specific MatLab<sup>®</sup> application is used to facilitate the selection, because it

Table 7: Dataset parameters for the simulation with three junction regions

Ng	Hidden Neurons	Output Neurons	Initial values of the weights	Training samples	Test samples	Learning rate	Classification rate
3	70	6	5	80%	20%	1	0.84074



allows to obtain a very fast evaluation of the neural network performances by modifying the most important parameters.

The index used to evaluate neural network performances is called classification rate, and it is defined as the ratio between the number of correctly classified samples and the total number of samples used in the test phase. Indeed, only a percentage of the dataset is used for training the neural network, while the other samples are used to verify the correct functioning of the system. Table 8 presents the best neural network configuration for the first simulation and the corresponding value of the classification rate. The classification rate shown in Table 8 represents the global index to evaluate the performance of the diagnostic system and it is obtained by considering the classification errors made on 540 samples. This means that 2160 samples are used to modify the weight values during the training phase, while the remaining 540 samples are used to verify the results during the test phase.

Table 8: Complex neural network configuration for the simulation with three junction regions

Ng	Nrs	Nc	Nt	Dataset Input	Dataset Output	$f_m$ [kHz]
3	100	27	2700	8	6	35

It must be observed that each neuron has a specific classification rate greater than 0.84074 and this means that the performance of the system on a single junction region is much better than the global one. Table 9 shows the results obtained for each neuron and each junction region in the first configuration. Therefore, considering each joint separately, there is a probability range of 91÷95 % that the health state is correctly classified. It is not possible to establish a standard relationship between the classification rate for each pair of neurons and for the corresponding joint, because the number of errors for each junction region depends on the oxidation neuron, the rupture neuron or both.

Table 9: Results obtained for each joint and each neuron in the simulation with three junction regions

Fault class	Neuron	Classification Rate for each neuron	Joint	Classification Rate for each joint
Oxidation	1	0.9741	1	0.9470
Partial Breakage	2	0.9796		
Oxidation	1	0.9685	2	0.9430
Partial Breakage	2	0.9977		
Oxidation	1	0.9444	3	0.9130
Partial Breakage	2	0.9778		

To get a more accurate evaluation of the system performance it is necessary to introduce the cross-validation method. In this case, the previously described training phase is repeated 5 times changing the learning data and the testing data. Therefore, the global classification rate is calculated on 2700 samples and represents the most reliable index. Cross-validation method is presented in Table 10 where the maximum and minimum classification rates represent the best and the worst result obtained at the end of each learning phase.

Table 10: Results obtained through the cross-validation method in the simulation with three junction regions

Fault class	Neuron	Joint	Maximum Classification Rate	Classification Rate
Oxidation	1	1	0.97037	0.9648
Partial Breakage	2		0.98889	0.9793
Oxidation	1	2	0.96481	0.9533
Partial Breakage	2		0.98333	0.9778
Oxidation	1	3	0.95000	0.9411
Partial Breakage	2		0.99259	0.9837
Global Classification Rate			0.8167	

The same simulation procedure is used to evaluate the health state of a line with four junction regions. In this case 6480 samples are used to modify the weight values during the training phase, while the remaining 1620 samples are used to verify the results during the test phase. Table 11 summarizes the situation and shows the classification rate used to choose the neural network set up.

Table 11: Dataset and complex neural network set up for the simulation with four junction regions

Ng	Nrs	Nc	Nt
4	100	81	8100
Hidden Neurons	Output Neurons	Initial values of the weights	Training samples
190	8	5	80%
Dataset Input	Dataset Output	$f_m$ [kHz]	Test samples
8	8	80	20%
Classification rate			
0.70247			

Table 12: Results obtained through the cross-validation method in the simulation with four junction regions

Fault class	Neuron	Joint	Maximum Classification Rate	Classification Rate
Oxidation	1	1	0.97716	0.9700
Partial Breakage	2		0.98889	0.9830
Oxidation	1	2	0.93025	0.9188
Partial Breakage	2		0.98333	0.9757
Oxidation	1	3	0.94383	0.9367
Partial Breakage	2		0.98210	0.9781
Oxidation	1	4	0.91667	0.9046
Partial breakage	2		0.98333	0.9793
Global Classification Rate			0.6899	

Also, in this case the cross-validation method is used to carry out the correct evaluation of the diagnostic system (Table 12). As mentioned above, by repeating the training phase 5 times, it is possible to evaluate all 8100 samples during the test phase. As shown in Table 12, the results obtained for each neuron are still excellent, while the global classification rate is lower than the first configuration. Obviously, increasing the number of the joints reduces the reliability of the complete system, which must be considered as a connection of four classifiers (one for each junction region). Therefore, if the main objective of the diagnostic system was to classify the exact combination of the health state of the joints, the system reliability is about 70%.

On the other hand, the working condition of each junction region can be classified with an accuracy level higher than 90%. This result is obtained considering each pair of output neurons separately (Table 13).

Table 13: Results obtained for each joint in the simulation with four junction regions

Joint	Classification Rate for each joint
1	0.9538
2	0.9006
3	0.9157
4	0.8854

The results presented in Table 13 can be used to compare the performance of the complex neural network with that of the other classifiers. Table 14 shows the comparison with one of the most used classifiers based on quadratic SVM (Support Vector Machine).

Table 14: Comparison between complex neural network and SVM classifier in the simulation with four junction regions

Joint	Classification for each joint with complex neural network	Classification rate for each joint with quadratic SVM
1	0.9538	0.9670
2	0.9006	0.8500
3	0.9157	0.8530
4	0.8854	0.8480

## 6. Conclusion

A diagnostic system based on a complex neural network has been presented and the procedure to obtain the state classification of the electrical joints has been verified through some different simulations. The results obtained for power lines characterized by three and four junction regions show that the system is able to identify the health state of each joint, with an accuracy level greater than 90%. The actual limit for the diagnostic method is represented by the line length, because the global classification rate obtained for a line with five joints is lower than the previous ones (Tables 15-16). Since the results obtained for the fifth junction region are not good, the system can be only used in a real application for small branches of the network. For this reason, future developments will concern the creation of a macroscopic method of locating failure

mechanisms: the section of the non-functioning line could be localized through classification methods based on load flow measurements and after, the system presented in this work could be integrated to identify the junction region in the worst condition. In this way it would be possible to analyze a great variety of networks. Finally, a further development of the prognostic method could concern the analysis of the errors. The division between false positive and false negative errors is currently calculated, but no operation results from this. In the future, the method could be modified to minimize false negatives, which are the most dangerous errors for a diagnostic system

Table 15: Dataset and complex neural network configuration for the simulation with five junction regions

Ng	Nrs	Nc	Nt
5	100	243	24300
Hidden Neurons	Output Neurons	Initial values of the weights	Training samples
280	10	5	80%
Dataset Input	Dataset Output	$f_m$ [kHz]	Test samples
8	10	80	20%
Classification rate			
0.098971			

Table 15: Results obtained for each neuron in the case of five junction regions

Fault class	Neuron	Joint	Specific Classification Rate
Oxidation	1	1	0.9383
Partial Breakage	2		0.9885
Oxidation	1	2	0.7043
Partial Breakage	2		0.7648
Oxidation	1	3	0.7342
Partial Breakage	2		0.8438
Oxidation	1	4	0.7792
Partial breakage	2		0.9023
Oxidation	1	5	0.6780
Partial Breakage	2		0.6926

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

Smart Energy Lab, University of Florence and Dept. of Information Engineering, University of Florence.

## References

- [1] M. Bindi, F. Grasso, A. Luchetta, S. Manetti, M. C. Piccirilli, "Smart Monitoring and Fault Diagnosis of Joints in High Voltage Electrical Transmission Lines," in 2019 6th International Conference on Soft

- Computing & Machine Intelligence (ISCM), Johannesburg, South Africa, 40-44, 2019. doi: 10.1109/ISCM147871.2019.9004307
- [2] Y. D. Liu, G. H. Sheng, Z. M. He, X. Y. Xu, X. C. Jiang, "Method of fault location based on the distributed traveling-wave detection device on overhead transmission line," in 2011 IEEE Power Engineering and Automation Conference, Wuhan, 136-140, 2011. doi: 10.1109/PEAM.2011.6134925
- [3] X. Dong, S. Wang, S. Shi, "Research on characteristics of voltage fault traveling waves of transmission line", in 2010 Modern Electric Power Systems, Wroclaw, 1-5, 2010. doi: 10.1109/MEPS16993.2010
- [4] L. Xun, L. Shungui, H. Ronghui, A. Jingwen, A. Yunzhu, C. Ping, X. Zhengxiang, "Study on accuracy traveling wave fault location method of overhead line — Cable hybrid line and its influencing factors," in 2017 Chinese Automation Congress (CAC), Jinan, 4593-4597, 2017. doi: 10.1109/CAC.2017.8243590
- [5] M. Bindi, F. Grasso, A. Luchetta, S. Manetti, and M.C. Piccirilli "Modeling and diagnosis of joints in high voltage electrical transmission line", in 2019 Proc. of Int. Conference on Energy, Electrical and Power Engineering (CEEPE 2019), IOP Conf. Series: Journal of Physics: Conf. Series 1304 (2019) 012006, 2019. doi:10.1088/1742-6596/1304/1/012006
- [6] R.K. Aggarwal, A.T. Johns, J.A.S.B. Jayasinghe, W Su, "An overview of the condition monitoring of overhead lines", Electrical Power Systems Research 53, 15-12, 2000. [https://doi.org/10.1016/S0378-7796\(99\)00037-1](https://doi.org/10.1016/S0378-7796(99)00037-1)
- [7] C. Jinbing, W. Guogang, L. Huijun, "Mechanisms analysis and countermeasure of electrical contact overheating of overhead line," in 2014 International Conference on Power System Technology, Chengdu, 1503-1508, 2014. doi: 10.1109/POWERCON.2014.6993786
- [8] G. Andersson, "Modelling and Analysis of Electric Power System", PhD Thesis, ETH Zurich, 2008.
- [9] M. Davoudi, "Sensitivity Analysis of Power System State Estimation Regarding to Network Parameter Uncertainties", PhD Thesis, Politecnico di Milano, 2012.
- [10] A. Cataliotti, V. Cosentino, D. D. Cara, G. Tinè, "Simulation of a power line communication system in medium and low voltage distribution networks" in 2011 IEEE International Workshop on Applied Measurements for Power Systems (AMPS), Aachen, 107-111, 2011. doi: 10.1109/AMPS.2011.6090429
- [11] A. Cataliotti, D. Di Cara, R. Fiorelli, G. Tine, "Power-Line Communication in Medium-Voltage System: Simulation Model and Onfield Experimental Tests," IEEE Transactions on Power Delivery, 27(1), 62-69, 2012. doi: 10.1109/TPWRD.2011.2171009
- [12] J. S. Monteiro, W. L. A. Neves, D. Fernandes, B. A. Souza and A. B. Fernandes, "Computation of transmission line parameters for power flow studies," in 2005 Canadian Conference on Electrical and Computer Engineering, Saskatoon, 607-610, 2005. doi: 10.1109/CCECE.2005.1557004
- [13] N. A. Jaffrey, S. Hettiwatte, "Corrosion detection in steel reinforced aluminium conductor cables," in 2014 Australasian Universities Power Engineering Conference (AUPEC), Perth, WA, 1-6, 2014. doi: 10.1109/AUPEC.2014.6966630
- [14] G. M. Hashmi, M. Lehtonen, "Covered-Conductor Overhead Distribution Line Modeling and Experimental Verification for Determining its Line Characteristics," in 2007 IEEE Power Engineering Society Conference and Exposition in Africa - PowerAfrica, Johannesburg, 1-7, 2007. doi: 10.1109/PESAfr.2007.4498072
- [15] F. Steier, "Tribologic analyses of a self-mated aluminium contact used for overhead transmission lines", in 2017 Proc. of 6th Int. Conf. on Fracture Fatigue and Wear, IOP Conf. Series: Journal of Physics: Conf. Series 843 (2017) 012057, 2017. doi: 10.1088/1742-6596/843/1/012057
- [16] Z. Liu, Y. Sun, "A solder joint crack — Characteristic impedance model based on transmission line theory," in 2014 10th International Conference on Reliability, Maintainability and Safety (ICRMS), Guangzhou, 237-241, 2014. doi: 10.1109/ICRMS.2014.7107178
- [17] F. de Paulis, C. Olivieri, A. Orlandi, G. Giannuzzi, F. Bassi, C. Morandini, E. Fiorucci, G. Bucci, "Exploring Remote Monitoring of Degraded Compression and Bolted Joints in HV Power Transmission Lines," IEEE Transactions on Power Delivery, 31(5), 2179-2187, 2016. doi: 10.1109/TPWRD.2016.2562579
- [18] M. H. Azarian, E. Lando, M. Pecht, "An analytical model of the RF impedance change due to solder joint cracking," in 2011 IEEE 15th Workshop on Signal Propagation on Interconnects (SPI), Naples, 89-92, 2011. doi: 10.1109/SPI.2011.5898847
- [19] D. Achiriloaiei, S.V. Galatanu and I. Vetres, "Stress and strain determination occurring in contact area of 450/75 ACSR conductor", in 2018 Proc. of 7th Int. Conf. on Advanced Materials and Structures AMS (Timisoara), IOP Conf. Series: Materials Science and Engineering 416 (2018) 012071, 2018. doi:10.1088/1757-899X/416/1/012071
- [20] G. Fontana, A. Luchetta, S. Manetti, M. C. Piccirilli, "A Fast Algorithm for Testability Analysis of Large Linear Time-Invariant Networks," IEEE Transactions on Circuits and Systems I: Regular Papers, 64(6), 1564-1575, 2017. doi: 10.1109/TCSI.2016.2645079
- [21] G. Fontana, F. Grasso, A. Luchetta, S. Manetti, M. C. Piccirilli, A. Reatti, "A symbolic program for parameter identifiability analysis in systems modeled via equivalent linear time-invariant electrical circuits, with application to electromagnetic harvesters", Int. J. of Numerical Modelling: Electronic Networks, Devices and Fields, 32, 2019. <https://doi.org/10.1002/jnm.2251>
- [22] G. Fontana, A. Luchetta, S. Manetti and M.C. Piccirilli, "An unconditionally sound algorithm for testability analysis in linear time invariant electrical networks" Int. J. Circ. Theor. Appl., 44, 1308-1340, 2016. doi: 10.1002/cta.2164
- [23] N. Sen and R. Saeks, "Fault diagnosis for linear systems via multifrequency measurements," IEEE Transactions on Circuits and Systems, 26(7), 457-465, 1979. doi: 10.1109/TCS.1979.1084659
- [24] F. Grasso, A. Luchetta, S. Manetti, M. C. Piccirilli, "A Method for the Automatic Selection of Test Frequencies in Analog Fault Diagnosis," IEEE Transactions on Instrumentation and Measurement, 56(6), 2322-2329, 2007. doi: 10.1109/TIM.2007.907947
- [25] I. Aizenberg, Complex-Valued Neural Networks with Multi-Valued Neurons. Berlin: Springer-Verlag Publishers, 2011.
- [26] I. Aizenberg, A. Luchetta, S. Manetti, "A modified learning algorithm for the multilayer neural network with multi-valued neurons based on the complex QR decomposition", Soft Computing, 16, 563-575, 2012. doi: 10.1007/s00500-011-0755-7
- [27] I. Aizenberg, "MLMVN with Soft Margins Learning", IEEE Transactions on Neural Networks and Learning Systems, 25, 1632-1644, 2014. doi: 10.1109/TNNLS.2014.2301802
- [28] E. Aizenberg, I. Aizenberg, "Batch linear least squares-based learning algorithm for MLMVN with soft margins," in 2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM), Orlando, FL, 48-55, 2014. doi: 10.1109/CIDM.2014.7008147
- [29] Anjia Mao, Jing Wen and Shasha Luo, "A power system analysis model with consideration of environmental factors," in 2010 5th International Conference on Critical Infrastructure (CRIS), Beijing, 1-5, 2010. doi: 10.1109/CRIS.2010.5617557
- [30] J. R. Santos, A. G. Exposito and F. P. Sanchez, "Assessment of conductor thermal models for grid studies," IET Generation, Transmission & Distribution, 1(1), 155-161, 2007. doi: 10.1049/iet-gtd:20050472
- [31] F. Grasso, A. Luchetta, S. Manetti, M. C. Piccirilli and A. Reatti, "SapWin 4.0 – a new simulation program for electrical engineering education using symbolic analysis" Computer Application to Engineering Education, 24, 44-57, 2016. doi: 10.1002/cae.21671
- [32] [www.sapwin.info](http://www.sapwin.info) Dept. of Information Engineering, University of Florence, Italy.

## Remote Control of Garden Plantation Water Pumps using Arduino and GSM Mobile

Beza Negash Getu\*, Mohamed Abdulkadir, Michael Tous

Department of Electrical, Electronics and Communications Engineering, American University of Ras Al Khaimah, Ras Al Khaimah, 10021, UAE.

### ARTICLE INFO

Article history:

Received: 23 June, 2020

Accepted: 15 July, 2020

Online: 09 August, 2020

Keywords:

Soil moisture

Light intensity

Temperature

Humidity

Arduino

Greenhouse

Automatic gardening

GSM module

### ABSTRACT

A remotely mobile phone controlled electronic system for supplying water to garden plantations in a greenhouse or similar environment was designed and experimentally implemented. The system monitors the garden environmental conditions such as soil moisture, temperature, humidity and sunlight and a remote user can send commands from his mobile phone such as to switch ON/OFF water pumps, supply water for the plants for certain duration and acquire the environmental status information of the plantation. The water pumps are controlled by the Arduino microcontroller that is the core part of the electronic system. The user has the ability to monitor the environmental conditions on his mobile phone. Such systems can facilitate monitoring, give flexibility of controlling, save time and human labour and increase productivity as a result of automation and remote controlling.

### 1. Introduction

In the agricultural sector, the traditional irrigation techniques like watering cans, drip and sprinkler irrigations require farmers to be physically present in their farm lands to regularly supply water to their crops, which is a manually controlled irrigation method that results in wastage of human power as well as time. To circumvent labour presence at the farm site, in [1] Dual Tone Multi-Frequency (DTMF) based remotely controlled electronic system was proposed to control the switching states of agricultural water pumps located in remote farm areas. On the other hand, nowadays water shortage is becoming one of the biggest problems in the world from domestic home use as well as in agricultural and industrial sectors. Many different methods are developed for conservation of water. As water is one of the scarce natural resources, it is important to properly use and manage our usage in different sectors and hence there is a need to design systems that are used to monitor and control our water usage including in homes, agriculture and industrial sectors [2], [3]. It is known that agriculture is one of the fields where water is required in tremendous quantity and there are many techniques to save or to

control wastage of water from agriculture. With the development of technology in water saving irrigation and automation, automatic irrigation is going to be more popular in the farms. In [4] an automatic control system is designed and implemented for garden plantations based on soil moisture, temperature and light sensors. The system monitors and controls the soil moisture, temperature and light intensity based on the three sensors and the Arduino module that automatically controls a water pump, a fan and a lighting fixture based on the sensor information in the greenhouse. It controls the water pump automatically by the Arduino and the built electronic system.

Automation of irrigation systems based on soil moisture sensors (SMS) was also implemented in [5] to optimize water usage efficiency by maintaining soil moisture at optimum levels. It was found that a significant percentage of water savings can be achieved based on the quality of the installed sensors in the farm land. An Automatic irrigation system based on Arduino microcontroller and a sensor sensing the dampness of the soil was also proposed in [6]. An Arduino based automated irrigation system based on soil moisture sensor was also proposed in [7]. In this system, the microcontroller is interfaced with the HC-05 Bluetooth module that can be connected to Android phone for

\*Corresponding Author: Beza Negash Getu, American University of Ras Al Khaimah, bgetu@aurak.ac.ae



controlling the water pumps in the farm. The drawback of this control is the limited range of the Bluetooth communication that can be from a distance of 10 m to 100 m from the transmitter [8]. In plant watering system that was proposed in [9], the soil moisture sensor checks the moisture level in the soil and if moisture level is low then the Arduino informs the user by sending message to the user cell phone then the user sends a command that switches on the water pump to provide water to the plants. However, this system is based on a PIC16F628 microcontroller that is different from Arduino and used two metal plate sensors for moisture sensing instead of standard commercially available sensors that limits the accuracy. A microcontroller based irrigation system was also proposed in [10] based on 8051 Microcontroller. However, a system based on Arduino simplifies the process of working with microcontrollers and it is inexpensive and widely used everywhere in recent days [11], [12].

The purpose of the work in this paper is to further improve the work done in [4] so that there is an ability to remotely control the water pumps depending on the farm environmental conditions received from the Arduino. The designed electronic system monitors the soil water level (moisture level) of the garden plantation and supplies water through the water pumps that are controlled by the Arduino microcontroller. The system is to be controlled by sending commands remotely from a mobile phone of the owner of the garden plantation or the farms. In addition, the system also needs to monitor the garden environmental conditions (moisture, temperature, light, humidity) and inform the user so that he can take the necessary actions. Such Global System for Mobile (GSM) based automatic irrigation system increases the productivity of farms and eliminates the cost of human labour used in manual irrigation methods. It provides convenience in accessing the system from anywhere at any time.

## 2. Methods and Materials

The designed system and the various components and standards used for this project are described in this section

### 2.1. System Design and Components

The designed system as shown in Figure 1 is implemented to provide the necessary supplies for a farm or a greenhouse. The system is designed to monitor and control the sensor elements through the Arduino microcontroller and providing the status of the land to the user through the GSM module. The sensors are connected to the Arduino to provide the microcontroller with the soil moisture, humidity and temperature and light status of the plantation. The microcontroller stores the sensors data information and then sends them to the user upon request. It also receives commands from the user through the GSM module and then supplies the land with water, light and air conditioning by switching ON/OFF the water pumps, the light and the fans respectively depending on the required needs of the plants.

The standards of the project are commercially available components brought together to form the project and make it a complete one. The design aspect is to integrate all the design components in a way to achieve our target and the target tasks are achieved through programming the Arduino microcontroller that is interacting with the various input and output components. Some of the main components of the system are listed below.

[www.astesj.com](http://www.astesj.com)

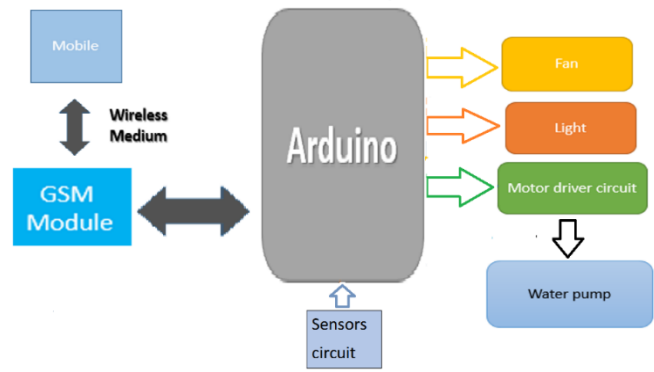


Figure 1: System Block Diagram

#### 2.1.1. Arduino Uno

Arduino is an open source microcontroller used for hardware and software integration. Arduino can be interfaced with sensor elements and control them in the real world. Arduino has a set of digital and analogue input/output pins that can be interfaced to distinctive developed boards (shields) and different sensor circuits [12]. Serial communication interfaces including Universal Serial Bus (USB) are required for loading programs from Personal Computer (PC) to the Arduino. The Arduino board as shown in Figure 2 provides an integrated development environment (IDE) for programming the microcontroller to process various types of projects. The Arduino IDE provides a simplified integrated platform which can run on regular PCs and allows users to write programs for Arduino using C or C++ language.



Figure 2: Arduino UNO Board

#### 2.1.2. GSM GPRS Quad Band Board SIM900 Shield for Arduino

GSM is a second generation (2G) digital cellular networks used by mobile phones. A GSM modem is a device that modulates and demodulates signals for communications. The modem used in this project is SIMCOM SIM900 [13], which is a quad-band GSM/General Packet Radio Service (GPRS) modem as it can operate at four frequencies (850 MHz, 900 MHz, 1800 MHz and 1900 MHz). GSM provides a wide range of things to remotely control from anywhere with your fingertips. SIM900 GSM module can be used to send and receive SMS by connecting it to a PC or Arduino when a Subscriber Identification Module (SIM) card is inserted. The GSM modem can send ATtention (AT) commands to send or receive Short Message Service (SMS) from PC via a

COM (serial or USB) port. The SIM900 GSM module used in this project as shown in Figure 3 consists of a Transistor-Transistor Logic (TTL) and an RS232 interface. The TTL interface allows you to interface directly with a microcontroller while the RS232 interface allows to communication with PC.

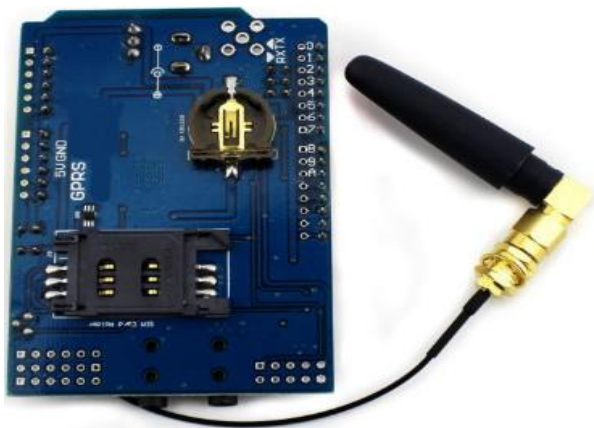


Figure 3: GSM/GPRS Quad Band Board

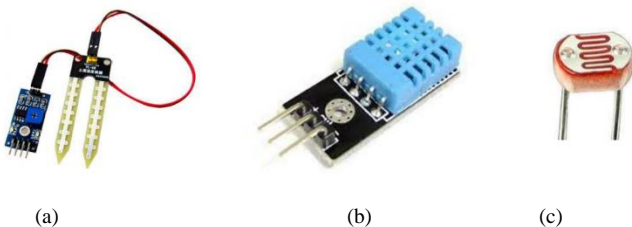


Figure 4: Plant Environmental Condition Sensors: (a) Soil moisture sensor, (b) The Arduino DHT11, (c) LDR

### 2.1.3. Plant Environmental Condition Sensors

The proposed system is intended to control and monitor four different parameters (soil moisture, temperature, humidity and light status) of a certain garden or greenhouse planation that are vital for the healthy growth of plants. Calculating the soil moisture using soil moisture sensors in the fields is important to help the farmers improve irrigation system efficiently. The quality of the crop production can be increased by management of soil moisture during critical stages of plant growth. Volumetric water content is calculated indirectly by soil moisture sensor using different parameters of the soil like dielectric constant, electrical resistance and neutrons interaction. Controlling and monitoring of the temperature around the plants is essential for maintaining their healthy growth. The system turns on the fans for cooling the plants if the temperature and/or humidity increases above a certain preset value. The threshold (preset) temperature and/or humidity is chosen depending on the requirements of the plants so that they are not affected by any excess heat which may affect their metabolism and healthy growth. Similarly, the system has a light sensor that detects the level of lighting in the plant surroundings. The sensor is a light dependent resistor (LDR) used for in light/dark sensor circuits. Normally, the resistance of an LDR is very high, can be in the order of Mega ohms ( $M\Omega$ ), but when it is illuminated with light, its resistance drops dramatically allowing more current to

pass through it. Figure 4 shows the soil moisture sensor, the temperature/ humidity composite sensor (DHT11 sensor) and the LDR light sensor components used for this project. The system is flexible in that the preset moisture level, temperature, humidity or light levels could be modified if needed by the user depending on the requirements. By changing the Arduino code, any desired level could be programmed by the user.

### 2.1.4. Arduino Compatible L298N Stepper Motor Driver

This 2x2A DC Motor drive for Arduino allows driving two channel DC motors. It uses a L298N chip, which deliveries output current up to 2A for each channel. The speed control is achieved through conventional Pulse Width Modulation (PWM) which can be obtained from Arduino's PWM output Pin 5 and 6. The enable/disable function of the motor control is signaled by Arduino Digital Pins 4 and 7. The Motor shield can be powered directly from Arduino or from external power source. It is strongly encouraged to use external power supply to power the motor shield. The Arduino Motor Shield as shown in Figure 5 is based on the L298, which is a dual full-bridge driver designed to drive inductive loads such as relays, solenoids, DC and stepping motors. In this project, this component is being used because it lets to drive the water pump motors connected to the Arduino board to control the speed and direction of each one independently.

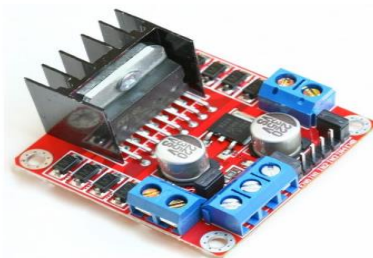


Figure 5: Arduino Compatible L298N motor drivers

## 2.2. Complete overall circuit

The system contains the following components: the microcontroller which is the Arduino Uno, GSM Module, the temperature and humidity sensor which is the DHT11, soil moisture sensors, a L298N motor driver, a 5V relay, water pumps, fans, LEDs, light sensor and rechargeable batteries as shown in the overall circuit in Figure 6. The micro controller used is the Arduino Uno based on Atmega328P, it has 14 electronic input/output pins, powered by a battery and connected to the GSM Module, sensors and hardware devices compatible with each sensors. All the components are tightly fixed on a wooden base and surrounded with wood and partly glass from the sides to allow light in the day time so that the plants get the amount of light it needs in the day time as seen in the prototype in [4]. When it is dark the LDR will give the readings to the Arduino Uno then the Arduino sends this information to the user through the GSM Module. Then the user will send back a message to the GSM that turns ON the lights. There are also four holes on the surrounding wood to fix the four fans. The fans will work when the user finds that the plants need cooling depending on the environmental conditions.

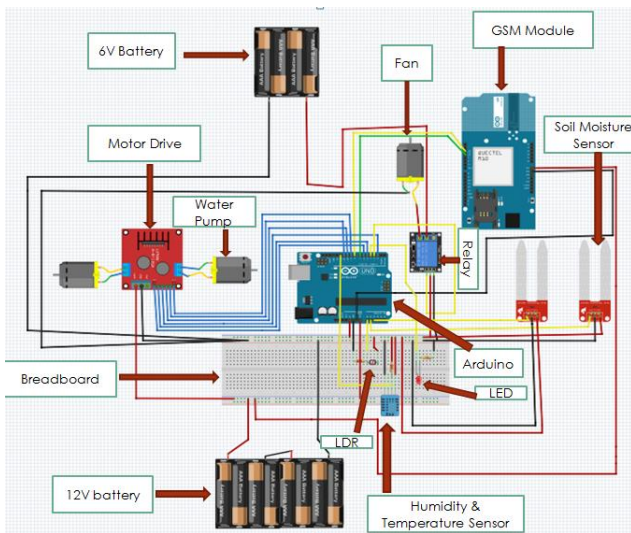


Figure 6: Complete Circuit Connections

### 3. Software Implementation

The Arduino Uno was programmed for preset values of moisture, temperature, humidity and light environmental conditions that can be simply changed in the program (Arduino software). For the testing purposes, when temperature is equal to or above 27 degrees and the humidity is above 70 %, the user will send a command message to turn on the fans and will stop it, if temperature is less than 27 degrees and the humidity is less than 70 %. These temperature and humidity threshold values can be easily changed in the Arduino code as per the user’s interest. The DH11 composite humidity-temperature sensor will be responsible to give the reading to the system. The water pumps work with the user’s permission. The speed of the water pump that supplies water to the land depends on the moisture level of the land. For the experiment, the moisture levels are divided into four levels that are **high** ( $65 \% \leq \text{moisture level} \leq 100 \%$ ), **medium** ( $50 \% \leq \text{moisture level} < 65 \%$ ), **medium-low** ( $25 \% \leq \text{moisture level} < 50 \%$ ) and **very low** ( $0 \% \leq \text{moisture level} < 25 \%$ ). This threshold values can be easily changed in the program. When the soil moisture is high the water pump is OFF. When the soil moisture is medium the speed of the water pump is medium. When the soil moisture is medium-low the water pump is in medium-low speed. However, when the soil moisture is very low the water pump speed is high so that it pumps high quantity of water to the plants.

For simplicity, we call the text message sent by the user to be “switching states”. This switching states provide information about the status of the two motor pumps, the lighting bulb (LED for the experiment) and the fan status. The set of the switching states consists of MOTORION, MOTORIOFF, MOTOR2ON, MOTOR2OFF, FANON, FANOFF, LIGHTON and LIGHTOFF. The user sends one of the mentioned switching states to turn ON/OFF, the motor pumps, the fan or the light in the farm. The other text message sent by the user is the STATUS message to get the environmental condition (moisture level, temperature, humidity and lighting conditions) of the farm or the greenhouse so that the user is aware of the status on the site and then sends the necessary commands to control the water pumps, the light and the fans accordingly. The overall working of the program that is

uploaded to the Arduino can be described in a flow-chart as shown in Figure 7.

After declaration of the variables and initializations, in the main program loop, the Arduino checks if the message is sent from a specific user before executing the SMS command from the sender. This will make the system secure so that only a specific user can control the system.

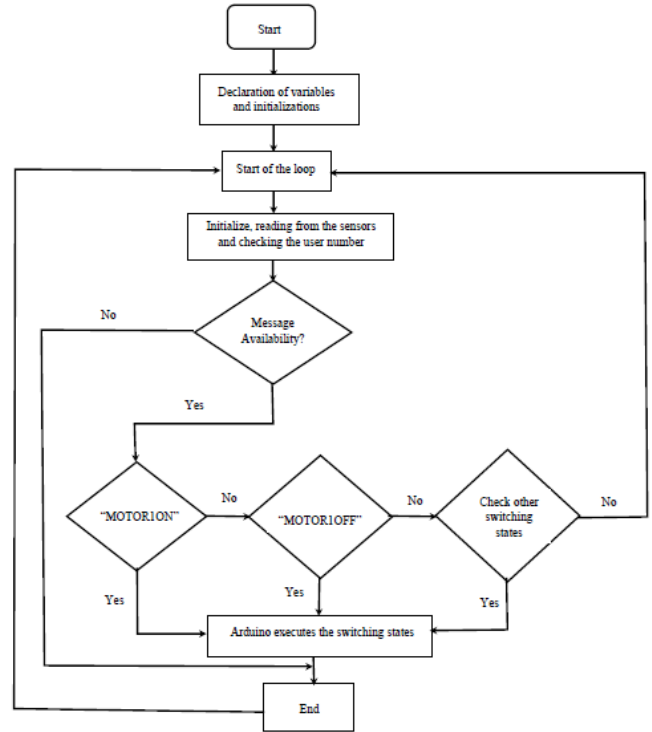


Figure 7: Flow chart of the Arduino program

A sample of the Arduino program (Algorithm 1) to switch on the motor for water pump 1 with high speed is given below. As per the implemented Arduino code, a moisture level between 0% and 25% is considered relatively dry and thus the water pump will run at high speed. A similar code is constructed to check the other three different moisture levels and the water pump will run with a speed corresponding to the proper range.

**Algorithm 1:** Arduino code to switch on the motor for water pump 1.

```

output_value1 = map(output_value1,1023,0,0,100); // Get the
reading value of moisture sensor 1 and map the value in the
range [0 100] where 100 means the maximum wet condition
if(SIM900.find("+971xxxxxxxx")>=0){ // Check whether the
message is sent from a specific mobile user
if(SIM900.available(>)>0) // Display any text that the GSM
shield sends out to the serial monitor
{
textMessage = SIM900.readString(); //Get the character
from the cellular serial port
Serial.print(textMessage); //Print the incoming character to
the terminal
if(textMessage.indexOf("MOTORION")>=0){ // Turn on
motor 1 and save current state

```



```

if(output_value1<25) //if the moisture level is less than 25%,
run the water pump at high speed
{
digitalWrite(in1, LOW);
digitalWrite(in2, HIGH);
analogWrite(enA, 255);
message = "Motor 1 is at high speed";
}
}
sendmsg(message); // send the message " Motor 1 is at high
speed" to the user
}}

```

The message “Motor 1 is at high speed” will be sent to the farmer or the user to let him/her know that the water pump is switched on and it is running at high speed due to very low moisture level and the high need of water for the plants. The sample of the Arduino code (Algorithm 2) for this subroutine for sending the message is given below.

**Algorithm 2:** Arduino code for sending message to the user confirming that Motor 1 is at high speed.

```

void sendmsg(String message)
{
SIM900.println("AT+CMGF=1"); //Set the GSM Module in
Text Mode
delay(100);
SIM900.println("AT+CMGS=\"+971xxxxxxxxx\"); // The
farmer or user mobile number to whom the message to be sent
delay(100);
SIM900.println(message); // The SMS text to be sent to the
user
delay(100);
SIM900.println((char)26); // ASCII code of CTRL+Z
delay(100);
SIM900.println();
delay(1000);
}

```

In a similar way, the subroutines corresponding to the rest of the switching states (MOTOR1OFF, MOTOR2ON, MOTOR2OFF, FANON, FANOFF, LIGHTON and LIGHTOFF) SMS command messages are incorporated in the overall program. The corresponding confirmation that provides switching status information (ON or OFF) will also be sent to the farmer or the user regarding the status of the water pumps, the fan and the light bulb. The user has also the option to obtain the environmental condition of the farm or the green house by sending the command “STATUS”. When the Arduino gets this message through the GSM module, the microcontroller collects all the sensors information (light intensity, temperature, humidity and moisture levels in the two sensors) and sends this information as environmental STATUS to the user. The sample of the Arduino code (Algorithm 3) for this subroutine is given below.

**Algorithm 3:** Arduino code used for sending STATUS information to the user.

```

Void sendSMS(){ // AT command to set SIM900 to SMS mode
SIM900.println("AT+CMGF=1"); //Set the GSM Module in
Text Mode
delay(100);

```

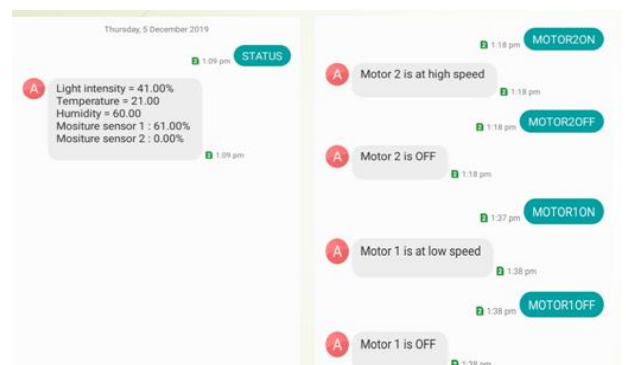
```

SIM900.println("AT+CMGS=\"+971xxxxxxxxx\"); //The
user mobile number that gets the STATUS information
delay(100);
SIM900.print("Light intensity = "); // Send the SMS
SIM900.print(val);
SIM900.println("%");
delay(100);
SIM900.print("Temperature = ");
SIM900.println(t);
delay(100);
SIM900.print("Humidity = ");
SIM900.println(h);
delay(100);
SIM900.print("Moisture sensor 1 : ");
SIM900.print(output_value1);
SIM900.println("%");
delay(100);
SIM900.print("Moisture sensor 2 : ");
SIM900.print(output_value2);
SIM900.println("%");
delay(100);
SIM900.println((char)26); // ASCII code of CTRL+Z
delay(100);
SIM900.println(); // Give module time to send SMS
delay(1000);
}

```

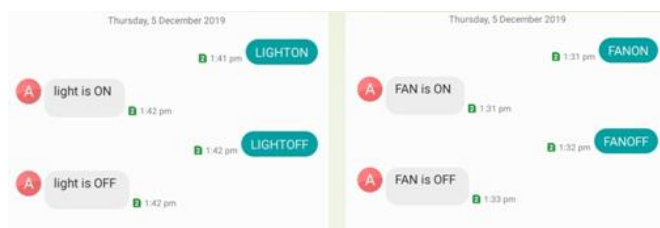
#### 4. Experimental Test, Results & Discussions

As discussed in the above sections, an automatic Arduino and GSM based control system for garden plantations or a greenhouse was designed, built and tested in the laboratory and it was found that the system works properly. Figure 8 shows a screenshot sample results of communication between the mobile user and the designed Arduino based system to show how remote controlling is achieved. As shown Figure 8 (a), the owner of the greenhouse sent the message “STATUS” to the system and the owner received the whole status of the greenhouse. This message is helpful because it lets the owner to decide what has to be switched ON or OFF. There are two moisture sensors in this project. Since the moisture sensor 2 has a 0 % moisture, the owner of the green house has decided to switch ON Motor 2. As a result, Motor 2 pumps the water at high speed since the soil level of sensor 2 is 0 %. Similarly, this goes for soil moisture sensor 1. Since the soil moisture level is at 61%, Motor 1 pumps the water at low speed depending on the level of the soil. Depending on the status received, the user can turn ON or OFF the lights, the fans when needed as shown in Figure 8 (b).



(a) Getting the STATUS information and controlling the water pumps.





(b) Controlling the LIGHT and the FAN in the green house

Figure 8. Sample communication between the mobile user and the designed Arduino & GSM based system

## 5. Conclusions

An Arduino and GSM based electronic system was designed and implemented for monitoring and controlling the level of soil water, temperature, humidity and light intensity for the plants in a garden plantations or greenhouse system. These environmental factors affect the proper and healthy growth of plants. The system is designed and implemented with the Arduino being the core microcontroller interfaced with the GSM module and the sensors providing input information to the Arduino. The system was tested in Laboratory and found working properly. The system can be easily implemented with low-cost budget to make it affordable for everyone as Arduino is a cheap microcontroller with open source software availability. This system can be integrated in the agricultural sector as GSM based irrigation system since mobile phone devices are nowadays used almost everywhere in the world. It can help the farmers to easily remotely control the site and thus save human labour by eliminating the physical presence of persons on the agricultural site. Moreover, it is beneficial to properly manage, conserve water and get increased productivity in agriculture and in the end adding value to the economic growth of a country. By implementing this system not only agricultural farms but also parks, gardens, golf courses can be irrigated. The system benefits the environment positively as it saves water the scarce natural resource available for human kind.

## Acknowledgment

The authors appreciate the support for the research provided by the American University of Ras Al Khaimah (AURAK).

## References

- [1] B.N. Getu et al., "Remote Controlling of an Agricultural Pump System Based on the Dual Tone Multi-Frequency (DTMF) Technique," *Journal of Engineering Science and Technology*, 10 (10), 1261 – 1274, 2015. [http://jestec.taylors.edu.my/Vol%2010%20Issue%2010%20October%202015/Volume%20\(10\)%20Issue%20\(10\)%201261-1274.pdf](http://jestec.taylors.edu.my/Vol%2010%20Issue%2010%20October%202015/Volume%20(10)%20Issue%20(10)%201261-1274.pdf).
- [2] B.N. Getu et al., "Automatic water level sensor and controller system" *International Conference on Electronic Devices Systems and Applications (ICEDSA)*, 1-4, 2016. DOI: 10.1109/ICEDSA.2016.7818550.
- [3] B.N. Getu et al., "Automatic Control of Agricultural Pumps Based on Soil Moisture Sensing. Proceedings of the IEEE AFRICON 2015 Conference, 667-671, 2015. DOI: 10.1109/AFRCON.2015.7332052.
- [4] B.N. Getu et al., "Design of Automatic Control System for Garden Plantations Based on Soil Moisture, Temperature and Light Sensing" *Journal of Engineering and Applied Sciences*, 12 (18), 4782-4786, 2017. DOI: 10.36478/jeasci.2017.4782.4786.
- [5] B.C. Lailhacar, M.D. Dukes, G.L. Miller, "Sensor-Based Control of Irrigation in Bermuda grass" *ASAE Annual International Meeting, Tampa, Florida*, 1-14, 2005. [https://abe.ufl.edu/faculty/mdukes/pdf/publications/SMS/ASAE\\_05-2180.pdf](https://abe.ufl.edu/faculty/mdukes/pdf/publications/SMS/ASAE_05-2180.pdf).

- [6] M. Nagarajapandian, "Automatic Irrigation System On Sensing Soil Moisture Content" *International Journal of innovative research in electrical, electronics, instrumentation and control engineering*, 3(1), 96-98, 2015. DOI 10.17148/IJIREEICE.2015.3120.
- [7] S. Rakshak, R.W. Deshpande, "Automated Irrigation System Based on Arduino Controller Using Sensors" *International Journal of Innovative Research in Computer and Communication Engineering*, 5 (7), 13394-13400, 2017. DOI: 10.15680/IJIRCCE.2017.0507084.
- [8] K.V.S. Sairam, N. Gunasekaran S.R. Redd, "Bluetooth in wireless communication" *IEEE Communications Magazine*, 40 (6), 90-96, 2002. DOI: 10.1109/MCOM.2002.1007414.
- [9] J. Uddin et al., "Automated Irrigation System Using Solar Power". 7<sup>th</sup> International conference on electrical and computer engineering, 228-231, 2012. DOI: 10.1109/ICECE.2012.6471527.
- [10] M. Ojha et al. "Microcontroller Based Plant Irrigation" *International Journal of Advanced Technology in Engineering and Science*, 3 (10), 71-79, 2015. [http://ijates.com/images/short\\_pdf/1446208511\\_P71-79.pdf](http://ijates.com/images/short_pdf/1446208511_P71-79.pdf).
- [11] L. Louis, "Working Principle of Arduino and using it as a tool for Study and Research" *International Journal of Control, Automation, Communication and Systems (IJACS)*, 1(2), 21-29, 2016. DOI: 10.5121/ijcacs.2016.1203
- [12] Arduino Board, 2019: <https://www.arduino.cc/en/main/arduinoBoardUno>.
- [13] SIM900, 2019. <https://simcom.ee/modules/gsm-gprs/sim900/>.

## Decision Making System for Improving Firewall Rule Anomaly Based on Evidence and Behavior

Suchart Khummanee\*, Phatthanaphong Chomphuwiset, Potchara Pruksasri

Department of Computer Science, Faculty of Informatics, Mahasarakham University, 44150, Thailand

---

### ARTICLE INFO

Article history:

Received: 27 June, 2020

Accepted: 23 July, 2020

Online: 09 August, 2020

---

Keywords:

Firewall rule anomaly

Decision support system

Firewall rule analysis

Probability of firewall rules

---

### ABSTRACT

Firewalls are controlled by rules which often incur anomalies. The anomalies are considered serious problems that administrators do not desire to happen over their firewalls because they cause more vulnerabilities and decrease the overall performance of the firewall. Resolving anomaly rules that have already occurred on the firewall is difficult and mainly depends on the firewall administrator's discretion. In this paper, a model is designed and developed to assist administrators to make effective decisions for optimizing anomaly rules using the probability approach (Bayesian). In this model, the firewall needs to add four property fields (Extra fields) to the firewall rules: frequency of packets matching against rules, evidence of creating rules, the expertise of rules creator and protocol priority. These fields are used to calculate the probability of each firewall rule. The probability for each rule is used while the rules conflict and administrators need to resolve them. The rule having the highest probability value indicates that it has the highest priority in consideration. Experimental results show that the proposed model allows firewall administrators to make significant decisions about solving anomaly rules. The data structure of this model is based on  $k$ -ary tree, therefore the speed of building tree, time complexity and space complexity:  $O(n)$ ,  $O(\log_m n)$  and  $O(m*n)$  respectively. Besides, the confidence of the proposed firewall for resolving firewall rule anomalies of the administrator increase by 29.6% against the traditional firewall, and the reliability value between the inter-raters also increase by 13.1%.

---

### 1. Introduction

A firewall is an indispensable system for today's computer networks. It plays an important role to prevent access to various resources on the networks, for example, networking devices, databases, web servers and etc. Besides, it can also prevent attacks and intrusions by malicious users from dark side networks. Basically, the firewall is commanded by a set of instructions, called the rule. The number of firewall rules depends on the complexity of the policies of each organization. As the number of rules increases, the number of anomaly rules is also enlarged. The anomalies arise from two or more rules overlapping but having different decisions. There are five types defined by [1]: the shadowing, correlation, generalization, redundancy, and irrelevancy anomaly. Recently, a new definition of anomalies has been advanced, which is the semantics loss of rules [2]. The rule anomalies have a great impact on the overall performance of a firewall. That is, they decrease the processing speed of rule

verifying or matching. Therefore, reducing the number of anomalies improves the speed of the rule verification as well. Analyzing, managing and resolving anomalies are big problems of firewall researches and attract much interest. The first major researcher about the anomalies was Al-Shaer, who presented five types of anomaly rules and also invented an effective method for detecting anomalies, called the finite state diagram [3]. Later, several researchers contributed various methods for detecting anomalies, for example, in [4] the authors revealed a powerful algorithm to relieve the root cause of anomalies, called SDD. However, the rules are only one type either "an acceptable status" or "an unacceptable status" appearing on the firewall. Next, in [5] they proposed a method to resolve the anomaly problems effectively by using the firewall decision diagram (FDD) and this is also the prototype of much other later research. The propositional logic model was presented by [6], they claimed that their model could remove anomalies that lead to a decrease in the number of rules without changing the policy. Next, in [7] authors demonstrated the anomaly management framework which encourages systematic detection and resolution of firewall policy

\*Corresponding Author: Suchart Khummanee, Department of Computer Science, Faculty of Informatics, Mahasarakham University, Thailand, Email: suchart.k@msu.ac.th

anomalies based on the average risk values, called FAME. The risk values are calculated from the Common Vulnerability Scoring System (CVSS) [8] which does not consider attacking vulnerabilities in the overview, but it is considered only the point. The next proposed solution for analyzing and managing the firewall policies was Lumeta [9] and Fang [10]. They represented the tools that are used to analyze firewall rules, but tools cannot completely verify the misconfigurations policy settings. In [11] the authors contributed an algorithm for detecting and resolving conflicts in packet filtering. Though, the algorithm can only detect certain specific conflicts. The firewall rule optimization based on Service-Grouping was proposed by [12]. The basis of this technique is resolving the conflicts segment by grouping rules from work behaviors. They claimed that the processing time and number of packet hits are better than the traditional and FIREMAN [13] firewall. FPQE [14] is an automated system to resolve rule anomalies, which does not require any admin intervention. It uses an automatic rule removal in the case of redundancy and contradiction anomaly, and uses an automatic rule permutation against shadowing and correlation. Besides, some techniques allow the firewall to automatically detect and analyze conflict rules such as [15] and [16], but they are not based on real tangible evidence. By most methods, the burden of resolving rule conflicts is often given to the administrator's discretion instead.

This paper contributes the model for optimizing firewall rule anomalies by applying the probability (Bayesian) together with evidence of each rule, i.e., frequency of packets matching against rules, evidence of creating rules, the expertise of rules and protocol priority. This model provides guidance to firewall administrators to resolve rule anomalies with probability values. In order for administrators to be confident that the decision is more accurate based on the actual evidence. This paper is organized as follows: Section 2 overviews the background and related work. Section 3 presents the key contributions. In Section 4 articulates our system design. Section 5 addresses the implementation of details and evaluations. Section 6 concludes this paper.

## 2. Background and Related Works

### 2.1. Rule Definition and Anomaly

Generally, the firewall rule consists of two parts: the condition and decision parts. Let  $R$  be a firewall rule,  $C$  as a condition part and  $A$  is a decision part, a firewall rule format:

$$R = C \rightarrow A \quad (1)$$

In fact, firewalls always have more than a single rule. Therefore, the first equation (1) needs to be revised to the second equation:

$$R_i = C_i \rightarrow A_i \quad (2)$$

where  $C_i$  and  $A_i$  are the condition and decision of rule  $R_i$  (Any firewall rule) by  $i \in [1, n]$ , and  $n$  is a non-negative integer. Given  $f_i$  representing the domain of positive integers is a finite range, denoted  $D(f_i)$ . For example, the domain of the source and destination address in an IP packet is  $[0, 2^{32} - 1]$  ( $D(f_1)$  and  $D(f_2)$ ), source and destination port is  $[0, 2^{16} - 1]$  ( $D(f_3)$  and  $D(f_4)$ ) and protocol is  $[0, 2^8 - 1]$  ( $D(f_5)$ ).  $C_i$  defines a set of packet fields over the fields  $f_1$  through  $f_d$  specified as  $f_1 \in F_1 \wedge f_2 \in F_2 \wedge \dots \wedge f_d \in F_d$  where  $F_i$  is a subset of  $D(f_i)$ .  $A_i$  is either *accept* or *deny* for each rule. If all conditions ( $f_i$ ) in  $C_i$  are true, the decision is either *accept* or *deny* depending on the specified administrators like:

$$R_1 = (f_1 \wedge f_2 \wedge f_3 \wedge \dots \wedge f_d)_1 \rightarrow \text{accept}_1 | \text{deny}_1$$

Given  $P_i$  as an IP packet over the  $d$  fields  $f_1, \dots, f_d$ ,  $P_i$  is a tuple of  $d(p_1, p_2, \dots, p_d)$  where each  $p_i (1 \leq i \leq d)$  is an element of  $D(f_i)$ . An IP packet  $(p_1, p_2, \dots, p_d)$  matches  $R_i$  if and only if the condition  $p_1 \in f_1 \wedge p_2 \in f_2 \wedge \dots \wedge p_d \in f_d$ . A set of rules  $(R_1, \dots, R_i)$  is valid when there is at least one rule in the set matching against  $p_i$ . To make sure that firewall rules are working properly, the condition of the final rule in the firewall is usually specified as  $f_1 \in D(f_1) \wedge \dots \wedge f_d \in D(f_d)$ , where every packet must be matched as shown in  $R_3$ , called the implicit rule. The set of rules below shows an example of three rules over the three fields of condition  $C(f_1, f_2, f_3)_i$  where  $D(f_1) = D(f_2) \in [1, 100]$  and  $D(f_3) \in [1, 50]$ .

$$R_1 = (f_1 \in [25, 50] \wedge f_2 \in [40, 60] \wedge f_3 \in [5, 25]) \rightarrow \text{accept}$$

$$R_2 = (f_1 \in [35, 70] \wedge f_2 \in [30, 90] \wedge f_3 \in [10, 25]) \rightarrow \text{accept}$$

$$R_3 = (f_1 \in [1, 100] \wedge f_2 \in [1, 100] \wedge f_3 \in [1, 50]) \rightarrow \text{deny}$$

$R_1$  and  $R_2$  are redundant because any packet can match both rules which have the same actions (*accept*). Furthermore,  $R_1$  and  $R_2$  also conflict with  $R_3$  because both  $R_1$  and  $R_2$  are subsets of  $R_3$  while they are different actions. One typical solution to resolve such conflicts, that is, firewalls choose a rule which matches with the packet  $p_i$  being considered first, called the first-match approach. The firewall rules anomalies can be classified into six [1], [2] types representing each anomaly by the theorems:

**Shadow anomaly:**  $R_x$  is shadowed by  $R_y$ , if and only if their intersection is equal to  $R_x$  and there are different actions illustrated in Figure 1(a).

$$R_x = C_x \rightarrow A_x, R_y = C_y \rightarrow A_y$$

$$R_x, R_y \in R_{db} \wedge \neg(A_x \leftrightarrow A_y) \wedge (C_x \cap C_y = C_x) \quad (3)$$

where  $R_{db}$  is a database of all rules, and  $R_y$  is the rule executed before  $R_x$ .

**Correlation anomaly:**  $R_x$  and  $R_y$  in  $R_{db}$  are correlated if their intersection is not equal to  $\emptyset$ ,  $R_x - R_y \neq \emptyset$ ,  $R_y - R_x \neq \emptyset$ , and they have different actions represented in Figure 1(b).

$$R_x = C_x \rightarrow A_x, R_y = C_y \rightarrow A_y$$

$$R_x, R_y \in R_{db} \wedge \neg(A_x \leftrightarrow A_y) \wedge (C_x \cap C_y \neq \emptyset) \wedge (C_x - C_y \neq \emptyset) \wedge (C_y \cap C_x \neq \emptyset) \quad (4)$$

where  $R_{db}$  is a database of all rules, and  $R_y$  is the rule executed before  $R_x$ .

**Generalization anomaly:**  $R_x$  is generalized by  $R_y$ , if and only if their intersection is equal to  $R_y$ , and there are different actions (Figure 1(c)), where  $R_y$  is the rule matched before  $R_x$ .

$$R_x = C_x \rightarrow A_x, R_y = C_y \rightarrow A_y$$

$$R_x, R_y \in R_{db} \wedge \neg(A_x \leftrightarrow A_y) \wedge (C_x \cap C_y \neq C_y) \quad (5)$$

**Redundancy anomaly:**  $R_x$  is redundant to  $R_y$ , if and only if their intersection is not equal to  $\emptyset$ , and they are have same actions (Figure 1(d)).

$$R_x = C_x \rightarrow A_x, R_y = C_y \rightarrow A_y$$

$$R_x, R_y \in R_{db} \wedge (A_x \leftrightarrow A_y) \wedge (C_x \cap C_y \neq \emptyset) \quad (6)$$

**Irrelevance anomaly:** Irrelevance anomaly occurs in the firewall if no packets can be matched against all rules in the firewall. This anomaly is caused by the administrator's misunderstanding about the network connections.

$$R_x = C_x \rightarrow A_x, R_y = C_y \rightarrow A_y, R_n = C_n \rightarrow A_n$$

$$R_1, R_2, \dots, R_n \in R_{db} \wedge P_i \not\subset (C_i, \dots, C_n) \quad (7)$$

where  $P_i$  is an IP packet executed by the firewall.

**Semantics loss anomaly:** The semantics loss represented by [2]

occurs when  $R_x$  and  $R_y$  are merged to  $R_z$  by the meaning of both old rules that have been changed or replaced by a new meaning. This anomaly is mostly caused by redundant rules as shown in Figure 1(f).

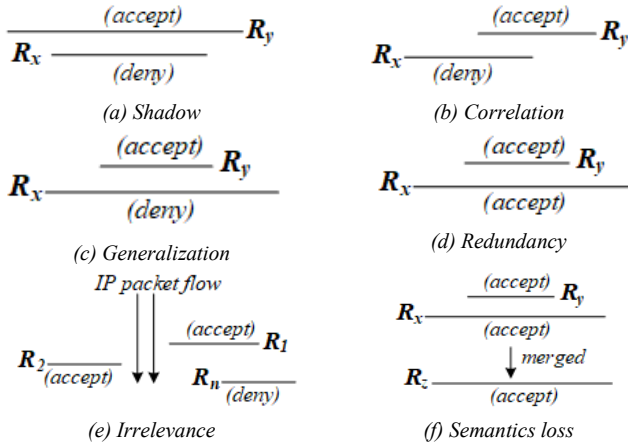


Figure 1: Firewall rule anomalies

Min-Max feature scaling [17] (also known as data normalization) is the standard method used to adjust the range of data. Since the range of data values may be very different, it is therefore a necessary step in data preprocessing before processing in the next step. It is normally used to resize any data range into the range  $[0, 1]$ , called unity-based normalization. Also, it can normalize the finite range of values in the dataset between any arbitrary points  $t_{max}$  and  $t_{min}$  as the following equation.

$$m' = \frac{m - r_{min}}{r_{max} - r_{min}} \times t_{max} - t_{min} + t_{min} \quad (8)$$

Let  $m'$  denotes the value being considered has been normalized by  $m \in [r_{min}, r_{max}]$ .  $r_{min}$  and  $r_{max}$  denote the minimum and maximum of the measurement range.  $t_{min}$  and  $t_{max}$  are the minimum and maximum of the target range to be scaled.

### 2.3. Bayes' Theorem

Let  $A_1, A_2, \dots, A_k$  be events that partition the sample space  $S$ , i.e.,  $S = A_1 \cup A_2 \cup \dots \cup A_k$  and  $A_i \cap A_j = \emptyset$  when  $i \neq j$  and let  $B$  be an event on that space for which  $P_r(B) > 0$ . Then Bayes' theorem is:

$$P_r(A_j|B) = \frac{P_r(A_j)P_r(B|A_j)}{\sum_{j=1}^k P_r(A_j)P_r(B|A_j)} \quad (9)$$

### 2.4. Moving Average (MA)

$$SMA = \frac{A_1 + A_2 + A_3 + \dots + A_n}{n} \quad (10)$$

where  $A_i$  is an average in period  $n$ , and  $n$  is the number of periods. EMA is a weighted average of the last  $n$  data, where the weighting decreases exponentially with each previous data per period. In other words, the formula gives greater weight to more recent data. The formula for the exponential moving average is the following.

$$EMA_t = [V_t \times \frac{s}{1+d}] + EMA_y \times [1 - \frac{s}{1+d}] \quad (11)$$

where  $EMA_t$  = EMA today,  $V_t$  = Value today,  $EMA_y$  = EMA yesterday,  $s$  = smoothing,  $d$  is the number of day.

### 2.5. Converting an IP Address to a non-negative Integer

The Internet Protocol Address (as known as IP Address) is a unique address that networking devices such as routers, switching, and computers use to identify themselves and communicate over other devices in computer networks. An IPv4 address (IP version 4) is equal to 32 bits, ranging from 0 to  $2^{32} - 1$  address space. It is usually divided into four parts, each part (8 bits = an octet) separated by a dot, e.g.,  $A_1.A_2.A_3.A_4$  where  $A_{(1-4)} \in [0, 255]$ . IPv4 address can be converted to any non-negative integer with the following equation.

$$IPv4' = A_1 \times 2^{24} + A_2 \times 2^{16} + A_3 \times 2^8 + A_4 \times 2^0 \quad (12)$$

where  $IPv4'$  is a new IP address to be converted, for example, 1.2.3.4 will be convert to:

$$IPv4' = 1 \times 2^{24} + 2 \times 2^{16} + 3 \times 2^8 + 4 \times 2^0 = 16,909,060$$

### 2.6. Arithmetic Mean and Kappa Statistics

We use the average method ( $\bar{x}$ ) to evaluate the administrator's satisfaction with the proposed firewall and use the Cohen's kappa coefficient ( $\widehat{K}_F$ ) [19] to measure the interrater reliability as the following equations.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (13)$$

where  $\bar{x}$  is an average (or arithmetic mean),  $n$  is the number of terms (e.g., the number of items or numbers being averaged), and  $x_i$  is the value of each individual item in the list of numbers being averaged.

$$\widehat{K}_F = \frac{\bar{P}_a - \bar{P}_e}{1 - \bar{P}_e} \quad (14)$$

where:  $\bar{P}_a$  denotes the relative observed agreement among raters,  $\bar{P}_e$  denotes the hypothetical probability of chance agreement as in the equation 15 and 16.

$$\bar{P}_a = \frac{1}{r} \sum_{i=1}^r Z_j, \bar{P}_e = \sum_{k=1}^q P_j^2 \quad (15)$$

$$Z_j = \frac{1}{m(m-1)} (\sum_{k=1}^q n_{kj}^2 - \sum_{k=1}^q n_{ij}), P_j = \frac{1}{rm} \sum_{i=1}^r n_{ij} \quad (16)$$

where  $m$  is the number of raters, the objective of the assessment is  $r$ ,  $q$  is the type of information that needs to be evaluated (e.g., most satisfied, very satisfied, ..., least satisfied), and  $n_{ij}$  denotes the observed cell frequencies.

## 3. Key Contributions

As rule anomalies occur over firewalls, the decision-making power to resolve the anomalies mainly depends on the administrator's discretion. However, the decisions made often result in errors or loopholes over the existing rules, if admins



cannot entirely understand the relationship between conflict rules. Therefore, it is necessary to develop a decision support system for admins to assist decision-making during real-time anomaly detection. The system consists of four procedures:

- 1) Firstly, preparing various information to be ready before processing,
- 2) Analyzing and detecting the rule abnormalities by the Path Selection Tree (PST),
- 3) Calculating the probability (Bayesian) of each rule based on the frequency of packets matched against rules, evidence of creating rules, expertise on creating rules, and protocol priority to help admins decide before optimizing the rules,
- 4) Lastly, optimizing anomaly or conflict rules based on the probability.

#### 4. The System Design

There are four steps in the system design as shown in Figure 2.

##### 4.1. Preparing Rule-Based Firewall (Step 1)

**The conditions ( $C_i$ ) and decision ( $A_i$ ) of each rule:** Referring to  $R_i$  in the equation (2), in general, the members of  $C_i$  have five fields ( $f_1 \wedge \dots \wedge f_5$ ), where  $f_1$  = source IP address (SIP),  $f_2$  = destination IP address (DIP),  $f_3$  = source port (SP),  $f_4$  = destination port (DP) and  $f_5$  = protocol (PRO) respectively as shown in Table 1. According to  $R_1$  of Table 1, the preparation process of firewall rules begins with converting the IP addresses of  $f_1$  and  $f_2$  into a range of positive integers by equation (12). Hence,  $f_1$  and  $f_2$  are then converted into the following numbers:  $f_1 \in [16909056, 16909066]$  and  $f_2 \in [1, 256]$ . The fields converted in the next order are  $f_3$  and  $f_4$  which contain the numbers ranging from 0 to  $2^{16} - 1$ :  $f_3 \in [0, 65535]$  and  $f_4 \in [80, 80]$ , where \* means all numbers in such domain. The field  $f_5$  is both TCP and UDP protocol, thus they are translated to:  $f_5 \in \{6, 17\}$ , where TCP = 6 and UDP = 17. In case of decision field ( $A_i$ ), it is changed to a positive integer either 0 or 1, such as  $A_i \in \{0, 1\}$ , where *accept* = 1 and *deny* = 0. As a result of all these calculations,  $R_1$  is converted to:

$$R_1: (f_1 \in [16909056, 16909066] \wedge f_2 \in [1, 256] \wedge f_3 \in [0, 65535] \wedge f_4 \in [80, 80] \wedge f_5 \in \{6, 17\}) \rightarrow 1$$

Table 1: The basic member fields of  $C_i$  and  $A_i$

$R_i$	$C_i(f_1 \wedge f_2 \wedge f_3 \wedge f_4 \wedge f_5)$					$A_i$
$R_i$	$f_1(SIP)$	$f_2(DIP)$	$f_3(SP)$	$f_4(DP)$	$f_5(PRO)$	decision
$R_1$	1.2.3.0-1.2.3.10	0.0.0.1-0.0.1.0	*	80	TCP, UDP	accept

**Calculating Probability of Extra Fields of Each rule:** To determine the probability of each rule in this model, there are four additional fields added including the frequency of packets matching against rules (FPM), evidence of creating rules (ECR), the expertise of rules creator (ERC) and protocol priority (PRI). Let  $P(e_1)$ ,  $P(e_2)$ ,  $P(e_3)$  and  $P(e_4)$  are the probability of FPM, ECR, ERC, and PRI respectively. Therefore, the sum of the probability of rule  $R_i$  is equal to the equation (17).

$$P(E_i) = \sum_{j=1}^4 P(e_j) \tag{17}$$

Where  $P(E_i)$  is the probability of  $R_i$ . For example, the information of extra fields of  $R_i$  as shown in Table 2. From  $R_1$  in

Table 2., matching rate (FPM :  $e_1$ ) between packets and  $R_1$  is equal to 2,125 times.  $e_2$ (ECR),  $e_3$ (ERC) and  $e_4$ (PRI) are 3, 2 and 4 respectively, explained more details in the next section. These four extra fields are calculated to be the probability pattern in which data is in the range from 0.0 to 1.0 ( $R_1$ ' by the Min-Max Feature Scaling in the equation (8).

**In the  $e_1$  case:** it is the frequency of the packets matching against any rules over the firewall, the counting process starts from the time at which a rule had been created and continues until the present time. For example, if the maximum and minimum number of matching of any rules in the firewall are 5,000 and 1,200 times respectively, then  $e_1'$  here is equal to:

$$e_1' = \frac{m - r_{min}}{r_{max} - r_{min}} \times t_{max} - t_{min} + t_{min}$$

$$= \frac{2,125 - 1,200}{5,000 - 1,200} \times (1.0 - 0.0) + 0.0 = 0.24$$

where  $m = 2,125$ ,  $r_{min} = 1,200$ ,  $r_{max} = 5,000$ ,  $t_{min} = 0.0$  and  $t_{max}$  is equal to 1.0. However, recording  $e_1$  in the firewall requires the equation (10 and 11) to make the data smoother since the recorded data may be a swinging data caused by network attacks, user behaviors, or network usage during rush hours, etc. The period for calculating data with EMA method will depend on the suitability of each organization. For this research,  $e_1$  is recorded every hour per day as in the following example:

Given  $e_1$  of  $R_1$  for each hour per a day: 1300, 1500, 1200, 1300, 1400, 1500, 1800, 4500, 6000, 6300, 5500, 1000, 2400, 2800, 2600, 2600, 2400, 1900, 1500, 1200, 1000, 800, 700, 600 times, then it can calculate the EMA of  $e_1$  using five hours in the past as follows.

$$SMA \text{ of } 5^{\text{th}} \text{ hour} = \frac{(1,300+1,500+1,200+1,300+1,400)}{5} = 1,340.0$$

$$EMA \text{ of } 6^{\text{th}} \text{ hour} = 1,340.0 + \frac{2}{(5+1)} \times (1,500 - 1,340.0) = 1,393.3$$

$$EMA \text{ of } 7^{\text{th}} \text{ hour} = 1,393.3 + \frac{2}{(5+1)} \times (1,800 - 1,393.3) = 1,528.8$$

$$EMA \text{ of } 8^{\text{th}} \text{ hour} = 1,528.8 + \frac{2}{(5+1)} \times (4,500 - 1,528.8) = 2,519.2$$

$$EMA \text{ of } 9^{\text{th}} \text{ hour} = 2,519.2 + \frac{2}{(5+1)} \times (6,000 - 2,519.2) = 3,679.5$$

:

$$EMA \text{ of } 24^{\text{th}} \text{ hour} = 1,094.2 + \frac{2}{(5+1)} \times (600 - 1,094.2) = 929.5$$

Calculating the EMA for twenty-four hours as shown in Figure 3, the results will be calculated with SMA again. In order to find the average value of each day, which the calculated value will be recorded as  $e_1'$  in Table 2. To calculate the average value of  $e_1$  of each day as follows.

$$e_1 = \frac{\sum_{i=5}^n EMA_i}{n - (\text{hours used in the past})}$$

$$= \frac{1,340.0 + 1,393.3 + \dots + 929.5}{24 - 5} = 2,463.0$$

where  $n$  is the number of hours per day subtracts by the number of hours used in the past.

Table 2: Four extra fields of  $R_i$

$R_i$	$C_i(f_1 \wedge f_2 \wedge f_3 \wedge f_4 \wedge f_5)$			
$R_i$	$e_1(FPM)$	$e_2(ECR)$	$e_3(ERC)$	$e_4(PRI)$
$R_1$	2,125	3	2	4
$R_i'$	$e_1'=0.24$	$e_2'=1.0$	$e_3'=0.66$	$e_4'=0.62$

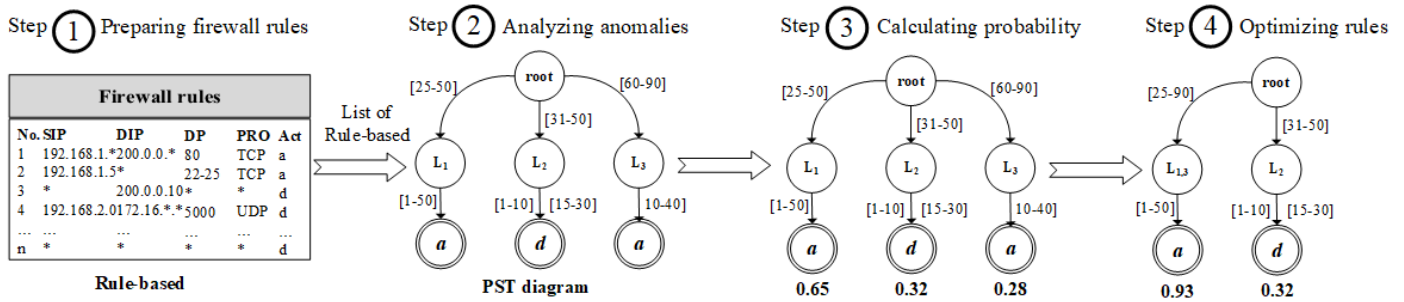


Figure 2: Overview of the system design

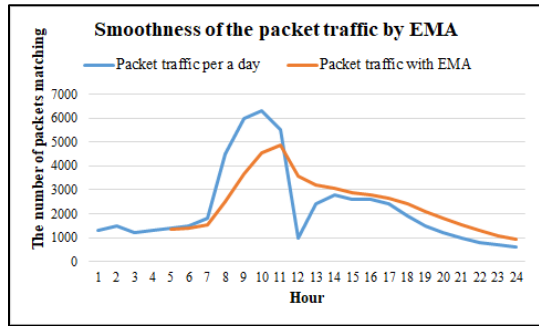


Figure 3: Adjusting raw packets traffic to be more smooth by EMA

**In the  $e_2$  case:** it refers to documents or pieces of paper used to confirm that such rules are approved to create them. In this paper, for example, the evidence for creating rules is divided into four levels: there is no evidence for approval, a firewall administrator is an approver, the head of the department is the approver, and approval is made by the owner of the organization. By dividing the weight of evidence according to the priority of the document approver: no evidence = 0, an administrator = 1, a head of the department = 2 and an owner of the organization = 3 respectively. If the weight of the document obtained is calculated by the Min-Max equation (8), the result will be the  $e_2'$ . Let the owner of the organization be approved to create the rule  $R_1$ , the result of the calculation is equal to:

$$e_2' = \frac{m - r_{min}}{r_{max} - r_{min}} \times t_{max} - t_{min} + t_{min}$$

$$= \frac{3 - 0}{3 - 0} \times (1.0 - 0.0) + 0.0 = 1.0$$

where  $m = 3$ ,  $r_{min} = 0$ ,  $r_{max} = 3$ ,  $t_{min} = 0.0$  and  $t_{max}$  is equal to 1.0.

**In the  $e_3$  case:** similar in the case of evidence for creating rules, the expertise in creating rules is also divided into four levels: newbie admins, normal admins with sufficient expertise, professional admins and very expert administrators. The newbie admins mean those who have recently been assigned to configure the firewall system and have the least experience. When they have configured the firewall for a while, they will be more proficient, which should have at least 3 - 5 years of working hours, called normal admins. For those who have a lot of experience and training or firewall customization, with working hours of 5 - 10 years, they will be professional administrators. Finally, those who have received a lot of training and certificates about firewalls will be considered expert admins. From the statistics, those who are very skilled will be able to analyze and design firewall rules to minimize errors as well. In this paper, therefore, determines the

weight of the following expertise  $e_3$ : newbie = 0, normal = 1, professional = 2 and very expertise administrator = 3. Let professional admins create the rule  $R_1$ , the result is calculated as.

$$e_3' = \frac{2 - 0}{3 - 0} \times (1.0 - 0.0) + 0.0 = 0.66$$

**In the  $e_4$  case:** the protocols communicating on the computer networks are always prioritized, such as video conferencing, and must be smooth throughout the meeting. On the other hand, sending electronic mail does not have an urgent need to be sent or received immediately. The protocol prioritization can be done depending on the policies of each organization. In this research, prioritization of the protocol is based on priorities from 3 GPP QoS Class Identification QCI categories [20] by IP Multimedia singling having the highest priority (1 = highest); Chat, FTP and P2P having the lowest priority (9 = lowest). From  $e_4$  in Table 2, it is a teleconference application with a priority of 4. When processing in the form of probability using the Min-Max Scaling, the result is equal to:

$$e_4' = \frac{6 - 1}{9 - 1} \times (1.0 - 0.0) + 0.0 = 0.625$$

Where  $m = 6$  (Teleconference = 4),  $r_{min} = 1$ ,  $r_{max} = 9$ ,  $t_{min} = 0.0$  and  $t_{max}$  is equal to 1.0. Notice that the priority of the protocols calculated must always reverse priorities, such as from 9 to 1 and from 1 to 9. For example, the priority of 4 is reversed to 6. Last, in Table 3, 4. represents examples of firewall rules consisting of all anomalies as previously mentioned, and these rules will be processed in the next step. Extra fields of each rule, when passing the data preparation process, it will produce the following results  $R_i \rightarrow R_i'$ .

#### 4.2. Analyzing and Detecting Anomalies (Step 2)

In this phase, the rules from the 1st step are used to build a tree structure, called the Path Selection Tree (PST), to analyze the anomalies. The algorithm begins with the creation of the root node of PST. After that, field  $f_1$  of the first rule is created as the first node on the tree, namely  $SIP_1$  as shown in Figure 4(a). In this node records source IP addresses of  $R_1$  to be  $\langle R_1 : [1, 100] \rangle$ , where  $f_1 \in [1, 100]$ . The next node ( $DIP_1$ ) stores the range of destination IP addresses ( $f_2$ ) of  $R_1$  ranging from 1 to 100. Next, it is the node that records source port ranges from 0 to 65535, called  $SP_1$ . The next node as  $DP_1$ , this node contains a group of destination ports  $f_4$  between 80 and 85 ( $\langle R_1 : [80, 85] \rangle$ ). The final field  $f_5$  of  $R_1$  as  $PRO_1$  which keeps the range of TCP and UDP protocols. In the decision field, a bottom rectangular box in the tree, contains an acceptance decision (1) of  $R_1$ . At the end of the tree, it records what rules members of this path are like  $\langle R_1 \rangle$ .

Table 3: Examples of firewall rules ( $R_i$ )

$R_1: f_1 \in [1, 100] \wedge f_2 \in [1, 100] \wedge f_3 \in [0, 65535] \wedge f_4 \in [80, 85] \wedge f_5 \in \{6, 17\} \rightarrow 1$
$R_2: f_1 \in [10, 50] \wedge f_2 \in [20, 60] \wedge f_3 \in [0, 65535] \wedge f_4 \in [80, 80] \wedge f_5 \in \{6, 17\} \rightarrow 0$
$R_3: f_1 \in [20, 40] \wedge f_2 \in [30, 70] \wedge f_3 \in [0, 65535] \wedge f_4 \in [80, 90] \wedge f_5 \in \{6, 17\} \rightarrow 0$
$R_4: f_1 \in [20, 30] \wedge f_2 \in [20, 30] \wedge f_3 \in [0, 65535] \wedge f_4 \in [80, 82] \wedge f_5 \in \{6, 17\} \rightarrow 0$
$R_5: f_1 \in [1, 100] \wedge f_2 \in [1, 100] \wedge f_3 \in [0, 65535] \wedge f_4 \in [30, 90] \wedge f_5 \in \{6, 17\} \rightarrow 1$
$R_6: f_1 \in [1, 100] \wedge f_2 \in [1, 100] \wedge f_3 \in [0, 65535] \wedge f_4 \in [0, 65535] \wedge f_5 \in \{6, 17\} \rightarrow 0$

Table 4: Examples of extra fields ( $R_i'$ )

$R_1: e_1 = 2,500, e_2 = 1, e_3 = 2, e_4 = 6 \rightarrow R_1': e_1' = 0.26, e_2' = 0.33, e_3' = 0.67, e_4' = 0.67$
$R_2: e_1 = 1,500, e_2 = 3, e_3 = 2, e_4 = 3 \rightarrow R_2': e_1' = 0.06, e_2' = 1.0, e_3' = 1.0, e_4' = 0.33$
$R_3: e_1 = 2,000, e_2 = 2, e_3 = 1, e_4 = 8 \rightarrow R_3': e_1' = 0.16, e_2' = 0.67, e_3' = 0.33, e_4' = 0.89$
$R_4: e_1 = 3,200, e_2 = 1, e_3 = 2, e_4 = 5 \rightarrow R_4': e_1' = 0.40, e_2' = 0.33, e_3' = 0.67, e_4' = 0.56$
$R_5: e_1 = 1,200, e_2 = 3, e_3 = 0, e_4 = 2 \rightarrow R_5': e_1' = 0.0, e_2' = 1.0, e_3' = 0.0, e_4' = 0.22$
$R_6: e_1 = 500, e_2 = 0, e_3 = 3, e_4 = 9 \rightarrow R_6': e_1' = 0.76, e_2' = 0.0, e_3' = 1.0, e_4' = 1.0$

In the next order, the second rule  $R_2$  is imported into PST as illustrated in Figure 4(b). Firstly,  $f_1$  of  $R_2 \subset f_1$  of  $R_1$ , thus  $R_2$  ( $f_1$ ) uses the same route as  $R_1$  ( $f_1$ ) and also records  $\langle R_2: [10, 50] \rangle$  into the  $SIP_1$  node. Likewise,  $R_2$  ( $f_2$ )  $\subset R_1$  ( $f_2$ ), it is also recorded to the same node ( $DIP_1 = \langle R_1: [1, 100], R_2: [20, 60] \rangle$ ), and travels over the same route as  $R_1$ . Similar to  $R_2$  ( $f_3$ ), it is equal to  $R_1$  ( $f_3$ ), hence  $R_2$  ( $f_3$ ) is appended in the  $SP_1$  node to be  $\langle R_1, R_2: [0, 65536] \rangle$ . In case of  $R_2$  ( $f_4$ ) and  $R_1$  ( $f_4$ ),  $R_2$  ( $f_4$ ) is a subset of  $R_1$  ( $f_4$ ), so the data of  $DP_1$  is updated as  $\langle R_1: [80, 85], R_2: [80, 80] \rangle$  as well as  $PRO_1$  updated to  $\langle R_1, R_2: \{6, 17\} \rangle$ . On the other hand, the decision of  $R_1$  and  $R_2$  are not the same, so the decision path must be separated from each other, where  $\langle R_1 \rangle = 1, \langle R_2 \rangle = 0$ . For inserting  $R_3$  (Figure 4(c)) into PST, there is not much difference from inserting  $R_2$ , it is slightly different in the position of the protocol level in the tree. Since  $R_3$  ( $f_4$ ) is a superset of  $R_1$  ( $f_4$ ) and  $R_2$  ( $f_4$ ), some destination ports of  $R_3$  ( $f_4$ ) have to be separated into another node of the tree, namely  $DP_2$ , which stores the destination ports ranging from 86 to 90 ( $R_3$  ( $f_4$ ) -  $R_1$  ( $f_4$ )) like  $\langle R_3: [86, 90] \rangle$ . Remaining destination ports are combined with  $DIP_1$  in the first path together with  $R_1$  and  $R_2$  as  $\langle R_1, R_3: [80, 85], R_2: [80, 80] \rangle$ . The decision of  $R_3$  is not allowed in both paths, where  $\langle R_3 \rangle = 0$ . Remaining firewall rules ( $R_4, R_5, R_6$ ) will be executed like the previous rules ( $R_1, R_2, R_3$ ). If all rules have been implemented successfully over PST, the results are shown in Figure 5.

In the process of checking the rule anomalies, the algorithm uses the information recorded on each node to detect anomalies by using the Cartesian product of all nodes separated from the protocol layer ( $PRO_i$ ) and looking back from the bottom to the root as follows.

**Group 1:** path number 1 and 2 under the node  $PRO_1$

$$CP(\langle R_1, R_5 \rangle) = (R_1, R_5) \quad (18)$$

$$CP(\langle R_2, R_3, R_4, R_6 \rangle) = (R_2, R_3), (R_2, R_4), (R_2, R_6), (R_3, R_4), (R_3, R_6), (R_4, R_6) \quad (19)$$

$$CP(\langle R_1, R_5 \rangle, \langle R_2, R_3, R_4, R_6 \rangle) = (R_1, R_2), (R_1, R_3), (R_1, R_4), (R_1, R_6), (R_2, R_5), (R_3, R_5), (R_4, R_5), (R_5, R_6) \quad (20)$$

**Group 2:** path number 3 and 4 under the node  $PRO_2$

$$CP(\langle R_3, R_6 \rangle) = (R_3, R_6) \quad (21)$$

$$CP(\langle R_3, R_6 \rangle, \langle R_5 \rangle) = (R_3, R_5), (R_5, R_6) \quad (22)$$

**Group 3:** path number 5 and 6 under the node  $PRO_3$

$$CP(\langle R_5, R_6 \rangle) = (R_5, R_6) \quad (23)$$

where  $CP$  is the Cartesian product. The results of each pair of the Cartesian product are calculated by equations (3) to (7) to find out what kind of anomalies they are. For example, in the equation (18) of group 1, ( $R_1, R_5$ ) has the same decision (decision = 1), so it is executed by the equation (6). The result of the execution is a redundant anomaly. Next example, in equation (19), they consist of ( $R_2, R_3$ ), ( $R_2, R_4$ ), ( $R_2, R_6$ ), ( $R_3, R_4$ ), ( $R_3, R_6$ ) and ( $R_4, R_6$ ) by every pair of rules has the same decisions, thus all is executed by the equation (6) as same as the equation (18). Results of Cartesian product in the equation (20): ( $R_1, R_2$ ), ( $R_1, R_3$ ), ( $R_1, R_4$ ), ( $R_1, R_6$ ), ( $R_2, R_5$ ), ( $R_3, R_5$ ), ( $R_4, R_5$ ) and ( $R_5, R_6$ ), and every pair has different decisions, so they are executed by equations (3), (4) and (5) respectively. The calculated results: ( $R_2, R_3$ ) = Redundancy and Semantics loss (Executed by equation (6)), ( $R_2, R_4$ ) = Redundancy and Semantics loss (6), ( $R_2, R_6$ ) = Redundancy and Semantics loss (6), ( $R_3, R_4$ ) = Redundancy and Semantics loss (6), ( $R_3, R_6$ ) = Redundancy and Semantics loss (6), ( $R_4, R_6$ ) = Redundancy and Semantics loss (6), ( $R_1, R_2$ ) = Shadowing (3), ( $R_1, R_3$ ) = Correlation (4), ( $R_1, R_4$ ) = Shadowing (3), ( $R_1, R_6$ ) = Generalization (5), ( $R_2, R_5$ ) = Generalization (5), ( $R_3, R_5$ ) = Generalization (5), ( $R_4, R_5$ ) = Generalization (5), ( $R_5, R_6$ ) = Generalization (5). The results obtained from the calculations of group number 2 and 3 in equation 21 to 23: ( $R_3, R_6$ ) = Redundancy and Semantics loss (Executed by equation (6)), ( $R_3, R_5$ ) = Generalization (5), ( $R_5, R_6$ ) = Generalization (5).

Losing the meaning of rules always occurs by redundant rules, thus all members in equations 18, 19 and 21 are possible to be the semantics loss as well.

#### 4.3. Calculating Probability of Each Path of PST (Step 3)

The PST obtained from the previous steps is used to calculate the probability of each path in order to advise administrators to make decisions about optimizing firewall rules effectively, which has the following steps. Let  $R$  as a firewall rule,  $e$  as an attribute field of a rule, and  $S$  is a sample space, then the conditional probability of  $R$  given  $e$  is the equation (24) and shown in Figure 6.

$$P(R|e) = \frac{P(R|e)}{P(e)} \quad (24)$$

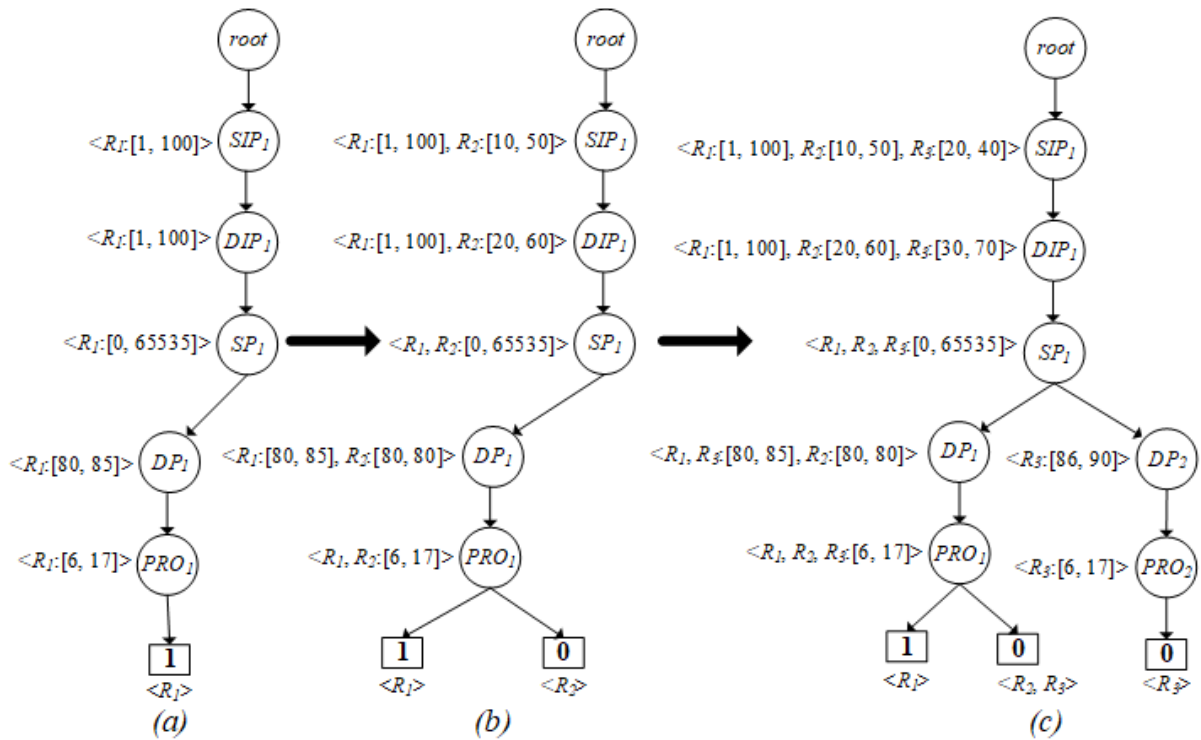


Figure 4: Creating rule  $R_1$  (a),  $R_2$  (b) and  $R_3$  (c) into PST

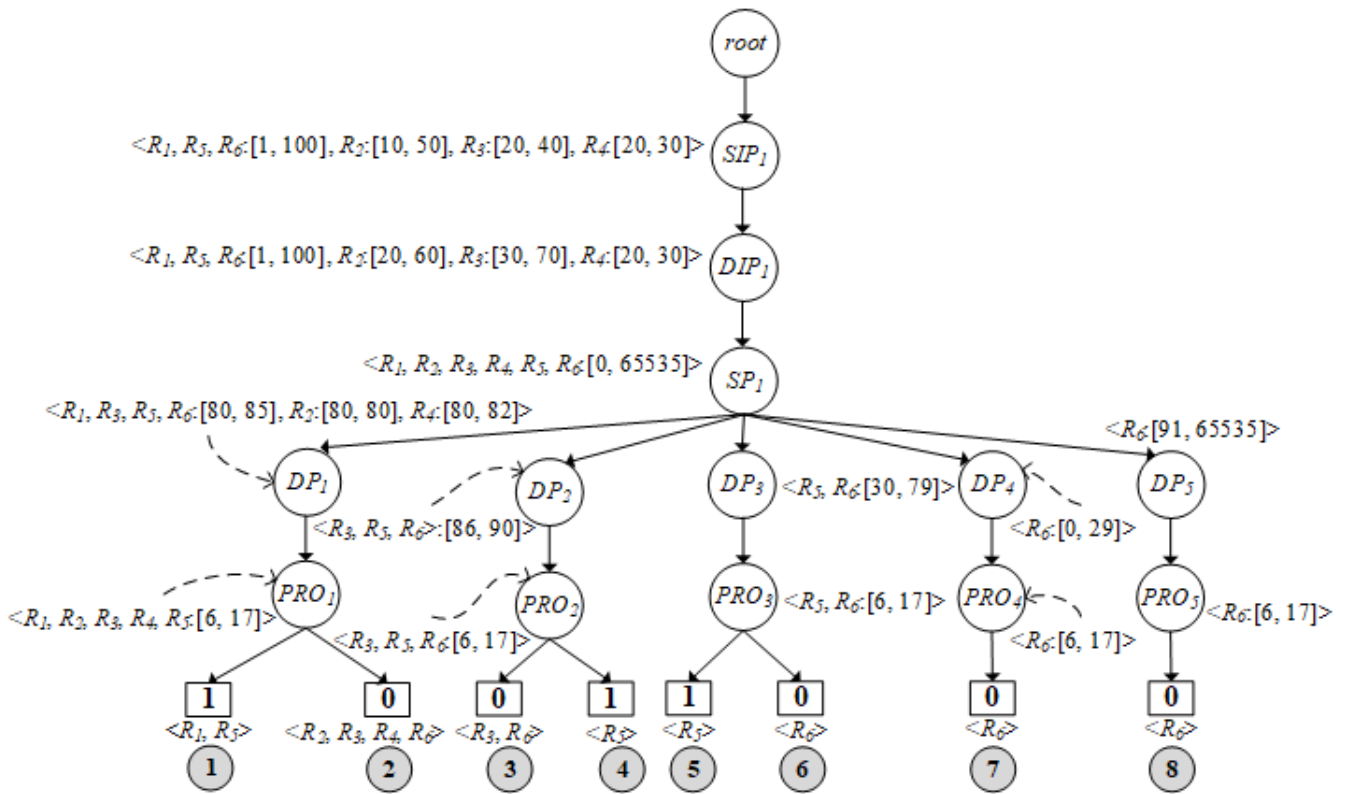


Figure 5: Complete PST structure after compiling all rules



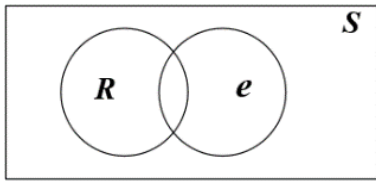


Figure 6: Conditional probability of R given e in Venn diagram

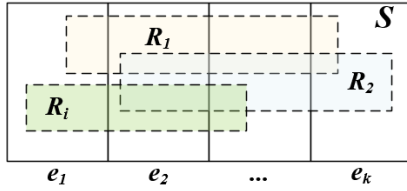


Figure 7: Conditional probability of Ri given ek

According to Figure 7., given  $R_i$  as any rule,  $e_k$  is any attribute (Extra field) of  $R_i$ :

$$e_i \cap e_k = \emptyset; \forall i, k; i \neq k$$

$$e_1 \cup e_2 \cup e_3 \cup \dots \cup e_k = S$$

$$P(R_i) = P(R_i \cap e_1) \cup P(R_i \cap e_2) \cup \dots \cup P(R_i \cap e_k) \quad (25)$$

From the equation (24),  $P(R|e) = \frac{P(R|e)}{P(e)}$ , so  $P(R|e): P(e)P(R|e)$  or  $P(R)P(e|R)$ . Since we already know the value of the  $P(e)$ , then we choose  $P(R \cap e) = P(e)P(R|e)$  and represent  $i$  and  $k$  into the equation as follows.

$$P(R_i \cap e_k) = P(e_k)P(R_i|e_k) \quad (26)$$

Applying the equation (26) instead of the equation (25):

$$P(R_i) = P(e_1)P(R_i|e_1) \cup P(e_2)P(R_i|e_2) \cup \dots \cup P(e_k)P(R_i|e_k) \quad (27)$$

$$P(R_i) = \sum_{i,k=1}^n (e_k)P(R_i|e_k) \quad (28)$$

Given  $e_k$  is any property considering when giving  $P(R_i)$  on the firewall. Finally, we can substitute this into Bayes' rule (Equation 9) from above to obtain an alternative version of Bayes' rule, which is used heavily in Bayesian inference:

$$P(e_k|R_i) = \frac{P(R_i \cap e_k)}{P(R_i)} = \frac{P(e_k)P(R_i|e_k)}{\sum_{i,k=1}^n (e_k)P(R_i|e_k)} \quad (29)$$

From examples of the property fields (Extra fields) in Table 4., there are four fields ( $e_1', e_2', e_3', e_4'$ ), where  $e_1'$  = the frequency of packets matching against rules (FPM),  $e_2'$  = evidence of creating rules (ECR),  $e_3'$  = expertise of rules creator (ERC) and  $e_4'$  = protocol priority (PRI). Assuming the weight of each feature is equal, so  $e_1', e_2', e_3'$  and  $e_4'$  are equal to 25% (0.25). Substituting various values in equations (28) and (29), the calculated results:

$$P(e_1') = 0.25, P(e_2') = 0.25, P(e_3') = 0.25, P(e_4') = 0.25$$

In case of  $P(R_1')$ :

$$P(R_1|e_1') = 0.26, P(R_1|e_2') = 0.33, P(R_1|e_3') = 0.67, P(R_1|e_4') = 0.67$$

$$P(R_i) = P(e_1')P(R_1|e_1') + P(e_2')P(R_1|e_2') + P(e_3')P(R_1|e_3') + P(e_4')P(R_1|e_4') + P(e_1')P(R_2|e_1') + P(e_2')P(R_2|e_2') + P(e_3')P(R_2|e_3') + P(e_4')P(R_2|e_4') + \dots + P(e_1')P(R_6|e_1') + P(e_2')P(R_6|e_2') + P(e_3')P(R_6|e_3') + P(e_4')P(R_6|e_4')$$

$$P(R_i) = ((0.25 \times 0.26) + (0.25 \times 0.33) + (0.25 \times 0.67) + (0.25 \times 0.67)) + ((0.25 \times 0.06) + (0.25 \times 1) + (0.25 \times 1) + (0.25 \times 0.33)) + \dots + ((0.25 \times 0.76) + (0.25 \times 0) + (0.25 \times 1) + (0.25 \times 1)) = 3.07$$

$$P(e_1'|R_1) = \frac{P(e_1')P(R_1|e_1')}{P(R_1)} = \frac{0.25 \times 0.26}{3.07} = 0.0211$$

$$P(e_2'|R_1) = \frac{P(e_2')P(R_1|e_2')}{P(R_1)} = \frac{0.25 \times 0.33}{3.07} = 0.0268$$

$$P(e_3'|R_1) = \frac{P(e_3')P(R_1|e_3')}{P(R_1)} = \frac{0.25 \times 0.67}{3.07} = 0.0545$$

$$P(e_4'|R_1) = \frac{P(e_4')P(R_1|e_4')}{P(R_1)} = \frac{0.25 \times 0.67}{3.07} = 0.0545$$

$$\therefore P(R_1') = P(e_1'|R_1) + P(e_2'|R_1) + P(e_3'|R_1) + P(e_4'|R_1) = 0.157$$

In case of  $P(R_2')$ :

$$P(R_2|e_1') = 0.06, P(R_2|e_2') = 1.0, P(R_2|e_3') = 1.0, P(R_2|e_4') = 0.33$$

$$P(e_1'|R_2) = 0.0488, P(e_2'|R_2) = 0.0814, P(e_3'|R_2) = 0.0814, P(e_4'|R_2) = 0.0268$$

$$\therefore P(R_2') = P(e_1'|R_2) + P(e_2'|R_2) + P(e_3'|R_2) + P(e_4'|R_2) = 0.194$$

In case of  $P(R_3')$ :

$$P(R_3|e_1') = 0.16, P(R_3|e_2') = 0.67, P(R_3|e_3') = 0.33, P(R_3|e_4') = 0.89$$

$$P(e_1'|R_3) = 0.0130, P(e_2'|R_3) = 0.0545, P(e_3'|R_3) = 0.0268, P(e_4'|R_3) = 0.0724$$

$$\therefore P(R_3') = P(e_1'|R_3) + P(e_2'|R_3) + P(e_3'|R_3) + P(e_4'|R_3) = 0.166$$

In case of  $P(R_4')$ :

$$P(R_4|e_1') = 0.40, P(R_4|e_2') = 0.33, P(R_4|e_3') = 0.67, P(R_4|e_4') = 0.56$$

$$P(e_1'|R_4) = 0.0325, P(e_2'|R_4) = 0.0268, P(e_3'|R_4) = 0.0545, P(e_4'|R_4) = 0.0456$$

$$\therefore P(R_4') = P(e_1'|R_4) + P(e_2'|R_4) + P(e_3'|R_4) + P(e_4'|R_4) = 0.159$$

In case of  $P(R_5')$ :

$$P(R_5|e_1') = 0.0, P(R_5|e_2') = 1.0, P(R_5|e_3') = 0.0, P(R_5|e_4') = 0.22$$

$$P(e_1'|R_5) = 0.0, P(e_2'|R_5) = 0.0814, P(e_3'|R_5) = 0.0, P(e_4'|R_5) = 0.0179$$

$$\therefore P(R_5') = P(e_1'|R_5) + P(e_2'|R_5) + P(e_3'|R_5) + P(e_4'|R_5) = 0.099$$

In case of  $P(R_6')$ :

$$P(R_6|e_1') = 0.76, P(R_6|e_2') = 0.0, P(R_6|e_3') = 1.0, P(R_6|e_4') = 1.0$$

$$P(e_1'|R_6) = 0.0618, P(e_2'|R_6) = 0.0, P(e_3'|R_6) = 0.0814, P(e_4'|R_6) = 0.0814$$

$$\therefore P(R_6') = P(e_1'|R_6) + P(e_2'|R_6) + P(e_3'|R_6) + P(e_4'|R_6) = 0.224$$

According to the weight of each rule property, admins can adjust the weight of each property as needed, such as  $P(e_1') = 35\%$  (0.35),  $P(e_2') = 15\%$  (0.15),  $P(e_3') = 25\%$  (0.25) and  $P(e_4') = 25\%$  (0.25) depending on each organization to give weight to their properties. After calculating all the probability values successfully, the algorithm inserts these probabilities to each path of PST as shown in Figure 8.

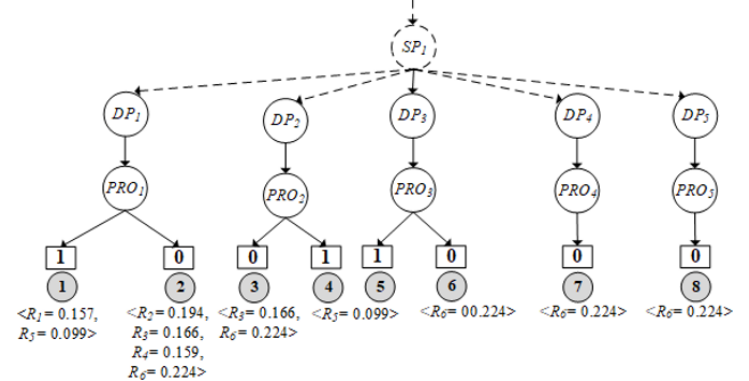


Figure 8: Inserting probability of each  $R_i$  into PST

4.4. Optimizing Rule Anomalies (Last step)

Anomalies occurred over firewall rules have different solutions, for example, the redundant anomaly is solved by merging the rules together. However, this method may result in semantics loss instead. Other anomalies such as shadowing, correlation, and generalization, should not use the merging method because their decisions are different. Sometimes, administrators choose to resolve problems by switching rules, but they are not sure what will happen in the future. Therefore, this research uses the calculated probability in each rule to help administrators decide how to proceed with anomalies to achieve maximum efficiency and reasonableness. For example, the path number 1 of Figure 5.,  $R_1$  and  $R_5$  are redundancy. If admins decide to combine the two rules together, the result is.

$$R_1 \Phi R_5 = R_1(f_1) \cup R_5(f_1) \wedge R_1(f_2) \cup R_5(f_2) \wedge \dots \wedge R_1(f_5) \cup R_5(f_5)$$

$$R_1 \Phi R_5 = ([1, 100] \cup [1, 100]) \wedge ([1, 100] \cup [1, 100]) \wedge ([0, 65535] \cup [0, 65535]) \wedge ([80, 85] \cup [80, 85]) \wedge ([6, 117] \cup [6, 17])$$

$$R_{new} = f_1 \in [1, 100] \wedge f_2 \in [1, 100] \wedge f_3 \in [0, 65535] \wedge f_4 \in [80, 85] \wedge f_5 \in \{6, 17\} \rightarrow 1$$

The property fields of  $R_1$ :  $e_1 = 2500, e_2 = 1, e_3 = 2, e_4 = 6$ ; and  $R_5$  are  $e_1 = 1200, e_2 = 3, e_3 = 0, e_4 = 2$ . Thus,  $R_1(e_i) \Phi R_5(e_i)$ :

$$R_1(e_i) \Phi R_5(e_i) = (R_1(e_1) + R_5(e_1)), \text{Max}(R_1(e_2), R_5(e_2)), \text{Max}(R_1(e_3), R_5(e_3)), \text{Max}(R_1(e_4), R_5(e_4))$$

$$R_1(e_i) \Phi R_5(e_i) = (R_1(2500) + R_5(1200)), \text{Max}(R_1(1), R_5(3)), \text{Max}(R_1(2), R_5(0)), \text{Max}(R_1(6), R_5(2))$$

$$R_{1new(e)} \rightarrow e_1 = 3700, e_2 = 3, e_3 = 2, e_4 = 6$$

where  $R_{1new}$  as a new merged rule,  $\Phi$  as a merging operator for the same decisions,  $\text{Max}$  is a function calculated the maximum value. In the same way as  $(R_2, R_3), (R_2, R_4), \dots$ , and  $(R_3, R_6)$ , which are a redundant conflict, so they can solve the problem by combining rules like  $(R_1, R_5)$ . The methods of resolving the remaining anomalies (Shadowing, Correlation, and Generalization) can be done in three ways: merging, swapping and removing rules. Nevertheless, admins must be highly skilled and aware of the consequences, almost all researchers do not recommend using these methods and often push the burden to the discretion of administrators instead. If the admins choose one of three methods, they can do by checking the probabilities of each rule. If the probability of any rule is the highest, it means that admins have the opportunity to resolve anomalies to be more effective. For example,  $(R_1, R_2)$  is the shadowing anomaly. If admins need to delete, merge or swap rules, they should give priority to  $R_2$  rather than  $R_1$  because  $R_2$  is a higher probability ( $R_1 = 0.157, R_2 = 0.194$ ) as shown in Figure 9.

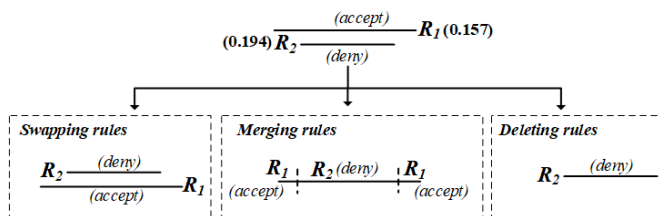


Figure 9: Resolving shadowing by swapping, merging and deleting rules

Updating property of  $R_1$  and  $R_2$  is not necessary in case of swapping and deleting rules, but in the case of merging, there are the following details.

$$R_2 \Theta R_1 = R_2(f_1) - R_1(f_1) \wedge R_2(f_2) - R_1(f_2) \wedge \dots \wedge R_2(f_5) - R_1(f_5)$$

$$R_2' = f_1 \in [10, 50] \wedge f_2 \in [20, 60] \wedge f_3 \in [0, 65535] \wedge f_4 \in [80, 80] \wedge f_5 \in \{6, 17\} \rightarrow 0$$

$$R_1' = f_1 \in [1, 9], [51, 100] \wedge f_2 \in [1, 19], [61, 100] \wedge f_3 \in [0, 65535] \wedge f_4 \in [81, 85] \wedge f_5 \in \{6, 17\} \rightarrow 1$$

where  $\Theta$  is a merging operator for different decisions. Although,  $R_2(f_3) - R_1(f_3)$  and  $R_2(f_5) - R_1(f_5)$  are equal to  $\emptyset$ . However, for this model, both are not equal to  $\emptyset$  because they share the same path in the tree. The property fields of  $R_1$ :  $e_1 = 2500, e_2 = 1, e_3 = 2, e_4 = 6$ ; and  $R_2$  are  $e_1 = 1500, e_2 = 3, e_3 = 2, e_4 = 3$ . Thus,  $R_1(e_i) \Theta R_2(e_i)$ :

$$R_2(e_i) \Theta R_1(e_i) = (R_2(1500) + R_1(2500)), \text{Max}(R_2(3), R_1(1)), \text{Max}(R_2(2), R_1(2)), \text{Max}(R_2(3), R_1(6))$$

$$R_2'(e_i) \rightarrow e_1 = 4000, e_2 = 3, e_3 = 2, e_4 = 6$$

$$R_1'(e_i) \rightarrow e_1 = 2500, e_2 = 1, e_3 = 2, e_4 = 6$$

Note that while updating each conflict rule each time, the PST structure will be changed, which means that the probability has to be recalculated whenever when resolving conflicts.

5. Implementation and Performance Evaluation

PST uses the k-ary tree structure (also known as m-ary or k-way tree) to develop, so the processing speed to build the tree is  $O(n)$ , where n is the number of nodes of the given k-ary tree. The number of levels of the existing k-ary tree is L, the depth of the k-ary in the worst case is  $N - 1$ , where N is the number of nodes in a tree. The k-ary tree can also be stored in breadth-first order as an implicit data structure in pointer-based, each node would have an internal array for storing pointers to each of its m children. So, the space complexity of k-ary tree structure is  $O(m \times n)$ . Traversing the k-ary tree is very similar to binary tree traversal. Besides, the worst-case time complexity is  $O(\log_m n)$ . In practically, PST is developed on Intel Core i7 2.3 GHz, 8 GB RAM. The software developed includes Python language version 3.7 (64-bit), Graphviz [21] and NetworkX version 1.11 [22] running on Linux kernel (Version 4.4). The proposed model is illustrated in Figure 10.

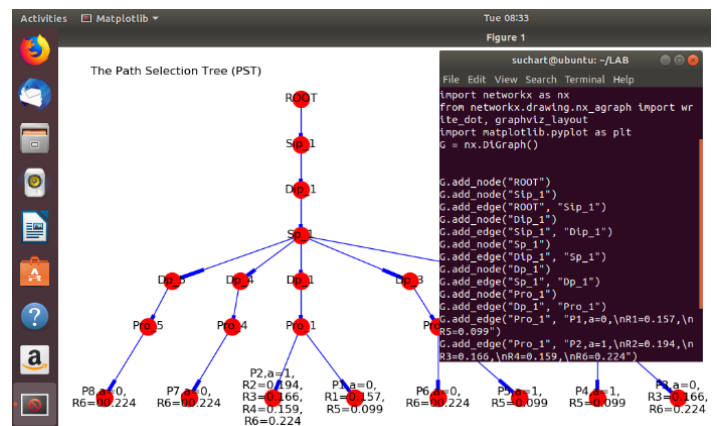


Figure 10: The developed PST structure over Linux operating system

In this paper, we used  $\bar{x}$  (Equation 13) to evaluate the satisfaction of the firewall administrators for resolving firewall rule anomalies of both traditional (No recommendation system) and our proposed firewall (Recommendation system with probability). In which of the confidence test consists of ten scenarios to resolve anomalies and the total number of testers (Firewall expert) is five as shown in Table 5.

Referring to Table 5, the average ( $\bar{x}$ ) of the five administrators' confidence for resolving ten scenarios of rule anomalies based on their skills for the traditional firewall is equal to 2.68; however, the average confidence of our proposed firewall which is a recommendation system based on probability is 4.16, which the confidence rate of the proposed firewall increased by 29.6% from the conventional firewall. In the case of evaluating reliability between raters, we apply Kappa statistics [19] in the equation (14) with the data from Table 5. The reliability value between the inter-raters of the conventional firewall is equal to 0.379, which means that the reliability is at a fair agreement as shown in Table 6. The reliability value between the assessors of the proposed firewall is equal to 0.510 (Moderate agreement), which means that the reliability increased from 37.9% to 51% significantly.

Table 5: Confidence tests to resolve anomalies of firewall rules

Firewall rule anomaly and resolving methods	The confidence of five admins for resolving rule anomalies									
	Traditional firewalls (No the recommendation system) Admins judged by their experience					Our proposed firewall (Use the recommendation system) Admins judged by probability				
	5	4	3	2	1	5	4	3	2	1
<b>Shadowing</b>										
- Merging rules	1	1	2	1	0	0	0	1	2	2
- Swapping rules	1	1	2	0	1	0	0	1	3	1
- Deleting rules	0	2	2	1	0	0	0	2	2	1
<b>Correlation</b>										
- Merging rules	1	2	2	0	0	0	0	1	2	2
- Swapping rules	0	2	2	1	0	0	0	1	2	2
<b>Generalization</b>										
- Merging rules	1	2	2	0	0	0	0	2	1	2
- Swapping rules	1	1	2	1	0	0	0	1	2	2
- Deleting rules	0	1	2	1	1	0	0	1	2	2
<b>Redundancy</b>										
- Merging rules	0	2	2	1	0	0	0	0	2	3
<b>Semantics loss</b>										
- Merging rules	1	1	2	1	0	0	0	1	2	2

Remarks: the level of satisfaction is divided in to five levels: 5, 4, 3, 2, and 1; 5 is the highest confidence.

Table 6: Interpretation of the reliability between Inter-raters [19]

Kappa statistic value	Inter-rater reliability description
0	Agreement equivalent to chance
0.1 - 0.20	Alight agreement
0.21 - 0.40	Fair agreement
0.41 - 0.60	Moderate agreement
0.61 - 0.80	Substantial agreement
0.81 - 0.99	Near perfect agreement
1	Perfect agreement

6. Conclusion

Practically, fixing anomalies of firewall rules is quite complex, depending on the administrator's perspective and experience. Correcting mistakes may lead to other anomalies. For example, when resolving the redundant anomaly, it may become

the semantics loss of rules. In order to reduce the impact of errors and to resolve anomalies of administrators, this paper has designed and developed a system to assist in the decision-making of administrators by using probability together with four additional features of rules: frequency of matching between packets and rules, evidence of creating rules, expertise of rules creator and protocol priority. For each rule, the probability is calculated based on their features. If the probability of any rule is high, it indicates that the rule has a high priority. While rules in the firewall are conflicts, the rule that has a high probability value is always considered first. As a result of system testing, administrators can make more accurate decisions about conflict rules in the firewall. For the overall efficiency of the system, the time complexity of creating a system (PST) is equal to  $O(n)$ , searching time over PST is  $O(\log_m n)$  and the space complexity is  $O(m \times n)$ . However, the system still has a limitation against the establishment of the tree structure. As resolving any anomaly of rules in each period, it needs to reconstruct the whole PST tree structure. For the evaluation of confidence for resolving firewall rule anomalies, the firewall that we have proposed on the basis of probability obtains a confidence value more than the traditional firewall by 29.6%, and the reliability between Inter-raters of proposed firewall is in the moderate agreement (0.51), which increased by 13.1% from the traditional firewall.

Acknowledgment

This research project was financially supported by Faculty of Informatics, Mahasarakham University 2020.

References

- [1] E. S. Al-Shaer and H. H. Hamed, "Modeling and management of firewall policies," IEEE Transactions on Network and Service Management, 1(1), 2–10, April 2004. <https://doi.org/10.1109/TNSM.2004.4623689>
- [2] S. Khummanee, "The semantics loss tracker of firewall rules," in IC2IT 2018 Recent Advances in Information and Communication Technology, 769, Jun. 2018, 220–231, 2018. [https://doi.org/10.1007/978-3-319-93692-5\\_22](https://doi.org/10.1007/978-3-319-93692-5_22)
- [3] E. Al-Shaer, H. Hamed, R. Boutaba, and M. Hasan, "Conflict classification and analysis of distributed firewall policies," IEEE Journal on Selected Areas in Communications, 23(10), 2069–2084, Oct 2005. <https://doi.org/10.1109/JSAC.2005.854119>
- [4] S. Khummanee, A. Khumseela, and S. Puangpronpitag, "Towards a new design of firewall: Anomaly elimination and fast verifying of firewall rules," in The 2013 10th International Joint Conference on Computer Science and Software Engineering (JCSSE), May 2013, 93–98, 2013. <https://doi.org/10.1109/JCSSE.2013.6567326>
- [5] A. X. Liu, "Formal verification of firewall policies," in 2008 IEEE International Conference on Communications, May 2008, 1494 – 1498, 2008. <https://doi.org/10.1109/ICC.2008.289>
- [6] M. Rezvani and R. Aryan, "Analyzing and resolving anomalies in firewall security policies based on propositional logic," in 2009 IEEE 13th International Multitopic Conference, Dec 2009, 1–7, 2009. <https://doi.org/10.1109/INMIC.2009.5383125>
- [7] H. Hu, G. Ahn, and K. Kulkarni, "Detecting and resolving firewall policy anomalies," IEEE Transactions on Dependable and Secure Computing, 9(3), 318–331, May 2012. <https://doi.org/10.1109/TDSC.2012.20>
- [8] S. R. Peter M. Mell, Karen A. Scarfone. (015) A complete guide to the common vulnerability scoring system version 2.0. [Online]. Available: <https://www.nist.gov/publications/complete-guidecommon-vulnerability-scoring-system-version-20>
- [9] A. Wool, "Architecting the lumeta firewall analyzer," in SSYM'01 Proceedings of the 10th conference on USENIX Security Symposium, 10, Aug 2001, p. 7.
- [10] A. Mayer, A. Wool, and E. Ziskind, "Fang: a firewall analysis engine," in Proceeding 2000 IEEE Symposium on Security and Privacy. S P 2000, May 2000, 177–187, 2000. <https://doi.org/10.1109/SECPRI.2000.848455>
- [11] A. Hari, S. Suri, and G. Parulkar, "Detecting and resolving packet filter conflicts," in Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer

- and Communications Societies (Cat. 00CH37064), 3, March 2000, 1203–1212 3, 2000. <https://doi.org/10.1109/INFCOM.2000.832496>
- [12] L. Zhang and M. Huang, “A firewall rules optimized model based on service-grouping,” in 2015 12th Web Information System and Application Conference (WISA), Sep. 2015, 142–146, 2015. <https://doi.org/10.1109/WISA.2015.47>
- [13] Lihua Yuan, Hao Chen, Jianning Mai, Chen-Nee Chuah, Zhendong Su, and P. Mohapatra, “Fireman: a toolkit for firewall modelling and analysis,” in 2006 IEEE Symposium on Security and Privacy (S P’06), May 2006, 15 – 213, 2006. <https://doi.org/10.1109/SP.2006.16>
- [14] A. Sa`adaoui, N. B. Y. B. Souayeh, and A. Bouhoula, “Formal approach for managing firewall misconfigurations,” in 2014 IEEE Eighth International Conference on Research Challenges in Information Science (RCIS), May 2014, 1–10, 2014. <https://doi.org/10.1109/RCIS.2014.6861044>
- [15] W. Krombi, M. Erradi, and A. Khoumsi, “Automata-based approach to design and analyze security policies,” in 2014 Twelfth Annual International Conference on Privacy, Security and Trust, July 2014, 306–313, 2014. <https://doi.org/10.1109/PST.2014.6890953>
- [16] A. Sa`adaoui, N. B. Y. B. Souayeh, and A. Bouhoula, “Automated and optimized fdd-based method to fix firewall misconfigurations,” in 2015 IEEE 14th International Symposium on Network Computing and Applications, Sep. 2015, 63–67, 2015. <https://doi.org/10.1109/NCA.2015.31>
- [17] C. C. Aggarwal, Neural Networks and Deep Learning. Boca Raton, FL, USA: Springer, 2018.
- [18] D. Morris and M. Koning, Bayes’ Theorem Examples: A Visual Introduction for Beginners. 80 Strand, London, WC2R 0RL UK: Blue Windmill Media, 2016.
- [19] P. W. M. J. Kenneth J. Berry, Janis E. Johnston, The Measurement of Association: A Permutation Statistical Approach. Boca Raton, FL, USA: Springer, 2018.
- [20] CelPlan Technologies. Gpp qos class identification qci categories 2019. [Online]. Available: <http://www.celplan.com>
- [21] Graphviz. Graph visualization software 2019. [Online]. Available: <https://www.graphviz.org>
- [22] NetworkX. Software for complex networks, 2019. [Online]. Available: <https://networkx.github.io>



## IoT Based Human Activity Recognition System Using Smart Sensors

Deepti Sehrawat\*, Nasib Singh Gill

Department of Computer Science & Applications, Maharshi Dayanand University, Rohtak, Haryana, 124001, India

---

### ARTICLE INFO

Article history:

Received: 24 May, 2020

Accepted: 19 July, 2020

Online: 09 August, 2020

---

Keywords:

Types of IoT Sensors

HAR Systems

Sensor Applications

IoT and Sensors

---

### ABSTRACT

Internet of Things provides a virtual view of real-life things by guiding challenges faced by persons in daily life. It is reforming our world with trillions of sensors and other IoT enabled devices by creating a smart environment. Effective use of IoT sensors in various smart IoT applications is very important. After analyzing different sensor applications, this paper presents various types of wearable sensors used for monitoring of human activities along with different locations optimal for their placement. This paper enlightens sensors suitable for any particular application. IoT has opened up a new avenue of research in the field of human activity recognition using wearable sensors. It provides an individual's valuable information regarding functional ability and lifestyle. Human activities are monitored automatically to provide personalized support to different individuals. Recently, various researchers presented different human activity recognition systems, a few are cumulated in this paper. Furthermore, a Human Activity Recognition (HAR) system is also proposed in this paper for a smart IoT environment that would be secure enough to use.

## 1. Introduction

The advent of Internet of Things (IoT) has opened a plethora of opportunities. In recent years, human activity recognition has received growing attention. This paper is an extension of work originally presented in the 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI) [1]. IoT not only connects animate/living things but inanimate/non-living things are also connected in a smart environment. This brings revolutionary changes in the world [2]. In any IoT based applications, sensors are very significant. Any physical/chemical change can be detected by these sensors which are then collected and processed to automate the application/devices to make it smarter in the form of smart application. IoT integrates several sensors of different types attached to the devices and nodes that can communicate without any human intervention [3].

Wearable sensors and the advances of wireless networks has opened the door to a world of novel interactive IoT applications. Different types of inexpensive and unobtrusive body-worn sensors like accelerometers, gyroscopes, and microphones have been often used in daily life activities [4]. Various researchers worked on detecting human activities by observing different parameters. It can be used for the monitoring of people of any age group like for infants, young people, disabled people, and elderly people in a

solitary life. These monitoring systems are usually smaller and are wearable [5].

The mental or physical health of a user can be monitored remotely and then the results can be sent to their remote family. Most of the earlier researchers in this area made use of multiple body-worn sensors worn/ placed at different body parts like chest, thigh, waist, ankle, knee, wrist, ear, eyes, and so on [6]. The purpose of those research work was to categorize a definite subset of activities for particular applications. For instance, for upper body movements, wrist and arm sensors were used whereas for simple lifestyle monitoring a single sensor would be enough [7]. Most of the human activity recognition systems use accelerometers for the classification and assessment of day-to-day activities. Some used single accelerometers while some used multiple accelerometers attached to different sites on one's body. Wearing multiple accelerometers gives a good level of accuracy for certain activities but wearing multiple sensors all the time is not feasible. Wearable sensors are mostly used to detect physical movements but these can also be used for monitoring academic performance, analysing sleep class/patterns, stress level, and mental fitness conditions. Researchers in [8] gathered data from 66 individuals and analysed behavioural and physiological patterns. They also identified factors affecting performance in academics, sleep quality, stress level, and mental health. In the process of human activity recognition first step is to collect the data and prepare a

---

\*Corresponding Author: Deepti Sehrawat, dips.scorpio@gmail.com

dataset which is then pre-processed and then the classification of data is carried out. This paper also presents various steps and techniques used in the process of recognition of human activities.

For the recognition of human activities from wearable, there are different classification techniques supervised and unsupervised (both) which are widely used by different researchers. Supervised techniques are considered more accurate when compared to unsupervised techniques [9].

IoT has numerous applications where different types of sensors are used in different types of applications. This paper studies different IoT applications and enlightens which smart application demands which type of sensor. Besides, various types of wearable sensors which are used for monitoring of human activities and application area where these are used are also presented in this paper.

This paper presents a detailed review of different locations where we can place wearable sensors and explains which one is ideal for a given group. It is also considered whether the sensor is comfortably wearable or not. For this, it should be small, lightweight, and able to provide the information content for which it is meant [10]. The positioning of wearable sensors along with the selection of relevant features for different activities pose challenges in this area. One of the limitations of wearable sensors in the prediction of optimal location on the body of the subject which can provide the most relevant features regarding activity classification [10].

Furthermore, various open areas are also explored in this paper so that a possible future research direction could be recognized and in the near future, we will come up with a good solution in this direction.

Section II of this paper presents various application areas of IoT where different types of sensors are used. Then, in section III, an analysis is made to identify the type of sensors that are used in

different IoT applications. This helps various researchers in this area to identify which sensor is suited best for any specific type of application. Sensors are also used for HAR systems. Section IV presents several HAR systems and then categorizes various types of activities that are recognized in different HAR systems. Finally, a new HAR system is proposed in section V that helps the users to identify different activities. And soon, we come up with the detailed design of this HAR system that would be beneficial for the society in general.

## 2. Sensor Applications for a Smarter World

There are various issues and challenges in implementing and operating smart environment applications. Advanced smart applications need sustainability, reliability, and efficiency. A smart environment commonly has various kinds of Sensor nodes that control the current environment condition and energy usage. Sensors give their measurements and observations to the control system and then intelligent algorithms are implemented by the control system in the network. For surveillance and monitoring of smart environment applications, sensors are implemented in wide areas. Various examples are Highway Data Collection, Electric Toll Collection System, Traffic Signal Monitoring, Structural Health Management System, Traffic Management System, etc. Figure 1 presents the major application areas of IoT where IoT is marking its presence to create a smart environment.

In this section, we have presented some of the applications of IoT sensors.

### 2.1. Intelligent Transportation:

It is one of the prominent domains of Smart City application. All aspects of VANET safety involve communication, decision making, observation, and action. Currently, self-driving cars are also high attention to intelligent transportation. And this is possible due to sensors embedded in them [11].

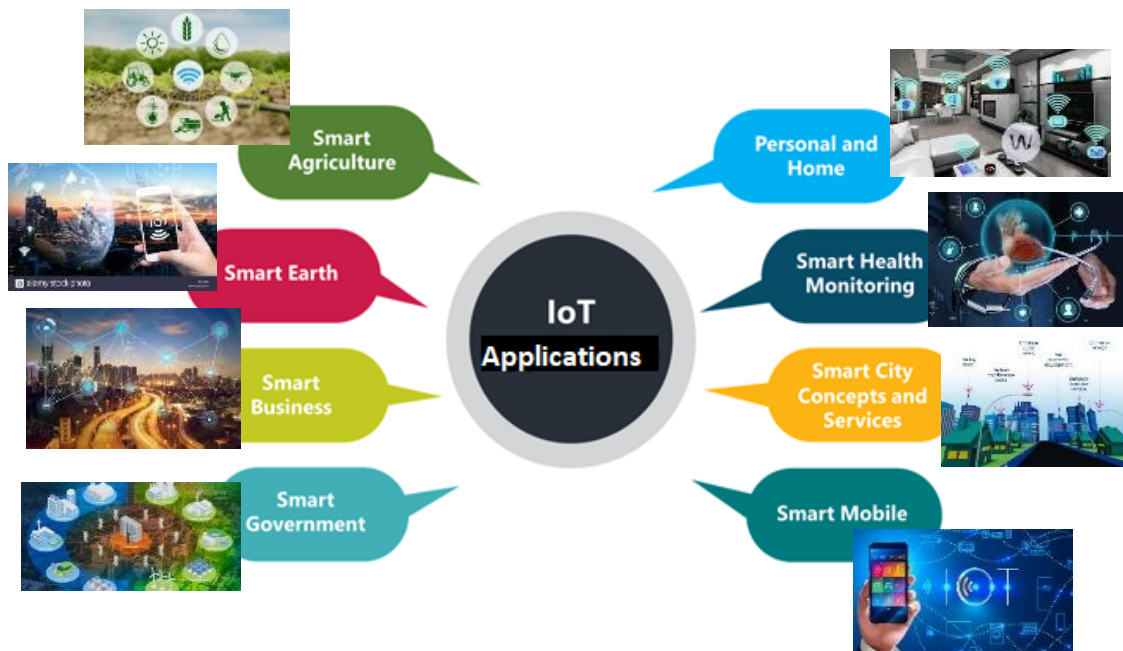


Figure 1: Major IoT Applications

## 2.2. Military Applications

In Defense Operation, Information sharing between soldiers is an important issue for effective combat. But, in the hostile sector, there is a very difficult process for building infrastructure, most of which, communication infrastructure has always become the first target of the enemy. Hence creation and maintaining communication infrastructure is the essential issue of a military operation. In defense Operations, Information sharing between soldiers is an important issue for minimal casualties and effective combats. Sensors are used in various areas in military like Reconnaissance of opposing forces and terrain, Battlefield surveillance, Battle damage assessment, Equipment, and ammunition, monitoring friendly forces.

## 2.3. Medical and Healthcare Applications

In the medical & Healthcare applications, the Sensors promises to solve various issues that are related to medical discoveries on incurable diseases. Most of the Healthcare devices are linked wirelessly through sensors [12]. In medical, IoT is used in identification management (IDM). In IDM each thing (device or user) needed to have a Unique Identification and it ensures identity and information safety. IDM also relays on sensors. Authors in [13], proposed a NIGHTCare platform that overnight detects the abnormal activities and events of patients in a health center. In this application, sensors are attached to the clothes of the patients. Another researcher in the same area presented a low-cost rural e-health monitoring model [14]. A similar smart health application by authors in [15] monitors the physiological parameters of the patients. Sensors that detect the temperature and humidity of a patient's body are used by the researchers in [16]. These sensors are placed inside the shoe insole.

## 2.4. Vehicular Network

VANET usually linked to the internet using APs using WLAN (wireless local area network) Technology such as Sensors, Bluetooth, WiMAX, WI-Fi. Hence sensors play an important role in this area. VANET application: Prevention of Disaster scenarios, Evacuation emergency, Proximity, and correlation analysis, Traffic flows, Congestion, Mobile sensor networks through vehicle communications, Road congestion avoidance, etc.

## 2.5. Smart Cities

Sensors aim to make things convenient and dynamic. Things can exchange data using wireless technology and provides many convenient services. We need new techniques to address various challenges in Smart City. We need to come up with a new approach, where the internet of thing is used to minimize the communication problem of the sensors. IoT deployed in a smart environment comprises IoT devices to transmit the packets to the sink. The Internet of Things encapsulates sensors as an essential component. For surveillance and monitoring of smart environment applications, sensors are implemented in wide areas. Various examples are Highway Data Collection, Electric Toll Collection System, Traffic Signal Monitoring, Structural Health Management System, Traffic Management System, etc [11] [17].

## 2.6. Heterogeneity of Network

Intelligent transportation and smart building involve Sensors to add the different components within the network. In these systems,

sensors need to be capable to co-exist. The network requires accurate mapping of different control information within the Network Stack.

## 2.7. Wired and Wireless Connectivity

Almost all the applications in the smart environment require wireless communication and wireless communication need sensors. The Internet of Things has a discrete topology that ties heterogeneous things/ devices together. Since in the Smart City, the data is inexhaustible and it has an expiration time associated with it means the information must reach the destination within a limited mention time and with good connectivity. Sensors play a vital role to provide good connectivity.

## 2.8. Smart Environment/ Smart Earth

Sensors are also used for monitoring the environmental conditions of the earth to provide a smart environment. It helps in the development of the country. This is not so simple and easy as it seems to be. Monitoring earth is a challenging area because of various constraints like harsh operating conditions and difficulty in maintenance. Also, highly developed and latest advancements make it more challenging [18].

## 2.9. Smart Water

Water is an essential part of life for every living being. It is necessary to monitor the quality of water in terms of pollution, minerals, its level from the ground, and many more. Sensors are widely used to monitor water quality. It's one widely used application is detecting leakages and water overflow in tanks. This not only saves water but also assist in water management [19]. Some of the smart applications of water are detecting water leakages, monitoring seawater pollution level, smart pools, monitoring river floods, dams and reservoirs levels, and many more. All this is possible due to the sensors that are attached to various sites.

## 2.10. Smart Security:

IoT is marking its presence in almost all areas. In this scenario, our day-to-day very personal data is over network all the time. So, it becomes more important to secure this data. Also, sensors are attached to the walls of the containers in various applications for security purposes like intrusion detection systems in homes and offices, monitoring gas levels in sensitive applications, smart monitoring radiation levels, and early detection of leakages. This not only saves a life but also helps in timely management. Sensors are also used for implementing smart perimeters around specific areas that restrict the entry of intruders at that place. A similar application is designed by the authors in [19] where an intelligent intrusion system attached at door, named ADXL 345 accelerometer is given. A similar embedded security framework for IoT is proposed by the researchers in [20].

## 2.11. Smart Agriculture

Sensors are extensively used in several agriculture applications. When attached to the soil, it detects moisture in the soil and by monitoring weather conditions like humidity and temperature it sprinkles water and pesticides in the fields according to the grown crops. An application model is proposed by the



researchers in [21] which detects the humidity, moisture in the soil, rainfall, pH value of soil, temperature, and seed recognition. A similar study in [22] used a drone for smart agriculture. Water sprinkling in IoT agriculture is presented in [23], it is a case of vertical farming.

### 2.12. Smart Homes

Sensors are widely used in creating a smart home. Sensors are attached to various household devices like refrigerator, doors, lights, fan, AC, washing machine, grass, water tank, kids' room, school bags, curtains, tv, and almost to every device or appliance. Its application for smart home includes intruder detection, switching on/off electric appliances, opening/closing doors, sending meeting information, smart cameras for guests, detecting the amount of ingredient in cooking and so many. This all needs sending alerts to provide a better lifestyle. Various researchers worked in this area. Researchers in [24] proposed an intrusion detection system that intimates the owner via messages whenever senses motion in the door.

## 3. Analysis of IoT Applications and Sensors

Different types of sensors are widely used in different IoT applications. After analyzing different applications of IoT, this section recommends which sensors are more suitable and efficient for a particular application. For example, smart city monitors movement on highways which requires position sensor, velocity sensor, accelerometer sensor; toll collection system requires proximity sensor, position sensor, and accelerometer sensor; traffic management in a smart city requires infrared sensor, GPS, and position sensor; and to monitor building structural health, it requires humidity, pressure and temperature sensors.

When talking about a smart environment, it requires monitoring of weather conditions for which a light sensor, temperature sensor, humidity sensor, and gyroscope are required. Whereas, to monitor the emission of harmful gases in the environment, there is a need for chemical sensors, biosensors, and also optical sensors.

Similarly, smart agriculture requires monitoring of soil via humidity and chemical sensors. For monitoring weather conditions for sprinkling water, it requires temperature, humidity sensors. When there is a need for monitoring of intruders in the field it requires proximity or position sensor. A water quality sensor helps in detecting the water level of the soil.

Smart health monitors a patient, so it requires different bio-related sensors that detect abnormal behavior in a patient like biosensors, chemical sensors, magneto sensors, and pressure sensors. To detect the physical location of patients or health providers there is a need of a position sensor.

Different types of sensors are used and required in different IoT applications. Figure 2 shows which type of sensor is used for a particular IoT application.

The analysis of various IoT applications and sensors shows that some applications require more than one type of sensor. Besides, selecting any sensor for an IoT application requires analyzing and identifying different types of activities for which we need sensors.



Figure 2: Types of sensors used in major IoT applications

## 4. Human Activity Recognition (HAR) System

HAR System detects human activities by observing various parameters through different sensors attached to the human body via different wearables. Various researchers worked in this field and proposed numerous HAR systems.





Figure 3: Various types of activities according to category

These monitoring systems are usually smaller and are wearable. Various researchers worked on detecting human activities by observing different parameters. It can be used for the monitoring of people of any age group like for infants, young people, disabled people, and elderly people in a solitary life. Researchers in [25] presented a wearable monitoring system for a patient that monitors various physiological parameters of the patient in healthcare. Notifications are sent to the concerned (doctor/ relative) in any abnormal event via a twitter account. Another similar study in [26] proposed a personalized health monitoring system. This system attaches a sensor inside the shoe that monitors the pressure, temperature, and humidity and sends the health data to the user via a hand-held device. A smart university campus is proposed by the authors in [27] to make smart rooms, smart parking, and smart education a reality. Wearable sensors are mostly used to detect physical movements but these can also be used for monitoring academic performance, analyzing sleep quality/patterns, stress level, and mental health conditions. Researchers in [8] gathered data from 66 individuals and analyzed behavioral and physiological patterns. They also identified factors affecting performance in academics, sleep quality, stress level, and mental health. HAR systems recognize various activities these are generally classified in different groups. Figure 3 enlists these types of activities according to the category.

## 5. Security Mechanism in IoT

As we know that in IoT enabled smart environments, everything is connected via the internet and data is communicated with other similar devices. In IoT based environment, there are approximately 6 times IoT enabled devices than people which are

connected in IoT. Every minute, every second our very personal and sensitive data is over the network. This is serious and couldn't be overlooked. So, this smart environment requires an even more strong security mechanism than ever before. To fulfil the needs of IoT security, strong cryptographic solutions are required. Numerous security algorithms are present now-a-days like AES, DES, RSA, and so many. But using these traditional security algorithms in the IoT environment is not possible because of various constraints. As sensors are the main elements of IoT which are attached to even very small devices. So, implementing traditional security mechanisms in IoT enabled devices is not possible as these are having high memory requirements that are not available in small sensors. Also, there is a different operating environment, low available power, and frequent maintenance in IoT enabled devices is not possible. So, there is a need for such a security mechanism which fulfils the needs of constraint devices. In this direction, there comes a new area of cryptography, which is known as lightweight cryptography. Numerous researchers in the area of IoT security has given different lightweight security algorithms. Moving in the same direction, we have also proposed three different lightweight security algorithms, BRIGHT [28], [29], [30], and UBRIGHT [31]. A total of 9 different security algorithms are proposed in these three security algorithms.

## 6. Proposed Work

In this section, a HAR System is proposed. This system enables users to identify various types of activities. Security is an essential part of any HAR system, so in the proposed system for enhancing the security, one of the newly designed lightweight cipher out of BRIGHT, UBRIGHT, and Expanded-BRIGHT is proposed to be used.



Figure 4: Generic data acquisition architecture for Proposed HAR systems

The first step in any HAR system is collecting raw data using sensors which are then passed through feature extraction and learning processes that recognize the activities. This is achieved through machine learning techniques that may be supervised or unsupervised or semi-supervised.

Figure 4 represents generic data acquisition architecture for the proposed HAR systems. First using different wearable sensors or accelerometer, data are gathered which is sent to some integrated devices where processing is carried out.

Feature Extraction is the next step where the sensors sampling rate is faster as compared to the time taken by human activities. A single sample at any specific instant of time is not sufficient to provide useful information. So, to successfully recognize any activity sampling must be considered for time windows rather than at any specific instant of time. The main underlying problem is that it is not possible to make a comparison between two-time windows. It's almost impossible for signals from the same subject performing the same activity to be identical. So, feature extraction comes into the picture and plays an important role in any HAR systems. There are two types of techniques for feature extraction from time-series data, these are statistical and structural. To extract features from the raw data, statistical methods make use of quantitative characteristics of the data. Fourier transform and the wavelet transform are examples of statistical methods. The methods in which interrelationship among data are considered into account are known as structural approaches.

Learning is a necessary step to build useful patterns that describe and analyze the data. Machine learning is carried out from examples or observations called a training set. Machine learning may be supervised or unsupervised. In unsupervised learning, it is quite not possible to distinguish the activities that are to be recognized. In supervised approaches, the labeling of different activities along with supervising collection process is carried out by some person whereas in an unsupervised learning technique no person is involved. Different researchers in this area are working on both supervised and unsupervised learning techniques. Then

this data is communicated to the concerned and finally stored on some integration device or remote server.

## 7. Conclusion

IoT has numerous applications where different types of sensors are used in different types of applications. This paper discusses different IoT applications and enlightens which smart application demands which type of sensor. Besides, various types of wearable sensors which are used for monitoring of human activities and application area where these are used are also presented in this paper. The positioning of wearable sensors along with the selection of relevant features for different activities pose challenges in this area. Our very personal and sensitive data is over the network in this smart environment. This is serious and couldn't be ignored. So, it requires a strong security mechanism in the form of lightweight cryptography. In this paper, a HAR system is also proposed that enables users to identify various types of activities. To provide high enough security to the proposed system, very efficient and secure crypto solutions are also used.

## References

- [1] D. Sehrawat, N. S. Gill, "Smart Sensors: Analysis of Different Types of IoT Sensors" in 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI) (523-528), 2019. IEEE. <https://doi.org/10.1109/ICOEI.2019.8862778>
- [2] D. Sehrawat, N. S. Gill, "Security Requirements of IoT Applications in Smart Environment", in: Proc. of 2nd International Conference on Trends in Electronics and Informatics (ICOEI 2018), 324-329, 2018. <https://doi.org/10.1109/ICOEI.2018.8553681>
- [3] M. Ambrosin, A. Anzanpour, M. Conti, T. Dargahi, S. R. Moosavi, A. M. Rahmani, P. Liljeberg, "On the Feasibility of Attribute-Based Encryption on Internet of Things Devices", IEEE Micro, **36**(6), 25-35, 2016. <http://doi.org/10.1109/MM.2016.101>
- [4] M. Sehrawat, N. S. Gill, "Deployment of IoT based Smart Environment: Key Issues and Challenges", International Journal of Engineering & Technology, **7**(2), 544-550, 2018. <http://dx.doi.org/10.14419/ijet.v7i2.9504>
- [5] H. Lee, K. Kim, J. Kwon, "A pervasive interconnection technique for efficient information sharing in social iot environment," Int. J. Smart Home, **10**(1), 9-22, 2016.
- [6] S. Kale, M. Satendra, P. Patil, "IOT based Wearable Biomedical Monitoring System," in International Conference on Trends in Electronics and

- Informatics ICEI 2017 IOT, 2017, 971–976. <https://doi.org/10.1109/ICOEI.2017.8300852>
- [7] J. Wilden, A. Chandrakar, A. Ashok, N. Prasad, “IoT Based Wearable Smart Insole,” in *Wireless Summit (GWS), 2017 Global IEEE.*, 2017, 186–192. <https://doi.org/10.1109/GWS.2017.8300466>
- [8] A. Sano, A. J. Phillips, Z. Y. Amy, A. W. McHill, S. Taylor, N. Jaques, N., ... & R. W. Picard, “Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones”, in 2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN) (1-6) , 2015. IEEE. <https://doi.org/10.1109/BSN.2015.7299420>
- [9] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, Y. Amirat, “Physical human activity recognition using wearable sensors” *Sensors*, 15(12), 31314-31338, 2015. <https://doi.org/10.3390/s151229858>
- [10] L. Atallah, B. Lo, R. King, G. Z. Yang, “Sensor placement for activity detection using wearable accelerometers” in 2010 International conference on body sensor networks, 24-29, 2010. IEEE. <https://doi.org/10.1109/BSN.2010.23>
- [11] A. Zanella, N. Bui, a Castellani, L. Vangelista, M. Zorzi, “Internet of Things for Smart Cities,” *IEEE Internet Things*, 1(1), 22–32, 2014. [10.1109/JIOT.2014.2306328](https://doi.org/10.1109/JIOT.2014.2306328)
- [12] D. Zissis, D. Lekkas, “Addressing cloud computing security issues”, *Futur. Gener. Comput. Syst.*, 28(3), 583–592, 2012. Available Online: <https://doi.org/10.1016/j.future.2010.12.006>
- [13] C. Occhiuzzi, C. Vallese, S. Amendola, S. Manzari, G. Marrocco, “NIGHT-care: A passive RFID system for remote monitoring and control of overnight living environment”, *Procedia Comput. Sci.*, 32, 190–197, 2014. <https://doi.org/10.1016/j.procs.2014.05.414>
- [14] R. Chanchal, J. Chaman, W. Arif, “HEMAN : Health Monitoring and Nous”, in *IEEE WiSPNET 2017 conference.*, 2017, 2115–2119. [10.1109/WiSPNET.2017.8300134](https://doi.org/10.1109/WiSPNET.2017.8300134)
- [15] S. Kale, M. Satendra, P. Patil, “IoT based Wearable Biomedical Monitoring System”, in *International Conference on Trends in Electronics and Informatics ICEI 2017 IOT*, 2017, 971–976. [10.1109/ICOEI.2017.8300852](https://doi.org/10.1109/ICOEI.2017.8300852)
- [16] J. Wilden, A. Chandrakar, A. Ashok, N. Prasad, “IoT Based Wearable Smart Insole”, in: *Wireless Summit (GWS), 2017 Global IEEE.*, 2017, 186–192. [10.1109/GWS.2017.8300466](https://doi.org/10.1109/GWS.2017.8300466)
- [17] P. Patil, “Smart IoT Based System For Vehicle Noise And Pollution Monitoring Piyush,” in *Proc. of International Conference on Trends in Electronics and Informatics*, 2017, 322–326. [1109/ICOEI.2017.8300941](https://doi.org/10.1109/ICOEI.2017.8300941)
- [18] M. T. Lazarescu, “Design of a WSN Platform for Long-Term Environmental Monitoring for IoT Applications,” *IEEE J. Emerg. Sel. Top. CIRCUITS Syst.*, 3, 45–54, 2013. [10.1109/JETCAS.2013.2243032](https://doi.org/10.1109/JETCAS.2013.2243032)
- [19] R. Khan, S. U. Khan, R. Zaheer, S. Khan, “Future internet: The internet of things architecture, possible applications and key challenges,” in: *Proc. of -10th Int. Conf. Front. Inf. Technol. FIT 2012*, April 2012, 257–260. [10.1109/FIT.2012.53](https://doi.org/10.1109/FIT.2012.53)
- [20] A. Stango, N. R. Prasad, J. Sen, “Proposed Embedded Security Framework for Internet of Things (IoT),” in: *Proc. of 2nd International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology (Wireless VITAE)*, IEEE, 2011, 1–5. [10.1109/WIRELESSVITAE.2011.5940923](https://doi.org/10.1109/WIRELESSVITAE.2011.5940923)
- [21] P. Patil, V. Sachapara, “Providing Smart Agricultural Solutions/Techniques By Using IoT Based Toolkit,” in: *Proc. of International Conference on Trends in Electronics and Informatics ICEI 2017 Providing*, 2017, 327–331. [10.1109/ICOEI.2017.8300942](https://doi.org/10.1109/ICOEI.2017.8300942)
- [22] A. K. Saha et al., “IoT-Based Drone for Improvement of Crop Quality in Agricultural Field,” *IEEE Internet Things J.*, 612–615, 2018. [10.1109/CCWC.2018.8301662](https://doi.org/10.1109/CCWC.2018.8301662)
- [23] M. Ikhwan, N. M. Thamrin, “IoT Implementation for Indoor Vertical Farming Watering System”, *IEEE Internet Things J.*, 89– 94, 2017. [10.1109/ICEESE.2017.8298388](https://doi.org/10.1109/ICEESE.2017.8298388)
- [24] S. Nazeem Basha, S. Jilani, M. Arun, “An Intelligent Door System using Raspberry Pi and Amazon Web Services IoT”, *Int. J. Eng. Trends Technol.*, 33(2), 84-89, 2016.
- [25] M. Ermes, J. Pärkkä, J. Mäntyjärvi, L. Korhonen, “Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions”, *IEEE transactions on information technology in biomedicine*, 12(1), 20-26, 2008. <https://doi.org/10.1109/TITB.2007.899496>
- [26] L. Bao, S. S. Intille, “Activity recognition from user-annotated acceleration data”, in *International conference on pervasive computing*, 1-17, 2004. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-540-24646-6\\_1](https://doi.org/10.1007/978-3-540-24646-6_1)
- [27] K. Kunze, P. Lukowicz, H. Junker, G. Tröster, “Where am i: Recognizing on-body positions of wearable sensors”, in *International Symposium on Location-and Context-Awareness*, 264-275, 2005. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/11426646\\_25](https://doi.org/10.1007/11426646_25)
- [28] D. Sehrawat, N. S. Gill, “BRIGHT - Proposed Family of Lightweight Block Ciphers for IoT-Enabled Smart Environment”, *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*. 8(9), 584-592, 2019. <https://www.ijitee.org/wp-content/uploads/papers/v8i9/I7778078919.pdf>
- [29] D. Sehrawat, N. S. Gill, “BRIGHT: A Small and Fast Lightweight Block Cipher for 32-bit Processor”, *International Journal of Engineering and Advanced Technology*, 8(5), 1549-1556, 2019. <https://www.ijeat.org/wp-content/uploads/papers/v8i5/E7302068519.pdf>
- [30] D. Sehrawat, N. S. Gill, “Performance Evaluation of Newly Proposed Lightweight Cipher, BRIGHT”, *International Journal of Scientific & Engineering & Systems*, 12(4), 71-80, 2019. <http://www.inass.org/2019/2019083108.pdf>
- [31] D. Sehrawat, N. S. Gill, “UBRIGHT: Ultra BRIGHT, a Tiny and Fast Ultra Lightweight Block Cipher for IoT”, *International Journal of Scientific & Technology Research*, 9(2), 1063-1068, 2020. <http://www.ijstr.org/final-print/feb2020/Ultra-Bright-A-Tiny-And-Fast-Ultra-Lightweight-Block-Cipher-For-Iot.pdf>



## Damage Accumulation Model for Cracked Pipes Subjected to Water Hammer

Zakaria Mighouar\*, Laidi Zahiri, Hamza Khatib, Khalifa Mansouri

SSDIA Laboratory, Hassan II University of Casablanca, ENSET of Mohammedia, Mohammedia, 28830, Morocco

### ARTICLE INFO

Article history:

Received: 30 June, 2020

Accepted: 22 July, 2020

Online: 09 August, 2020

Keywords:

Damage accumulation

Fatigue life

Loading sequence

Pipe

Semi-elliptical crack

Variable amplitude loading

Water hammer

### ABSTRACT

During service, oil and gas pipelines may be exposed to cyclical loads during service, which may lead to a structural failure. Failure is due in most cases to cracks in the structure. Mechanics of the propagation of surface cracks pose a serious challenge and, therefore, models are required to help analyze it. In this study, a nonlinear model is proposed to estimate the accumulation of fatigue damage in the case of defected pipes subjected to a water hammer. The studied pipes are in the presence of a semi-elliptical longitudinal surface crack. This numerical model allows the load sequence to be considered when the structure is under variable amplitude loading. The validated model is used in a parametric analysis, the purpose of which is to determine the influence of the fluid transported and the defect parameters on the evolution of the accumulated damage. The results allow the conclusion of the parameters that have the most impact on the harmfulness of the crack defect as well as the most dangerous cracks in the case of a pipe subjected to water hammer.

### 1. Introduction

Oil and gas pipelines may be exposed to cyclic loads during service which could provoke failure. Failure of pipes could endanger human life, which is why manufacturers comply with very strict safety requirements. Nonetheless, the exploitation of welded structures still presents the risk of catastrophic failures. Failure is in most cases due to the existence of cracks at stress concentration zones such as inclusions, dents and weld defects [1].

Surface cracks, as well as inclusions, dents, corrosion damage, and cracks of the welding bead region, result in the most frequent defects in the welded structures. The propagation of surface cracks presents a substantial problem. Its mechanics need therefore to be studied after its initiation [2]-[4]. Possible crack propagation mechanisms of the welded structures, through the weld metal (WM), the base metal (BM) and the heat-affected-zone (HAZ), should be analyzed for safe operations [4]-[5].

Mechanical elements subjected to variable amplitude loading are susceptible to failure during operation because of fatigue. Load interaction takes place in this case [6]-[7]. When contrasted with the constant amplitude case, the applied sequence load could be responsible for accelerated or delayed crack growth [8]-[11]. Therefore it is necessary, in these conditions, to use an accurate and fast prediction method of fatigue life [12]-[13].

If sudden changes in pressure of the fluid occur in a pressurized system, then the defect of the pipe becomes even more dangerous [14]. When the flow velocity is changed sharply, the kinetic energy is forcibly converted into pressure energy that causes large and rapid changes in pressure inside the pipe. In the pipeline network, this pressure disturbance is propagated at high speed in a form of pressure wave [14]-[16].

The change in pressure caused by this effect, known as the water hammer, may lead to serious failures or leaks, particularly dangerous in the case of large transport systems [16]-[18]. The pipeline operators must characterize the level of risk associated with the threat once the defect has been identified, with the objective of making the right decision quickly [19]. This paper aims to propose a simple but effective nonlinear model to estimate the fatigue damage accumulation in longitudinal welded pipe structures subjected to water hammer wave pressures. This numeric model considers the effect of the loading sequence when the structure is under variable amplitude loading.

After validation of the model against results in the literature, an analysis is carried out to assess the effect of several parameters on the evolution of the fatigue damage accumulation. This study takes into account the location of a semi-elliptical longitudinal surface crack (base metal, weld metal and heat affected zone), its dimensions, its position (internal and external) and the transported fluid (water, light oil and heavy oil). This study will also shed light on the most harmful crack defects.

\*Corresponding Author: Zakaria Mighouar, [zakaria.mighouar@gmail.com](mailto:zakaria.mighouar@gmail.com)



## 2. Fatigue failure

Fatigue is a process that, in the event of stresses or deformations varying over time, alters the local properties of a material. This may be magnified by surface irregularities, defects or structural discontinuities [20]. The major phases of structural fatigue failure are the formation of a crack called initiation, the steady propagation of crack followed by a rapid and unstable fracture of the structure [1], [3], [7].

In order to develop predictive capabilities, various theories have been advanced to characterize the fracture process, such as Linear Elastic Fracture Mechanics (LEFM) and Elastic-Plastic Fracture Mechanics (EPFM) [21]. LEFM is more appropriate for predicting material failure when the load response is elastic and the fracture response is brittle [22].

Fracture mechanics can be useful in predicting the life of components that are subject to time-dependent crack growth mechanisms. The cracking rate can be correlated with fracture mechanics parameters, such as the stress-intensity factor (SIF) and the critical crack size ( $a_c$ ). Failure can be estimated if the fracture toughness of the material ( $K_{IC}$ ) is known.

Paris et al. [23] initially suggested that the crack growth rate ( $da/dN$ ) could be associated with the stress intensity factor range ( $\Delta K$ ) if the material is subjected to variable loading of constant amplitude (1). The material coefficients ( $C$ ) and ( $m$ ) are obtained experimentally.

$$\frac{da}{dN} = C \Delta K^m = C (K_{MAX} - K_{MIN})^m \quad (1)$$

$C, m$ : Paris law constants which depends on the material  
 $K_{MIN}$ : Minimum stress intensity factor ( $MPa \cdot m^{0.5}$ )  
 $K_{MAX}$ : Maximum stress intensity factor ( $MPa \cdot m^{0.5}$ )

Several experiments have shown that the length of the crack varies exponentially as a function of the number of cycles [1], [4], [9]. From this, it can be concluded that crack growth is rather slow until the final phase of fatigue life, where a relatively short number of cycles could lead to rapid crack growth resulting in failure. The initial crack length appears to be a very significant parameter for fatigue life assessment.

Numerous crack growth models have been suggested to predict the fatigue life of the structures. These models relate ( $da/dN$ ) to the amplitude of stress or the maximum stress that can be expressed by the stress intensity factor.

Paris' Law is the simplest fatigue crack growth equation that gives the rate of fatigue crack growth [23]. A modified form which uses the effective stress to account for the crack closure [24-25] can be written as in (2).

$$\frac{da}{dN} = C \left( 0.25 K_{MAX} + 0.5 K_{MIN} + 0.25 \frac{K_{MIN}^2}{K_{MAX}} \right)^m \quad (2)$$

Crack closure is a fatigue loading phenomenon where, even if an external load is acting on the material, the contact is maintained between the opposite faces of the crack [7]. With an increase in load, a critical value of stress intensity factor ( $K_{OP}$ ) is achieved when the crack is "open" (Figure 1).

Previous work by the authors [25] outlines the process of calculating the SIF where the correction factors for the analytical model have been introduced (2). These factors are used to enhance the results obtained in the case of high internal pressure and help to take into account the position of the crack (internal or external).

This process will be used in this paper to propose a method for assessing damage caused by a longitudinal semi-elliptical crack present in a metallic pipe. This pipe is considered to be subject to variable amplitude loading (internal pressure) caused by the water hammer phenomenon.

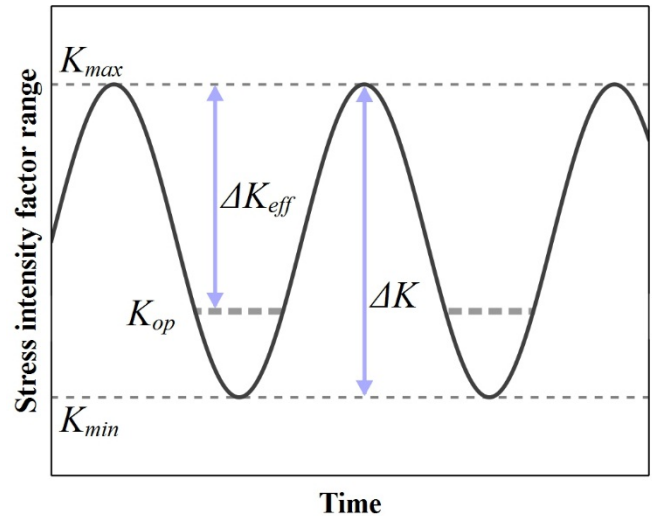


Figure 1: Crack closure effect

## 3. Damage accumulation model

Fatigue loading of pipeline structures varies by nature. Figure 2 shows a typical operating pressure spectrum for a pipeline. During fatigue loading, the pipeline material undergoes alterations resulting in damage [26]-[28]. It is therefore essential that models explicitly take into account the variability of fatigue loading.

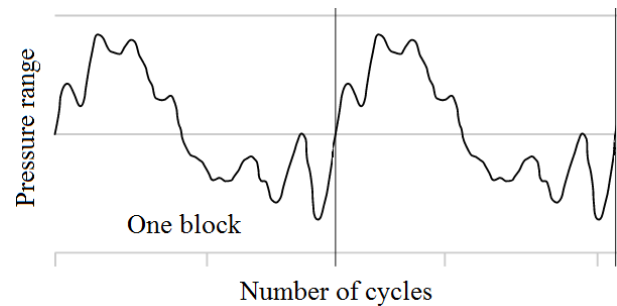


Figure 2: Typical operation pressure spectrum

The models used to describe fatigue-related damage accumulation can currently be classified into two categories: linear and non-linear approaches. Palmgren-Miner rule (3) is a linear damage accumulation approach that is widely used in engineering equipment because of its simplicity [29]. Damage ( $D$ ) per stress level is estimated to be the ratio of the number of cycles ( $n$ ) to the estimated lifetime of fatigue ( $N$ ). Total damage is given by summing all estimated damage values. The summation should be equal to 1 or a given critical value as a failure criterion.

$$D = \frac{n}{N} \quad (3)$$

It is well known that the Palmgren-Miner Rule does not take into account the load sequences, load conditions and load interaction effects [30]. Many methods of fatigue damage accumulation have been suggested to fix the downside of this rule, and a large number of these models are based on nonlinear accumulation laws [31]-[33]. Load sequences and interaction effects are, in general, very important issues in the fatigue damage accumulation.

The main purpose of this paper is to propose a nonlinear model for fatigue damage accumulation based on fracture mechanics parameters. In this numerical model, the effect of the load interaction is considered. This damage model is described in equation (4). The stress intensity factor ( $K_I$ ) and the crack depth ( $a$ ) for a given stress level are estimated using a model developed by authors in [25]. This model allows the calculation of damage induced by a single pressure cycle and is defined as the maximum value between the ratio ( $a/t$ ) and ( $K_I/K_{IC}$ ), where ( $t$ ) is the thickness of the pipe wall.

The damage accumulation values vary between a minimum value and 1 (or a specified critical value). The minimum value corresponds to the damage value calculated with the initial crack size. The value 1 signifies the rupture of the pipe.

$$D = \max\left(\frac{a}{t}; \frac{K_I}{K_{IC}}\right) \quad (4)$$

In the case where the value of 1 is caused by the ratio ( $a/t$ ), it can be concluded that the propagation of the crack induces a leak in the structure. Otherwise, we will conclude that there has been a break ( $K_I/K_{IC}$  ratio equal to 1).

The model (4) alone does not allow to estimate the damage induced by an entire pressure spectrum. This model is intended to be used in an algorithm described in Figure 3. This allows the load interaction effect to be taken into account when the pipe is subjected to static or variable loading.

The algorithm for calculating accumulated damage developed hereby is based on the work of Thun et al. [34]. They estimate the damage accumulation using an analytical model. In this paper, the approach is based on numerical calculation of the damage accumulation and is illustrated in Figure 3. In this Figure, the pressure peak ( $p_i$ ) is applied for ( $n_i$ ) cycles. This algorithm is developed using C language and permits to estimate the damage accumulation after the application of ( $N_b\_Blocks$ ) cycles. Each cycle comprises ( $k$ ) pressure peaks. The pressure peaks correspond to the sudden variation of the internal pressure in the pipe due to the phenomenon of water hammer.

The first step of the algorithm consist on calculating the damage ( $D_1$ ) due to the first pressure peak ( $p_1$ ) and the corresponding number of cycles ( $n_1$ ) is estimated using the proposed damage model (4). This damage ( $D_1$ ) is used to compute ( $N_{2,eq}$ ) that induce the same damage but this time using the second pressure peak ( $p_2$ ) in this case). This step converts the previous damage into equivalent cycles.

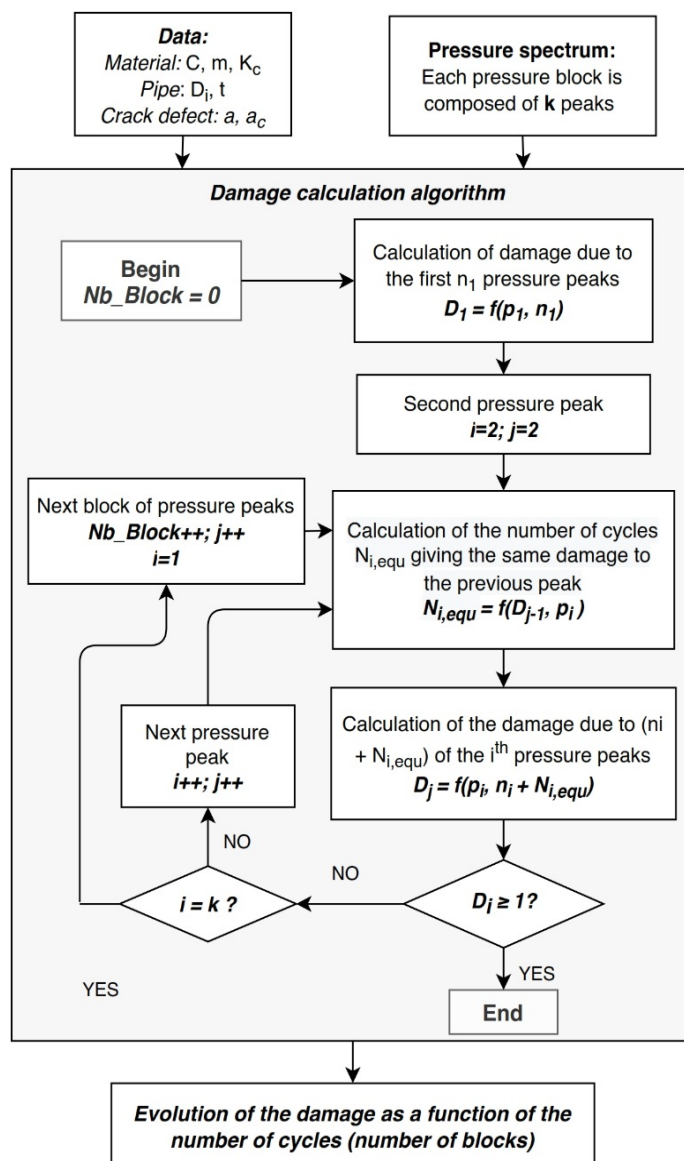


Figure 3: Damage accumulation algorithm

To calculate the damage ( $D_2$ ) after the second pressure peaks ( $p_2$ ) (and the previous ones), the number of cycles ( $N_2$ ) for the second stress level is added to ( $N_{2,eq}$ ). The damage model (4), taking the dimensions of the crack and the constant amplitude pressure  $p_2$  for ( $n_2+N_{2,eq}$ ) cycles as parameters, is used to estimate  $D_2$ .

The same steps apply for the following pressure peaks until reaching the last peak of the block. As long as the value of the calculated damage does not exceed the critical value (here 1), the algorithm continue by resuming to the first pressure peaks and incrementing the number of cycles by 1 (one block equal to one cycle). The value of the estimated damage takes the previous damages into consideration on the basis of this concept.

It should be noted that, this model takes as input the pressure spectrum due to water hammer. This can be achieved by solving a system of equations that describes the transient flow of fluid in the pipe using the Method of Characteristics (MOC) [14, 17, 18, 25]. A more detailed study is led by authors in [25]. Only the results needed to build the study will be presented in this paper.

#### 4. Water hammer

The water hammer is one of the most significant examples of a transient flow in pressurized pipelines. The history of the study of this phenomenon dates back to the 19th century [35]. The first approaches to mathematical formulation resulted in the most common formula that relates the variation in maximum pressure ( $\Delta p$ ) to the change in fluid velocity before and after a sudden valve shift ( $\Delta v$ ) [35] as mentioned in (5). Where ( $\rho$ ) is fluid density, and ( $c$ ) is wave propagation celerity.

$$\Delta p = \rho \cdot c \cdot \Delta v \tag{5}$$

The above method is a useful description of the phenomenon from a practical point of view. In view of its simplicity, equation (5) has made it possible for transient flow cases in elastic pipes, to estimate the maximal pressure range due to the water hammer. This has been seen to be satisfactory for many years. While this formulation is still common, more complex and accurate descriptions of the phenomenon have been developed.

The mathematical description of transient flow in pipelines is expressed by a set of two partial differential equations [14, 25, 35], i.e. the continuity equation (6), the momentum equation (7). It is applicable in the case of elastic pipe and describes the analytical pattern of the water hammer model for compressible fluid flow.

$$\frac{\partial p}{\partial t} + \rho c^2 \frac{\partial V}{\partial x} = 0 \tag{6}$$

$$\frac{\partial V}{\partial t} + \frac{1}{\rho} \frac{\partial p}{\partial x} + \frac{\lambda V |V|}{2D_i} = 0 \tag{7}$$

Where:

- p: fluid pressure (MPa)
- V: Instantaneous fluid velocity (m/s)
- t: Time (s)
- x: One-dimensional axis (m)
- $\lambda$ : Darcy-Weisbach friction coefficient
- $D_i$ : Internal pipe diameter (m)

The method of characteristics (MOC), characterized by rapid convergence and high accuracy of estimation outcomes, can be used to easily resolve this partial differential equation system [36-37]. In the computation, the initial conditions used are  $p_0=56.05MPa$  and  $V_0=1.2m/s$ . The value of  $p_0$  corresponds to the nominal pressure and  $V_0$  corresponds to the fluid velocity at the inlet of the pipe.

The obtained curves, using the MOC, describe the pressure variation of the fluid within the pipe as time-based. Fatigue life can only be predicted through the proposed numerical approach using pressure variation against the number of cycles. In an earlier study [37], it is detailed about how to convert time-based pressure spectrum to pressure spectrum according to the number of cycles by applying filters.

In this way, the pressure spectrum caused by the water hammer phenomenon can be estimated when a valve is suddenly closed/opened. Thus, the effect of this dynamic load on crack defects in the metallic pipe is simulated and the time for calculation is greatly optimized.

Three separate fluids are included in this study, their characteristics are shown in Table 1. Figure 4 shows the pipe dimensions. The pipe is made from the material API X52 with a Young modulus of 200GPa and a Poisson ratio of 0.3. As the pressure considered in this study does not exceed the specified minimum yield strength (SMYS), a linear elastic behavior of the material is considered. The friction factor  $\lambda$  is 0.019.

Table 1: Characteristics of the studied fluids [25]

Fluid	$\rho$ (kg/m <sup>3</sup> )	C (m/s)
Water	998,2	1366,03
Light Oil	813	1463,21
Heavy Oil	962	1105,13



Figure 4: Pipe dimensions

Figure 5 indicates the outcome of that process for the three separate fluids when the valve opens and closes. The pressure drops to 26.85, 26.36 and 25.98 MPa during the sudden closing of the valves for water, light oil and heavy oil respectively.

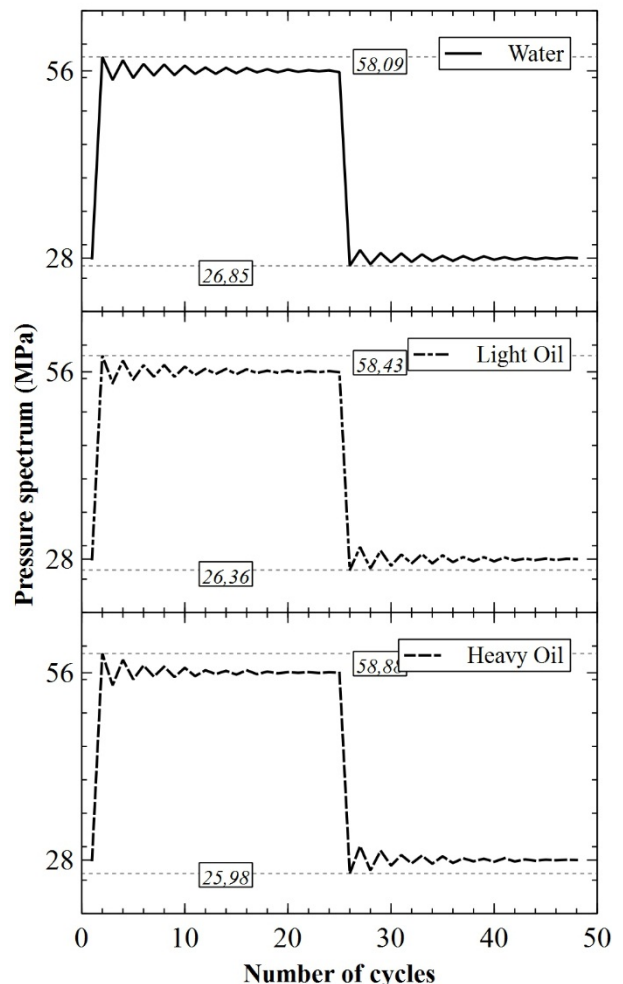


Figure 5: Filtered water hammer pressure waves for three different fluids (One block of pressure spectrum)

The maximum pressures reached are 3.73, 4.34 and 5.14% higher than the nominal pressure (56.05 MPa) of the fluid for water, light oil and heavy oil respectively. The simple equation (5) predicts increases in pressure due to the effect of water hammer of 2.43, 2.12 and 1.89% for water, light oil and heavy oil respectively.

This shows that taking more proper method to estimate pressure variation in case of water hammer when estimating the life of pipes will allow more reliable results. These pressures spectrums represent one block of variable amplitude pressure range. This can be used to assess the effect of water hammer in the case of different fluids on the integrity of the pipe structure.

The method used to calculate the pressure difference induced by water hammer provides results superior to those predicted by (5). Note that (5) does not provide any information on the cycles count until water hammer waves are attenuated. It is therefore important for industrialists to use a more accurate method, taking into account the effect of the water hammer. This is particularly important for the analysis of defects in highly pressurized pipeline structures.

### 5. Validation study

In order to validate the damage estimation approach proposed in this paper, we will compare the results obtained from our model with those published by Zoran [38, 39]. Pipe and fluids properties will be adapted for this purpose.

The proposed approach will be compared in the case of the two types of loading: constant (Figure 6) and variable (Figure 7) amplitude loading. In both cases, the pipe contains an internal semi-elliptical crack in longitudinal orientation. Initial crack size is 0.4 mm for all cases with a ratio (crack depth / crack width) equal to 0.4. In the case of constant amplitude loading, the evolution of the propagation of the crack is calculated for four internal pressures: 40, 50, 65 and 75MPa. The results are shown in Figure 6. One can see from the Figure 6 that the evolution of the number of cycles as a function of the crack depth obtained from the proposed model follows that of the literature.

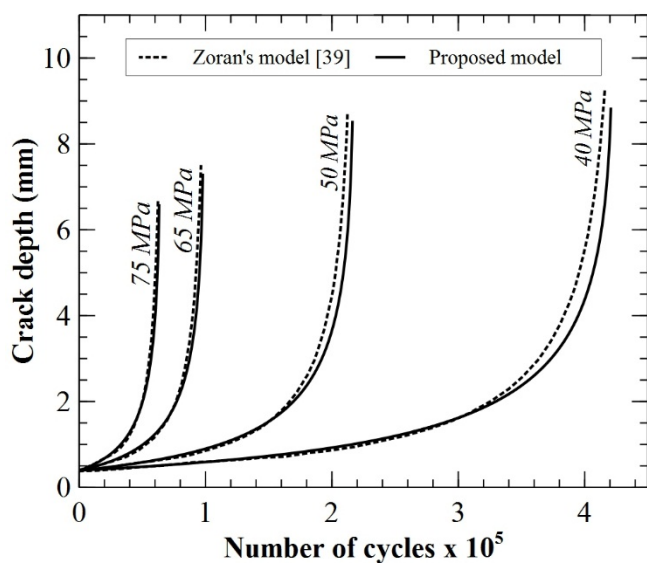


Figure 6: Crack propagation vs. number of cycle in the case of constant amplitude pressure spectrum

In the case of variable amplitude loading, Zoran [39] defined two types of loadings, one heavy and one medium. Five standardized pressure values (20, 35, 45, 60 and 70 MPa) are considered for each type of loading.

In Figure 7, the references EP70, EP60, EP45, EP35, and EP20 are kept the same from that of the literature. For example, in the codification EP70, 70 represents the equivalent pressure range of the pipe [39].

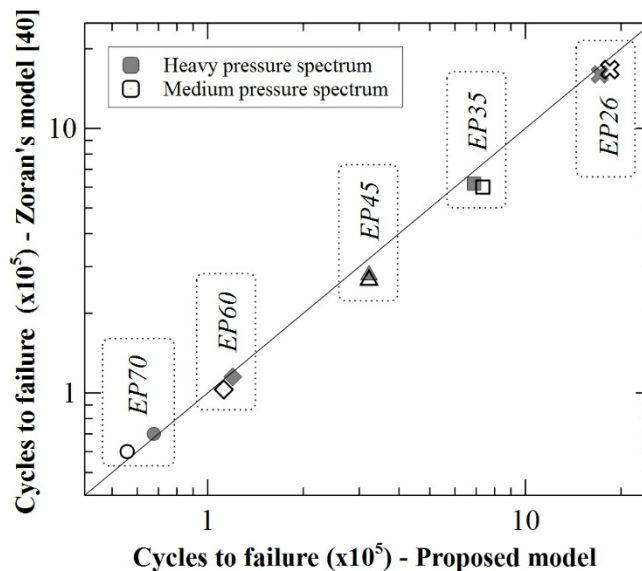


Figure 7: Crack propagation vs. number of cycle in the case of variable amplitude pressure spectrum

As shown in Figures 6 and 7, the deviance between the results obtained from the model developed and those of the literature remain acceptable. Which validated the proposed model.

### 6. Parametric study

In order to study the effect of the water hammer on the severity of the crack defect in the pipe, 54 cases will be investigated. In all cases, the initial crack depth is the same and equal to 3.25 mm. A semi-elliptical surface crack is the defect being studied.

The proposed algorithm (Figure 3) takes as input the pressure spectrum presented in Figure 5 to simulate a pipe transporting water, light oil and heavy oil. The crack dimensions (a/c) are 0.4, 0.6 and 0.8. The ratio (a/c) corresponds to the ratio crack depth to crack width.

The crack is considered to be located in the base metal, the welding metal and the heat-affected-zone. The parameters of the Paris Law (C and m) along with the critical stress intensity factor used for each zone are shown in Table 2. The process for obtaining the Paris Law and  $K_{IC}$  parameters can be found in [40] and [41] respectively. The position of the crack (internal or external) is also being studied.

Table 2: Material properties

Material	C <sup>[40]</sup>	m <sup>[40]</sup>	K <sub>IC</sub> <sup>[41]</sup> (MPa.m <sup>0.5</sup> )
Base metal	3.3 e-09	2.74	53.36
HAZ	1.13 e-09	3.25	53.36
Weld metal	1.04 e-09	3.28	61.02



## 7. Results and discussion

Figures 8, 9 and 10, along with Table 3, present the results of the numerical damage estimation model developed and validated hereby. The curves show an exponential evolution of the damage until the rupture or leakage of the structure (when the damage accumulation is equal to unity). This is typical for the propagation of longitudinal cracks in the metal pipe under cyclic pressure.

In order to obtain readable curves, Figure 8 groups together the evolutions of the damage accumulated only for the internal surface cracks present in pipes transporting water. This curve will allow us to evaluate the effect of the crack dimension on the evolution of the damage accumulation. Figure 9 contains the results of the accumulated damage for cracks of dimension (a/c) of 0.4 present in pipes transporting water. This will allow us to evaluate the effect of the crack position (internal or external) on the evolution of the damage. Figure 10 regroups the evolutions of the accumulated damage for the internal surface cracks of dimensions a/c of 0.4. This will allow us to evaluate the effect of the fluid transported by the pipeline on the evolution of the damage of the crack defect. Table 3 presents the number of cycles to fatigue for all of the studied cases.

All the curves start with a damage value of 0.237, which corresponds to the initial crack depth (3.25mm) divided by the thickness of the pipe (13.7mm). This shows that for all the cases studied in this paper and for the first peak of the pressure spectrum, we obtain a value of (a/t) greater than (K/K<sub>IC</sub>).

In the following, the lower value of the crack dimension is used as a reference for the calculation of the deviations (the results for a/c equal to 0.8 to calculate the deviation for the cracks a/c equal to 0.6 and likewise for a/c equal to 0.6 and 0.4).

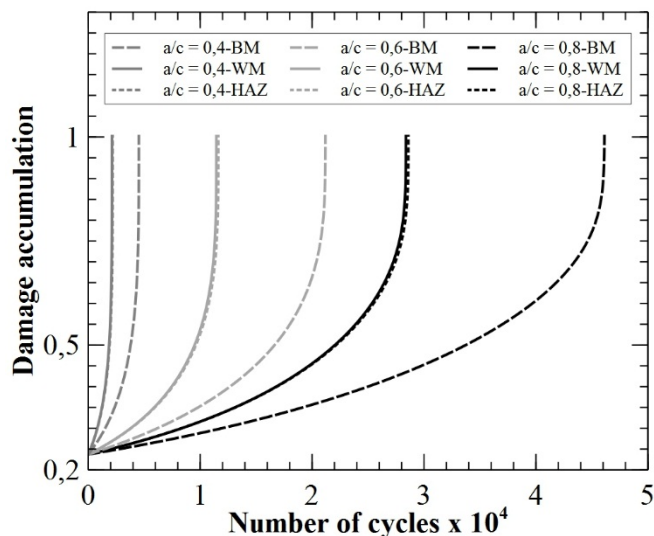


Figure 8: Effect of crack dimensions on the damage accumulation for water and internal cracks

It can be noted from reading the Figure 8 that, for the cracks located in the base material, a drop of -54% in term of fatigue life if crack dimension a / c evolves from 0.8 value to 0.6. From a/c = 0.6 to 0.4, a drop of -78.6% is observed. For the cracks located in the weld material, one notices a drop in fatigue life of -59.7% and -81.5% for respectively the crack with a/c dimension of 0.6 and

0.4. In the case of crack located in the heat affected zone, one notices a fall of -59.4% and -81.3% for respectively the crack with a/c dimension of 0.6 and 0.4. These values of deviations are very close to those observed in the case of a crack located in the welding material.

From these values of deviations, one can clearly see the important effect of the dimension of the crack and more precisely the ratio (a/c) on the fatigue lifetime. In fact, the lower the (a/c) ratio, the more harmful the crack.

From Figure 9, and if we take this time as reference of deviations the values obtained for the internal surface cracks, we observe a difference of + 26.8%, + 31.5% and + 39.4% compared to external surface cracks for defects located in the base material, in the weld material and in the heat-affected area, respectively. This confirms the fact that internal surface cracks are more harmful than external cracks (with same dimensions), regardless of the type of load or the material in which it is located.

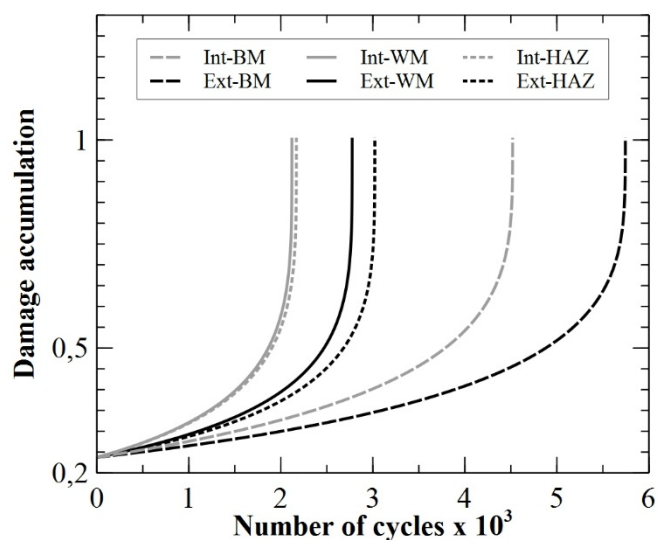


Figure 9: Effect of the position of the crack on the damage accumulation for crack with a/c of 0.4 and water

From a simple reading of Figure 10, we can notice that the evolution of the accumulated damage in the case of a cracked pipe transporting light oil is very close to that obtained for the pipe transporting heavy oil.

Let us take as reference of calculation of the deviations the results obtained in the case of crack present in a pipe transporting water. There is a difference of almost -6.9%, -8.15% and -8.1% in terms of fatigue life between the values obtained for the pipe transporting water and the pipes transporting oil (light or heavy) for cracks located in the base material, the weld material and in the heat-affected area, respectively.

This shows that the transported fluid has an effect on the evolution of the damage and therefore on the propagation of the crack. Admittedly, the effect of the fluid is not as pronounced as the other parameters studied. But if one seeks to obtain a precise evaluation of the damage induced by a defect on the structure, this parameter should not be neglected.

Let us consider the case of an internal crack of dimension (a/c) of 0.4 present in a pipe transporting heavy oil. From Figures 8, 9

and 10 as well as Table 3, the crack defect in the weld material has a deviation of -54.5% in terms of fatigue life compared to a crack in the base material. By taking the same reference, the crack located in the heat affected zone presents a deviation of -53.39%.

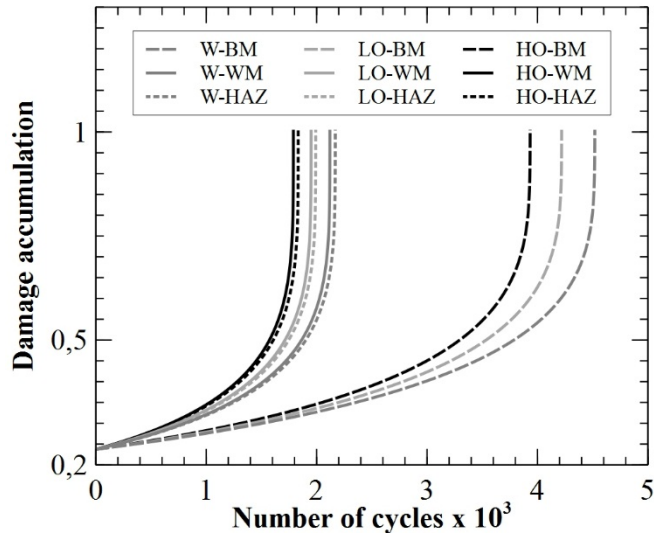


Figure 10: Effect of transported fluid on the damage accumulation for internal crack with dimension of 0.4

It is concluded that the crack is much more harmful when it is located in the weld material or the HAZ than when it is located in the base material. The most harmful crack remains that located in the welding material.

Table 3: Number of cycles to fatigue for all the studied cases

		Number of cycles to fatigue (cycles)						
		Base Metal		Weld Metal		HAZ		
		a/c	Int.	Ext.	Int.	Ext.	Int.	Ext.
Water	0,4		4531	5744	2118	2785	2168	3023
	0,6		21194	28211	11447	15846	11618	17009
	0,8		46117	64454	28382	41250	28600	43965
Light Oil	0,4		4218	5453	1944	2607	1991	2831
	0,6		18790	25506	10006	14127	10161	15170
	0,8		38939	55498	23628	35024	23822	37344
Heavy Oil	0,4		3931	5093	1787	2409	1832	2618
	0,6		16677	22688	8760	12432	8904	13360
	0,8		32914	47016	19701	29354	19881	31322

Int.: Internal cracks. Ext.: External cracks

### 8. Conclusion

A nonlinear model was proposed to estimate the accumulation of fatigue damage in cracked pipes subjected to water hammer. This model allows for the load sequence to be taken into consideration when the structure is subjected to variable amplitude loading, but can also be used in the case of constant amplitude loading.

The model uses the pressure spectrum calculated using the method of characteristics as input. It is based on an algorithm that allows the load sequence to be taken into account. This algorithm uses a mathematical model of the stress intensity factor that incorporates correction factors. One factor is used to improve the

accuracy of the results in the case of high stress, and the other allows the crack location (internal or external) to be taken into account.

After validation of the model against results from the literature, it is used in a parametric study. The objective of this study is to evaluate the effect of certain parameters on the evolution of accumulated damage with the number of fatigue cycles. These parameters are the fluid transported by the pipe (water, light oil and heavy oil), the position (external and internal), the dimension (expressed by the ratio depth of the crack to the width of the defect) and the location of a semi-elliptical crack (base metal, weld metal and heat affected zone). These are the parameters that have the greatest impact on the harmfulness of the crack defect.

On the basis of the results achieved, the significant effect of the crack dimension and, more precisely, the ratio (a/c) on accumulation damage can be clearly seen. In fact, the lower the ratio (a/c), the more harmful the crack is. Internal surface cracks are found to be more harmful than external cracks, regardless of the type of load or material in which they are located. When located in the welding material or in the heat-affected zone, the surface crack is far more harmful than when located in the base material. The most harmful crack is the one in the welding metal.

The effect of the transported fluid is not as pronounced as the other parameters studied, but this parameter should not be neglected if the objective is to obtain a precise assessment of the damage caused by a structural defect.

This model will be used in further studies to analyze the effect of other types of crack defects on the integrity of a pipe structure subjected to a water hammer or even a high internal pressure.

### References

- [1] L. Zahiri, Z. Mighouar, H. Khatib, K. Mansouri, B. Salhi, "Fatigue life analysis of dented pipes subjected to internal pressure" *International Review of Mechanical Engineering*, **11**(8), 587–596, 2017. <https://doi.org/10.15866/irem.v11i8.12089>
- [2] M. A. N. Beltrão, E. M. Castrodeza, F. L. Bastian, "Fatigue crack propagation in API 5L X-70 pipeline steel longitudinal welded joints under constant and variable amplitudes" *Fatigue & Fracture of Engineering Materials & Structures*, **34**(5), 321–328, 2011. <https://doi.org/10.1111/j.1460-2695.2010.01521.x>
- [3] K. Pawan, M. Mamookho, S. Srijan, D. Aniket, "Prediction of the Propagation of Fatigue Cracks in Part-Through Cracked Pipes with CASCA and FRANC2D" *Transactions of the Indian Institute of Metals*, 2020. <https://doi.org/10.1007/s12666-020-01886-z>
- [4] S. A. Barter, L. Molent, P. White, B. Dixon, "Recent Australian full-scale F/A-18 fatigue tests" *Structural Integrity and Life*, **9**(2), 89–112, 2009.
- [5] M. Manjgo, A. Sedmak, B. Grujić, "Fracture and fatigue behaviour of NIOMOL 490K welded joint" *Structural Integrity and Life*, **8**(3), 149–158, 2008.
- [6] B. K. C. Yuen, F. Taheri, "Proposed modifications to the Wheeler retardation model for multiple overloading fatigue life prediction" *International Journal of Fatigue*, **28**(12), 1803–1819, 2006. <https://doi.org/10.1016/j.ijfatigue.2005.12.007>
- [7] K. D. Singh, K. H. Khor, I. Sinclair, "Finite element and analytical modeling of crack closure due to repeated overloads" *Acta Materialia*, **56**(4), 835–851, 2008. <https://doi.org/10.1016/j.actamat.2007.10.046>
- [8] M. A. Meggiolaro, J. T. P. de Castro, "On the dominant role of crack closure on fatigue crack growth modeling" *International Journal of Fatigue*, **25**(9-11), 843–854, 2003. [https://doi.org/10.1016/S0142-1123\(03\)00132-4](https://doi.org/10.1016/S0142-1123(03)00132-4)
- [9] H. Dirik, T. Yalçinkaya, "Fatigue Crack Growth under Variable Amplitude Loading through XFEM" *Procedia Structural Integrity*, **2**, 3073–3080, 2016. <https://doi.org/10.1016/j.prostr.2016.06.384>
- [10] K. Kurmoiartseva, P. Trusov, "Multilevel description of damage accumulation in titanium and titanium alloys" in *AIP Conference Proceedings*, **2053**, 030032, 2018. <https://doi.org/10.1063/1.5084393>

- [11] L. Muys, J. Zhang, N. Micone, W. De Waele, S. Hertelé, "Cycle-by-cycle simulation of variable amplitude fatigue crack propagation" *International Journal Sustainable Construction & Design*, **8**(1), 2017. <https://doi.org/10.21825/scad.v8i1.6808>
- [12] A. Vojdani, G. Farrahi, "Reliability assessment of cracked pipes subjected to creep-fatigue loading" *Theoretical and Applied Fracture Mechanics*, **104**, 1–12, 2019. <https://doi.org/10.1016/j.tafmec.2019.102333>
- [13] J. R. Mohanty, B. B. Verma, P. K. Ray, "Prediction of fatigue crack growth and residual life using an exponential model. Part II: mode-I overload induced retardation" *International Journal of Fatigue*, **31**(3), 425–432, 2009. <https://doi.org/10.1016/j.ijfatigue.2008.07.018>
- [14] A. Kodura, P. Stefanek, K. Weinerowska-Bords, "An Experimental and Numerical Analysis of Water Hammer Phenomenon in Slurries" *Journal of Fluids Engineering*, **139**(12), 1–9, 2017. <https://doi.org/10.1115/1.4037678>
- [15] C. Cristoffanini, M. Karkare, M. Aceituno, "Transient Simulation of Long Distance Tailings and Concentrate Pipelines for Operation Training" in *SME Annual Meeting/Exhibit*, Salt Lake City, UT, 210–214, 2014.
- [16] M. Kandil, A. M. Kamal, T. El-Sayed, "Effect pipes material on water hammer" *International Journal of Pressure Vessels and Piping*, **179**, 2019. <https://doi.org/10.1016/j.ijpvp.2019.103996>
- [17] T. Wang, J. Jiang, G. Lan, "Research on Accumulator for Water Hammer Protection of Long-Distance Slurry Transportation Pipelines" in *Sixth International Symposium on Fluid Machinery and Fluid Engineering (ISFMFE)*, Wuhan, China, 1–6, 2014.
- [18] W. Wan, W. Huang, "Water hammer simulation of a series pipe system using the MacCormack time marching scheme" *Acta Mechanica*, **229**, 3143–3160, 2018. <https://doi.org/10.1007/s00707-018-2179-2>
- [19] T. Vilkys, V. Rudzinskas, O. Prentkovskis, J. Tretjakovas, N. Visniakov, P. Maruschak, "Evaluation of Failure Pressure for Gas Pipelines with Combined Defects" *Metals*, **8**(5), 346, 2018. <https://doi.org/10.3390/met8050346>
- [20] A. Okab, K. Alkhazraji, A. Gatta, "Influence of Pressure and Thermal Parameters on Stresses Analysis of Pressurized and Cracked Pipes" *Oriental Journal of Chemistry*, **35**(6), 1640–1646, 2019. <https://doi.org/10.13005/ojc/350604>
- [21] H. F. B. de Oliveira Dias, *Failure Assessment on Effects of Pressure Cycle Induced Fatigue on Natural Gas Pipelines*, Instituto Superior Técnico, 2014.
- [22] A. A. Griffith, "The Phenomena of Rupture and Flow in Solids" *Philosophical Transactions of the Royal Society of London*, **221**(582–593), 163–198, 1921. <https://doi.org/10.1098/rsta.1921.0006>
- [23] P. C. Paris, F. A. Erdogan, "Critical Analysis of Crack Propagation Laws" *Journal of Basic Engineering*, **85**(4), 528–533, 1963. <https://doi.org/10.1115/1.3656901>
- [24] W. Elber, "Fatigue crack closure under cyclic tension" *Engineering Fracture Mechanics*, **2**, 37–45, 1970. [https://doi.org/10.1016/0013-7944\(70\)90028-7](https://doi.org/10.1016/0013-7944(70)90028-7)
- [25] Z. Mighouar, L. Zahiri, H. Khatib, K. Mansouri, Z. El Majid, "Effect of Water Hammer on Pipes Containing a Crack Defect" *International Journal of Mechanical & Mechatronics Engineering*, **18**(3), 25–31, 2018.
- [26] A. Benin, S. Nazarova, A. Uzdin, "Designing Scenarios of Damage Accumulation" in *International Scientific Conference Energy Management of Municipal Facilities and Sustainable Energy Technologies*, 600–610, 2019. [https://doi.org/10.1007/978-3-030-19868-8\\_57](https://doi.org/10.1007/978-3-030-19868-8_57)
- [27] J. M. Torrenti, G. Pijaudier-Cabot, J. M. Reynouard, *Mechanical Behaviour of Concrete*, John Wiley and Sons, Inc., 185–223, 2013. <https://doi.org/10.1002/9781118557587>
- [28] A. Belegundu, S. Nayak, J. Loverich, M. Grissom, "Vibration-Based Damage Accumulation Modeling" in *28th Conference on Mechanical Vibration and Noise*, Charlotte, North Carolina, USA, 2016. <https://doi.org/10.1115/DETC2016-59106>
- [29] M. A. Miner, "Cumulative damage in fatigue" *Journal of Applied Mechanics*, **12**(3), 159–164, 1945.
- [30] S. P. Zhu, H. Z. Huang, Z. L. Wang, "Fatigue life estimation considering damaging and strengthening of low amplitude loads under different load sequences using fuzzy sets approach" *International Journal of Damage Mechanics*, **20**(6), 876–899, 2011. <https://doi.org/10.1177/1056789510397077>
- [31] B. Isojeh, M. El-Zeghayar, F. Vecchio, "Concrete Damage under Fatigue Loading in Uniaxial Compression" *ACI Materials Journal*, **114**, 225–235, 2017. <https://doi.org/10.14359/51689477>
- [32] H. Gao, H. Huang, S. Zhu, Y. Li, R. Yuan, "A Modified Nonlinear Damage Accumulation Model for Fatigue Life Prediction Considering Load Interaction Effects" *The Scientific World Journal*, 2014. <https://doi.org/10.1155/2014/164378>
- [33] S. R. Maitra, K. S. Reddy, L. S. Ramachandra, "Stress Intensity Factor Based Damage Prediction Model for Plain Concrete under Cyclic Loading" *Journal of Materials in Civil Engineering*, **30**(7), 2018. [https://doi.org/10.1061/\(ASCE\)MT.1943-5533.0002289](https://doi.org/10.1061/(ASCE)MT.1943-5533.0002289)
- [34] H. Thun, U. Ohlsson, L. Elfgren, "A deformation Criterion for Fatigue of Concrete in Tension" *Structural Concrete*, **12**(3), 187–197, 2011. <https://doi.org/10.1002/suco.201100013>
- [35] B. E. Wylie, V. L. Streeter, L. Suo, *Fluid Transients in Systems*, Englewood Hills, Prentice Hall, NJ, 1993.
- [36] M. Dallali, M.A. Guidara, M.A. Bouaziz, C. Schmitt, E. Haj-Taieb, Z. Azari, "Accuracy and security analysis of transient flows in relatively long pipelines" *Engineering Failure Analysis*, **51**, 69–82, 2015. <https://doi.org/10.1016/j.engfailanal.2015.03.001>
- [37] Z. Mighouar, L. Zahiri, H. Khatib, K. Mansouri, "Numerical Modeling Of Water Hammer Pressure Waves In Steel Pipes" in *1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, Meknes, Morocco, 2020. <https://doi.org/10.1109/IRASET48871.2020.9092171>
- [38] D. P. Zoran, "Fatigue Crack Growth Prediction For Pipes With Known Initial Crack" *Trends in the Development of Machinery and Associated Technology*, 461–464, 2013.
- [39] D. P. Zoran, "Assessment Of Fatigue Crack Growth In Pipes Subjected To Variable Loading" *Trends in the Development of Machinery and Associated Technology*, 109–112, 2014.
- [40] L. Zahiri, Z. Mighouar, H. Khatib, K. Mansouri, B. Salhi, "Fatigue Behavior Of Longitudinal Welded Pipes Subjected To Cyclic Internal Pressure, Containing Welding Defects" *International Journal of Mechanical Engineering & Technology*, **9**(3), 560–569, 2018.
- [41] A. Belalia, A. Rahmani, G. B. Lenkey, G. Pluvinage, Z. Azari, "Dynamic characterization of API 5L X52 pipeline steel" *Key Engineering Materials*, **498**, 15–30, 2012. <https://doi.org/10.4028/www.scientific.net/KEM.498.15>

## Development of an Adaptive HVAC Fuzzy Logic Controller for Commercial Facilities: A Case Study

Hamidi Meryem\*, Bouattane Omar, Raihani Abdelhadi, Khalili Tajeddine

Signals, Distributed Systems and Artificial Intelligence Laboratory, ENSET, University Hassan II, Mohammedia, 28000, Morocco

### ARTICLE INFO

Article history:

Received: 11 June, 2020

Accepted: 14 July, 2020

Online: 09 August, 2020

Keywords:

Fuzzy logic

Commercial building

Electrical consumption

EMS

HVAC

### ABSTRACT

*This paper is a case study of the integration of an energy management system (EMS) in a commercial building. A detailed analysis of this EMS's impact on energy consumption is presented. High energy demand is very common in commercial and industrial facilities. With this in mind, the present work aims to improve the energy consumption rate within a commercial facility by proposing an energy management system based on a central fuzzy logic controller. Thus, the central controller adapts the energy management to a pre-established schedule taking into consideration primarily the Heating, Ventilation and Air Conditioning (HVAC) of the targeted building.*

## 1. Introduction

The energy consumption is growing considerably in recent years making global warming and climate change an everlasting threat. Commercial activities and individual lifestyles directly impact the energy demand [1], [2]. Therefore, it is very important to exploit energy in a more efficient way without altering the desired comfort [3]-[7]. Heating, Ventilation and Air-Conditioning systems (HVAC) maintain a good indoor air quality through adequate ventilation and provide thermal comfort. However, HVAC systems are among the most energy demanding components within buildings [1], [8]. These elements require about 40% of the overall energy demand causing an excessive and rapid increase in energy consumption leading to a harmful impact on the global environment [3], [9]. In addition, energy consumption is more substantial in residential and commercial facilities [10]. The energy consumption is around 18% in commercial buildings while it is 22% in residential ones for the same components architecture [2], [11]. This energy requirement is due to the poor design and thermic insulation of existing classic commercial buildings and population general behavior [2], [8], [12]. Consequently, new energy efficient technologies and applications are emerging as a key solution capable of facing the building performance challenge [7], [11]-[13]. Furthermore, researchers recently, focus on the management and control of energy in buildings using active control systems within the buildings [8]. Their primary interest is to realize sustainable energy

savings, while improving the energy efficiency and maintaining a proper comfort conditions [14-19]. In this scope, the current work is a case study where we propose a new approach to optimally manage energy in a typical commercial facility. In this building, we establish a very specific controlling algorithm in order to reduce the energy consumption as much as possible. The integrated system targets basically the management of loads following various parameters. In addition to the electrical loads, the proposed controller takes account of other parameters such as the building architecture, customers flow, internal and external temperature, energy consumption behavior. The building architecture and load structure is first described. Then, the HVAC components and specification are briefly presented. Moreover, the energy saving methodology is described and the design and modeling of the Fuzzy logic controller is well depicted in addition to the simulation results and analysis.

## 2. Building description and component models

### 2.1. Outline description

In a commercial facility, the two main energy-consuming loads are light and the HVAC system. Figure 1 summarizes an analysis study conducted on commercial buildings of the same type with a surface between 1,000 m<sup>2</sup> and 3,000 m<sup>2</sup> and present the average loads consumption of these commercial buildings. As shown in Figure 1, the HVAC system consumes about 58% of the total energy; a huge amount compared with the other categories. The

\*Corresponding Author: HAMIDI Meryem, [meryem.ham@gmail.com](mailto:meryem.ham@gmail.com)  
+212 619 152 755

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050463>



lighting system consumes about 17%, followed by office devices with 15% consumption and only 10 % for all the other instruments.

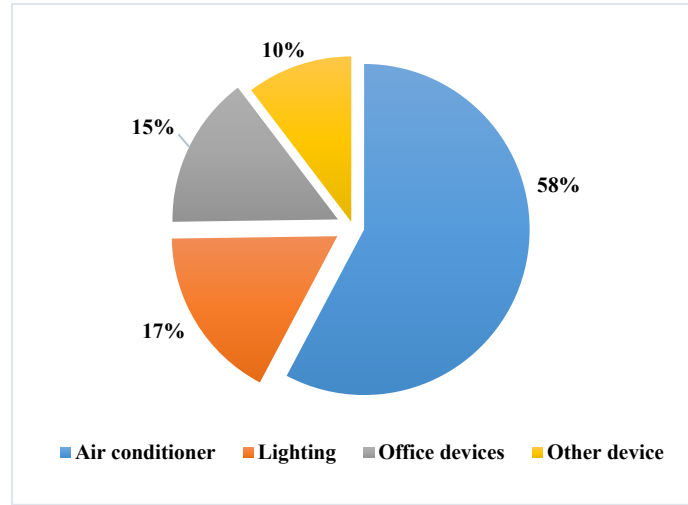


Figure 1: Electrical loads distribution in commercial buildings

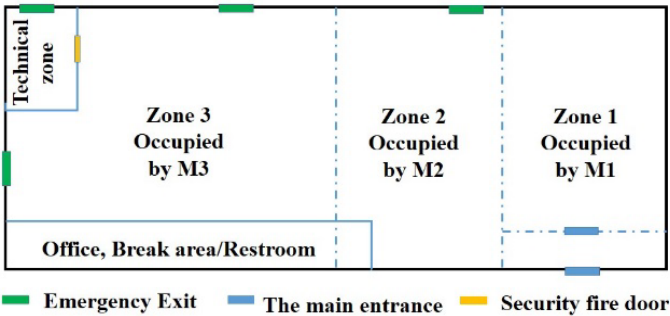


Figure 2: zoning of the HVAC system

The present case study studies a commercial building located at Oujda, Morocco (34° 40' 48" N, 1° 54' 36" W). Oujda city has a Mediterranean climate with a cold, rainy winter and a hot summer. The average annual temperatures are between 15°C and 20 °C. The average temperature in summer is 35 °C and can reach 40 °C. Unlike winter, the average temperature is 18 °C and minimal temperature can drop under 0°C. The building was opened in December 2016, extending over an area of 1600 m<sup>2</sup>, the monthly average electrical consumption is 11879 kWh. The commercial building is equipped with LED lighting system. The HVAC system is divided to three Modules M1, M2 and M3 as presented in Figure 2. The first and second modules control 25 % of the buildings each, the third one is responsible for the rest. The electrical power of the HVAC system is 72KW in average, total power can be divided as follows:

$$\begin{aligned}
 M1 &= M2 = 18 \text{ KW} \\
 M3 &= 2 * M2 = 2 * M1 = 36 \text{ KW}
 \end{aligned}
 \tag{1}$$

In a typical commercial store environment, the parameters such as customer flow, collaborators number may be taken into consideration beside internal and external temperature. Indeed, these parameters can be used to assess more details related to the energy consumption. Different works have studied their effect on the overall user's comfort and building profitability [8, 15].

## 2.2. Modeling approach

So far, several reviews have been published on modeling methods [20-23], control approaches [24 - 26] and optimization techniques [26]. Besides HVAC systems simple modeling and practical energy savings, recent works target the evaluation of overall system performance [27]. The HVAC modeling techniques are widely discussed in the review study conducted in [28]. The presented models have been classified into three different categories physics-based model, data-driven model and gray-box model. Authors have concluded that each model category has both strengths and weaknesses and none of those models could be commonly used as HVAC system modeling. However, authors recommend to select the most suitable modeling technique based on the case study. In the present study, the modeling approach was carried out according to the commercial building specifications and the required energy ( $E_{req}$ ) to meet the loads demand of the building. The global data taken at consideration consist of internal parameters (people, HVAC, lighting) and external ones (sun, air, moisture). Thus, the energy need can be written as:

$$E_{req} = E_{loss} - E_{gain} \tag{2}$$

Where  $E_{loss}$  represents the energy losses through envelope components of the building and  $E_{gain}$  depicts the ventilation/infiltration losses. The energy losses consist of energy transfer outwards from the building in terms of conduction ( $E_{cond}$ ) through envelope components (roof, walls, glazing, ground), ventilation by the HVAC equipment ( $E_{equ}$ ), and infiltration process ( $E_{inf}$ ). This energy can be defined as:

$$E_{loss} = E_{cond} + E_{equ} + E_{inf} \tag{3}$$

While, energy gain is the useful energy to the building from occupants ( $E_{occ}$ ), lighting equipment ( $E_{lig}$ ) and sun ( $E_{sol}$ ). This can be written as:

$$E_{gain} = E_{occ} + E_{lig} + E_{sol} \tag{4}$$

Based on the studied building specifications, some energy losses could be neglected for simplicity purposes. Therefore, the building conditioning system can be described by energy balance model [29] as follows:

$$T_{n+1}^i = \mu T_n^i + (1 - \mu)(T_n^o \pm \eta e_n / A) \tag{5}$$

Where  $T_n^i$  and  $T_n^o$  are respectively the building inside and outside temperature at the time sequence  $n$ ,  $T_{n+1}^i$  is the inside temperature at the time sequence  $n + 1$ ,  $e_n$  represents the electric power input at  $n$ .  $\mu$ ,  $\eta$  and  $A$  represent respectively the inertia factor, the overall thermal conductivity and the thermal conversion efficiency. The term “ $\pm$ ” in “equation (5)” means that the operator “-“ is used when the HVAC is on cooling mode and “+” when heating.

## 3. Energy saving Methodology

### 3.1. Energy repartition in typical buildings

Authors in [2] propose a review on large-scale energy saving studies. The work deals with the governments energy policies capable of generating noticeable energy savings. Authors also propose a scientific approach for quantifying the potential energy savings by improving the customer operations planning in

commercial buildings. The study associates between energy modeling techniques, studies on human behavior in buildings exploitation activities, surveying and sampling methods.

In commercial buildings, less than 50% of office equipment are switched off during off-work hours [1], [2]. In the present case, we also investigate the application of eco-energetic protocols in both operating and non-operating hours. The present study confirmed that the use of traditional eco-energy methods and protocols could save up to 60% of the monthly electric energy consumption. Moreover, authors in [3] confirmed that in HVAC systems, the average energy saving is generally 20–60%, lighting (20–50%), water heating (20–70%) and refrigeration (20–70%). Consequently, it is estimated that in new buildings, an energy economy of up to 50% can be achieved. In almost all cases, HVAC shows the highest share of energy savings, especially in small buildings where HVAC typically accounts for the major portion of the energy use [8], [2].

3.2. Operations time slots

The general electricity cost is not solely related to the total electricity consumption; the daily consumption periods noticeably affect the energy consumption budget [30]. The electricity consumption during a time slot is related to the number of costumers during that period and the classic exterior climate conditions.

In this work, the energy consumption fees are divided into three time slots, each one corresponds to a specific cost. According to [31], there are generally three classic time slots: Off-Peak hours, Peak hours, and Full hours. We have closely studied each period in order to understand the evolution of the energy consumption in each slot. This allows us to propose a schedule for the operation of the HVAC system. Table 1 illustrate the difference between the three time slots, and their consumption cost.

Table 1: General time slots in low voltage power subscription

Time slots	Period	Billing	Subscription cost (MAD/KVA)
	October - April- March - September	Cost (MAD/kWh)	
Peak-Off hours (T1)	22h-7h	23h-7h	C1 0,648
Peak hours (T2)	7h-17h	7h-18h	C2 0,886
Full hours (T3)	17h-22h	18h-23h	C3 1,241

Figure 3 presents records of the monthly consumption in the studied building. The available records in the study are dispatched on 6 months only: from December 2016 to May 2017.

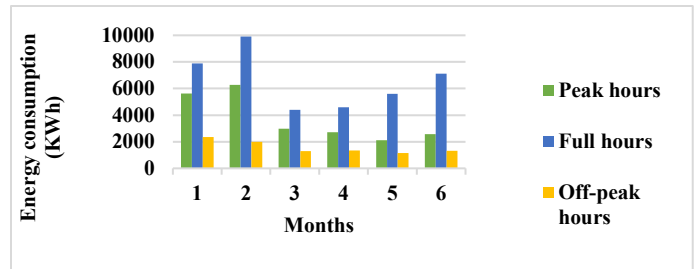


Figure 3: Energy consumption by slot in KWH

During winter season, we have an early sunset which causes a rapid drop in temperature in the evening. Figure 3 illustrates the impact of winter season on the consumption during the designed time slots. As the figure illustrates, the power consumption is at its peak during the Full hours between December and January. This change in consumption is due to the alteration in customers' flow which increases between 5 pm and 9 pm. On the other hand, during the Off-Peak hours, the energy consumption is very low because during closing time, the building consumes a minimum of the energy. The function cost in this system is essentially based on three variables and can be expressed as:

$$Cost = f(x_1, x_2, x_3) \tag{6}$$

Where  $x_1$  is the actual operating time slot,  $x_2$  is related to the number/costumer flow change and  $x_3$  is related to the external temperature. The control system aims to reduce as much as possible the cost function (6). In order to grasp the whole concept, the cost function can be expressed in more details as follows:

$$Cost = T_1 * C_1 + T_2 * C_2 + T_3 * C_3 \tag{7}$$

Where  $T_i$  and  $C_i$  are variables related to the cost function factors. These variables are subject to seasonal changes and power generation cost in the region. "equation (6)" and "equation (7)" are related by the following equation:

$$x_i = T_i * C_i \quad i \in \{1, 2, 3\} \tag{8}$$

Figure 4 and 5 illustrate the variation of the costumers flow and external temperature from 8 am to 10 pm and during a whole 24h period respectively.

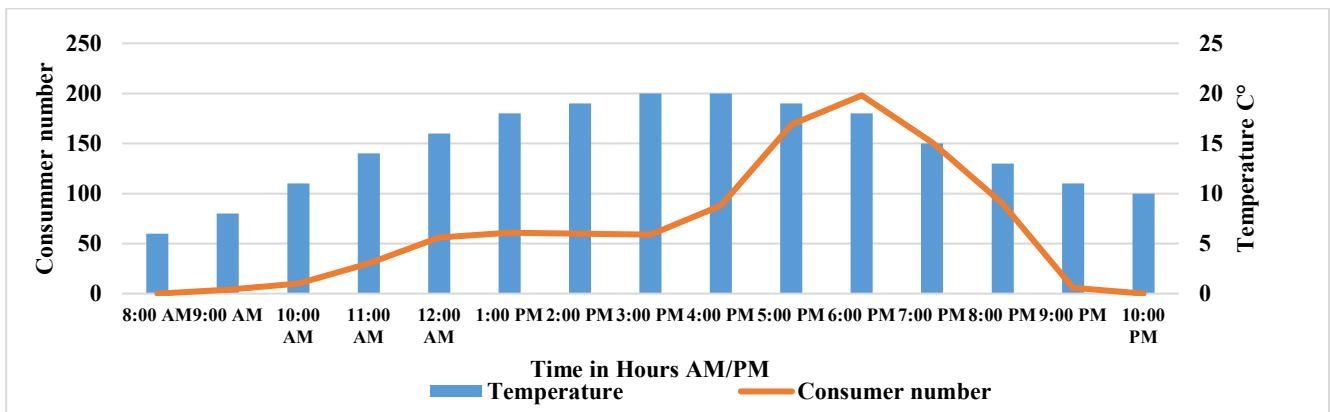


Figure 4: variation of costumers flow and external temperature from 8 am to 10 pm

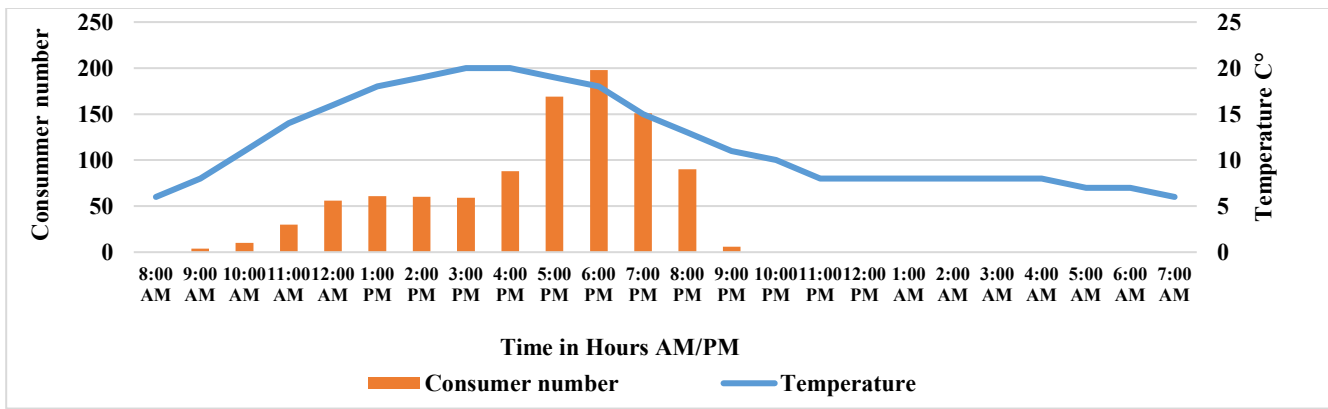


Figure 5 : variation of costumer flow and external temperature during a 24h period.

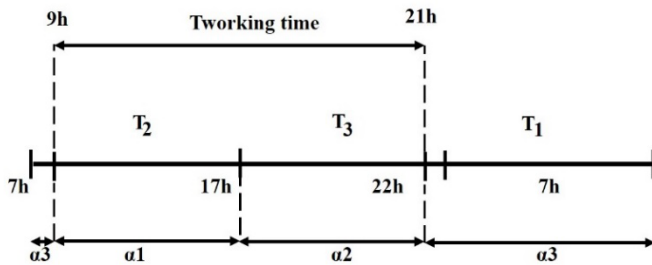


Figure 6: standard working day cost variables

The difference in costumers’ flow in Figure 4 and 5 shows that controlling the HVAC system during off-work periods is not necessary and can guarantee a considerable energy savings. Furthermore, the work conditions, and costumers’ availability are season dependent, thus, the total working time and periods of high activity change gradually along the year. Figure 6 shows the general segmentation of the working time and cost function factors for the studied case.

$\alpha_i$  are directly related to the cost function and represent the portion of time when energy can be represented as entirely consumed inside  $T_i$  slots.

$$Cost = g(\alpha_1, \alpha_2, \alpha_3) \tag{9}$$

On Figure 6 the total working time can be written as:

$$T_{workingtime} = \alpha_1(T_2) + \alpha_2(T_3) \tag{10}$$

Standard conditions in the present case study where  $\alpha_1(T_2) = 8h$ ,  $\alpha_2(T_3) = 4h$  and  $\alpha_3(T_1) = 12h$ , the total working time  $T_{workingtime} = 12h$ . At this point, it is clear that the working time, population condition and external conditions should be controlled depending on time segment. Authors in [ 32] the use of PID as control strategy for HVAC systems which resulted in in a high energy consumption. Whereas, in the same study, authors found that the use of HVAC control method presents the most promising results in terms of energy efficiency.

### 3.3. HVAC Controllers principle

The HVAC system being the key component to energy saving in commercial buildings, studies investigate multiple control strategies in order to encourage their implementation [8,28, 33]. Numerous recent studies have given an emphasis on improving the HVAC system through advanced control algorithms rather

than classic control methods [28, 34]. Thus, some control strategies manage the HVAC system based on a day-ahead electricity price and outdoor temperature forecast [30]. Furthermore, existing controllers are generally classified into two major categories: predictive and adaptive controllers [16]. The study in [ 28] offers a critical overview of the modeling technics currently used in HVAC systems with regard to their applicability and acceptance in real practice, the study also summarizes the weaknesses, applications and services of the modeling technics. Authors in [27] present a study where they examine the performance of each HVAC unit, while evaluating the energy consumption and efficiency of each control technic (FLC, PID and simple ON/OFF time slots).

Control policies, also, can be categorized into two major types: model based and data driven [ 35]. Model-based control strategies require mathematical models as much as the building and the HVAC system specifications to generate minimal control rules. The data driven HVAC control strategies operate by adjusting the management parameters. In [28] authors offer a fairly global review on the different algorithms used to improve the HVAC energy management systems. Most used ones are the data mining algorithms, the Artificial Neural Network (ANN) and the Support Vector Machine (SVM). The work in [27] compared between classic ON/OFF method, PID, Fuzzy logic, Agent based and neural network control. Related to the present case study, we have chosen to base the exploited control strategy on the fuzzy logic system given the promising results obtained [23-27]. Moreover, since the operation of HVAC systems is not linear, classic controllers not suitable for exploitation in such facilities. The common components of Fuzzy Logic Controller (FLC) attached to HVAC systems are well illustrated as block diagrams in various works [27, 36, 37].

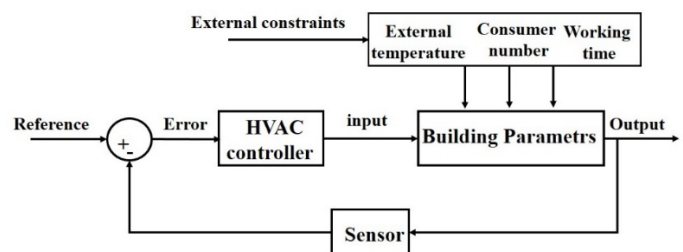


Figure 7: HVAC general control model

The implementation of fuzzy logic based methods requires a qualitative distribution of the inputs and outputs ranges by establishing a direct link to the different comfort zones sought for the clients visiting the facility. Several research works have analyzed this particular point, hence pushing us to adapting the same approach based on the facility’s schedule in present case study [38, 39]. Figure 7 presents a control model of the studied building that depends on the external temperature, consumer number, working time. The three zones of the building are controlled via the Fuzzy logic program taking for each section the main three parameters described before.

**4. The proposed design for the Fuzzy logic controller**

*4.1. Overall description of fuzzy logic controller*

The temperature control is achieved by integrating several parameters such as temperature and humidity inputs[38, 39]. In this case study, we implemented two inputs in the designed model; temperature and number of clients. During the test period, we consider that time slots configuration remains unchanged [40- 42]. Figure 8 illustrates the design adopted in present case study. As it can be understood from the figure, temperature variations and clients’ density inside the building are the inputs while the heating or cooling actions represent the outputs. This design is applied to each module zone  $M_i$  depending on costumers flow and the frequency of temperature variation.

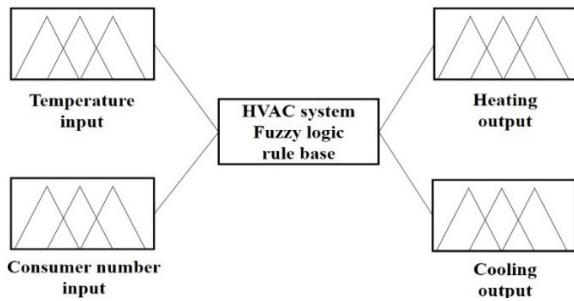


Figure 8: the proposed design of the fuzzy logic controller

In order to enhance the management strategy each of the customers’ flow, opening hours and the temperature is regularly updated by keeping a close feed on present season and any other sudden variations such as vacation periods and time zone change.

*4.2. Input variables’ design*

*a) Membership function for the temperature management*

In order to establish the membership of temperature variation inputs we used the local weather data registered over the previous year. The sections of temperature vary between very cold conditions when temperatures are less than 14°C and hot conditions when temperatures are above 38°C in Oujda city.

For the temperature membership function, it is divided into five ranges as described in Table 2. The “Too cold range” is reserved for any value below 14°C. For the “Cold” range between 7 and 23°C, “Warm” is the range of 15 and 30°C, the “Hot” range between 23 and 38 °C and the final range “Too Hot” above 38 °C. This segmentation was the key to designing the Ranges of temperature inputs in the fuzzy logic model as presented in Figure 9.

The definition or this partition of the membership function ranges, depends on the region and its climatic conditions. That means if we have another case of study in another region the value of the range may change in term of the variation of the local weather data.

Table 2 : Temperature membership functions

Membership functions	Range (C°)
Too cold	Below 14
Cold	7-23
Warm	15-30
Hot	23-38
Too Hot	Above 38

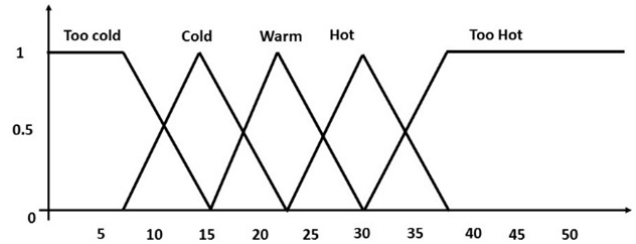


Figure 9 : Ranges of temperature inputs

*b) Membership function for the costumers’ flow*

In this case of study costumers flow is a primary factor in the proposed commercial building EMS. In this perspective, Table 3 shows that the number of membership functions is organized into five categories of clients’ density. The same process as the temperature input was applied for the consumer flow input.

Where the “Very low” range is reserved for 1 to 15 persons followed by the “Low” range between 10 and 30 persons, “Medium” range between 20 and 80 costumers, the “High” range between 60 and 150 peoples and the final range “Very high” is reserved for 120 to 250 costumers. The definition or this partition depends on the region, the building size and the most import the data of the costumers’ flow of the previous years to define the average number of people in each single range. In Figure 10 presents the ranges of clients' density inputs implemented in fuzzy logic model.

Table 3 : Clients' density membership functions

Membership functions	Range (number)
Very low	1-15
Low	10-30
Medium	20-80
High	60-150
Very high	120- 250

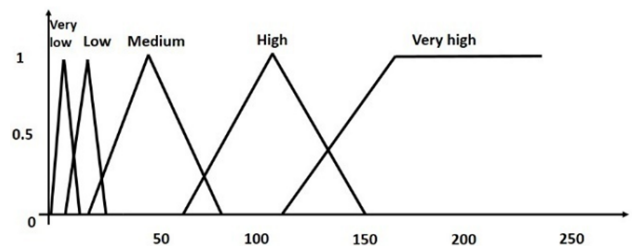


Figure 10: Ranges of clients' density inputs



4.3. Output variables' design

Membership functions and ranges are supposed to be similarly designed for each of the units M1, M2 and M3 constituting the HVAC system. Each unit is endowed with its own sensors and aeration control. The author in [17] describes the design of room temperature and humidity controller using Fuzzy Logic. He considered almost the same model of the distribution of the membership functions except that it was categorized in different ranges for cooling and heating functions.

Table 4 : Controller Output membership functions

Membership functions	Range
OFF	0-0.05
Slow	0-0.45
Medium	0.35-0.65
Fast	0.6-1

For this case of study, the Output membership function is organized in four ranges (OFF, Slow, Medium and Fast) into the interval of 0 and 1, either for the cooling or the heating controller output. The used control output in this case study is achieved by adopting two outputs with the same interval functions as illustrated on Table 4 and Figure 11.

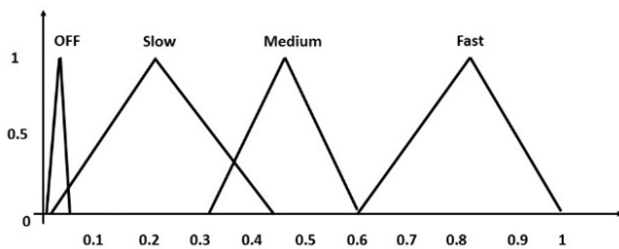


Figure 11: Fuzzy representation of the output membership

4.4. Rule base

The rule base describes the laws and conditions governing the activation of the heating/ventilation system in all the sections M<sub>1</sub>, M<sub>2</sub> and M<sub>3</sub>. Table 5 presents the fixed rules obtained after several operation tests in HVAC system. The control's rule base in Table 5, shows different possible decision making by the fuzzy logic system for different inputs range, and the considered decision that

Table 5 : Control's rules base

Number of consumers	Temperature C°				
	Too cold	Cold	Warm	Hot	Too-Hot
Very low	M1=OFF	M1=OFF	M1=OFF	M1=OFF	M1=OFF
	M2= Heating slow	M2=OFF	M2=OFF	M2=OFF	M2=OFF
	M3= Heating Slow	M3=OFF	M3=OFF	M3=OFF	M3=OFF
Low	M1=OFF	M1=OFF	M1=OFF	M1=OFF	M1=OFF
	M2=OFF	M2=OFF	M2=OFF	M2= Cooling medium	M2=OFF
	M3= Heating fast	M3= Heating fast	M3=OFF	M3=OFF	M3= Cooling medium
Medium	M1=Heating slow	M1=OFF	M1=OFF	M1=OFF	M1=OFF
	M2=OFF	M2= Heating medium	M2=OFF	M2=OFF	M2=OFF
	M3= Heating medium	M3=OFF	M3=OFF	M3=OFF	M3=OFF
High	M1=OFF	M1=OFF	M1=OFF	M1=OFF	M1= Cooling slow
	M2= Heating medium	M2= Heating slow	M2=OFF	M2=OFF	M2= Cooling medium
	M3=OFF	M3=OFF	M3=OFF	M3= Cooling medium	M3= Cooling fast
Very high	M1=OFF	M1=OFF	M1=OFF	M1=OFF	M1= Cooling slow
	M2=OFF	M2=OFF	M2= Cooling slow	M2= Cooling medium	M2= Cooling fast
	M3=OFF	M3=OFF	M3=OFF	M3= Cooling fast	M3= Cooling medium

the fuzzy logic control will made based on the previous input and output membershi

For example, if the Number of consumers is high (between 120-250 consumer) and the temperature is Too Hot (above 39 °C), the decision of the fuzzy logic controller for the three units will be:

$$M1 = \text{Cooling slow.}$$

$$M2 = \text{Cooling fast .}$$

$$M3 = \text{Cooling medium.}$$

Figure 12 presents a planning example based on the fuzzy logic controller decision. Figure 12 presents a one planning example based on a recorded set of data during summer. During morning working hours M1, M2 and M3 are generally all occupied by workers and store staff, this is why they are all initially activated in order to guarantee minimum comfort conditions. Indeed, between 10h and 12h in the example above, M3 is activated using medium cooling mode because temperature is above 38°C and clients are between 30 and 50 individuals.

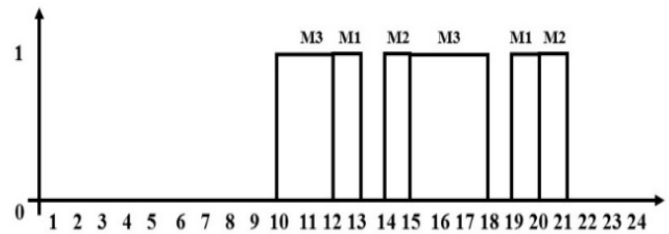


Figure 12 : Example of a HVAC control Schedule

From 12h to 13h, we notice that M3 is still activated but with a change in operating mode (cooling fast) and M1 is initiated in cooling medium. This decision is primarily related to the temperature and clients' number, however M1 will be rapidly switched off, because the controller anticipates the thermal exchange between the units of the building. The same procedure is followed throughout the day; the controller adapts the internal temperature to the external temperature and clients' number.

5. Study results and analysis

5.1. Impact of the proposed schedule on the fuzzy logic controller

Figure 13 illustrates the impact of adopting the fuzzy logic controller on electric consumption fees of the building, the first month corresponds to December 2016 and number 20 to July 2018.

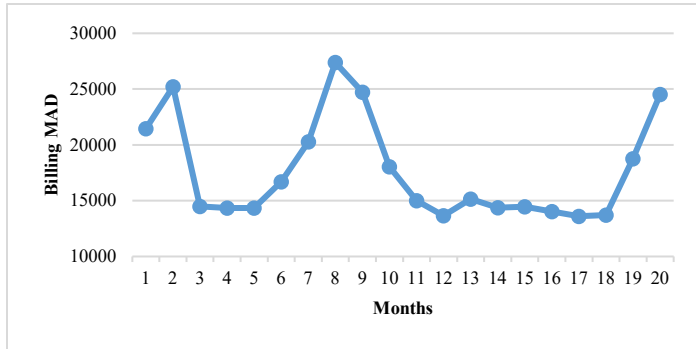


Figure 13: Monthly energy consumption fees in MAD

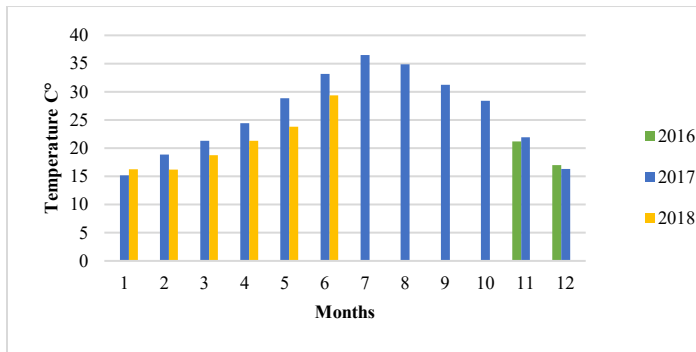


Figure 14: Collected temperature data of the city of Oujda [18].

The control model has been commissioned on June 2017, however it can be seen that there is an increase in power consumption in the month of July 2017 because of the high ambient temperature during the season period and laps of temporary suspension of the controller for technical adaptation. Moreover, the high energy consumption during summer season still poses a technical difficulty even after commissioning the proposed controller. Despite this stiffness we established that a steady decrease of 9.71% energy consumption has been achieved compared with the same period of the previous year. In Figure 14

month corresponds to January and number 12 to December of each corresponding year.

Related to the present case study, we can undoubtedly assume that promising and palpable results have been noticed after only two months from the first introduction of the fuzzy controller. After the first testing period it has been reported to us that the facility has achieved a 30% energy gain. Table 6 presents the recorded energy profit during the testing periods of 2017 and 2018. The recorded gains were calculated based on the following equations:

$$Profit E_b = \frac{\Delta E_b}{E_b} \tag{11}$$

$$Profit E_c = \frac{\Delta E_c}{E_c} \tag{12}$$

Where  $E_c$  is the energy consumption in  $KWh$  and  $E_b$  is the energy fees in  $MAD$ . As it can be seen on Table 6 substantial energy gains have been made which proves the efficiency of the developed controller.

5.2. System simulation and behavior

The ingenuity of the proposed controller is its ability to manage the three areas of the commercial building separately, while taking into consideration the thermal exchange between the three units.

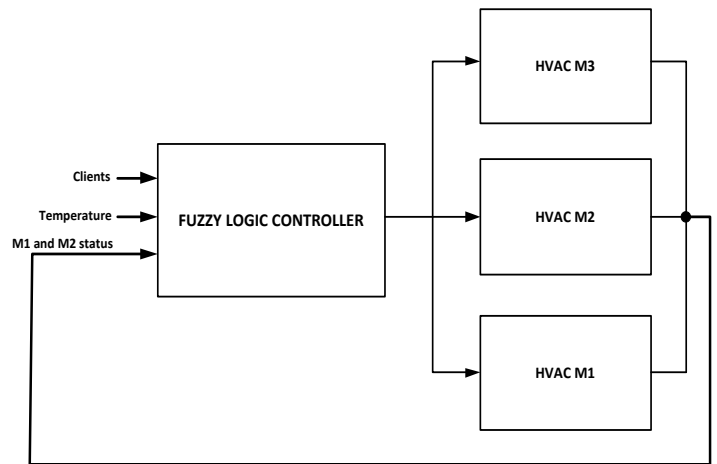


Figure 15: Simulation principle of the fuzzy logic controller

Table 6 : Recorded energy profit during testing periods of 2017 and 2018

Month	Energy consumption KWh		Profit	
	2017	2018	Profit (Ec)	Profit (Eb)
January	18178	8606	+ 52,66%	+43%
February	8673	8718	- 0,52%	+ 0,15%
March	8649	8338	+ 3,6%	+ 2,13%
April	8862	8128	+ 8,28%	+ 5,2%
May	11033	8387	+ 23,98%	+ 17,9%
June	14783	12651	+ 14,42 %	+ 7,42%
July	20494	17818	+ 13,06%	+ 10,41%

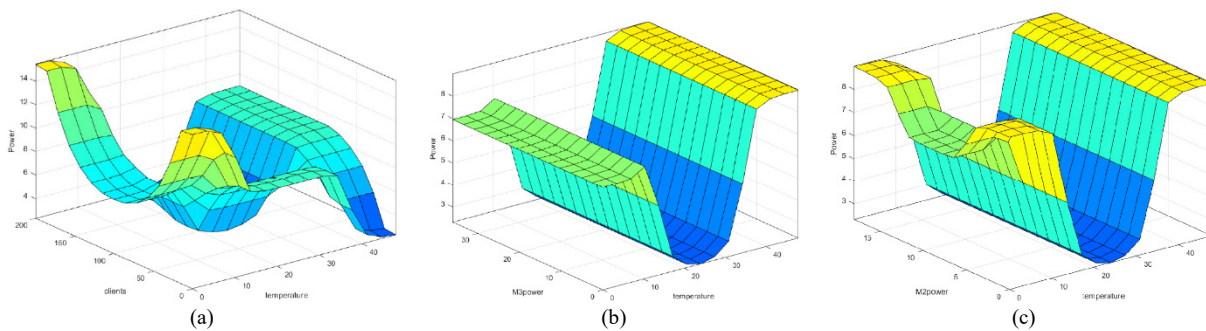


Figure 16: Variation of M1 HVAC power according to (a) ambient temperature and client's movement, (b) temperature and M3 HVAC power, (c) temperature and M2 HVAC power

Thus, the system, adapts the needed power for air conditioning purpose while exploiting at the same time the already deployed power in the other units. M3 requires the largest share of energy, this is why, during long periods of harsh temperature difference between the inside and the outside of the building, this unit is the main preserver of comfort inside the building. Figure 15 presents the simulation principle of the proposed controller. The controller doesn't simply adapt the inside ambient temperature based on defined set-points and external conditions, but it also exploits the possibility of under-powering M1 or M2 when the conditions are beneficial. Figure 16 presents all the possible cases for powering M1 based on the external temperature, clients' movement inside the store and power status of the HVAC system in M2 and M3.

As it can be seen on Figure 16 above, M1's HVAC system is only activated in extreme cases of temperature and clients' movement Figure 16 (a). Since the controller exploits the thermal exchange between the building units, power required in the middle of the day is greatly decreased. Furthermore, it can be noticed on Figure 16 (b) and (c) that during normal operating conditions, dissipated power of M1's HVAC system is greatly reduced, since M2 and M3 can help stabilize temperature inside the building.

## 6. Conclusion

In this paper, the design and implementation of a new Fuzzy logic controller was introduced into a commercial facility in Morocco. The case study presents how the building was functioning before and how the HVAC system was the primary drainer of energy during infrastructure analysis. The article describes how an appropriate modeling approach was designed in order to assess and quantify the energy loss during each month of the year. Furthermore, the article founded the energy saving methodology on existing research works. Thus, the working time was distributed into operating time slots, each with its own working characteristics defined by workers and visitor's dynamics. Particularly, in the present case study, special variables were introduced to cover the season time change and personnel interaction with the visitors and clients in the facility. This design allows the HVAC system to adapt automatically to the season without undergoing any change by the management team. During the design phase, input and output variable were carefully selected to minimize the number of uncontrolled events, especially during season shifts.

This case study showed very promising results in terms of energy saving and smoothness in HVAC control. For instance, the

integration of this particular controller granted the facility a 20% steady energy savings during the first official test period. The presented controller can also help HVAC oversizing errors caused by classic design methodologies, thus limiting the maximum power subscription. Besides, by taking additional measures such as population flow control and energy efficiency ethics among working staff, a more substantial gain can be achieved. In the upcoming work, we intend to investigate how categorizing areas in the facility, each with its own control strategy, can affect the HVAC system in terms of energy economy.

## References

- [1] E. Delzendeh et al., "The impact of occupants' behaviours on building energy analysis: A research review" *Renewable and Sustainable Energy Reviews*, 80, 1061–1071. doi:10.1016/j.rser.2017.05.264.
- [2] A. Elie, et al., "A Comprehensive Framework to Quantify Energy Savings Potential from Improved Operations of Commercial Building Stocks", *Energy Policy*, 67, avril 2014, 459-72. DOI.org, doi:10.1016/j.enpol.2013.12.031.
- [3] V.S.K.V. Harish, A. Kumar, "A Review on Modeling and Simulation of Building Energy Systems" *Renewable and Sustainable Energy Reviews*, 56, avril 2016, 1272-92. DOI.org, doi:10.1016/j.rser.2015.12.040.
- [4] P. Jianli, "A Survey of Energy Efficiency in Buildings and Microgrids using Networking Technologies", *IEEE Communications Surveys & Tutorials*, 16(3), 2014, 1709-31. Doi:10.1109/SURV.2014.060914.00089.
- [5] R. Rajeev et al., "Improving the Energy Efficiency of the Existing Building Stock: A Critical Review of Commercial and Institutional Buildings", *Renewable and Sustainable Energy Reviews*, 53, janvier2016, 1032-45. Doi:10.1016/j.rser.2015.09.084.
- [6] C. Jinbao et al., "Improving Energy-Efficiency in Public Buildings in China: Challenges and Solutions", 2009 International Conference on Computational Intelligence and Software Engineering, IEEE, 2009, 1-5. Doi:10.1109/CISE.2009.5362887.
- [7] E.G. Ochieng, "Integration of Energy Efficient Technologies in UK Supermarkets", *Energy Policy*, 67, avril 2014, 388-93. Doi:10.1016/j.enpol.2013.12.002
- [8] L. Pérez-Lombard, "A review on buildings energy consumption information. *Energy and Buildings*", 40(3), 394–398. doi:10.1016/j.enbuild.2007.03.007
- [9] A. Simelyt, "Promotion of renewable energy in Morocco. *Energy Transformation Towards Sustainability*", 249–287, 2020. doi:10.1016/b978-0-12-817688-7.00013-6
- [10] S. D'Oca, T. Hong, J. Langevin, "The human dimensions of energy use in buildings: A review", *Renewable and Sustainable Energy Reviews*, 81, 731–742, 2017. doi:10.1016/j.rser.2017.08.019
- [11] L. Yanfei, Z. O'Neill, "A Critical Review of Fault Modeling of HVAC Systems in Buildings", *Building Simulation*, 11(5), October 2018, 953-75, 2018. DOI.org, doi:10.1007/s12273-018-0458-4.
- [12] A.M. Waseem, et al., "Computational Intelligence Techniques for HVAC Systems: A Review", *Building Simulation*, 9(4), août 2016, 359-98. Doi:10.1007/s12273-016-0285-4.
- [13] H.S. Mehdi, A Novel Intelligent Control of HVAC System in Smart Microgrid, *Journal of Electrical Systems and Information Technology*, 4(2), septembre 2017, 299-309. Doi:10.1016/j.jesit.2017.01.005.
- [14] N. Soares et al., "A Review on Current Advances in the Energy and Environmental Performance of Buildings towards a More Sustainable Built Environment", *Renewable and Sustainable Energy Reviews*, 77, September 845-60, 2017. DOI.org (Crossref), doi:10.1016/j.rser.2017.04.027.

- [15] G.Y. Yasin et al., Design and Development of Advanced Fuzzy Logic Controllers in Smart Buildings for Institutional Buildings in Subtropical Queensland” *Renewable and Sustainable Energy Reviews*, 54, février 2016, 738-44, 2016. DOI.org, doi:10.1016/j.rser.2015.10.105.
- [16] A.I. Dounis, Advanced control systems engineering for energy and comfort management in a building environment—A review. *Renewable and Sustainable Energy Reviews*, 13(6-7), 1246–1261, 2009. doi:10.1016/j.rser.2008.09.015.
- [17] Woods, Ruth, et al. «The Influence of Users Behaviour on Energy Use in European ShoppingCentres: Influence Of User Behaviour on Energy Use » *Sustainable Development*, 25 (1), janvier 2017, 11-24. Doi:10.1002/sd.1638.
- [18] O.A. Nisiforou, et al. Behaviour, Attitudes and Opinion of Large Enterprise Employees with Regard to their Energy Usage Habits and Adoption of Energy Saving Measures, *Energy and buildings*, 55, decembre 2012, 299-311. Doi: 10.1016/j.enbuild.2012.08.034.
- [19] S. Janet, et al., The Energy Cultures Framework: Exploring the Role of Norms, Practices and Material Culture in Shaping Energy Behaviour in New Zealand ». *Energy Research & Social Science*, 7, 117-23, 2015. Doi:10.1016/j.erss.2015.03.005.
- [20] Invest in Morocco - Power Rates. <http://www.invest.gov.ma/index.php?id=34503&lang=en>. Consulted the 17th of November, 2020.
- [21] L.A. Zadeh, *Fuzzy Sets, Information and Control*, 338–353, (1965).
- [22] S. Adnan, Electric vehicle power electronics cooling system pump control using fuzzy logic” *International Journal of Advanced Computer Research*, 7(31), juin 2017, 111-20. Doi:10.19101/IJACR.2017.730019.
- [23] A. Agarwal, A. Mishra, and M. Dixit, “Design of an Improved Fuzzy Logic based Control System for Washing Machines.” *51(8)*, 5–10, 2016
- [24] M. Mongkolwongrojn, “Implementation of fuzzy logic control for air conditioning systems, In: *Proceedings of 8th International Conference on Control, Automation and Systems*, 313–321, 2005.
- [25] G.Y. Yasin, et al., Recent Developments of Advanced Fuzzy Logic Controllers Used in Smart Buildings in Subtropical Climate, *Energy Procedia*, 61, 2014, 1021-24. Doi:10.1016/j.egypro.2014.11.1015.
- [26] M.M. Gou, et al., Thermal Comfort Based Fuzzy Logic Controller, *Building Services Engineering Research and Technology*, 22(4), 237-53, 2001. DOI.org (Crossref), doi:10.1177/014362440102200403.
- [27] K. Mahmoud, Investigation of energy saving in HVAC systems : Modeling, simulation, and measurement using fuzzy logic controller ». 2018 IEEE International Conference on Industrial Technology (ICIT), IEEE, 445-50, 2018. DOI.org, doi:10.1109/ICIT.2018.8352218.
- [28] A. Zakia, Modeling Techniques Used in Building HVAC Control Systems: A Review ». *Renewable and Sustainable Energy Reviews*, 83, 64-84, 2018. DOI.org, doi:10.1016/j.rser.2017.10.044.
- [29] C. Schwegge, Estia: A real-time consumer control scheme for space conditioning usage under spot electricity pricing. *Computers & Operations Research*, 18(8), 751–765, 1991. doi:10.1016/0305-0548(91)90013-h
- [30] Z. Yanyu et al., Multi-objective optimal control algorithm for HVAC based on particle swarm optimization ». *Fifth International Conference on Intelligent Control and Information Processing*, IEEE, 417-23, 2014. DOI.org, doi:10.1109/ICICIP.2014.7010290.
- [31] A. Abdel-Hamid et al., Fuzzy Logic Control of Air-Conditioning System in Residential Buildings, *Alexandria Engineering Journal*, 54(3), 395-403, 2015. Doi:10.1016/j.aej.2015.03.023.
- [32] M. Abbas, M. Saleem Khan, Fareeha Zafar, « Autonomous Room Air Cooler Using Fuzzy Logic Control System, *International Journal of Scientific & Engineering Research* 2(5), 2011.
- [33] M. Jamieson, “\$3 Billion Opportunity: Energy Management in Retail Operations” , Schneider Electric executive summary, 1-6, 2014.
- [34] R. Nikitha et al.,Token Based Scheduling for Energy Management in Building HVAC Systems ». *Applied Energy*, 173, 67-79, 2016. DOI.org, doi:10.1016/j.apenergy.2016.04.023.
- [35] G. Ali et al., HVAC System Energy Optimization Using an Adaptive Hybrid Metaheuristic, *Energy and Buildings*, 152, 149-61, 2017. DOI.org, doi:10.1016/j.enbuild.2017.07.053.
- [36] S. Jaw-Kuen et al., Fuzzy Controller for a Voltage-Regulated Solar-Powered MPPT System for Hybrid Power System Applications, *Energies*, 8(5), 3292-312, 2015. DOI.org, doi:10.3390/en8053292.
- [37] L. Bortoni-Anzures, et al., Fuzzy Controller for Automatic Steering in Heavy Vehicle Semi-Trailers, *Ingeniería, Investigación y Tecnología*, 14(1), 1-9, 2013. DOI.org, doi:10.1016/S1405-7743(13)72220-0.
- [38] B. Ben et al., Apportioning Energy Consumption in the Workplace: A Review of Issues in Using Metering Data to Motivate Staff to Save Energy, *Technology Analysis & Strategic Management*, 26(10) 1196-211, 2014. Doi:10.1080/09537325.2014.978276.
- [39] A. Christina, Co-Simulation of Fuzzy Control in Buildings and the HVAC System Using BCVTB, *Advances in Building Energy Research*, 12(2), 195-216, 2018. Doi:10.1080/17512549.2017.1279077.
- [40] Z. Ma, An Assessment for Energy Management in Retail Stores, *Journal of Energy and Power Engineering*, 12(5), 2018. Doi:10.17265/1934-8975/2018.05.002.
- [41] S. Steve, Reducing Energy Demand: A Review of Issues, Challenges and Approaches, *Renewable and Sustainable Energy Reviews*, 47, 74-82, 2015. Doi:10.1016/j.rser.2015.03.002.
- [42] O. Darvelle, “ Énergie et commerces, conception et rénovation énergétiques des bâtiments”, *Les Cahiers de l’Urbanisme* N° 70, Université catholique de Louvain Architecture et Climat Chargés de recherche p92-96, 2008.



## Human-Robot Multilingual Verbal Communication – The Ontological knowledge and Learning-based Models

Mohammed Qbadou<sup>\*1</sup>, Intissar Salhi<sup>1</sup>, Hanaâ El fazazi<sup>1</sup>, Khalifa Mansouri<sup>1</sup>, Michail Manios<sup>2</sup>, Vassilis Kaburlasos<sup>2</sup>

<sup>1</sup>Laboratory Signals, distributed systems and artificial intelligence, Mathematics & Computer Science, ENSET, Hassan II University of Casablanca, Casablanca, 28000, Morocco

<sup>2</sup>HUMAN-MACHINES INTERACTION (HUMAIN) Lab, Department of Computer Science, International Hellenic University (IHU), Kavala Campus, GR-65404 Agios Loukas, Kavala, 65201, Greece

### ARTICLE INFO

Article history:

Received: 30 June, 2020

Accepted: 09 July, 2020

Online: 09 August, 2020

Keywords:

Natural language Processing  
multilingual interaction  
human-robot interaction  
general ontology  
LSTM Neural Networks

### ABSTRACT

In their verbal interactions, humans are often afforded with language barriers and communication problems and disabilities. This problem is even more serious in the fields of education and health care for children with special needs. The use of robotic agents, notably humanoids integrated within human groups, is a very important option to face these limitations. Many scientific research projects attempt to provide solutions to these communication problems by integrating intelligent robotic agents with natural language communication abilities. These agents will thus be able to help children suffering from verbal communication disorders, more particularly in the fields of education and medicine. In addition, the introduction of robotic agents into the child's environment creates stimulating effects for more verbal interaction. Such stimulation may improve their ability to interact with pairs. In this paper, we propose a new approach for the human-robot multilingual verbal interaction based on hybridization of recent and performant approach on translation machine system consisting of neural network model reinforced by a large distributed domain-ontology knowledge database. We have constructed this ontology by crawling a large number of educational web sites providing multi-lingual parallel texts and speeches. Furthermore, we present the design of augmented LSTM neural Network models and their implementation to permit, in learning context, communication between robots and children using multiple natural languages. The model of a general ontology for multilingual verbal communication is produced to describe a set of linguistic and semantic entities, their properties and relationships. This model is used as an ontological knowledge base representing the verbal communication of robots with children.

## 1. Introduction

The great evolution of the theories and tools of artificial intelligence in addition to the technological achievements in the field of industrial robotics have contributed to the advent of intelligent robots endowed with physical as well as intellectual capacities. The efficient and rational integration of this kind of intelligent robots in real working environments in the presence of human beings is always a challenge and an important objective for the engineer as well as for the scientist. Several fields and sectors of application are directly concerned by this integration and

consider it very promising for greater productivity and efficiency. Among these areas are mainly the industrial sectors and to a lesser but growing extent the education and health sectors. Especially for the education and health sectors, the most important aspect that determines the success or failure of this integration is the capacity for verbal interaction between humans and these intelligent robots. Indeed, verbal exchanges for these two sectors are essential for students' learning activities and for exchanges with patients during medical procedures. Verbal exchanges are even more important in the case of students or patients with special needs and especially in the case of autistic children.

\* Corresponding Author: Mohammed QBADOU, 340 Cité Anfa, 28000, Mohammedia, Morocco, +212667514296 qbmedn7@gmail.com

This research work is part of this context and aims to produce generic models representing a solution to the problem of verbal interaction of intelligent robots with several groups of children who use different natural languages. We also provide an implementation and experimentation of these models by considering the case of three natural languages most formally used in the geographical area of North Africa in this case Arabic, French and English. Our models will allow a smart robot to be able, in real time, to detect and communicate in the natural language of the children in front. This will be of great use for teachers or tutors to carry out collaborative learning activities in which the robot will have the role of teacher, tutor or even of student.

The paper is outlined in a way that Section 2 gives a survey on human-robot interaction systems. The section 3 explains our architecture model that illustrate the main three transformation modules the robot must have to interact with humans in multiple natural languages. The three modules that compose the whole system have also been discussed. Section 4 give the details on the first module especially the description of the data structure used, a scraping approach we developed to extract the expected data of parallel text and audio corpus mainly as JSON file format. The end of this section provides the Extract-Transform-Load model describing the mapping operations of data and metadata from the corpus to the ontology database. In section 5, we show how we can improve the performance of the translation machine algorithm based on our ontological knowledge base. Section 6 provides a conclusion on the different aspects of our models and ends with a critical discussion in order to identify new avenues for future research.

## 2. Survey on Human-Robot interaction

### 2.1. Classes of Human-Robot interaction

In a common work environment, an intelligent robot can have two types of interaction with humans [1]: verbal or non-verbal. The first interaction form covers the production of emotions or meaning without use of words or speech but using only facial or body gestures. On the other hand, the second form necessarily involves the use of natural language expressions in the orally and/or written forms.

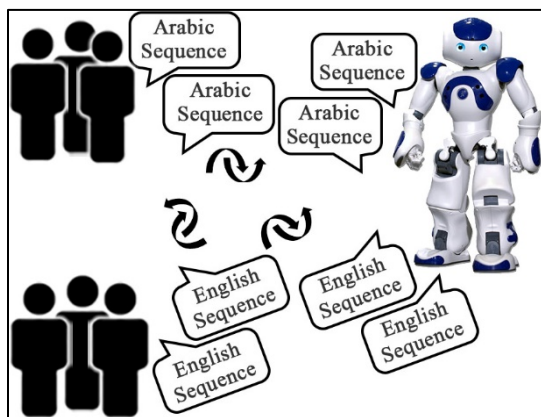


Figure 1: Human-Robot multilingual verbal communication

Figure 1 illustrates the key idea of our work where an intelligent robot is called to allow real-time verbal exchanges between two or more groups not speaking the same language and

having to share a common experience of learning or therapy. The robot thus plays the role of mediator and communication interface between the different groups.

### 2.2. Overview of research findings on oral interactions between robotic agents and humans

In the last decade, several high-quality research studies have been conducted on problems related to the oral communication possibilities between robots and humans. This period was marked by a significant number of articles indexed in this area. The study of these references revealed the main guidelines relating to our work and the following three research areas namely: (1) smart robotics and cognitive [2]-[6], (2) knowledge systems and (3) ontological databases [7]-[10] and artificial intelligence and machine learning techniques [11]-[16].

These research works highlight the inability of conventional robotic systems to respond to rapid changes in technical and functional production requirements. They present as a solution to these problems, the introduction of cognitive robots in the production lines. These robots while being able to adapt quickly, they are also able to interact and collaborate with human operators.

## 3. General architecture for robot-human communication in multi-language context

### 3.1. Block diagram of our proposal system

Figure 2 describes our general-purpose showing that robot-human interactions pass essentially through four modular units: the control unit, the perception unit, the knowledge data management system and the multilingual communication interface including text and speech sequences. This last unit contains all the transformations applied to the input signals collected by the robot's exchange unit. The knowledge base unit gives the semantic support of the different transformations for the synthesis of the robot reactions. This last unit contains all the transformations applied to the input signals collected by the exchange units of the robot. The knowledge management unit provides the semantic support for these transformations for the synthesis of the robot's reactions.

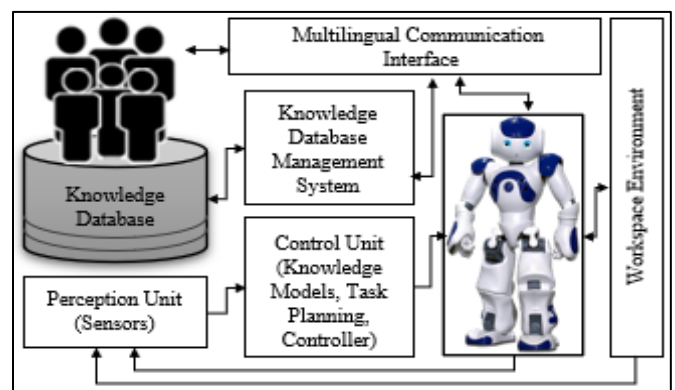


Figure 2: Block diagram of the proposed Human-Robot interaction system

### 3.2. Systemic model of verbal communication of robots based on an ontological knowledge database

The diagram in the figure 3 provides an integrated systemic representation of our approach which is built up into three large

independent parts, namely: (A) the devoted part for the construction and updating the text and audio parallel corpus links to the expandable list of languages, (B) the second part consisting in decisive phase which constitutes a driving force of construction and extension of our ontological knowledge database by the discovery of the entities and the relationships which connect them as well as their properties[10]. This engine also has the role of extracting all the occurrences associated with the entities found. (C) the third part constitutes the system of language identification, speech recognition and text translation.

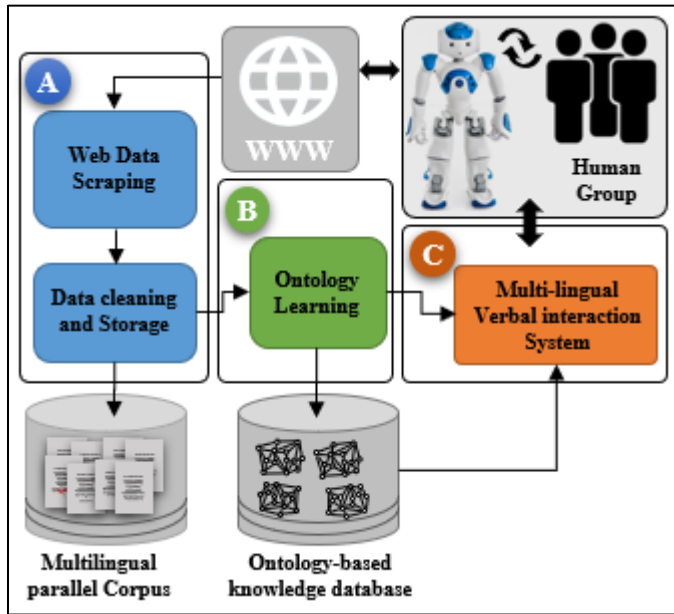


Figure 3: Architectural model of Human-Robot interaction system

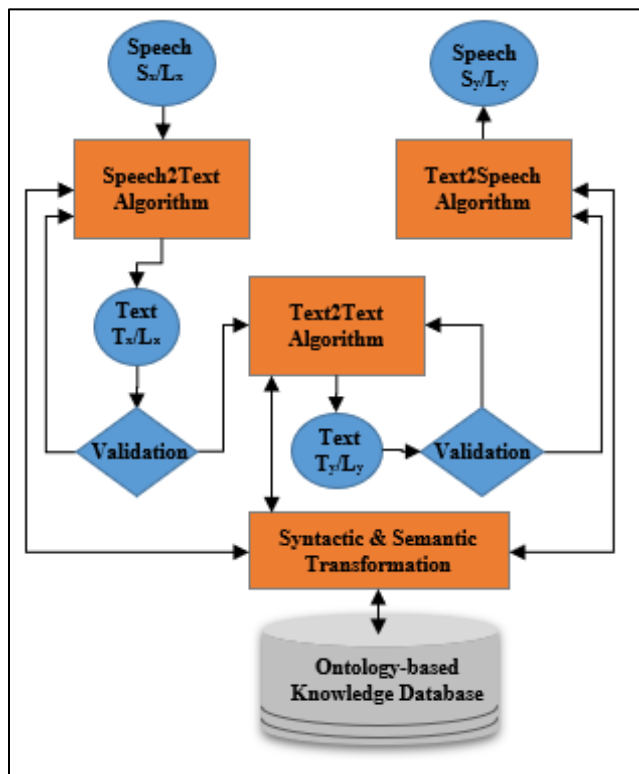


Figure 4: General model for Human-Robot multilingual verbal interactions

### 3.3. Specification diagram of Speech2speech algorithm using knowledge database

The objective of this module is the transformation of a given speech  $S_x$ , expressed in a language  $L_x$ , into an equivalent speech  $S_y$  expressed in another language  $L_y$ . This transformation is necessary to guarantee the exchange in a multi-language context. We have decomposed this transformation into three successive operations: (1) "Speech2Text": the transformation of the speech  $S_x$  into an equivalent text  $T_x$  and expressed in the same language  $L_x$ . (2) the transformation of the text  $T_x$  into an equivalent text  $T_y$  expressed in the language  $L_y$  and, (3) the transformation of the text  $T_y$  into an equivalent speech  $S_y$  expressed in  $L_y$ .

Figure 4 describes the sequence of the three operations and shows their relationship to the ontological knowledge management system. These operations are all based on the use of machine learning algorithms trained and validated by a high-quality dataset of parallel speeches and texts for the Arabic, French and English languages. The knowledge system also serves as a storage base [17] for all conversations collected or generated by the system.

### 3.4. Model of speech recognition and ontology-based tagging

Figure 5 describes the most complicated transformation in our system. It is carried out in five phases: firstly phase (1) dedicated to the pre-processing of the input speech in order to eliminate possible crudes by the application of filters and to segment it into pieces thanks to the application of techniques detection of breaks, framing and windowing. A second phase (2) consists of the application of an algorithm for identifying the language of speech. The results of these two phases are used by phase (3) which transforms each segment into a sequence of words thanks to the use of a neural network of the "multi-layer perceptron" type. semantic and contextual integration represents phase (4) and makes it possible by analysis and discovery of the semantic relationships provided by the ontological knowledge base. this completeness of the generated text also allows phase (5) to associate a set of tags with the speech that has just been processed.

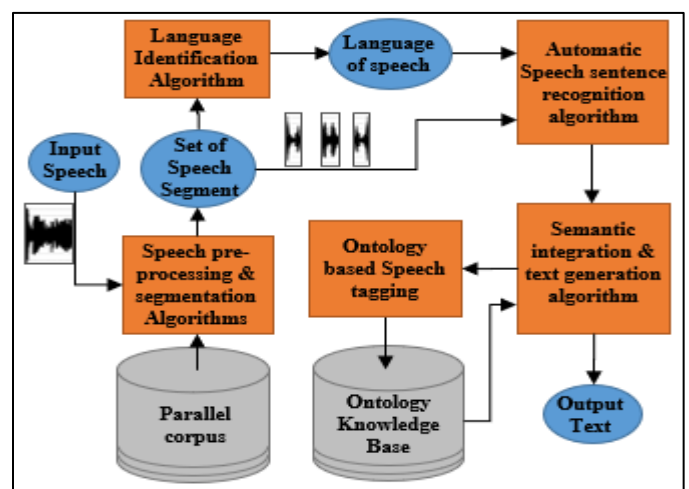


Figure 5: Speech recognition and text generation models

### 3.5. Language identification algorithm

The identification of the language of a speech is often done by selecting and analyzing a piece of this speech as reported in [8] and



[18]. As reported in [11]-[13] and [19], this identification may be done by a classification algorithm based on the automatic learning of a "deep neural network" trained using a multi-language speeches dataset recorded for several speakers. For its simplicity and the precision of its results, we chose the use of a classification algorithm based on the learning of a recurrent neural network (RNN) [20]. This learning is done on a large number of features generated by an extraction function based on the knowledge base.

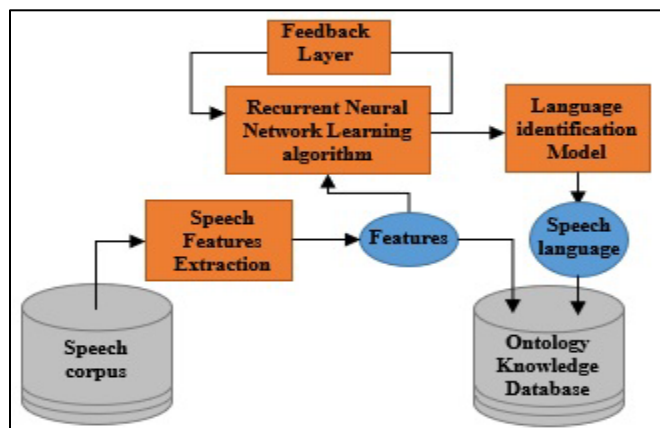


Figure 6: Ontology-based identification of the speech language

learning based on LSTM neural networks model with variants obtained by adding additional layers in order to improve the translation accuracy. Among recent approaches, we can notably cite "tree LSTM based methods with attention [14] and LSTM based methods with transformers [15]-[16]. As depicted in figure 8, our approach is based on the latest model whose learning is improved by the addition of a "labeling" module and reduction of the training dataset thanks to our ontological knowledge base.

The figure 7 describes the translation machine process in two steps and their interaction with the ontological knowledge database through the ontological engine. Thus, the input sequence message expressed in language  $L_i$  is first encoded based on Long Short-Term Memory (LSTM) networks augmented algorithm [22] in a common representation that is independent to the source and the target languages. The second step performs the decoding in the target language according to the second step of LSTM algorithm.

#### 4. Ontology-based knowledge data representation

##### 4.1. Ontology capabilities and usage

The formal modeling of knowledge on a field of activity and the representation of dependency relationships between its elements has always been a means expected by both engineers and scientists [23]. This modeling makes it possible to structure knowledge to serve as a knowledge base allowing to carry out new analytical studies and validate theories [10]. Among the models widely adopted by researchers stand out the ontological models. moreover, ontological models have a great capacity to represent semantic and contextual aspects relating to a domain [24]-[25]. Ontologies have an extensive use in Web applications where they are been considered as source of semantics with a rich and formal representations.

Thanks to these representations a large possibility in terms of reasoning mechanisms and, manipulation of data and metadata are possible [4]. For all these advantages, the ontological models are adopted in the fields of cognitive robotics and that of artificial intelligence. These possibilities also allow the management of knowledge relating to collaboration between robots and humans through verbal communication [7]. In this research work, we exploit all these possibilities through a domain ontology, as illustrated in figure 8, to serve as a model of a knowledge base for Human-Robot interaction system. This model should allow domain hierarchization into interest sub-domains (education, health, specific education, therapy, ...) [26]. Our model also provides the linguistic entities for natural language processing operations we have defined in section 3. We adopted OWL language to create and use our ontological knowledge database.

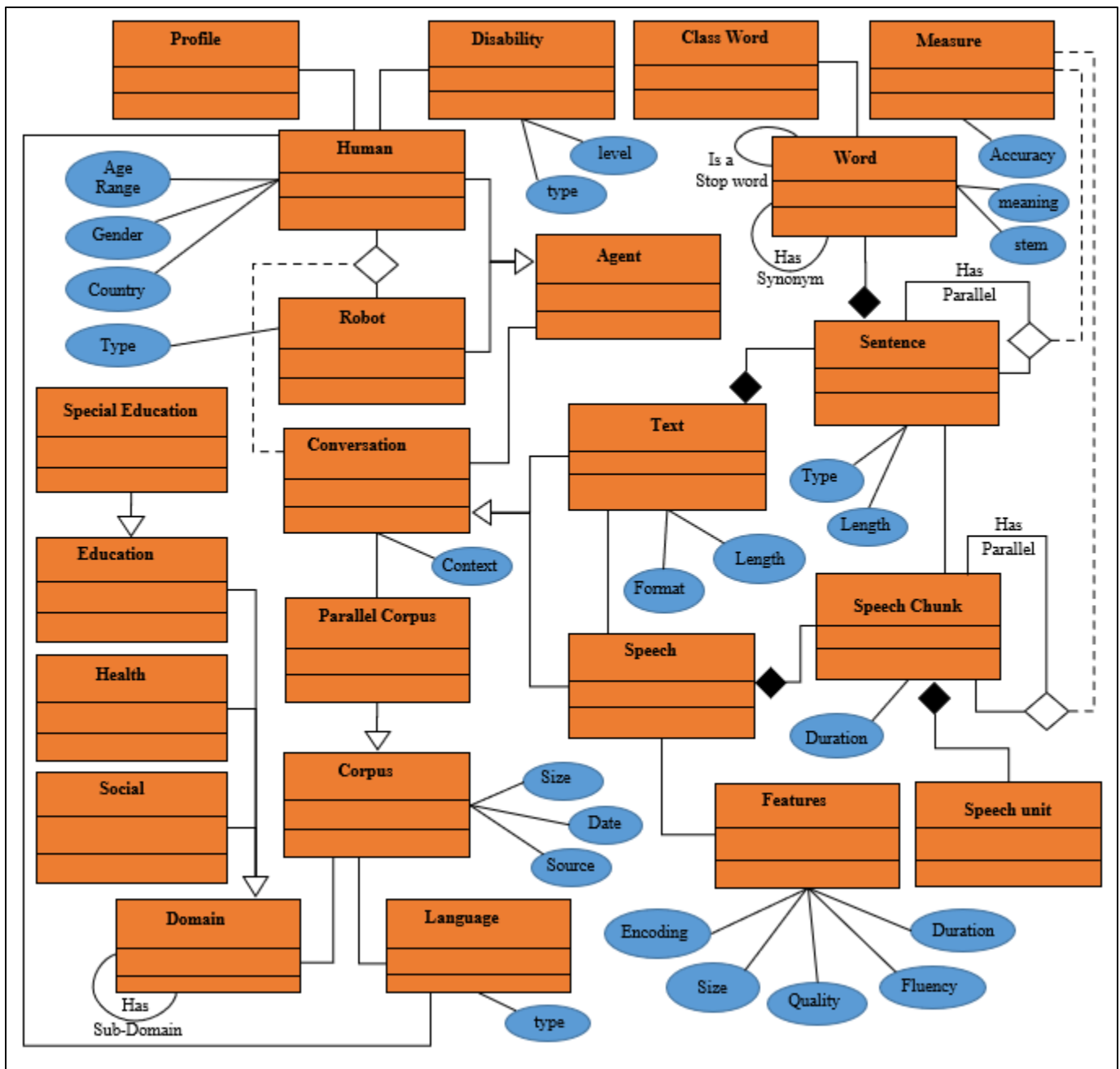
Figure 8 corresponds to the ontological model we designed. It represents a general ontology that represents several fields of knowledge, such as education, health and rehabilitation as examples. In connection with these domains, our ontological model also models the set of conversations that may occur between human and/or robotic agents. the detailed representation of the linguistic components of the conversations is also given by this model. the model represents a conversation as a set of textual or oral sequences. The ontology expresses the relationships between all the entities and their main properties.

To implement and test the models proposed in figures 5 and 6, we chose to create components programmed in Python and this for the range of these standard libraries in the fields of numerical calculation, machine learning and the field of processing of natural languages [21]. To define the corpus of speeches and parallel texts for the three languages considered, we identified a large list of websites offering free multilingual resources such as multilingual newspapers, online and multilingual tv channels, podcasts in lines in addition to forums and tweets of thematic conversations.

##### 3.6. Text to Text Translation

Several works on the problem of automatic translation machine have been published. The most recent have adopted machine





#### 4.2. Ontology use for tagging sequences and reducing dataset

The ontology represented at figure 8 is used as a knowledge management system for managing information related to the human-robot verbal interactions and for help in capturing important multilingual interactions' features for a specific domain or sub-domain [9]. Indeed, building on a hierarchy of domains and sub domains of interactions and their links to a set of linguistic and semantic concepts, our ontology knowledge base allows extraction of relevant, reduced and contextual datasets that facilitate learning significantly. It also allows a relevant labeling of these datasets producing semantic learning attributes. The figure 9 shows the principle followed to favor the reduction of training datasets taking into account the hierarchy of domains.

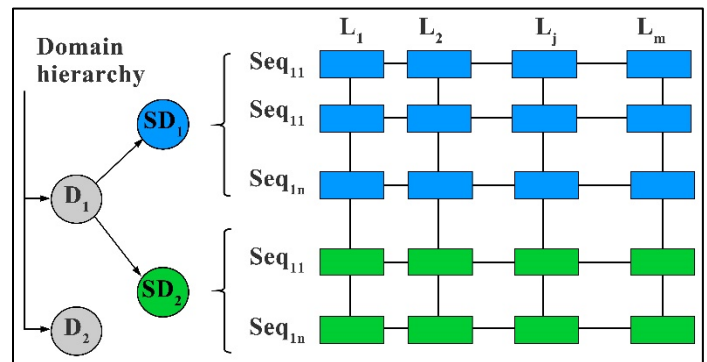


Figure 9: Dataset reduction and sequence tagging based on ontology

The exploitation of the domain hierarchy provided by our ontology allows to generate reduced and relevant dataset (table 1) according to the considered domain (education, therapy, ...)

Furthermore, an ontology-based feature optimization method is used to reduce significantly dimensionality of feature space [10]. Thus, we are able to extract dimension-reduced datasets of multilingual parallel texts for the training and validating the LSTM-based learning model. This method is conducted in two steps:

- As a first one, concepts of general ontology are generated by a transformation of the terms of vector space model, then the frequency weights of these concepts are calculated by the frequency weights of the terms.
- and according to the structure of the general ontology, similarity weights are associated then with the concept features.

### 5. Experiments and Results

#### 5.1. The used dataset

The results published in this paper are obtained on a parallel dataset constructed for Arabic, French and English (Table 2). The choice of these languages reflects our objective of deploying the results in the Moroccan context where these three languages are dominant. The table groups together the main data characterizing the distribution of sentences in the dataset used.

Figure 10 provides a graphical representation of these results. This graph shows a very strong similarity of these results for the three languages considered.

#### 5.2. Model training

The figure 11 shows the learning result of our LSTM neural network model. It shows in particular the list of layers created and the order in which they are connected with an LSTM layer and at the end the Dense layer which comes from the LSTM layer and which is responsible for producing the prediction of the outputs. The reshape functions have been used for successive dataset transformation from 2D to 3D format.

Table1: Parallel text dataset reduction based on domain hierarchies

Dom	SubD	N	En	Fr	Ar
D <sub>1</sub>	SubD <sub>1</sub>	1	Help	Aidez-moi	النجدة
		3	Jump	Saute	اقفز
		4	Run	Cours	اركض
		5	Stop	Arrête	قف
		...	...	...	...
	SubD <sub>2</sub>	1	The verbal interaction	L'interaction verbale	التفاعل اللفظي
	2	Computer components	Composants de l'ordinateur	مكونات الكمبيوتر	
		...	...	...	

Table2: Multilingual Training dataset - Length sentences distribution

English		Arabic		French	
Count	120.0000	Count	120.0000	Count	120.0000
Mean	123.2437	Mean	123.2437	Mean	123.2437
Std	143.1298	Std	152.1779	Std	152.1779
Min	1.0000	Min	7.0000	Min	2.0000
25%	31.5000	25%	31.0000	25%	31.0000
50%	45.0000	50%	44.0000	50%	46.0000
75%	164.5000	75%	172.0000	75%	167.0000
Max	518.0000	Max	523.0000	Max	526.0000

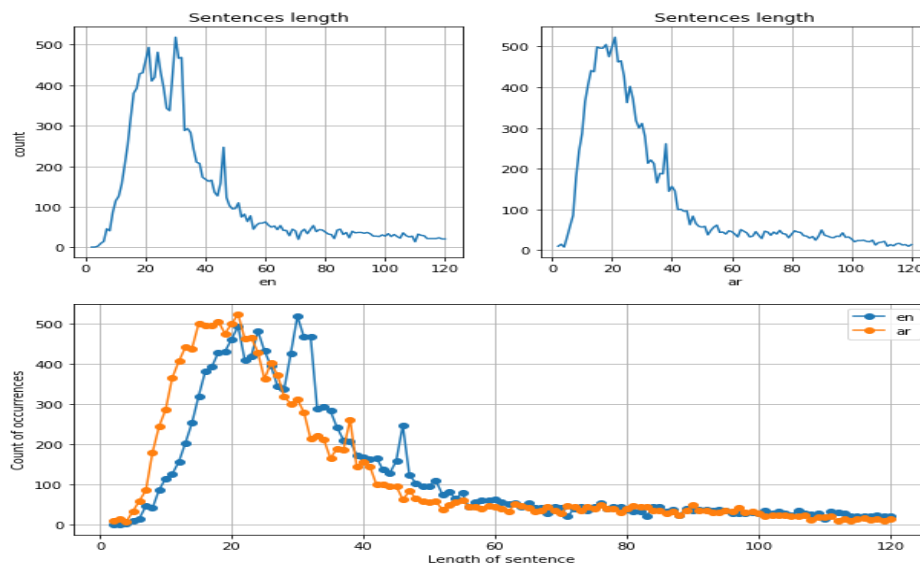


Figure 10: Distribution of the number sentence occurrences by length.

Layer (type)	Output Shape	Param #	Connected to
input_25 (InputLayer)	(None, None, 36)	0	
input_26 (InputLayer)	(None, None, 47)	0	
reshape_25 (Reshape)	(None, 120, 36)	0	input_25[0][0]
reshape_26 (Reshape)	(None, 120, 47)	0	input_26[0][0]
lstm_25 (LSTM)	[(None, 256), (None, 300032		reshape_25[0][0]
lstm_26 (LSTM)	[(None, 120, 256), ( 311296		reshape_26[0][0] lstm_25[0][1] lstm_25[0][2]
dense_13 (Dense)	(None, 120, 47)	12079	lstm_26[0][0]
activation_13 (Activation)	(None, 120, 47)	0	dense_13[0][0]
Total params: 623,407			
Trainable params: 623,407			
Non-trainable params: 0			

Figure 11: LSTM Model training life-cycle – model summary

## 6. Conclusion and future works

In this paper, we presented several models and algorithms for a new approach of the human robot multilingual verbal interaction. The strength of our method lies in the combination of the strength of ontological models in terms of semantic representation and their possibilities of reasoning and inference to the capacity of machine learning models based on the LSTM model. We perform the implementation of these models using the machine learning package named Keras and NLTK python framework. These implementations have been tested in a multi-agent environment made up of several robot and human agents which interact verbally according to a number of conversations in the three languages.

The development of new upgrades for our system will be the focus of our future research, we are also aiming to take action integrating human-robot interactions in the fields of therapy and special education with the purpose of accumulating data in these fields as well as emotional data, that would allow our robots to mimic human emotion expressed in facial expressions, choice of words and voice intonation as well as detect the emotional state of patients and act accordingly. In addition, we will consider implementing the proposed models in real situations of interaction between NAO robots and children, especially in the case of the two following areas: Therapeutic rehabilitation of children with specific needs and the educational system for this category of children.

### Acknowledgment

This research has been partially supported by the project "CybSPEED: Cyber-Physical Systems for Pedagogical Rehabilitation in Special Education" H2020-MSCA-RISE-2017. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 777720.

## References

- [1] N. Mavridis, "A review of verbal and non-verbal human-robot interactive communication" Robotics and Autonomous Systems, Volume 63(1), 22-35, 2015. <https://doi.org/10.1016/j.robot.2014.09.031>
- [2] I. Kafiev, P. Romanov and I. Romanova, "Intelligent Mobile Robot Control in Linguistic Evaluation of Solution Options" in 2019 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon), Vladivostok, Russia, 2019. <https://doi.org/10.1109/FarEastCon.2019.8934814>
- [3] R.R. Galin, R.V. Meshcheryakov, "Human-Robot Interaction Efficiency and Human-Robot Collaboration" Kravets A. (eds) Robotics: Industry 4.0 Issues & New Intelligent Control Paradigms. Studies in Systems, Decision and Control, 272, Springer, Cham, 55-63, 2020. [https://doi.org/10.1007/978-3-030-37841-7\\_5](https://doi.org/10.1007/978-3-030-37841-7_5)
- [4] A. Angleraud et al., "Human-Robot Interactive Learning Architecture using Ontologies and Symbol Manipulation" In Proceedings of the 27th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN, Nanjing, 2018. <https://doi.org/10.1109/ROMAN.2018.8525580>
- [5] K. Gaur, G.S.Virdi, "Smart Manufacturing with Robotic automation and Industry 4.0-A Review" in International Journal of Innovative Research in Science, Engineering and Technology, 7(3), 2481-2486, March 2018. DOI:10.15680/IJRSET.2018.0703082
- [6] A. Perzylo et al., "SMERobotics: Smart Robots for Flexible Manufacturing" in IEEE Robotics & Automation Magazine, 26(1), 78-90, March 2019. <https://doi.org/10.1109/MRA.2018.2879747>
- [7] L. Howard, G. Paulo, F. Sandro, J. I. Olszewska, "Ontology for autonomous robotics" in 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Lisbon, 189-194, 2017. <https://doi.org/10.1109/ROMAN.2017.8172300>
- [8] S. Kobayashi, S. Tamagawa, T. Morita and T. Yamaguchi, "Intelligent Humanoid Robot with Japanese Wikipedia Ontology and Robot Action Ontology" in 2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Lausanne, 417-424, 2011. <https://doi.org/10.1145/1957656.1957811>
- [9] A. Zeyer, P. Bahar, K. Irie, R. Schlüter and H. Ney, "A Comparison of Transformer and LSTM Encoder Decoder Models for ASR" in 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), SG, Singapore, 8-15, 2019. <https://doi.org/10.1109/ASRU46091.2019.9004025>
- [10] M.M. Ali et al., "Ontology-based approach to extract product's design features from online customers' reviews" Computers in Industry, 116, 103175,2020. <https://doi.org/10.1016/j.compind.2019.103175>

- [11] D. Cruz-Sandoval et al., "Towards a Conversational Corpus for Human-Robot Conversations" in HRI '17 Companion March 06-09, Vienna, Austria, ACM, 99-100, 2017. <http://dx.doi.org/10.1145/3029798.3038344>
- [12] K. Sugiura and K. Zetsu, "Rospeex: A cloud robotics platform for human-robot spoken dialogues" 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, 2015, 6155-6160, 2015. <https://doi.org/10.1109/IROS.2015.7354254>
- [13] T. Kawahara "Spoken Dialogue System for a Human-like Conversational Robot ERICA" in D'Haro L., Banchs R., Li H. (eds) 9th International Workshop on Spoken Dialogue System Technology. Lecture Notes in Electrical Engineering, **579**. Springer, Singapore, 65-75, 2019. [https://doi.org/10.1007/978-981-13-9443-0\\_6](https://doi.org/10.1007/978-981-13-9443-0_6)
- [14] Z. Huang et al., "TRANS-BLSTM: Transformer with Bidirectional LSTM for Language Understanding" in arXiv:2003.07000 [cs.CL], 2019. <https://arxiv.org/abs/2003.07000v1>
- [15] H. Li, A. Y. C. Wang, Y. Liu, D. Tang, Z. Lei and W. Li, "An Augmented Transformer Architecture for Natural Language Generation Tasks" in 2019 International Conference on Data Mining Workshops (ICDMW), Beijing, China, 1-7, 2019. <https://doi.org/10.1109/ICDMW48858.2019.9024754>
- [16] W. Yu, M. Yi, X. Huang, X. Yi and Q. Yuan, "Make It Directly: Event Extraction Based on Tree-LSTM and Bi-GRU" in IEEE Access, **8**, 14344-14354, 2020. <https://doi.org/10.1109/ACCESS.2020.2965964>
- [17] M.S.H. Ameer, Farid Meziane, Ahmed Guessoum, "ANETAC: Arabic Named Entity Transliteration and Classification Dataset" in arXiv:1907.03110v1 [cs.CL] 2019, <https://arxiv.org/abs/1907.03110v1>
- [18] Y. Tanimoto et al., "Feature extraction and Classification of Learners using Multi-Context Recurrent Neural Networks" in 2019 IEEE 11th International Workshop on Computational Intelligence and Applications (IWCIA), Hiroshima, Japan, 77-82, 2019. <https://doi.org/10.1109/IWCIA47330.2019.8955064>
- [19] M. Swain, A. Routray, P. Kabisatpathy, "Databases, features and classifiers for speech emotion recognition: a review," in International Journal of Speech Technology, **21**, 93–120, 2018. <https://doi.org/10.1007/s10772-018-9491-z>
- [20] M.A.A. Albadr, S. Tiun, F.T. "Spoken language identification based on the enhanced self-adjusting extreme learning machine approach" PLOS ONE, **13**(4), 2018. <https://doi.org/10.1371/journal.pone.0194770>
- [21] M. H. Nehrir, C. Wang, Deep Learning with Python, Manning Publications Co, 2018.
- [22] H.L. Yang HL et al., "A Personalization Recommendation Framework of IT Certification e-Learning System" in: Knowledge-Based Intelligent Information and Engineering Systems, **4693**, 50-57, 2007. [https://doi.org/10.1007/978-3-540-74827-4\\_7](https://doi.org/10.1007/978-3-540-74827-4_7)
- [23] J.J. Jung, H.G. Kim, G.S. Jo, "Alignment-Based Preprocessing of Personal Ontologies on Semantic Social Network" in Knowledge-Based Intelligent Information and Engineering Systems, **4693**, 255-262, 2007. [https://doi.org/10.1007/978-3-540-74827-4\\_33](https://doi.org/10.1007/978-3-540-74827-4_33)
- [24] S.U. Ya-ru et al., "Agricultural Ontology Based Feature Optimization for Agricultural Text Clustering" Journal of Integrative Agriculture, **11**(5), 752-759, 2012. [https://doi.org/10.1016/S2095-3119\(12\)60064-1](https://doi.org/10.1016/S2095-3119(12)60064-1)
- [25] K. Munir, M. Sheraz Anjum, "The use of ontologies for effective knowledge modelling and information retrieval" Applied Computing and Informatics, **14**(2), 116-126, 2018. <https://doi.org/10.1016/j.aci.2017.07.003>
- [26] H. I. Krebs and N. Hogan, "Therapeutic Robotics: A Technology Push" in Proceedings of the IEEE, **94**(9), 1727-1738, 2006. <https://doi.org/10.1109/JPROC.2006.880721>



# Development Trends of Smart Cities in the Future - Potential Security Risks and Responsive Solutions

Vu Nguyen Hoa Hong, Luong Tuan Anh \*

People's Security Academy, Ha Noi City, 151090, Viet Nam

## ARTICLE INFO

Article history:

Received: 05 June, 2020

Accepted: 13 July, 2020

Online: 19 August, 2020

Keywords:

Security

Risks

Smart cities

Development Trends

## ABSTRACT

Smart cities are an indispensable development of the ongoing urbanization process worldwide. Currently, smart cities are being built and deployed by countries in the first steps. Based on research on the current state of developing smart cities in the world, the author forecasts the future development trends of such cities, pointing to one of the major challenges for cities. It is a problem that responds to potential security risks. Identifying potential risks for coping solutions is essential for the future of smart sustainable cities. Main solutions include policy and law solutions, technology solutions and raising the peoples standard were given in the article.

## 1 Introduction

Before the industrial revolution, most of people in the world lived in rural areas. In 1800, only 3% of the world's population lived in cities. By 1900, there were 12 cities with a population of over 1 million but most people lived far away from urban centers [1]. Wellington Eweb - the former mayor of Denver city, Colorado, emphasized: "The 19<sup>th</sup> century belonged to empires; the 20<sup>th</sup> century belongs to nations. The 21<sup>st</sup> century will belong to cities". Nowadays, in developed countries, 70% of the population lives in cities, meaning more than 50% of the world's population lives in urban areas. By 2050, 70% of the worlds population will have lived in urban centers [2]. The unprecedented urbanization is taking place globally that creates tremendous pressure on infrastructure, houses, environmental services, health care, energy ... to meet the living needs of current and future generations. In the context of the 4th industrial revolution that has been taking place, the construction and development of smart cities is the indispensable result towards valuable, livable, resilient, environmentally friendly and sustainable cities.

Smart cities are places where information technology is combined with infrastructure, architecture, daily appliances and even the human body to solve economic, social and environmental issues [3, 4]. Smart cities mean the maintaining continuity of security, privacy, system robustness and the availability of services. The problem is that most of the smart cities are created by the use of real and virtual connectivity technologies, so smart cities will become the potential target of cyberattacks. Ensuring the safety of

resources and privacy of citizen data is the top concern in all smart city projects. In addition, smart cities also face many other security threats such as terrorism, population growth rate, monetary and financial security, increase of social inequality, human security due to the development of artificial intelligence.

Currently, experts and scientists have been paying much attention to research on smart cities even though the construction and development of new smart cities is at the very first stage of formation. However, such studies mainly focus on building and designing smart city models, foundations, principles, applications in economic development, social balance, environmental protection. Studies on the security of smart cities are mainly mentioned in general. The issue of ensuring smart city security is mostly studied in terms of cyber security ensuring, meanwhile in fact, although information technology is the basic foundation for building a successful city, operating a city safely requires abilities to cope with many other risks that may threaten the life and safety of people in the context of globalization.

The writers aim at providing an overview of smart cities, the development trend of smart cities in the future, pointing out the risks that threaten security of smart cities. On that basis, the authors recommend a number of solutions. To obtain research results, the authors have consulted with many experts in the field of security science and information technology experts. In fact, the study of building smart cities based on the approach and thinking of information technology experts is different from that of experts in the field of security. Therefore, combining different opinions of such experts together with the process of researching, analyzing, synthesizing

\*Corresponding Author: Luong Tuan Anh, People's Security Academy, Ha Noi City, 151090, Viet Nam & ongvietmamnon@gmail.com

related documents and data as well as participating in some international conferences, the writers have summarized and presented the results of the research.

The subject of the study is the building and development of smart cities all over the world. From studying this process, the authors point out future development trends of smart cities, and assess the potential risks related to national security in the construction of smart cities. The goal of the study is to provide appropriate solutions to deal with potential security risks to smart cities in the future. To achieve the results of the research, the authors focus on using the expert method along with the consultation and research on documents of many experts in the field of security science and information technology. In fact, the study of building smart cities based on the approach and thinking of information technology experts is different from that of experts in the field of security. Therefore, combining different opinions of experts together with researching, analyzing, synthesizing many related documents and data, participating in some international conferences and conferences, the author has summarized and presented the results of the research. At the same time, the authors use a case study method based on the survey of the implementation and construction of some smart cities in the world. Specific examples of building and implementing smart cities around the world will clarify the research problem.

## 2 Awareness of Smart Cities

Mankind is entering the early stages of the 4<sup>th</sup> industrial revolution with an unprecedented speed of development and outstanding achievements such as artificial intelligence, IOT, big data, sensors anywhere and anytime, robots, real - virtual connections. Making a qualitative change in all fields. Today, more than half of the worlds population lives in cities, and this number will continue to grow as cities have always been a motivation of economic growth. The speed and extent in which cities deploy technology with flexible policies will be highly competitive. Therefore, towards building and developing smart cities is an inevitable trend of the 4.0 revolution era.

The concept of smart cities has been known since 2009 originated from the proposal of IBM Corporation (USA), but until now most countries in the world have still been in the early stages of developing smart cities because of many technological factors, modifying the model according to development needs and depending on financial resources to comprehensively develop this model. Every year, there are still charts of smart cities in the world and the ranking position changes annually depending on the development of each region, continent as well as each city. However, “what is a smart city?” is the question that does not have a common answer [5]. Bruno Pete, IBI Group Smart City Task Force Lead in the “Top 10 Smart city success factor”, defines: “In general, a smart city is a city or region that can increase its competitiveness and quality of life, efficiently use resources, and support economic sustainability by using technology and creativity to raise the IQ of the built environment. Further, it is a city that uses innovation to address the needs and desires of the community, puts the citizens first and ultimately breaks down the barriers between agencies and departments, and between people and government” [5]. Researcher Nguyen Van Thanh - Deputy

Minister of Public Security of Vietnam said that in order to have a full concept of a smart city in our minds, we need to recognize the city as an ecosystem, so he gives a definition: Smart cities are cities that use integrated system control tools to connect the real world and the virtual world, mainly systematic thinking. Their means are information technology and the media; their objective is to build a valuable, energetic, resilient, and competitive city; their measure is the satisfaction of community; their criteria are security, welfare and safety indicators of social, economic and environmental factor of high standards [6]. Many other authors give different definitions due to different approaches, while also determining the factors that identify what the standards of a so-called smart city are [7, 8]. Common factors include accessibility to public transport, smart buildings, energy efficiency, environmental friendliness, policy effectiveness and governance, and high-tech presence [9, 10]. In this regard, the authors agree with Brett King in “Augmented: Life in the Smart Lane” (one of the two books that Chinese President Xi Jinping read to get ideas for the national AI strategy of China). He did not give a definition of a smart city, but he thought that the comprehensive definitions should include the following factors:

- Urban areas are using information and communication technologies to improve urban service efficiency, reduce costs and resource consumption, encourage citizens to participate in such efforts.
- Specific industries are being transformed by technologies including public services, transport management, energy, healthcare and air and water pollution reduction.

Smart cities are judged not only by their services to citizens, but also by their ability to respond to serious problems like natural disasters and the ability to make use of daily resources. A truly smart city will have to be redesigned from scratch, in an urban environment which makes use of technology to improve people’s lives, create jobs, transport systems and smart lives, in order to create a positive impact on the environment.

## 3 Development Trends of Smart Cities

IESE Business School, Navarra University, Spain conducted a survey on 165 cities in 80 countries and analyzed, evaluated based on 9 aspects: human resources, social cohesion, economy, environment, management, urban planning, international access, technology, mobility and transportation. As a result, the smartest cities have been selected, including: New York, London, Paris, Reykjavik, Singapore, Seoul, Toronto, Hong Kong, Amsterdam [11]. Although the development of new smart cities is in the early stage, many cities have implemented effective new services on data platforms. Klaus Schwab - the founder and executive chairman of the World Economic Forum, emphasizes that this is the 10th out of the 21 major shifts of the world under the impact of the 4th industrial revolution [12]. The development trend of smart cities in the near future can be generalized as follows:

**Firstly**, building and developing smart cities will become more and more popular and will be the goal of many countries in the world.



Figure 1: Illustration of services of a typical smart city [13]

Urbanization has been boosting economic growth, but it has also exposed many cities to housing, infrastructure and service challenges to meet the demand of a growing population. With more and more challenges, the application of new technologies to solve the problems is becoming more and more popular. To solve such problems, cities are required to be smarter. Therefore, smart cities will be a popular trend of many countries around the world. They will bring high quality living standards for everyone with a focus on several key aspects such as smarter environment, smarter mobility, smarter connectivity and smarter management.

**Secondly**, typical smart services will be provided by smart cities in the future.

The goal of smart cities is to create a sustainable urban environment by taking advantage of solutions based on information and communication technology, thereby providing better services to the people. The concept of smart services in smart cities remains abstract [14]. The article focuses on the trend of using typical smart services of smart cities in the future as follows:

- Smart management

New AI tools for future smart cities will soon improve the transparency and efficiency of city government as well as urban planning and design processes. Citizens' participation through social media and personal smart devices is indispensable for smart cities. Smart cities will set up dynamic communication channels. Planning will be changed from a linear process directed by the government to a process of collaboration, cooperation with citizens, leading to more open design, able to meet the participation of the general public.

- Smart traffic system

It is impossible to have a smart city without developing a

smart transportation system. Smart transport is the cornerstone of a smart city. With advances in sensors, optics and integrated processors, improvements in pedestrian safety and rudimentary vehicles will promote the use of public transport, reduce congestion and pollution, improve health, make travelling faster, more predictable, and less expensive.

In terms of public transport: Trolleys and small self-operated cabins will go around school campus and shopping areas to take people to trip-sharing locations or public transit stations. Public transport will be optimized based on demands, events, weather and other factors, all of which will be reacted in real time and controlled by AI. Parking lots and garages will gradually disappear as they become obsolete and operating internal combustion engines become more expensive. Buses and trains will become automated because of AI as well as integrated dynamic scheduling and routing capabilities. When AI starts driving buses and trains, they can work longer without getting tired or asking for overtime. Electric transport systems and automated electric vehicles will also save maintenance and operation costs. Public transport systems will use solar energy, reducing carbon emissions for the city. Costs paid to drivers are eliminated and electric motors will require less maintenance than internal combustion engines and associated steering systems. Intelligent transportation in the future will not only meet the needs of transportation and travelling of people but also greatly reduce the costs compared to the present [1, p. 313]. Traffic management system will be smart to manage traffic lights, roundabouts and traffic density. Sensors will be used in the parking lot, garage and roadside to report vacancies in real time via websites. This technology has now been applied very effectively in the Netherlands and some other countries. Street ecosystems will change to facilitate self-driving electric cars.

In terms of personal transport: The design of cities for self-

driving cars will be a popular trend in the near future. Cities will need less parking lots when people start to share car ownership. Even some cities will restrict manned vehicles or charge people because they will not be as safe as AI drivers, and cities will be in charge of injuries related to transportation.

- Smart Electricity and Water System

The future smart grid system deploys networked sensors and measuring devices to collect real-time system data from power generation, transmission and distribution systems. A central analysis tool receives this data through the communication network and sends control commands to smart electronic devices on the scene [15]. New generation LED street lights can serve as the basis for many sensor technologies to collect data on weather, pollution levels, seismic activities, movement of vehicles and people. By connecting such smart street light system into a network, we will know what is happening to the whole city in real time and provide lighting solutions in the areas of public safety [16].

The intelligent water management system integrates sensors, controllers and analytical elements to ensure that water is circulated only when and where it is needed. It also carries out real-time monitoring of water quality [15]. Applying Internet to water pipes, using sensors in water systems to monitor the flow thereby managing the entire cycle, providing sustainable water supply for the needs of people and ecosystems.

- Smart Energy

Renewable energy plays an important role in reducing air pollution, reducing costs, and increasing the ability to restore reserve energy. Potential alternative energy systems for future smart cities include: distribution networks that can withstand terrorist attacks and natural disasters; fuel cell storage production system; modular small thorium reactors; renovating high-rise buildings with transparent solar batteries; coastal cities with living embankments which are able to produce energy from tidal forces [1, p. 315].

- Smart Healthcare System for Citizens

Smart healthcare systems are becoming more and more popular in urban areas, especially when obesity is on the rise in rich countries such as the United States and the United Arab Emirates, which leads to an outbreak of diabetes and expensive health care costs. An intelligent healthcare system will provide health services by taking advantage of sensor infrastructure in smart cities. Instead of promoting the treatment of various diseases and symptoms, the smart health care system will focus on the use of technologies such as genotypes, biometric monitoring, microcirculation and testing based on chips in laboratories, assessing risk for each individual to help prevent diseases or stimulate behavior changes.

The smartest cities of the future will build a healthcare system in parallel with sensor-based factories and AI-based private health databases to reduce the burden of long-term healthcare. Technologies such as gene therapy, personal drugs and smart drugs will undoubtedly become part of the treatment.

- Smart Pollution Reduction

Water and air pollution is a major challenge for modern cities. New smart sensor technologies can be designed to be compact, connected to mobile devices, helping municipalities monitor pollution levels and reduce costs related to health and well-being by pollution. Smart cities will monitor pollution levels and take action to achieve a safe state. When connected to a smartphone, air pollution sensors create an environmental map from crowdsourced data and provide it to citizens for free. In addition, new types of water purifiers for households or the community can eliminate most dangerous chemicals and parasites. Smart cities can also equip necessary equipment in case of water pollution due to natural and man-made disasters. Smart materials and nanotechnology will allow us to create an efficient filtration system that cleans air and water with a much lower investment than that of today.

- Smart emergency response system

Cities often cope with many emergency situations or natural disasters. Therefore, smart cities can address emergencies by requiring relevant agencies to have smart devices with high resilience to deal with risks. Future smart cities will use a predictive model to alert city leaders about imminent problems instead of just reacting after the incident has taken place. The system can be pre-programmed to send notifications, enable tasks, automatically send orders and task lists to relevant organizations, as well as check the reception, processing and completion of these tasks. Moreover, the system will receive messages and other input from citizens about the state of emergency, which allows any citizen to actively contribute to the safety and functioning of the city.

**Thirdly,** typical future smart city models

Each country, each city chooses a different smart city model based on geographical conditions, history and related factors. Some typical smart city models in the future can be summarized in the Table 1.

Table 1: Classification of smart cities

Future smart city models	Cities successfully apply the models
City Development Strategy	Amsterdam, London
Green City	Reykjavik
Eco - City	Toronto
Garden City	Dubai
Ecological as Economic City	Singapore
Sustainably developed city	Paris
Livable City	New York
Resilience City	Tokyo
Digitalized urban	Hong Kong, Seoul

## 4 Potential Security Risks of Smart Cities

Smart cities are the totality of integration, collection, and information sharing through smart devices and sensors. The city's smart information system is a combination of wireless and wired infrastructure. On the basis of data collected, smart cities use smart



software tools to improve management efficiency and provide intelligent services. However, smart infrastructure of the smart cities opens up a huge attack surface, which is a collection of all known and unknown vulnerabilities, including the existing security controls on all subsystems and networks [14, 17]. The overall connectivity of a smart city helps criminals conditionally launch large-scale attacks that violate the security and privacy of individuals, organizations and city governments. The risks of insecurity first come from the fact that many application products are not capable of security (for example, automation control software vulnerable to hacker attacks). Meanwhile, managers do not or are not able to test security features of products and systems, resulting in a gap that still exists for hackers to exploit and attack on a large scale when solutions are connected on the Internet. Typically, there are security threats with some key smart city services.

Smart traffic with the development of self-driving cars could make these cars vulnerable to theft and attack from afar. Attackers may control one or more self-propelled vehicles to cause them to collide. Unlike general-purpose computing devices, installing security software for self-driving cars is often more expensive and complicated. Features such as Internet access while on the move significantly increase the vehicle's attack surface. The use of mobile technology and bluetooth in modern vehicles increases the risk of remote attacks. Devices in self-driving cars that can receive signals from outside are susceptible to signal interference, attackers can intercept and modify security-related data because it is connected via a network. Actions such as sudden braking, speeding, changing lanes, or changing direction when being attacked may result in a collision with another vehicle.

Intelligent water and electricity systems will be dependent on smart infrastructure. The entire city's electricity system, if attacked, could have serious consequences. At that time, the traffic lights will not work, self-driving vehicles will lose control and endanger human life. These infrastructures have holes which are once exploited can make the operation of the whole system affected. Automatically-measured parameter readers which are widely used in smart water and electricity systems without basic security measures can directly affect the confidentiality, integrity and authenticity of data. Moreover, some of them periodically transmit its data about energy use via inherently insecure wireless links. Attackers can capture such data by using simple techniques to identify uninhabited neighborhoods or identify citizen habits. Monitoring and data collection control systems are widely used in water distribution and sewage treatment plants are primarily designed to maximize functionality, thus security issues are often taken into account. Attackers can use simple techniques to capture or modify traffic between central control units and field units. Besides, but another threat to the smart water and electricity system is the issue of sensitive information related to the sitemap, details of chemical processes, web security plans, etc. for water and electricity control networks that are prone to service disruption, leading to public safety problems.

Smart management systems and many other smart services can lead to a great deal of citizens' information being collected, jeopardizing their privacy. Typically, smart healthcare services can make it possible to collect an unprecedented amount of information. Moreover, based on the data collected, one can guess the habits of citizens, their social status, basic information, a lot of sensitive

information. And when they combine with medical information, the issue of security and personal safety will become even more serious [18, 19].

It can be said that the issue of ensuring security and safety of smart cities is a difficult challenge and a difficult problem that governments need to solve when building these cities. All are tied together on the Internet. If there is no effective security strategy, cities will be quickly paralyzed due to vulnerabilities hacked by hackers.

## 5 Proposed Solutions

The process of urbanization will make existing cities uninhabitable in the coming time, therefore, despite the challenges, the development and construction of smart cities should not be delayed any more. All of these security challenges have put pressure on urbanization to adjust to form a number of more livable and more sustainable cities. These challenges make urbanization addressed by smart solutions [20, 21]. Basic issues related to future smart cities' security which need to be ensured are generalized in the figure 2.

A city which can carry out these solutions is becoming smarter. From the studies, the authors propose some following solutions:

**Firstly**, solutions for policy and law making in the field of security and safety for smart cities.

The number of cyberattacks increasing every day create favorable conditions for criminals to launch large-scale attacks to violate security of resources and data. Therefore, policy making, planning and implementation of security measures are extremely important for a smart city. Smart city governance in general and policy and law making, implementation of management measures to ensure security and safety for smart cities are different from traditional governance. Policies, laws on governing and security measures of smart cities require stability, transparency and high predictability to adapt to drastic changes of technology.

These policies and measures require a reasonable balance between the costs and benefits of each technology, and socio-economic development and security of city. It is very important to study and analyze the requirements of the security of smart cities based on their common infrastructure with the interdependence among different branches exist in that city, hence, identify security risks as a basis for designing intelligent security measures to help detect and prevent risks. Smart cities can handle risks by avoiding them, applying control measures or transferring these risks to third parties such as insurance, software creation, etc. [14].

Security policies and measures need to be constantly changed in accordance with infrastructure changes. New holes and new threats need to be promptly detected. Vulnerabilities may be technical and technological flaws that exist due to loopholes of softwares and hardwares as well as management flaws due to the inconsistency of security policies and procedures. Identifying the vulnerabilities and their severity allows people to prevent them. Such vulnerabilities often lead to security incidents. Therefore, it is important to detect, report and activate appropriate control measures to minimize and recover the impact of the incident caused by the vulnerabilities. Citizens should be encouraged to report incidents that need to be handled. It is necessary to have a complete form of relevant

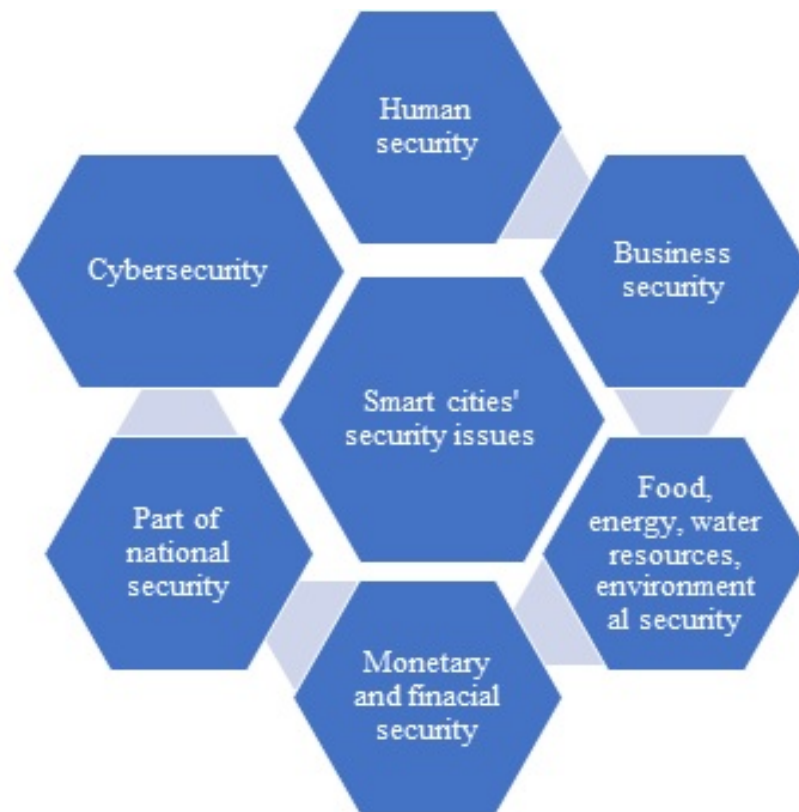


Figure 2: Basic Issues Related to Future Smart Cities' Security which need to be Ensured

information to create detailed and convenient records for remediation. Attention should be paid to the development of cybersecurity policies and procedures, and the selection of appropriate hardware and software tools for security ensuring. The procedures should be guided in detail with documented forms.

It is essential to create specific cybersecurity parameters, suitable to the conditions of each smart city. This issue was mentioned by Nguyen Van Thanh - Deputy Minister of Vietnam Ministry of Public Security and Do Quang Hung in Building and developing smart cities to ensure security, safety and safety indicator in the context of the 4th industrial revolution. The security indicators the author pointed out include: human security, business security, financial and monetary security, block chain, food security, national security (including cities), cybersecurity, big data analysis, ethnic, religious, and social security. The two authors also pointed out the indicators related to security assurance: The social security index includes sustainable economic development, improvement of resilience, life quality and social equality ensuring, application of digital technology, health care, education, culture, friendly environment, access to urban services, and public transport; safety indicators include crime reduction, medical safety, food safety, traffic safety, environmental safety, water safety, air safety, personal property safety, personal information safety, production safety, business, competition and safety against fire, explosion, and rescue [6].

In addition, it is necessary to build and manage a flexible and certain legal framework to receive new technologies and business models, and protect the rights of relevant parties. Due to the transna-

tional nature of non-traditional security threats, when building a smart city, nations and ethnic groups need to accept the common rules or have to adjust their legal systems towards internationalization. Cities need legislative policies that facilitate innovation, promote smart technologies and services, but at the same time minimize risks. It is necessary to strengthen cybersecurity management and strictly and promptly handle violations on the cities' security.

**Secondly**, technology solutions for security ensuring of smart cities. Smart city technologies are not only related to people who settle in cities but also have a very important impact on the prosperity of an area since smart cities are the driving force for economic development. As the global population grows and more and more people migrate to cities, the security forces and the police cannot be sufficient to ensure security effectively. Therefore, to help them, law enforcement agencies around the world need to use increasingly advanced applications of modern technology to keep confidentiality.

Smart cities which desire to succeed and develop sustainably need to apply technology such as artificial intelligence in public security and safety. This is a great lever in ensuring the safety of public assets, and more importantly the safety of people in the city. It is necessary to improve interoperability of network infrastructure, thereby improving connectivity among government departments; between governments and businesses and citizens. Here are some ways that smart cities can apply to ensure security and safety for the city:

Pay attention to expanding smart camera system, data management, and intelligent video surveillance. The use of surveillance cameras is a solution that brings tremendous economic and social

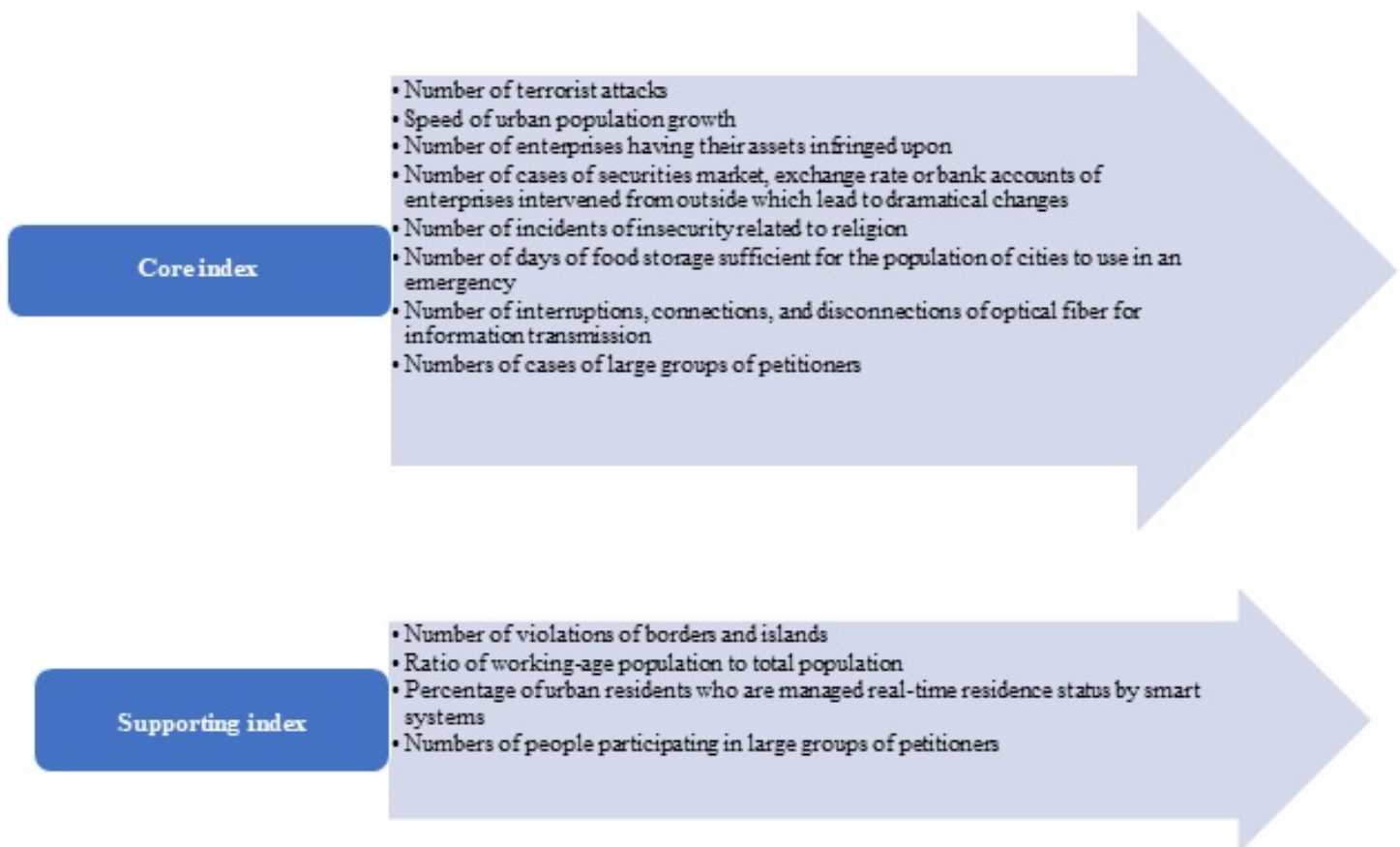


Figure 3: Smart cities' security indexes [6]

benefits by preventing the daily occurrence of events or scheduled or extraordinary events. They will help the cities react quickly to any situations, ensure smooth traffic and many other issues. The priority locations for installing cameras in smart cities should include: Central areas; hot spots and important events; public parking lots; parks, squares. One problem is how to have enough people to monitor the entire surveillance cameras throughout the city to promptly handle and act against arose situations. The limitations can be easily overcome with the help of video analysis, in which real-time footage is analyzed with deep learning AI. Any irregularities can be identified by this system and reported to nearby staff immediately. This can help to respond quickly to crimes or even prevent many offenses.

Use different types of IoT sensors to monitor the crowd, especially in shopping areas, busy intersections, to detect potential security incidents. For example, scanning devices can detect objects made of certain materials and distinguish their shape, even if they are hidden from view. AI algorithms can analyze this scanned footage in real time to determine if these objects are harmful. If any potentially lethal object is identified, security personnel can have further inspection. In addition, crowd monitoring systems powered by IoT and AI can also track in important places, such as footsteps. These systems can ensure that bridges are always optimized and safe to those who use them. IoT-enabled systems can alert those present as well as the governments when such places have a tonnage beyond the allowed threshold and begin to take appropriate action.

Restrict access to high security areas such as airports, banks, power plants, government data centers and military bases by using technologies such as IoT and AI. Security systems, security gateways provided by IoT can ensure that only authorized persons can access the main government facilities. Multi-factor authentication, which may include biometric scanning, can be used to add even more security layers to the most sensitive parts of those facilities.

Make transportation more secure through the use of IoT in safe ticketing by embedding RFID technology. In fact, RFID activation tickets have been used for the 2018 FIFA World Cup, which helps speed up the ticket inspection process despite a large number of fans without compromising security. In addition, tickets can be stored digitally in everyone's smartphones, which are able to automatically connect with security systems and automatic turnstiles at airports to allow only the right people to enter without security guards. This allows more effective use of security personnel and keep them wait for more serious situations.

Increase use of patrol robots and drones. Use robots as security guards which can patrol the streets of smart cities and identify potential threats with the help of advanced sensors and cameras. They can prevent crimes with their mere presence and can perform basic security tasks such as crowd control and search, sometimes more effectively than humans. Security personnel can stand still and participate only when situations escalate. Similarly, drones can be deployed to monitor public places and report to nearby security personnel in case of a potential public safety risk. This has been

effectively applied by a number of countries in checking people wearing masks and other violations to prevent the spread of the Covid-19 epidemic, such as the use of drones in China and the use of robots in Thailand [22, 23].

**Thirdly**, solutions to improve people's intellectual and technological levels to build and operate smart cities well. Building a smart city requires not only the application of smart technology products, but more importantly building a civilized community so that smart governance measures and services can be used effectively. The people are the ones who are served. At the same time, it is also one of the most important actors when building a smart city.

Indispensable condition is peoples ability to receive "smart" factors of the city. People must ensure enough financial sources to invest in technology and equipment to absorb the intelligence of the city. The regional people's awareness must be at a high level to use the system of technology and equipment invested, along with voluntary cooperation, action and sharing. The second condition is to have experts loyal to the common interests to operate the systems of the smart city. Besides, the city needs to have "intelligent government" and "smart leadership", hence, having the right and appropriate investment and mobilizing all investing sources as well as operate smart cities, creating a connection in the implementation of the set goals [24].

## 6 Conclusion

The world has been speeding up the development towards smart cities. Smart cities bring many benefits to the community, environment and society. People's life quality in smart cities will be enhanced thanks to the benefits when cities eliminate greenhouse gas emissions; find out solutions to crime prevention, intelligent and time-saving transport, housing and energy-saving options; deploying intelligent medical solutions to improve life expectancy; create jobs; to save costs and facilitate the implementation of sustainable development targets. The author's research points out that smart cities development will become more and more popular and be the goal of many countries in the world. Simultaneously, the authors also emphasize some typical smart services provided by smart cities in the future including smart management, smart traffic systems, smart water and electricity systems, smart energy, smart healthcare systems for urban residents, smart emergency response systems. The research also synthesizes some typical future smart city models including: City Development State; Green City; Eco City; Garden City; Ecological as Economic City; Sustainably Developed City; Livable City; Resilience City; Digitalized Urban.

It is indispensable for the process of changing to a smart city to have security measures and solutions built to adapt and support its operation. In the entire process, information technology is only a tool to help a city solve one or more specific problems rather than the development goals of the city. The authors have suggested a number of basic solutions to cope with the security threats of smart cities including: Solutions on policy and law making in ensuring security and safety of smart cities; technology solutions to ensure security for smart cities; solutions to improve people's intellectual and technological levels to effectively build and operate smart cities. Therefore, the development and construction of a smart city with se-

curity ensuring is not only based on technological solutions, but also in a synchronous and unified manner for immediate and long-term solutions to ensure sustainable development of the city, meeting the inevitable future demands of the drastic urbanization process that has been taking place all over the world.

**Conflict of Interest** The authors declare no conflict of interest.

**Acknowledgment** This work is supported by scientists of People's Security Academy and Ministry of Public Security of Vietnam.

## References

- [1] B. King, "Augmented: Life in the Smart Lane", Marshall Cavendish International (Asia) Pte Ltd, June 7th 2016.
- [2] UNICEF/Periscope Study, 2020.
- [3] A. Toffle, "The Third Wave", (page 17), Bantam, 1984.
- [4] A.M. Townsend, "Smart Cities; Big Data, Civic, Hackers, and the Quest for the New Utopia New York", New York, W.W.Norton and Company, Inc., 2013.
- [5] D.L. Chang, J. Sabatine-Marques, E.M. da Costa, P.M. Selig, and T. Yigitcanlar, "Knowledge-based, Smart and Sustainable Cities: A Provocation for a Conceptual Framework," *Journal of Open Innovation: Technology, Market, and Complexity* 4(5), 117, 2018. DOI: <https://doi.org/10.1186/s40852-018-0087-2>
- [6] Nguyễn Văn Thành, D.Q. Hung, "Building and developing smart cities to ensure security, safety and safety indicator in the context of the 4th industrial revolution", National Political Publishing House, 2018.
- [7] D. Washburn, U. Sindhu, S. Balaouras, R.A. Dines, N.M. Hayes, L.E. Nelson, "Helping CIOs understand "smart city" initiatives: defining the smart city, its drivers, and the role of the CIO". Forrester Research. Inc., Cambridge, 2010.
- [8] H. Partridge, "Developing a human perspective to the digital divide in the smart city", In *Proceedings of the Biennial Conference of Australian Library and information Association, Queensland, Australia, Sep 21-24, 2004*.
- [9] V. Albino, U. Beradi, R.M. Dangelico, "Smart cities: Definitions, dimensions, performance, and initiatives", *Journal of urban Technology*, 2015. DOI: <https://doi.org/10.1080/10630732.2014.942092>
- [10] K. Kourtit, P. Nijkamp, "Smart cities in the innovation age. Innovation", *The European journal of social science research*, 2012. DOI: <https://doi.org/10.1080/13511610.2012.660331>
- [11] B. Cohen, "The Smartest Cities in The World", *Fast Company* (November 20, 2014), <https://www.fastcompany.com/3038765/the-smartest-cities-in-the-world> [Accessed November 29, 2018].
- [12] K. Schwab, "The Fourth Industrial Revolution", *World Economic Forum, Currency* pp. 11 (January 3, 2017).
- [13] Data, Energy, And The Smart City: A Conflicting Relationship, 2019, Online Available: <https://www.smartdatacollective.com/data-energy-and-the-smart-city-a-conflicting-relationship/>
- [14] A. Bartoli, J. Hernandez-Serrano, M. Soriano, M. Dohler, A. Kountouris, D. Barthelemy, "Security and privacy in your smart city", *Proceedings of the Barcelona smart cities congress* 292, 2011.
- [15] H. Song, R. Srinivasa, T. Sookoor, S. Jeschke, "Smart Cities: Foundation, Principles and Applications", John Wiley and Sons, Inc, 2017.
- [16] World Economic Forum, global conference on cities future, 10 best initiatives in urban renovation of October, 2015.
- [17] A. S. Elmaghaby, M. M. Losavio, "Cyber security challenges in smart cities: Safety, security and privacy", *Journal of advanced research*, 5(4):491497, 2014. DOI: <https://doi.org/10.1016/j.jare.2014.02.006>
- [18] D. Rebollo-Monedero, A. Bartoli, J. Hernández-Serrano, J. Fornè, M. Soriano, "Reconciling privacy and efficient utility management in smart cities", *Transactions on Emerging Telecommunications Technologies*, 25(1) 94-108, 2014. DOI: <https://doi.org/10.1016/j.jare.2014.02.006>
- [19] Z. Khan, Z. Pervez, A. Ghafoor, "Towards cloud based smart cities data security and privacy management", In *Proceedings of IEEE/ACM 7th International Conference on Utility and Cloud Computing (UCC)*, pages 806811. IEEE, 2014. DOI: 10.1109/UCC.2014.131



- [20] W. Lum, "Cybersecurity challenges: A view from the public sector", April 2016.
- [21] S. Ijaz, M. A. Shah, A. Khan, M. Ahmed, "Smart cities: A survey on security concerns", *International Journal of Advanced Computer Science and Applications* 7(2), 612625, 2016.
- [22] B. C. M. Fung, T. Trojer, P. C. K. Hung, L. Xiong, K. Al-Hussaeni, R. Dssouli, "Service-oriented architecture for high-dimensional private data mashup", *IEEE Transactions on Services Computing (TSC)*, 5(3):373386, 2012. DOI: 10.1109/TSC.2011.13
- [23] F. Bosco, "Critical infrastructure threat landscape: Understanding and reacting", April 2016.
- [24] A. Martínez-Ballesté, T. A. Pérez-Martínez, A. Solanas, "The pursuit of citizens privacy: a privacy-aware smart city is possible", *IEEE Communications Magazine*, 51(6):136141, 2013. DOI: 10.1109/MCOM.2013.6525606

# Application of EARLYBREAK for Line Segment Hausdorff Distance for Face Recognition

Chau Dang-Nguyen<sup>\*1,2</sup>, Tuan Do-Hong<sup>1,2</sup>

<sup>1</sup>Faculty of Electrical and Electronics Engineering, Ho Chi Minh University of Technology (HCMUT), 268 Ly Thuong Kiet Street, Ho Chi Minh City, 72555, Vietnam

<sup>2</sup>Vietnam National University Ho Chi Minh City, Ho Chi Minh City, 71308, Vietnam

## ARTICLE INFO

Article history:

Received: 16 June, 2020

Accepted: 27 July, 2020

Online: 19 August, 2020

Keywords:

Hausdorff distance

EARLYBREAK

Similarity measure

Runtime analysis

## ABSTRACT

The Hausdorff distance (HD) is defined as MAX-MIN distance between two geometric objects for measuring the dissimilarity between two objects. Because MAX-MIN distance is sensitive with the outliers, in face recognition field, average Hausdorff distance is used for measuring the dissimilarity between two sets of features. The computational complexity of HD, and also average HD, is high. Various methods have been proposed in recent decades for reducing the computational complexity of HD computing. However, these methods could not be used for reducing the computational complexity of average HD. Line Hausdorff distance (LHD) is a face recognition method, which uses weighted average HD for measuring the distance between two line edge maps of face images. In this paper, the Least Trimmed Square Line Hausdorff Distance method, LTS-LHD, is proposed for face recognition. The LTS-LHD, which is a modification of the weighted average HD, is used for measuring the distance between two line edge maps. The state-of-the-art algorithm, the EARLYBREAK method, is used for reducing the computational complexity of the LTS-LHD. The experimental results show that the accuracy of proposed method and LHD method are equivalent while the runtime of proposed method is 68% lower than LHD method.

## 1 Introduction

The Hausdorff distance (HD) is an useful parameter for measuring the degree of resemblance between two point sets. The advantage of HD is that there is no requirement of point-to-point correspondence. In many science and engineering fields, HD has particular attention of researcher, such as face recognition [1–5], image matching [6–8], image segmentation for medical image [9–11], image retrieval [12, 13], shape matching [14–16].

The HD is defined as MAX-MIN distance between two point sets, measuring how far two point sets are from each other. The HD computing contains two loops, the outer loop for maximization and the inner loop for minimization. Due to the sensitivity of the MAX-MIN HD with the outlier, average Hausdorff distance or partial Hausdorff distance (PHD) is used instead MAX-MIN HD in image matching or face recognition applications. Huttenlocher et al. The PHD was first proposed [17] for comparing the similarity between two shapes. In [18], the robust Hausdorff distance for face recognition was proposed, which uses PHD for measuring the distance between two sets of feature of face gray images. However,

PHD is just an effective distance when the pollution of noise points is low. In face recognition research, average Hausdorff distance is commonly used.

The modified Hausdorff distance (MHD) was first presented [19] for image matching, where the directed Hausdorff distance is the mean of all distance from points to other sets. The 'doubly' modified Hausdorff distance (M2HD), which is the improved of the MHD, was proposed for face recognition [20]. Various algorithms using average weighted HD, an extension of MHD, were proposed for face recognition with the difference in weighted function., i.e., spatially weighted Hausdorff distance (SWHD) [21], spatially weighted modified Hausdorff distance (SW2HD) [22], spatially eigen-weighted Hausdorff distances (SEWHD) [23], edge eigenface weighted Hausdorff distance (EEWHD) [24, 25].

The extension of M2HD, similarity measure based on Hausdorff distance (SMBHD), was proposed for face recognition [3]. Another version of MHD, called modified Hausdorff distance with normalized gradient (MHDNG), was proposed for face recognition [8]. A new modified Hausdorff distance (MMHD) was presented for face recognition [26], which used average weighted Hausdorff dis-

\*Corresponding Author: Chau Dang-Nguyen, Faculty of Electrical and Electronics Engineering, Ho Chi Minh University of Technology (HCMUT), Ho Chi Minh City, Vietnam, (84-8)38.647.256 Ext: 5700 & Email: chaudn@hcmut.edu.vn

tance for measuring the dissimilarity between two sets of dominant points of edge maps of face images. Based on edge map of face image, a novel face feature representation was proposed [27], line segment edge map (LEM). The line segment Hausdorff distance (LHD), which is the average weighted of distances between two line segments, was proposed for face recognition. Based on LHD, the extension of LHD, called spatial weighted line segment Hausdorff distance (SWLHD), was presented for face recognition [28]. In our previous work [1], we proposed a modification of LHD for face recognition, called Robust Line Hausdorff distance RLHD.

Supporting  $P$  and  $Q$  are the number of points, or elements, in two sets. For computing the directed distance of average HD, the distances from each point in the first set to all points in the second set have to be calculated to find the minimum value, which is the distance from the point in first set to its nearest neighbor in second set. The directed distance of average HD is the mean of all distances from points in first set to their nearest neighbors in second set. The computational complexities of methods using average HD are always  $O(PQ)$ , the same as MAX-MIN HD. In recent decades, many methods have been proposed for reducing the computational complexity of the HD computing, which is known very high. The key of reducing the computational complexity is reducing the average number of the inner loop. All of these methods uses a temporary HD  $cmax$  for quickly finding the non-contributed points in the inner loops and outer loops for final HD computing. However, these methods, which is proposed with the purpose of reducing the running time of MAX-MIN HD computing, can not be used for reducing computing complexity the average HD computing because the temporary HD  $cmax$  does not exist in average HD computing. Due to the high computing complexity of average HD, the face recognition methods, which use average HD for measuring the distance between two sets of feature, are restricted from face recognition applications.

Here we propose an extension of the method in [1], called Least trimmed square Line Hausdorff distance (LTS-LHD) for face recognition. The LTS-LHD, an extension of weighted average HD, is used for measuring the dissimilarity between two line edge maps (LEMs). The LTS-LHD is the average of largest distances instead of all distances as average HD. The experimental results shows that the face recognition accuracy of the proposed method, LTS-LHD, is equivalent to the accuracy of LHD method, which used average HD for measuring distance between two LEMs, with a suitable parameter. Moreover, with LTS-LHD, the temporary HD  $cmax$  exist, and the methods, which are proposed with the purpose reducing the computational complexity of MAX-MIN HD computing, could be used for reducing the runtime of the proposed method. The EARLYBREAK [29], which is known as the state-of-art algorithm for reducing the computational complexity of HD computing, is used for reducing the computational complexity of proposed method. The runtime of proposed method is 68% lower than the LHD method.

The rest of the paper is structured as follows. The brief review of methods, which were proposed for reducing the computational complexity of the HD, is presented in Section 2. Section 3 presents the proposed method for face recognition, which uses LTS-LHD for measuring the distance between two LEMs. Section 3 also presents how to apply EARLYBREAK method for reducing the

computational complexity of proposed method. Section 4 evaluates the performance of the proposed method and the performance of proposed method is also compared with the performance of LHD and RLHD method. Finally, conclusion is presented in Section 5.

## 2 Related works

Given two nonempty point sets  $M = \{m_1, m_2, \dots, m_p\}$  and  $T = \{t_1, t_2, \dots, t_q\}$ , the directed Hausdorff distance  $h(M, T)$  between  $M$  and  $T$  is the maximum distance of a point  $m \in M$  to its nearest neighbor  $t \in T$ . The directed Hausdorff distance from  $M$  to  $T$  as a mathematical formula is

$$h(M, T) = \max_{m \in M} \min_{t \in T} \|m - t\| = \max_{m \in M} \text{mindist}(m, T) \quad (1)$$

where  $\|\cdot, \cdot\|$  is any norm distance metric, e.g. the Euclidean distance. Note that, in general case,  $h(M, T) \neq h(T, M)$  and thus directed Hausdorff distance is not symmetric. The Hausdorff distance between  $M$  and  $T$  is defined as the maximum of both directed Hausdorff distance and thus it is symmetric. The Hausdorff distance  $H(M, T)$  is defined as

$$H(M, T) = \max(h(M, T), h(T, M)) \quad (2)$$

With two point sets  $M$  and  $T$ , if the Hausdorff distance between  $M$  and  $T$  is small, two point sets are partially matched, and if the Hausdorff distance is zero, two point sets are exactly matched.

Computing HD is challenging because its characteristics contain both maximization and minimization. Many efficient algorithms, in recent decades, have been proposed for reducing the computational complexity of the HD. We refer reader to the survey [17, 30] for general overview of this field. The efficient computing HD algorithms can be generally divided into two categories, approximate HD and exact HD. In the first category, which is approximation of HD, the algorithms try to efficiently find an approximation of the Hausdorff distance. These algorithms have been widely used in runtime-critical applications. On the other hand, the algorithms of the second category aim to efficiently compute the exact HD for point sets or special types of point sets like polygonal models or special curves and surfaces. Depending on data type of two sets, the HD algorithms can also be classified as polygonal models, curves and mesh surfaces, point sets.

With data type is polygonal models, a linear time algorithm for computing HD between two non-intersecting convex polygons was presented in [31]. The algorithm has a computational complexity of  $O(m + n)$ , where  $m$  and  $n$  are the vertex counts. In [32], an algorithm for computing the precise HD between two polygonal meshes with the complexity of  $O(n^4 \log n)$  was presented. Due to the high computational complexity of exact HD calculation, approximate HD methods have been proposed. In [14], a method, that has the complexity of  $O((m + n) \log(m + n))$ , used Voronoi diagram to approximate the HD between simple polygons was presented. Another method for approximating the HD between complicated polygonal models was presented in [33]. By using Voronoi subdivision combined with cross-culling, non-contributing polygon pairs for HD are discarded. The method is very fast in practice and can reach interactive speed.

Many efficient algorithms were also proposed for calculating HD between mesh surfaces or curves. An efficient algorithm for calculating HD between mesh surfaces was presented in [34]. This algorithm is built on the specific characteristics of mesh surfaces, where the surface consists of triangles. To avoid sampling all points in the compared surfaces, the algorithm samples in the regions where the maximum distance is expected. The method for calculating the HD between points and freeform curves was presented in [35]. In [36], the improved method of [35] was presented for computing HD between two B-spline curves. For approximating the HD between curves, in [37], an algorithm was proposed, that converts the problem of computing HD between curves to the problem of computing distance between grids. In [38], an algorithm to compute the approximate HD between curves was presented. By approximating curves with polylines, the problem of computing HD between curves is converted to the problem of computing HD distance between line segments.

However, above methods are not general because the algorithms are based on specific characteristic of data types. Some general methods were proposed for point sets. An algorithm for finding the aggregate nearest neighbor (ANN) in database was proposed [39], that uses the R-Tree for optimizing the searching for ANN. The extension of [39], incremental Hausdorff distance (IHD) was proposed in [40] for efficiently calculating HD between two point sets. The algorithm uses two R-Trees for the same time, each for one point set, to avoid the interaction of all points in both sets. The aggregate nearest neighbor is simultaneously determined in both directions. However, complex structure of above algorithms makes the computation cost increase and the R-Tree is not suitable for general point sets.

In [29], a fast and efficient algorithm for computing exact Hausdorff distance between two point sets, which is known as a state-of-art algorithm, was proposed. The algorithm has two loops, with the outer loop for maximization and the inner loop for minimization. The inner loop can break as soon as the distance is found that is below the temporary HD (called *cmx*) because the rest iterations of inner loop do not make the value of *cmx* change, and the outer loop continues with the next point. Moreover, for improving performance, random sampling is also presented in this algorithm to avoid similar distances in successive iterations. Based on EARLY-BREAK [29], an efficient algorithm, namely local start search (LSS) or Z-order Hausdorff distance (ZHD), for computing exact HD between two arbitrary point sets was presented [41]. The LSS method uses Morton curve for ordering points. The main idea of the LSS algorithm is that if the break occurs in current loop at the point  $x$ , it is quite possible that the break will occur at the position near  $x$  in the next loop. In the LSS algorithm, the variable *preindex* is used for preserving the location of break occurrence. In the next outer loop, the inner loop starts from *preindex* and scan its neighbor for finding the distance below the value *cmx*. In [42], an efficient framework, which contains two sub-algorithms Non-overlap Hausdorff Distance (NOHD) and Overlap Hausdorff Distance (OHD), for computing the HD between two general 3D point sets was proposed. For 3D point sets, [43] presented diffusion search of efficient and accurate HD computation between 3D models. This proposed method contains two algorithms for two types of 3D model, the ZHD for sparse point sets and the OHD for dense point sets.

In this study, we proposed the Least Trimmed Square Line Hausdorff Distance (LTS-LHD) for measuring the dissimilarity between two line edge maps (LEMs), which are the sets of line segments. The Hausdorff distance between two point set is based on the spatial locality of two point sets. Thus, the structure R-Tree of ANN and IHD or Z-order of LSS are just suitable for ordering the point sets [41]. However, the Hausdorff distance between two sets of line segments is based on both the spatial locality and the direction of line segment. Therefore, the structure R-Tree of ANN and IHD or Z-order of LSS are not suitable for the set of line segments as LEM, and thereby, the EARLYBREAK is used for reducing the complexity of computing the LTS-LHD.

### 3 Proposed method

#### 3.1 LTS-LHD for face recognition

The original HD, the MAX-MIN distance, uses the distance of most mismatched points for measuring the distance between two sets. When the set is attacked with noise points, the original HD can not be used. The PHD was proposed for solving this problem by sorting the distance and taking the  $K$  ranked maximum value. However, PHD is an effective distance when the pollution of noise points is low. The MHD was also proposed for solving the problem of sensitivity of original HD with noise by taking the mean distance. As in [44], the LTS-HD was proposed for combining the advantage of the PHD and the MHD. The definition of directed LTS-HD from  $M$  to  $T$  is as follow

$$h_{LTS-HD}(M, T) = \frac{1}{K} \sum_{i=1}^K \left( \min_{t \in T} \|m - t\| \right)_{(i)} \quad (3)$$

where the  $(\min \|m - t\|)_{(i)}$  represents the  $i^{th}$  distance in the sorted sequence  $(\min \|m - t\|)_{m \in M}$ . The LTS-HD takes the mean of  $K$  minimum distances for measuring the distance between two sets.

In this paper, a new HD, called Least Trimmed Square Line Hausdorff Distance (LTS-LHD), for face recognition is proposed. Supporting  $M^l = \{m_1^l, m_2^l, \dots, m_P^l\}$  and  $T^l = \{t_1^l, t_2^l, \dots, t_Q^l\}$  are the LEMs of model and test images respectively;  $m^l$  and  $t^l$  are the line segments in the LEMs;  $P$  and  $Q$  are the number of line segments in model and test LEMs, respectively. The directed distance of the LTS-LHD from LEM  $M^l$  to LEM  $T^l$ ,  $h_{pLTS-LHD}(M^l, T^l)$ , is defined as

$$h_{pLTS-LHD}(M^l, T^l) = \frac{1}{\sum_{i=K^{th}}^P l_{m_i^l}} \sum_{i=K^{th}}^P l_{m_i^l} \left( mindist(m_i^l, T^l) \right)_{(i)} \quad (4)$$

where  $(mindist(m_i^l, T^l))_{(i)}$  presents the  $i^{th}$  value in the sorted sequence in ascending order  $(mindist(m^l, T^l))_{(m^l \in M^l)}$ ; and  $K^{th} = f * P$ , where  $f$  is a given fraction. The  $mindist(m_i^l, T^l)$  denotes the distance from the line segment  $m^l$  to its nearest neighbor in  $T^l$ , and it is defined as below

$$mindist(m^l, T^l) = \min_{t^l \in T^l} d(m^l, t^l) \quad (5)$$



where  $d(m^l, t^l)$  is the distance between two line segments  $m^l$  and  $t^l$ . Here, the directed distance of the LTS-LHD is used for measuring how far from LEM  $M^l$  to LEM  $T^l$ . So, different from the LTS-HD as in (3), where the directed distance is the average of smallest distances, the distance as in (4) is the weighted average of largest distances, which are greater or equal to the  $K^{th}$  ranked value, from line segment  $m^l$  to its nearest neighbor  $mindist(m^l, T^l)$ . However, the directed distance of the LTS-LHD as in (4) still has weakness. Supporting  $m_1^l$  and  $m_2^l$  are two line segments in  $M^l$ ,  $d_1$  and  $d_2$  are two distances from these segments to their neighbors. Assuming  $d_1$  is greater than the  $K^{th}$  ranked value and  $d_2$  is less than the  $K^{th}$  ranked value of the sorted sequence  $(mindist(m_i^l, T^l))_{(m^l \in M^l)}$ . As in (4),  $d_1$  is used for the computing of  $h_{pLTS-LHD}(M^l, T^l)$ . However, it is possible that  $l_{m_1^l} \cdot d_1 \ll l_{m_2^l} \cdot d_2$ , because line segment  $m_2^l$  is much longer than line segment  $m_1^l$ . The miss match of the long line segment is more serious than the short line segment. So, the line segment  $m_2^l$  is much more important than the line segment  $m_1^l$  for computing the directed distance of LTS-LHD. Here, we modify the (4) for proposing the directed distance of the LTS-LHD as follow

$$h_{LTS-LHD}(M^l, T^l) = \frac{1}{\sum_{i=K^{th}}^P l_{m_i^l}} \sum_{i=K^{th}}^P (l_{m_i^l} \cdot mindist(m_i^l, T^l))_{(i)} \quad (6)$$

where  $(l_{m_i^l} \cdot mindist(m_i^l, T^l))_{(i)}$  presents the  $i^{th}$  value in the sorted sequence in ascending order  $(l_{m_i^l} \cdot mindist(m_i^l, T^l))_{(m^l \in M^l)}$ . The directed distance of the LTS-LHD is the weighted average of largest values of the product between length of line segment and the distance from that line segment to its nearest neighbor  $l_{m_i^l} \cdot mindist(m_i^l, T^l)$ .

In general,  $h_{LTS-LHD}(M^l, T^l) \neq h_{LTS-LHD}(T^l, M^l)$ . Thus, the primary LTS-LHD is defined as

$$H_{pLTS-LHD}(M^l, T^l) = \max(h_{LTS-LHD}(M^l, T^l), h_{LTS-LHD}(T^l, M^l)) \quad (7)$$

In our previous work [1], a new data structure of LEM was proposed. Due to the angle between line segments and the horizontal axis, the line segments in LEM are grouped into  $N$  groups and  $180/N$  degrees for each group. For example, in this paper, we use  $N = 18$ . An example of new data structure of LEM is shown in Fig 1.

The distance between two line segments  $d(m^l, t^l)$  is defined as

$$d(m^l, t^l) = \begin{cases} \|m^l - t^l\| & \text{where } g_{t^l} - k \leq g_{m^l} \leq g_{t^l} + k \\ V & \text{otherwise} \end{cases} \quad (8)$$

where  $V$  is the value that is larger than the largest possible value of distance between two line segments;  $\|m^l - t^l\|$  is the distance between two line segments and defined as

$$\|m^l - t^l\| = \sqrt{d_{pa}^2(m_i^l, t_j^l) + d_{pe}^2(m_i^l, t_j^l) + d_{\theta}^2(m_i^l, t_j^l)} \quad (9)$$

where  $d_{pa}$  is the parallel distance, which is the vertical distance between two lines;  $d_{pe}$  is the perpendicular distance, which is minimum displacement to align either the left end points or the right end

points of two lines; and  $d_{\theta}(m_i^l, t_j^l) = \theta^2(m_i^l, t_j^l) / W$  is the orientation distance.  $\theta^2(m_i^l, t_j^l)$  is the smallest intersection angle between two lines and  $W$  is a weight, which could be determined by a training process.

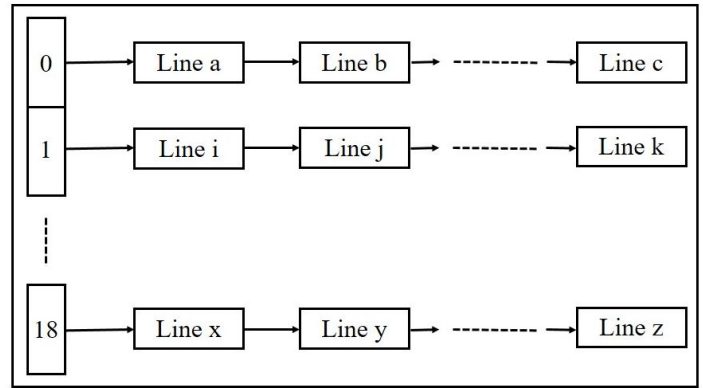


Figure 1: A novel data structure for LEM

It is possible that the line segment  $m^l$  could take a line  $t^l$ , that intersection angle between  $m^l$  and  $t^l$  is large, as it nearest neighbor. However, the line segment reflects the structure of human face, two corresponding line segments can not have large angle variation. For alleviating the undesire mismatch, the line segment finds its nearest neighbor if the group index ( $g_{m^l}$  and  $g_{t^l}$ ) of two line segments are slightly different as in (8). On the other hand, the distance between two line segments takes a large value  $V$ .

The number of corresponding line pairs between the model and the test LEM can be used as other measure similarity. Because test and model images are aligned and scaled at the same size by preprocessing before matching, if the line segment  $m^l$  finds that line segment  $t^l$  is the nearest neighbor in  $T^l$  and line segment  $t^l$  locates in the position neighborhood  $N_p$  of  $m^l$ , such line segment  $m^l$  could be named as *high confident line*. A *high confident ratio*, as in [27], of an image could be defined as the ratio between number of high confident line segments ( $N_{hc}$ ) and number of total line segments in the LEM of face image ( $N_{total}$ ), as follow

$$R = \frac{N_{hc}}{N_{total}} \quad (10)$$

Hence, the number disparity between two LEMs has mathematical formula as follow

$$D = 1 - \frac{R_M + R_T}{2} \quad (11)$$

The complete version of Hausdorff distance between two LEMs is defined as

$$H_{LTS-LHD} = \sqrt{H_{pLTS-LHD}^2 + W_n^2 D^2} \quad (12)$$

where  $W_n$  is a weight that be determined by a training process.

### 3.2 EARLYBREAK for LTS-LHD

The directed distance of the LTS-LHD as in (6) is the weighted average of  $(P - K^{th})$  largest values of the product between length of

line segment and distance to its nearest neighbor  $l_m.mindist(m^l, T^l)$ . For computing the directed distance of the LTS-LHD, the distances from line segment  $m^l$  to all of line segments in  $T^l$  must be calculated for finding the minimum value, which is the distance from line segment  $m^l$  to its nearest neighbor  $mindist(m^l, T^l)$ . This process could be named as the inner loop. The inner loop must be performed with all line segments  $m^l \in M^l$  and we call this is the outer loop. Assuming that  $(P - K^{th})$  temporary largest values are found and the minimum of these values is assigned to  $cmax$ . If a line segment  $m^l$  in the outer loop find out a line segment  $t^l$  in inner loop that makes the product between length of line segment  $m^l$  and distance  $d(m^l, t^l)$  between two line segments is below the value of  $cmax$ , such line segment  $m^l$  is non-contributed line segment for the computing of the directed distance LTS-LHD. So, the computing distance from line segment  $m^l$  to the remaining line segments  $t^l$  is not necessary. Therefore, the computing of directed distance LTS-LHD could break and continuing with the next line segment in the outer loop as soon as a non-contributed line segment be found. Thus, the number of iterations of the inner loop is reduced. The lower of average number of inner iteration is, the lower of computational complexity of the LTS-LHD computing is. Here, we proposed a method using EARLYBREAK for reducing the computational complexity of the LTS-LHD by reducing the average number of inner iterations.

The Algorithm 1 describes our proposed method. Line 5 and line 13 are the outer loop and the inner loop, respectively. The function  $DIST(.,.)$  in the Algorithm 1 is used for calculating the distance between two line segments as in (9).

The main steps of the Algorithm 1 are summarized as follows:

- A matrix  $h$  is created for saving the length of line segment  $m^l$  and the value of the product between length of line segment  $m^l$  and distance to its nearest neighbor.
- Adding line segments  $t^l$  having group index  $g_t$  into  $list$ .
- If there is at least one line segment in  $list$ , the inner loop will be executed. For each line segment  $m^l \in M^l$ , initializing distance to nearest neighbor  $cmin = \infty$ .
  - For each line segment in  $list$ , the distance from  $m^l$  to  $t^l$  is calculated. If a distance makes the product between it and length of line segment below the value of  $cmax$ , the algorithm will **break** and continue with the next line segment in the outer loop. In the other hand, this distance is used for updating the value of  $cmin$ .
  - The product between  $cmin$  and length of line segment will be used for updating the matrix  $h$ .
- On the other hand, if there is no line segment in  $list$ , the matrix  $h$  will be updated with length of line segment  $m^l$  and the large value  $V$ , which is the distance from line segment  $m^l$  to its nearest neighbor.
- The matrix  $h$  will be sorted in ascending order for each interaction of outer loop according to the values of the first row of matrix  $h$ .

In the Algorithm 1, during first  $K_M$  iterations of the outer loop, the value of  $cmax$ , which is the minimum value of matrix  $h$ , is 0.

The condition as line 15 in the Algorithm 1 does not appear, thus, the early break does not occur in first  $K_M$  iterations of the outer loop. The first  $K_M$  iterations of the outer loop, the value of  $K_M$  elements in the matrix  $h$  are updated. In the next iterations of the outer loop, the early break occurs if the value of product between the length of line and its distance to the nearest neighbor is below  $cmax$ .

---

#### Algorithm 1 : EARLYBREAKING for LTS-LHD

---

```

1: Inputs: Edge map  $M^l$  and  $T^l$ , fraction  $f$ 
2: Outputs: Directed Hausdorff distance  $h_{LTS-LHD}(M^l, T^l)$ 
3:  $K_M = (1 - f) * P$ 
4:  $h = zeros(2, K_M)$ 
5: for each line segment  $m^l$  in edge map  $M^l$  do
6:    $cmax = h(1, 1)$ 
7:   Get the group index  $g_m$  of line segment  $m^l$ 
8:    $cmin = \infty$ 
9:   for  $g_t = g_m - k : g_m + k$  do
10:     Insert line segments  $t^l$  has group index  $g_t$  into  $list$ 
11:   end for
12:   if  $list$  is not empty then
13:     for each line segment  $t^l$  in  $list$  do
14:        $d = DIST(m^l, t^l)$ 
15:       if  $d * LM \leq cmax$  then
16:          $cmin = 0$ 
17:         break
18:       end if
19:       if  $cmin > d$  then
20:          $cmin = d$ 
21:       end if
22:     end for
23:     if  $(cmin * LM) > cmax$  then
24:        $h(1, 1) = cmin * LM$ 
25:        $h(2, 1) = LM$ 
26:     end if
27:   else
28:      $h(1, 1) = V * LM$ 
29:      $h(2, 1) = LM$ 
30:   end if
31:   sort  $h$  in ascending order of the value of first row
32: end for
33:  $h_{LTS-LHD}(M^l, T^l) = sum(h(1, :)) / sum(h(2, :))$ 

```

---

By using the face database of Bern university [45] for training process, we obtain that:  $W = 45$ ,  $Np = 6$  and  $Wn = 13$ .

### 3.3 Analysis of computational complexity

Supporting  $P$  and  $Q$  are the number of line segments in LEM  $M^l$  and in LEM  $T^l$ , respectively. In the LHD method [27], each line segment  $m^l \in M^l$  calculate the distance to all line segments in  $T^l$  for finding its nearest neighbor. The directed distance of the LHD is the weighted average of the distances from line segments  $m^l \in M^l$  to their nearest neighbors in  $T^l$ . So, the complexity of computing directed distance of the LHD is  $O(PQ)$ .

The directed distance LTS-LHD computing as in (6) has computational complexity  $O(PQ)$  in the worst case. In the worst case,

all of line segments of two LEMs are located in the same group, and each line segment has to calculate distance to all line segments in other LEMs for finding nearest neighbor. On the other hand, with the best case, if all line segments of each LEM are located in one group and two groups of two LEMs are much different, the computational complexity of directed Hausdorff distance is  $O(P)$ . If the line segments of LEM  $T^l$  are equally divided into 18 groups, the computational complexity of directed Hausdorff distance would be  $O(PQ(2k + 1)/N)$ , where  $k$  is the difference of group index as defined in (8). The computational complexity of the LTS-LHD is always better than the LHD because the lower bound of the LTS-LHD computational complexity is the LHD computational complexity.

The Algorithm 1, in which the EARLYBREAK is applied for the LTS-LHD, has the computational complexities  $O(P)$  and  $O(PQ)$  for the best case and the worst case, respectively. In general case, assuming that line segments of LEM are equally divided into groups, the Algorithm 1 has a computational complexity of  $O((1 - f)PQ(2k + 1)/N + fPX)$ , where  $X$  is used to denote the average number of iterations in the inner loop. The lower value of  $X$  is, the lower of computational complexity of the method is and vice versa. The question is, in general, how high of the value of  $X$  is? In the formal way, the value of  $X$  in general case could be found through the analysis of probability theory.

Considering picking a random line segment  $t^l$  in the inner loop of the Algorithm 1, the distance  $d$  measured between line segment  $t^l$  and line segment  $m^l$  in current outer loop is a random variable. The event that meeting the condition that  $d$  is over  $cmax$  is denoted as  $e$ . The probability of that event is that  $P(e) = q$ . The event  $e$  means non-appearance of the break in the algorithm. Obviously, the event  $\bar{e}$ , that  $d$  is less than  $cmax$ , occur with probability  $P(\bar{e}) = p = 1 - q$ .

Assuming that the inner loop has been implemented for  $X$  times before the break occurs. This is equivalent to that  $(X - 1)$  distances from line segment  $m^l$  in the outer loop to the line segments  $t^l_1, t^l_2, \dots, t^l_{X-1}$ , namely  $d_1, d_2, \dots, d_{X-1}$ , are over  $cmax$  and one distance  $d_X \leq cmax$ . The probability density function of  $X$  is given by

$$f(x) = P(d_1 > cmax, \dots, d_{x-1} > cmax, d_x \leq cmax) = q * q * \dots * q * p = q^{x-1} p \tag{13}$$

Fig. 2 shows the probability distribution  $f(x)$ . The expectation of average number iterations of the inner loop  $X$  is equivalent to the expected value of  $f(x)$

$$E[X] = \sum_{x=1}^{\infty} x f(x) = \sum_{x=1}^{\infty} x q^{x-1} p \tag{14}$$

The Eq. (14) could be rewritten in the form of a polynomial as follows

$$E[X] = p + 2qp + 3q^2p + 4q^3p + \dots \tag{15}$$

By multiplying both side of (15) with  $q$  and subtracting the resulting equation from (15), a simpler formula of (15) is given as

$$E[X](1 - q) = p(1 + q + q^2 + q^3 + \dots) = p \frac{1}{1 - q} \tag{16}$$

Then, by using  $p = 1 - q$ , the expectation of number of iterations in the inner loop  $X$  could be found as

$$E[X] = \frac{1}{1 - q} = \frac{1}{p} \tag{17}$$

Equation (17) means that the number of iterations in the inner loop until the break depends on the value of  $p$ , which is the probability that a distance  $d$  is less than  $cmax$ . The higher  $p$  means the lower number of tries before early break and vice versa. The value of  $p$  depends on the value of  $cmax$ . For a larger value of  $cmax$ , it is easier for picking a line segment making the distance  $d$  to the current line segment of the outer loop less than  $cmax$ . Fig. 3 illustrates the relation between  $cmax$  and the probability  $p$ . Here, it is assumed the distance  $d$  is a random variable with normal distribution for illustration. The value of  $p$  does not depend on the size of set, but rather on the value of  $cmax$  and the distribution of pairwise distance  $d$ .

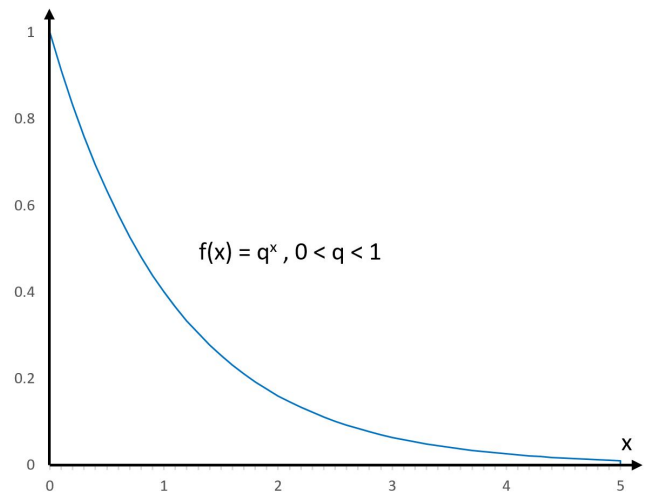


Figure 2: Probability density function  $f(x)$

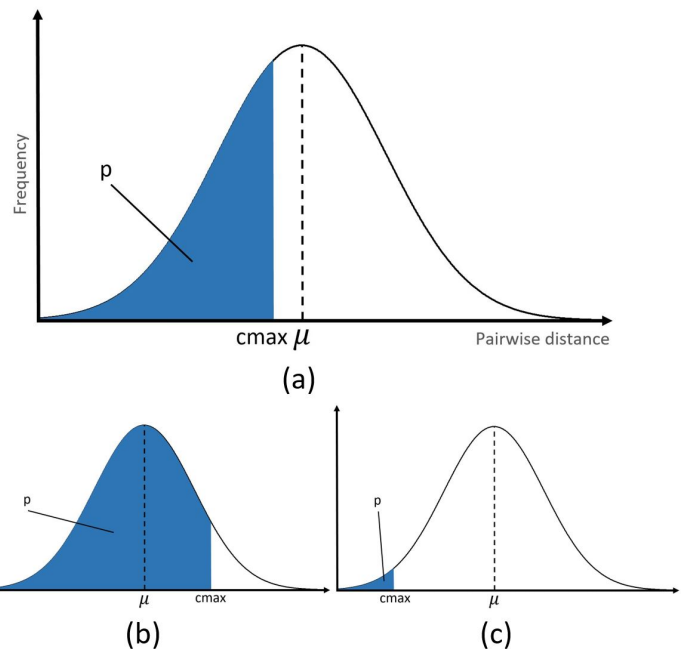


Figure 3: Normal distribution of pairwise distance is assumed. (a) Relation between  $cmax$  and probability  $p$ . (b) Large value of  $cmax$ . (c) Small value of  $cmax$ .

## 4 Experiments

In this section, the performance of the proposed method, LTS-LHD, is evaluated for face recognition. It is done by measuring the distance from test image to all model images for finding the smallest distance. The recognition rate, which is the ratio of number of images correctly classified to the total number of images in the test set, is used for evaluating.

In this study, the face database from the University of Bern [45] and the AR face database from Purdue University [46] are used. Bern university face database contains frontal views of 30 people. Each person has 10 gray images with different head pose variations: two frontal pose images, two looking to the right images, two looking to the left images, two looking upward image and two looking downward images. The AR face database contains 2599 color face images of 100 people (50 men and 50 women), there are 26 images for each person and be divided into 2 sessions separated by two weeks interval. Each session has 13 images are the frontal view faces with different facial expressions, illumination conditions, and occlusions (sun glasses and scarf). However, one of frontal face image is corrupted (W-027-14.bmp) and only 99 pairs of face image are used for examining the performance of system for face recognition under normal conditions. In our experiment, a preprocessing before recognition process is used for locating the face. All image are normalized such that the two eyes were aligned roughly at the same position with a distance of 80 pixels. After that, all images are cropped with size  $160 \times 160$  pixels. The experiments are conducted on the PC with 3GHz CPU and 4GHz RAM.

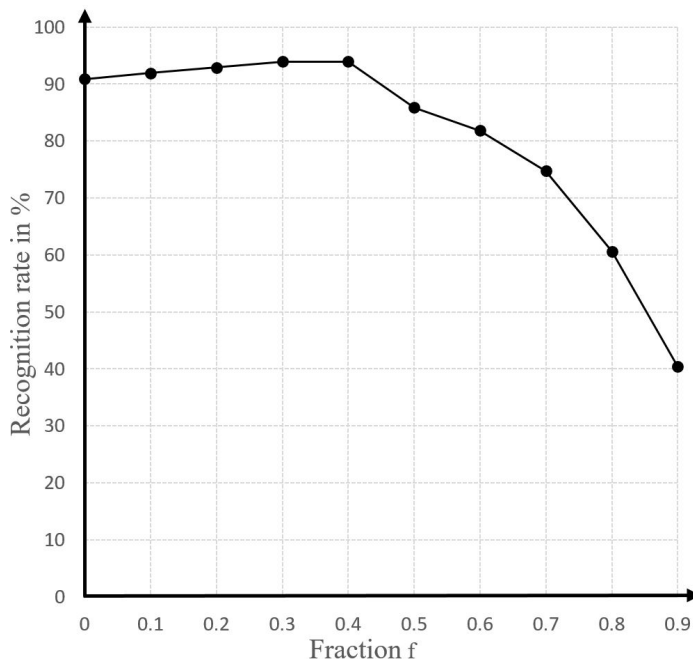


Figure 4: Influence of fraction  $f$  on recognition rate for AR database

### 4.1 Influence of fraction $f$ on the performance

The recognition rate of the proposed method is expected to be low for both low and high values of fraction  $f$ . With high value of fraction  $f$ , a low number of line segments, which have largest values of

the product between their length and distance to their neighbor, is used for computing the directed distance of the LTS-LHD. However, the outliers are commonly the line segments have largest values of that product. The high value of  $f$  means most of line segments used for computing the directed distance LTS-LHD are outliers, as in (6). On the other hand, when the value of fraction  $f$  is too low, a high number of line segments is used for computing the directed distance of the LTS-LHD. This means high number of line segments in similar regions of faces is used for the calculation of  $h_{LTS-LHD}$ . And thus, the contribution of line segments that discriminate the faces becomes low.

Fig. 5 shows the recognition rate of the proposed method for the AR database with various values of fraction  $f$ . It must be noticed that the value  $f$  of 1 is meaningless, because the value of directed distance of the LTS-LHD  $h_{LTS-LHD}$  is zero for all face image pairs. The proposed method has highest recognition rate at the fraction  $f$  values 0.3 and 0.4.

The computational complexity of the proposed method is based on the average number of iterations in the inner loop, and thus is based on the value of  $c_{max}$ . The higher value of  $c_{max}$  is, the lower of runtime is and vice versa. The low value of fraction  $f$  means the large number of line segments is used for computing  $h_{LTS-LHD}$ , and the value of  $c_{max}$  becomes low. On the other hand, the high value of fraction  $f$  means the high value of  $c_{max}$ , and thus the computational complexity of the method is low. So, the value of fraction  $f$  is chosen at 0.4.

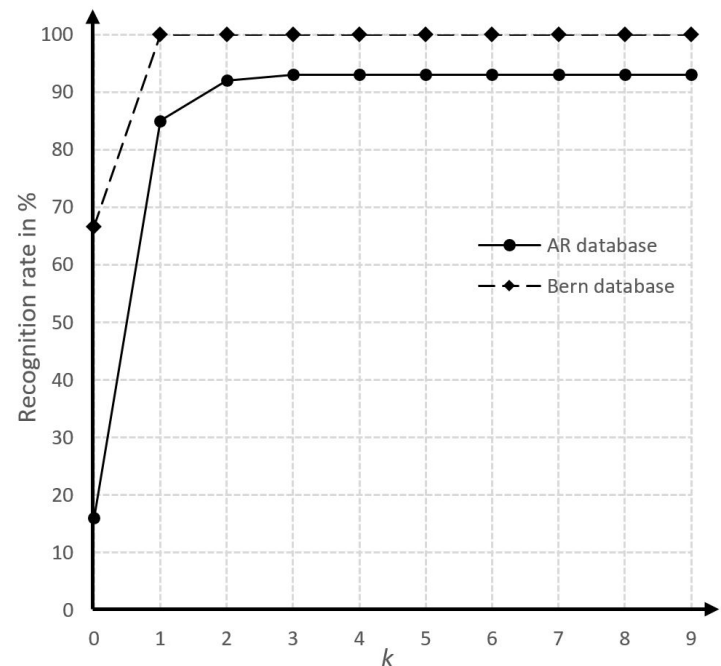


Figure 5: Influence of parameter  $k$  on recognition rate of proposed method

### 4.2 Influence of parameter $k$ on the performance

The recognition rate of the proposed method is expected to be low for low value of parameter  $k$  and vice versa. As in algorithm 1, the value of  $k$  determines the number of line segments in the inner loop. The low value of  $k$  has strong effect on system performance. Supporting  $m^l$  is the line segment in current outer loop. The low value



of  $k$  means  $m^l$  has to find its nearest neighbor in a few number of line segments in  $list$ , as in algorithm 1. And thus it is possible that  $list$  does not contain the corresponding line segment of  $m^l$ , and  $m^l$  could take another non-corresponding line segment as its nearest neighbor. On the other hand, the higher value of  $k$  is not necessary because too much line segments, which are too far from  $m^l$ , are added into  $list$ . Such non-contributed line segments make the number of iterations of inner loop increase and thus the runtime of method increases. Fig. 5 shows the recognition rates of the proposed method with various values of  $k$  for both the Bern university database and the AR database. The recognition rate does not change for the value of  $k$  higher than 2 for the Bern university database and 3 for the AR database. So, the chosen value of  $k$  is 3.

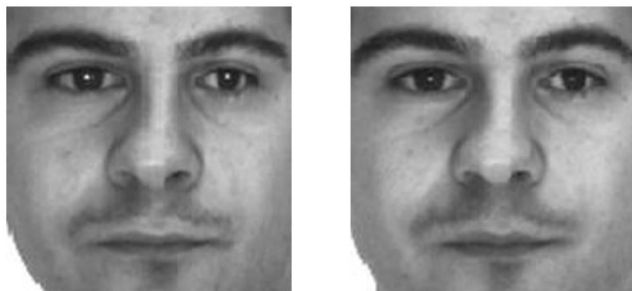
In the rest of this section, the recognition rate of the proposed method is compared with the LHD method in [27] and the RLHD method in [1], using average HD for measuring the dissimilarity between LEMs.

#### 4.3 Face recognition under normal conditions

The frontal face images in normal conditions in the Bern university database and the AR database are used for evaluating the performance of the proposed method. Each person has two images, one for test set and one for the model set. The example of images in this experiment is shown in Fig. 6. The recognition rates of different methods are given in Table. 1. The recognition rates of all methods with Bern university database are higher than those with the AR database. The reason is the different between two images of each person in AR database is larger than Bern database. The illumination of model image and test image in AR database are also different. The recognition rate of the proposed method is equal to the RLHD method.



(a) Bern University Database



(b) AR Database

Figure 6: Example of two pairs frontal face image in normal lighting condition in the face database of Bern University and the AR face database

Table 1: Face recognition result

Method	Recognition rate	
	Bern database	AR database
LHD	100%	93%
RLHD	100%	94%
Proposed method	<b>100%</b>	<b>94%</b>

The matching time for the Bern university database of different methods are given in Table. 2. The proposed method has the runtime 68% lower than the LHD method and 17.5% lower than RLHD method. The improvement in runtime is achieved by using EARLYBREAK for reducing the average number of iterations in the inner loop.

Table 2: Matching time of different methods

Method	LHD	RLHD	Proposed method
Matching time (second)	337	131	108

#### 4.4 Face recognition under varying lighting conditions and poses

The performance of the proposed method is also compared with the ones of the LHD and the RLHD methods for face recognition in the non-ideal conditions, e.g. face image with different poses or different lighting conditions. The AR database is used for evaluating the performance of different methods with varying lighting conditions of face image. Frontal face images of 100 people are used as model set. The face images with a light source on left side of face, with a light source on right side of face and with light sources on both sides of face are divided into three test sets with 100 images for each set. The recognition rates of different methods are given in Table. 3. The non-ideal lighting conditions make the recognition rate of all methods approximate 10% decrease. The face recognition accuracy of proposed method is 1%, on average, higher than the LHD and the RLHD methods. The interesting point of the experiment is that all three methods, in the condition of left light on, give the same recognition rates as the normal lighting condition in Table. 1 while the recognition rate in right light on condition is 6% - 9% lower than the ideal lighting condition. This could be due to the fact that the illumination of the right light is stronger than the left light. When both light on, the recognition rates of all methods are 12% lower than recognition rates in ideal lighting condition. The over-illumination has strong effect on recognition rates of all methods.

Table 3: Face recognition with varying lighting conditions

Lighting conditions	LHD	RLHD	Proposed method
Left light on	93%	94%	<b>94%</b>
Right light on	87%	85%	<b>88%</b>
Both light on	71%	72%	<b>72%</b>
Average	83.67%	83.67%	<b>84.67%</b>

The Bern university database is used for evaluating the performance of different methods with different poses of face image. The model set contains 30 frontal face images of 30 people. The test

set contains images of 30 people with different poses, e.g. looking to the left and right, looking up and down, for each person. The recognition rates of different methods are summarized in Table. 4. The pose variations have strong effect on recognition rates of all methods, where the recognition rates decrease 40% - 50% in comparing with the results in Table. 1. This could be explained that there are portions of face missing in comparing with frontal face. The recognition rate of the proposed method is lower than the RLHD method with the looking right image and higher than the RLHD method in other conditions. On average, the proposed method has the recognition rate 2% higher than the RLHD method and 3% higher than the LHD method.

Table 4: Face recognition with varying poses

	LHD	RLHD	Proposed method
Looking left	46.67%	53.33%	<b>63.33%</b>
Looking right	53.33%	46.67%	<b>46.67%</b>
Looking up	66.67%	63.33%	<b>66.67%</b>
Looking right	60%	70%	<b>63.33%</b>
<b>Average rate</b>	<b>56.67%</b>	<b>58.33%</b>	<b>60%</b>

## 5 Conclusion

The Hausdorff distance, which is used for measuring the degree of resemblance between two geometric objects, has been widely used in various science and engineering fields. The computational of HD computing is high because the computing contain both maximization and minimization. Many methods have been proposed in recent decades for reducing the computational complexity of MAX-MIN HD computing. However, the proposed methods for reducing the computational complexity of MAX-MIN HD computing can not be used for reducing the computational complexity of average HD computing. In face recognition, average HD is widely used for measuring the distance between two sets of features, instead of MAX-MIN HD, which is known as a sensitive measure with noise. The computational complexity of average HD computing as high as the MAX-MIN HD computing. The high computational cost restricts face methods using the average HD from real time applications.

The LHD and the RLHD use average HD for measuring the dissimilarity between two LEMs. In this paper, a modification of RLHD, called LTS-LHD was proposed for face recognition. The LTS-LHD uses only  $K_M$  line segments, not all of line segments as in RLHD, for calculating the directed distance. With a suitable parameter  $K_M$ , or suitable fraction  $f$ , the proposed method, LTS-LHD, has the performance slightly higher than the RLHD method, which is the average HD.

Moreover, in this paper, the EARLYBREAK is used for reducing the computational complexity of the proposed method. The early breaking can speed up the LTS-LHD by reducing the average number of iterations in the inner loop. The experimental results show that the runtime of proposed method is 68% lower than the LHD method and 17.5% lower than the RLHD method.

**Acknowledgment** We acknowledge the support of time and facilities from Ho Chi Minh City University of Technology (HCMUT),

VNU-HCM for this study.

## References

- [1] C. Dang-Nguyen, T. Do-Hong, "Robust line hausdorff distance for face recognition, in 2019 International Symposium on Electrical and Electronics Engineering (ISEE), HoChiMinh, Vietnam, 2019. <https://doi.org/10.1109/ISEE2.2019.8921218>
- [2] J. Wang and Y. Tan, "Hausdorff distance with k-nearest neighbors, in Proceedings of the Third International Conference on Advances in Swarm Intelligence - Volume Part II, ser. ICSII2. Berlin, Germany, 2012. <https://doi.org/10.1007/978-3-642-31020-1>
- [3] Y. Hu, Z. Wang, "A similarity measure based on hausdorff distance for human face recognition, in 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, China, 2006. <https://doi.org/10.1109/ICPR.2006.174>
- [4] S. Chen and B. C. Lovell, "Feature space hausdorff distance for face recognition, in Proceedings of the 2010 20th International Conference on Pattern Recognition (ICPR'10), Istanbul, Turkey, 2010. <https://doi.org/10.1109/ICPR.2010.362>
- [5] J. Paumard, "Robust comparison of binary images," Pattern Recognition Letters, **18**(10), 1057-1063, 1997. [https://doi.org/10.1016/S0167-8655\(97\)80002-5](https://doi.org/10.1016/S0167-8655(97)80002-5)
- [6] H. Zhu, T. Zhang, L. Yan, and L. Deng, "Robust and fast hausdorff distance for image matching," Optical Engineering, **51**(1), 2012. <https://doi.org/10.1117/1.OE.51.1.017203>
- [7] X. Li, Y. Jia, F. Wang, and Y. Chen, "Image matching algorithm based on an improved hausdorff distance," in 2nd International Symposium on Computer, Communication, Control and Automation (3CA 2013), Singapore, 2013. <https://doi.org/10.2991/3ca-13.2013.61>
- [8] C.H. Yang, S.H. Lai, and L.W. Chang, "Hybrid image matching combining hausdorff distance with normalized gradient matching," Pattern Recognition, **40**(4), 1173-1181, 2007. <https://doi.org/10.1016/j.patcog.2006.09.014>
- [9] D. Karimi and S. E. Salcudean, "Reducing the hausdorff distance in medical image segmentation with convolutional neural networks," IEEE Transactions on Medical Imaging, **39**(2), 499-513, 2020. <https://doi.org/10.1109/TMI.2019.2930068>
- [10] H. Yu, F. Hu, Y. Pan, and X. Chen, "An efficient similarity-based level set model for medical image segmentation," Journal of Advanced Mechanical Design, Systems, and Manufacturing, **10**(8), 01000108, 2016. <https://doi.org/10.1299/jamdsm.2016jamdsm0100>
- [11] F. Morain-Nicolier, S. Lebonvallet, E. Baudrier, S. Ruan, "Hausdorff distance based 3d quantification of brain tumor evolution from mri images," 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, 2007. <https://doi.org/10.1109/IEMBS.2007.4353615>
- [12] M. Grana, M. A. Vezagzones, "An endmember-based distance for content based hyperspectral image retrieval," Pattern Recognition, **45**(9), 3472-3489, 2012. <https://doi.org/10.1016/j.patcog.2012.03.015>
- [13] Y. Gao, M. Wang, R. Ji, X. Wu, Q. Dai, "3-D object retrieval with hausdorff distance learning," IEEE Transactions on Industrial Electronics, **61**(4), 2088-2098, 2014. <https://doi.org/10.1109/TIE.2013.2262760>
- [14] H. Alt, B. Behrendts, J. Blomer, "Approximate matching of polygonal shapes," Annals of Mathematics and Artificial Intelligence, **13**, 251-265, 1995. <https://doi.org/10.1007/BF01530830>
- [15] L. B. Kara and T. F. Stahovich, "An image-based, trainable symbol recognizer for hand-drawn sketches," Computer Graphics, **29**(4), 501-517, 2005. <https://doi.org/10.1016/j.cag.2005.05.004>
- [16] M. Field, S. Valentine, J. Linsey, and T. Hammond, "Sketch recognition algorithms for comparing complex and unpredictable shapes," in Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Three (IJCAI '11), Barcelona, Spain, 2011.
- [17] D. P. Huttenlocher, G. A. Klanderman, W. J. Rucklidge, "Comparing images using the hausdorff distance," IEEE Transactions on Pattern Analysis and Machine Intelligence, **15**(9), 850-863, 1993. <https://doi.org/10.1109/34.232073>
- [18] V. E. P and N. Sudha, "Robust hausdorff distance measure for face recognition," Pattern Recognition, **40**(2), 431-442, 2007. <https://doi.org/10.1016/j.patcog.2006.04.019>
- [19] M. P. Dubuisson and A. K. Jain, "A modified hausdorff distance for object matching," in Proceeding 12th International Conference on Pattern recognition, Jerusalem, Israel, 1994. <https://doi.org/10.1109/ICPR.1994.576361>

- [20] B. Takcs, "Comparing face images using the modified hausdorff distance," *Pattern Recognition*, **31**(12), 18731881, 1998. [https://doi.org/10.1016/S0031-3203\(98\)00076-4](https://doi.org/10.1016/S0031-3203(98)00076-4)
- [21] B. Guo, K. Lam, W. Siu, S. Yang, "Human face recognition using a spatially weighted hausdorff distance, in The 2001 IEEE International Symposium on Circuits and Systems (ISCAS 2001), Sydney, Australia, 2001. <https://doi.org/10.1109/ISCAS.2001.921027>
- [22] K. Lin, B. Guo, K. Lam, W. Siu, "Human face recognition using a spatially weighted modified hausdorff distance," in Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing. (ISIMP 2001), Hong Kong, China, 2001. <https://doi.org/10.1109/ISIMP.2001.925437>
- [23] K.H. Lin, K.M. Lam, and W.C. Siu, "Spatially eigen-weighted hausdorff distances for human face recognition," *Pattern Recognition*, **36**(8), 1827 - 1834, 2003. [https://doi.org/10.1016/S0031-3203\(03\)00011-6](https://doi.org/10.1016/S0031-3203(03)00011-6)
- [24] H. Tan and Y. Zhang, "Computing eigenface from edge images for face recognition based on hausdorff distance," in Fourth International Conference on Image and Graphics (ICIG 2007), Sichuan, China, 2007. <https://doi.org/10.1109/ICIG.2007.70>
- [25] H. Tan, Y.-J. Zhang, W. Wang, G. Feng, H. Xiong, J. Zhang, and Y. Li, "Edge eigenface weighted hausdorff distance for face recognition," *International Journal of Computational Intelligence Systems*, **4**(6), 1422-1429, 2011. <https://doi.org/10.1080/18756891.2011.9727894>
- [26] Y. Gao, "Efficiently comparing face images using a modified hausdorff distance," *IEE Proceeding Vision, Image and Signal Processing*, **150**(6), 346350, 2003. <https://doi.org/10.1049/ip-vis:20030805>
- [27] Y. Gao and M. K. H. Leung, "Face recognition using line edge map," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **24**(6), 764779, 2002. <https://doi.org/10.1109/TPAMI.2002.1008383>
- [28] C. Du, G. Su, X. Lin, "Face recognition using a modified line segment hausdorff distance," in Proceedings of the 2003 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.03EX693), Xi'an, China, 2003. <https://doi.org/10.1109/ICMLC.2003.1260095>
- [29] A. A. Taha and A. Hanbury, "An efficient algorithm for calculating the exact hausdorff distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**(11), 21532163, 2015. <https://doi.org/10.1109/TPAMI.2015.2408351>
- [30] H. Alt and L. J. Guibas, "Discrete geometric shapes: Matching, interpolation, and approximation, Handbook of computational geometry," Elsevier, 2000.
- [31] M. J. Atallah, "A linear time algorithm for the hausdorff distance between convex polygons," *Information Processing Letters*, **17**(4), 207-209, 1983. [https://doi.org/10.1016/0020-0190\(83\)90042-X](https://doi.org/10.1016/0020-0190(83)90042-X)
- [32] M. Barton, I. Hanniel, G. Elber, and M.-S. Kim, "Precise hausdorff distance computation between polygonal meshes," *Computer Aided Geometric Design*, **27**(8), 580-591, 2010. <https://doi.org/10.1016/j.cagd.2010.04.004>
- [33] M. Tang, M. Lee, Y. J. Kim, "Interactive hausdorff distance computation for general polygonal models," *ACM Transactions on Graphics*, **28**(3), 2009. <https://doi.org/10.1145/1531326.1531380>
- [34] M. Guthe, P. Borodin, R. Klein, "Fast and accurate hausdorff distance calculation between meshes," *Journal of WSCG*, **13**(2), 41-48, 2005.
- [35] H. Alt and L. Scharf, "Computing the hausdorff distance between curved objects," *International Journal of Computational Geometry and Applications*, **18**(4), 307320, 2008. <https://doi.org/10.1142/S0218195908002647>
- [36] X.D. Chen, W. Ma, G. Xu, and J.-C. Paul, "Computing the hausdorff distance between two b-spline curves," *Computer-Aided Design*, **42**(12), 1197-1206, 2010. <https://doi.org/10.1016/j.cad.2010.06.009>
- [37] E. Belogay, C. Cabrelli, U. Molter, R. Shonkwiler, "Calculating the hausdorff distance between curves," *Information Processing Letters*, **64**(1), pp. 1722, 1997. [https://doi.org/10.1016/S0020-0190\(97\)00140-3](https://doi.org/10.1016/S0020-0190(97)00140-3)
- [38] Y.B. Bai, J.H. Yong, C.Y. Liu, X.M. Liu, and Y. Meng, "Polyline approach for approximating hausdorff distance between planar free-form curves," *Computer-Aided Design*, **43**(6), 687698, 2011. <https://doi.org/10.1016/j.cad.2011.02.008>
- [39] D. Papadias, Y. Tao, K. Mouratidis, and C. Hui, "Aggregate nearest neighbor queries in spatial databases," *ACM Transactions on Database Systems*, **30**(2), 529576, 2005. <https://doi.org/10.1145/1071610.1071616>
- [40] S. Nutanong, E. H. Jacox, and H. Samet, "An incremental hausdorff distance calculation algorithm," *Proceedings of the VLDB Endowment*, **4**(8), 506517, 2011. <https://doi.org/10.14778/2002974.2002978>
- [41] Y. Chen, F. He, Y. Wu, and N. Hou, "A local start search algorithm to compute exact hausdorff distance for arbitrary point sets," *Pattern Recognition*, **67**, 139-148, 2017. <https://doi.org/10.1016/j.patcog.2017.02.013>
- [42] D. Zhang, F. He, S. Han, L. Zou, Y. Wu, and Y. Chen, "An efficient approach to directly compute the exact hausdorff distance for 3d point sets," *Integrated Computer-Aided Engineering*, **24**(3), 261277, 2017. <https://doi.org/10.3233/ICA-170544>
- [43] D. Zhang, L. Zou, Y. Chen, and F. He, "Efficient and accurate hausdorff distance computation based on diffusion search," *IEEE Access*, **6**, 13501361, 2018. <https://doi.org/10.1109/ACCESS.2017.2778745>
- [44] G. Chen, Q. Dai, X. Tang, and Z. Xu, "An improved least trimmed square hausdorff distance finger vein recognition," in 2018 5th International Conference on Systems and Informatics (ICSAI), Nanjing, China, 2018. <https://doi.org/10.1109/ICSAI.2018.8599439>
- [45] Bern University Database. <http://www.fki.inf.unibe.ch/databases/iam-faces-database/FullFaces.tgz>
- [46] A. Martinez and R. Benavente, "The AR face database," *CVC Technical Report*, **24**, Jun 1998.

## Organizational Agility Assessment of a Moroccan Healthcare Organization in Times of COVID-19

Fadoua Tamtam\*, Amina Tourabi

Systems Engineering and Decision Support Laboratory, National School of Applied Sciences, University IBN ZOHR, Agadir, 80000, Morocco

### ARTICLE INFO

Article history:

Received: .05 June, 2020

Accepted: 23 July, 2020

Online: 19 August, 2020

Keywords:

COVID-19

Organization agility

Assessment model

### ABSTRACT

Since its appearance, COVID-19 has severely impacted the healthcare sector all over the world. The healthcare organizations should be agile in order to cope with this new health crisis. Indeed, organization agility was highly recommended as an essential basis for flexibility, innovation, speed, as well competitiveness. Different research provided different conceptual models suitable to evaluate the organization agility. In this sense, this paper presents an assessment model, which by defining different agile enablers, criteria and attributes, aims at identifying the least and the most suitable enablers influencing the healthcare organization agility. To realize it practically, this paper uses the fuzzy logic approach which provides the improvement directions for enhancing the organization agility. Subsequently, the data gathered from a Moroccan healthcare organization was substituted in this assessment model and the level and the suggestions improvement for agility were derived. In this way, the organization will integrate the successful combination of the agility enablers in this dynamic environment.

## 1. Introduction

The story of the pandemic “COVID-19” began in 2019 when the first case were identified from Wuhan, China [1]. Since its first appearance, COVID-19 has been receiving an increasing attention by academic and executive specialists and many researches have been developed on it in order to provide a general definition of the virus. In the beginning, COVID-19 has created a global healthcare crisis, and then it disrupted other sectors: economic, environmental and social [2]. But perhaps the most significant pressure was for the healthcare organizations which strengthened their medical system [3] in order to enhance their responsiveness, adaptability, flexibility, which explains the importance of agility implementation in the healthcare sector through the outbreaks of COVID 19.

Agility concept was presented as the effective exploration of different competitive bases by including the suitable resources and practices in order to cope with the changing environment [4, 5]. Later, different proposals of agility definitions have been derived and which presented a general consensus [5]: It means the organization capacity to react quickly [5]–[11] to the varied changes in market demand [8]–[11] in terms of cost, specification, quality, quantity and delivery [11, 12]. Despite being defined in different ways and from different perspectives, agility has sometimes been used interchangeably to refer to concepts such as

adaptability, flexibility, speed, intelligence or sharpness. In contrast to this point of view, several authors have expressed the difference between these concepts, which justifies our choice to use the word "Agility" many times in our paper.

In order to evaluate the agility of an organization [5], several approaches such as system approach [5, 13], graph theory [5, 14], multi-grade fuzzy logic [15], regression analysis [5, 16] and other artificial intelligence techniques, such as neural network [15], neuro-fuzzy [15], have been used [5, 13, 14, 16–19]. A main objective of this study is to help the healthcare organization to implement an easier and less complicated practical tool in order to evaluate their agility [11]. The above purpose suggests an assessment model in which we evaluated the enablers influencing the adoption of agility [15].

Our paper is organized as follows: In the next paragraph, we review previous researches related to agile enablers. By presenting the fuzzy logic approach, we presented the required steps to apply this methodology to a real case. Moreover, the results provided are discussed and the limits of the study and suggestions for future research are finally presented [11].

## 2. Literature review: Agile enablers

According to different conceptual models of agility presented in literature, companies can benefit from different enablers [11] in order to achieve agility. These enablers, also known as providers

\*Corresponding Author: Fadoua Tamtam, [fadoua.tamtam@gmail.com](mailto:fadoua.tamtam@gmail.com)



or levers [20], were introduced by Gunasekaran [21, 22] in order to identify the required features of the agile organization [20]. In his study, he identified seven agile enablers: virtual corporation formation tools/metrics, physically distributed teams and manufacturing, quick partnership formation, concurrent engineering, integrated information system, quick prototyping tools and E-commerce [22]. In 1999, Yusuf et al. [4] presented different enablers under ten groups: the introduction of new products, the formation of partnerships, continuous improvement, short conception/production of deadlines, decentralized decision-making, response to market requirements ...etc [20]. Later, Sharifi and his colleagues proposed four enablers from four different areas: organization, people, technology and innovation [9, 23]. Based on their sample, Tolf et al. identified five essentials enablers for an agile organization: transparent and transient inter-organizational links at all levels, market sensitivity and customer focus, management by support for self-organizing employees, organic structures and flexible human and resource capacity for timely delivery [24]. In their paper [25], Lin and his colleagues suggested four agility enablers: collaborative relationships, process integration, information integration and customer sensitivity [26]. Other enablers were identified by Eshlaghy et al., as organizational structure, virtual organization, information technology, organizational culture, leadership, team working, empowerment and improvement, motivation system and planning and evaluation performance [27].

From this literature review, we can notice that there is no single list of agility enablers [20] which is due to the varied requirements of each organization [28]. However, all the enablers should have some criteria and attributes that make them agile. For example, the criterion called “Organizational structure” should be flexible to accept changes, this means that the different attributes of the organizational structure should be easily adaptable [20], while promoting a fluid flow of information [15], communication [29] and knowledge [30], which makes it possible to accept the interchangeability of employees [15] and focus on teamwork [20, 27, 30, 31]. For the other criterion “Processes”, it should be flexible [20, 30], promote and concentrate on external environment developments [20, 30, 32]. According to Sherehiy et al. [30], human resource agility, as an enabler of the agile organization [20], should be flexible [33], multi-skilled [15, 33], adaptable, resilient [20, 30, 32], able to cooperate [15, 20, 30], take personal initiative and cope well with changes [20, 30, 32]. The technology enabler should also be flexible like other enablers, modular and easily scalable [20].

Summarizing the above literature, different enablers, as listed in Table 1, are chosen as necessary conditions for organizational agility [33]. Table 1 suggest an assessment model in which we defined, firstly, the agile enablers that should be implemented by organizations; secondly, for each enabler different agile criteria are listed and finally agile attributes are identified in order to achieve the required agile criteria [15].

Table 1: Organizational agility enablers (Adapted from [12, 15, 27, 29–34])

Agile enablers	Agile criteria	Agile attributes	
<b>Management responsibility agility</b> (E <sub>1</sub> )	Organizational structure (E <sub>11</sub> )	Flattened, horizontal organizational structure that promotes innovation, training and having an open information, communication and knowledge policy (E <sub>111</sub> )	
		Fluid information flow (E <sub>112</sub> )	
		Staff interchangeability (E <sub>113</sub> )	
		Collaborative and team work (E <sub>114</sub> )	
	Devolution of authority (E <sub>12</sub> )	Clear definition of staff responsibility and authority (E <sub>121</sub> )	
		Training to create self-managed and multi-functional teams (E <sub>122</sub> )	
		Decentralized decision-making, knowledge and control (E <sub>123</sub> )	
		Loyalty and commitment to a project or a group (E <sub>124</sub> )	
		Authority change when tasks change (E <sub>125</sub> )	
	Nature of management (E <sub>13</sub> )	Participative management style (E <sub>131</sub> )	
		Clearly known management purpose (E <sub>132</sub> )	
		Management participation and support (E <sub>133</sub> )	
		Motivation of profit associated with a humanitarian approach (E <sub>134</sub> )	
		Regular conduct of employer–employees meetings (E <sub>135</sub> )	
		Quick evaluation and implementation of employee suggestions (E <sub>136</sub> )	
	<b>Manufacturing management agility</b> (E <sub>2</sub> )	Patient response adoption (E <sub>21</sub> )	Dominance of the culture of continuous improvement (E <sub>211</sub> )
			Communication media to collect responses (E <sub>212</sub> )
			Incorporating patient feedback into services (E <sub>213</sub> )
Staff empowerment to resolve patient issues (E <sub>214</sub> )			
Efficient information system and technology (E <sub>215</sub> )			
Change in business and technical processes		Flexible business system (E <sub>221</sub> )	
		Application of business process reengineering to reinvent and reorganize the organization (E <sub>222</sub> )	

	(E <sub>22</sub> )	Positive employee attitude towards change, new ideas and technology (E <sub>223</sub> )	
		Risk management (E <sub>224</sub> )	
	Outsourcing (E <sub>23</sub> )	Adopting supply chain management concepts to improve the efficiency of outsourcing (E <sub>231</sub> )	
		Exploitation of information technology (IT) in supply chain management (E <sub>232</sub> )	
		Involvement of suppliers and different agents in product/service development (E <sub>233</sub> )	
		Working with fewer qualified suppliers (E <sub>234</sub> )	
	Processes sensing (E <sub>24</sub> )	Promoting and concentrating on external environment developments (E <sub>241</sub> )	
	Processes responding (E <sub>25</sub> )	Reconfigurable process (E <sub>251</sub> )	
		Scalable process (E <sub>252</sub> )	
		Simple process to implement (E <sub>253</sub> )	
	Concurrent engineering (E <sub>26</sub> )	Process design (E <sub>261</sub> )	
		Intelligent Engineering Design Support System (E <sub>262</sub> )	
		Integrated multidisciplinary teams of customers and suppliers (E <sub>263</sub> )	
Continuous reengineering of the organization and business processes based on benchmarking (E <sub>264</sub> )			
<b>Human resource agility</b> (E <sub>3</sub> )	Employee status (E <sub>31</sub> )	Flexible employees to accept the adoption of new technologies (E <sub>311</sub> )	
		Multi-skilled and flexible staff (E <sub>312</sub> )	
		Implementation of job rotation system (E <sub>313</sub> )	
		Education and training for all the existing and new employees (E <sub>314</sub> )	
	Employee involvement (E <sub>32</sub> )	Employee cooperation (E <sub>321</sub> )	
		Employee empowerment (E <sub>322</sub> )	
	Human resource management practices (E <sub>33</sub> )	Entrepreneurial organizational culture (E <sub>331</sub> )	
		Reward programs to encourage innovation and based on financial and non-financial measures (E <sub>332</sub> )	
		Multi-skill training improving organizational agility (E <sub>333</sub> )	
		Multi-functional, developed and trained employees (E <sub>334</sub> )	
		Development of differentiation and diversity (E <sub>335</sub> )	
	Human resources capacities (E <sub>34</sub> )	Anticipation of problems linked to change and resolution of these problems (E <sub>341</sub> )	
		Personal initiative (E <sub>342</sub> )	
		Interpersonal and cultural adaptability (E <sub>343</sub> )	
		Resiliency (E <sub>344</sub> )	
	Coordination (E <sub>35</sub> )	Personal, informal, goal-oriented and spontaneous coordination (E <sub>351</sub> )	
		Network communication (E <sub>352</sub> )	
		Management-employee cohesion (E <sub>353</sub> )	
	Human knowledge and skills (E <sub>36</sub> )	Knowledge and skills management systems (E <sub>361</sub> )	
		Protection of sensitive information (E <sub>362</sub> )	
		Knowledge acquisition from internal and external sources (E <sub>363</sub> )	
	<b>Technology agility</b> (E <sub>4</sub> )	Manufacturing set-ups (E <sub>41</sub> )	Flexible manufacturing setups (E <sub>411</sub> )
			Less time to change machine settings (E <sub>412</sub> )
			Modernization of machines (E <sub>413</sub> )
Usage of collapsible set-ups, Jigs and Fixtures (E <sub>414</sub> )			
Usage of automated tools (E <sub>415</sub> )			
Active policy to keep work areas clean and tidy (E <sub>416</sub> )			
Product life cycle (E <sub>42</sub> )		Specification of product life to the patient (E <sub>421</sub> )	
		Company encourages patient to switch to new product (E <sub>422</sub> )	
		Products superior field performance for a stipulated period with least maintenance cost (E <sub>423</sub> )	
Product service (E <sub>43</sub> )		Products designed for easy serviceability (E <sub>431</sub> )	
		Products incorporated with a modular design (E <sub>432</sub> )	

		Service centers well equipped with spares (E <sub>433</sub> )
		Minimum time required to execute the planning and to restore the defective product to its original performance (E <sub>434</sub> )
	Production methodology (E <sub>44</sub> )	Management's interest towards investment on flexible manufacturing system (FMS) concepts (E <sub>441</sub> )
		Application of Lean manufacturing principles for waste elimination (E <sub>442</sub> )
		Development of products whose components are all outsourced and assembled in-house (E <sub>443</sub> )
		IT application for better supplier management (E <sub>444</sub> )
	Manufacturing planning (E <sub>45</sub> )	Execution of short range planning (E <sub>451</sub> )
		Organization's procurement policy based on time schedule (E <sub>452</sub> )
		Strategic network in supply chain management to exercise zero inventory system (E <sub>453</sub> )
		Improved manufacturing technology (E <sub>454</sub> )
		Structured and flexible manufacturing processes (E <sub>455</sub> )
	IT integration (E <sub>46</sub> )	IT utilities incorporated with reengineered pattern of working (E <sub>461</sub> )
		Electronic commerce [27] (E <sub>462</sub> )
	Customization (E <sub>47</sub> )	Rapid introduction of new products/services (E <sub>471</sub> )
		Responding to changing market requirements (E <sub>472</sub> )
		Products with high added value (E <sub>473</sub> )
First-time correct design (E <sub>474</sub> )		
Manufacturing strategy agility (E <sub>5</sub> )	Status of quality (E <sub>51</sub> )	Products/services exceeding patient expectations (E <sub>511</sub> )
		Carrying out surveys/studies to guarantee the quality status (E <sub>512</sub> )
		Usage of total quality management tools (E <sub>513</sub> )
	Status of productivity (E <sub>52</sub> )	Improved productivity in all functions (E <sub>521</sub> )
		Reduction of non value-adding costs (E <sub>522</sub> )
		Quality is not infused at the cost of productivity (E <sub>523</sub> )
	Cost management (E <sub>53</sub> )	Costing and product pricing system focused on value-added and non-value-added activities (E <sub>531</sub> )
		Costing system enabling the evaluation of future resource consumption (E <sub>532</sub> )
		Product cost fixed according to the pricing of the customer (E <sub>533</sub> )
	Time management (E <sub>54</sub> )	Scheduled activities (E <sub>541</sub> )
		IT based communication system (E <sub>542</sub> )
		Adoption of time compression technologies (E <sub>543</sub> )

### 3. Fuzzy logic methodology to evaluate organizational agility

In order to enhance organizational agility in practice, the use of different methods and tools were recommended in literature [11, 23]. Focusing on methodological articles [11], the fuzzy logic approach has been used to assess the current agility level and identify the weaker attributes that need a particular attention to enhance the organizational agility. This approach is preferred over other methodologies because it can take the linguistic data as input, then convert linguistic expressions into corresponding fuzzy intervals and finally express the results back in linguistic terms with the help of Fuzzy Agility Index (FAI) [5].

Many studies in literature have used fuzzy logic to measure agility level of the healthcare organization (e.g. [5, 35]). Taking cues from these papers, this study uses this approach to evaluate the agility of a Moroccan healthcare organization.

### 4. Numerical illustration of fuzzy logic approach

#### 4.1. About the healthcare organization

Our study has been done at a public hospital (referred as HealthOrg), located in Morocco and where patients can carry out

the diagnosis of COVID-19. In order to cope with the new dynamic environment, HealthOrg aims to strengthen its agility level. However, it found it difficult to identify enablers that influence its agility, in particular the weaker ones which need to be improved [15]. In this context, we aimed to evaluate the agility of HealthOrg.

Table 2 provides an illustration of different steps to apply the fuzzy logic approach [11].

Table 2: Steps required applying the fuzzy logic methodology (Adapted from [5, 32])

Steps
Identify a list of agile enablers that influence the organizational agility.
Define the linguistic variables for evaluating performance rating and importance weights of agile attributes.
Approximate the linguistic terms by the corresponding fuzzy intervals.
Calculate the FAI of the organization.
Match the FAI with the appropriate linguistic level.
Calculate Fuzzy Performance Importance Index (FPPI).

4.2. Fuzzy logic application

- Step1: Identification of agile enablers, criteria and attributes [32]: By identifying a list of five agile enablers from the literature [5], twenty-six criteria and ninety-eight attributes were identified (Table 1).
- Step 2: Definition of the linguistic variables for evaluating performance rating and importance weights of agile attributes [32]: Following this list, five experts (E1, E2,..., E5) from HealthOrg were asked to provide the weights in terms of linguistic variables ranging from “Very low (VL)” to “Very High (VH)” and ratings in terms of linguistic variables ranging from “Worst (W)” to “Excellent (E)” [5] (Table 3).

Table 3: Importance weight and performance rating of agile attributes

Agile attributes	Importance weight					Performance rating				
	E1	E2	E3	E4	E5	E1	E2	E3	E4	E5
E111	H	FH	M	H	FH	E	VG	VG	G	F
E112	H	M	FH	FH	H	P	P	F	G	G
E113	FH	H	FH	H	FH	VP	W	G	F	F
E114	H	M	H	FH	FH	E	G	G	F	F
E121	H	M	FH	M	FH	F	F	G	VG	VG
E122	H	FH	M	H	FH	P	F	P	G	F
E123	FH	H	FH	M	FH	W	W	P	F	P
E124	H	H	M	FH	H	W	W	VP	F	F
E125	M	H	H	FH	H	G	VG	E	E	G
E131	VH	H	H	VH	H	G	F	G	G	VG
E132	H	FH	H	FH	FH	E	E	G	F	E
E133	H	FH	H	FH	H	G	G	G	VG	G
E134	FH	M	FH	M	FH	F	G	F	F	G
E135	H	M	M	M	H	G	G	G	VG	F
E136	VH	H	H	VH	VH	VP	F	P	P	G
E137	H	FH	M	FH	H	W	F	G	F	G
E211	H	H	H	FH	H	G	VG	G	F	F
E212	FH	M	FH	FH	M	W	W	F	G	VP
E213	FH	FH	FH	FH	FH	VP	P	P	F	P
E214	VH	H	H	VH	H	F	G	E	VG	E
E215	H	FH	H	FH	FH	VP	P	W	F	F
E221	FL	M	FL	FL	M	VG	F	E	F	G
E222	H	FH	H	M	FH	W	VP	G	E	G
E223	H	M	M	FH	H	VP	F	F	W	E
E224	FH	H	H	H	FH	F	G	E	VG	G
E231	VH	H	VH	H	H	F	F	G	F	VG
E232	H	M	FH	H	FH	VP	P	G	F	W
E233	H	H	H	H	H	F	W	VP	F	G
E234	FH	H	M	FH	FH	E	VG	F	G	G
E241	H	M	FH	FH	FH	VP	F	G	G	F
E251	H	FH	M	H	M	F	VG	VG	G	VG
E252	FH	FH	FH	FH	FH	VP	VG	F	G	F
E253	FH	M	FH	FH	M	W	W	VP	F	VP
E261	H	FH	H	FH	M	G	F	G	VG	E
E262	H	M	FH	H	H	W	W	G	F	G
E263	FH	M	H	FH	FH	VP	P	W	P	F
E264	H	FH	H	FH	M	G	F	P	VP	G
E311	FH	FH	H	FH	H	P	W	F	G	P
E312	H	H	H	H	H	W	W	F	F	P
E313	VH	H	VH	VH	H	F	G	P	P	P
E314	FH	H	FH	M	FH	E	VG	G	E	F
E321	H	H	H	H	H	E	E	VG	F	F
E322	FH	M	H	H	H	P	F	F	G	VG
E331	H	FH	H	H	M	E	VG	G	F	G
E332	H	H	H	H	H	E	F	G	F	F
E333	H	M	M	M	M	W	W	F	G	F
E334	FH	FH	H	H	FH	W	P	VP	G	F
E335	H	FH	FH	H	H	E	E	E	G	E

E341	H	M	FH	M	H	VP	P	F	G	F
E342	H	FH	FH	FH	M	F	G	P	F	G
E343	H	FH	H	FH	M	F	E	VG	G	F
E344	H	FH	FH	H	H	G	F	F	E	G
E351	H	H	FH	H	M	VP	F	F	F	F
E352	FH	M	M	FH	FH	E	G	G	F	F
E353	H	FH	H	FH	FH	G	E	F	F	VG
E361	M	M	M	FH	FH	W	W	W	F	P
E362	H	M	FH	FH	H	E	E	E	E	VG
E363	H	FH	H	H	H	E	VG	F	P	F
E411	M	H	M	M	FH	VG	G	F	E	E
E412	FH	FH	M	M	H	W	G	P	F	F
E413	H	M	M	M	FH	G	F	G	E	F
E414	FH	H	H	H	H	G	F	F	F	VG
E415	M	H	FH	H	H	E	G	F	P	VP
E416	FH	H	M	H	FH	W	VP	P	VP	F
E421	H	VH	H	H	H	E	F	G	E	VG
E422	FH	FH	FH	M	H	E	E	G	G	G
E423	H	FH	M	M	FH	E	VG	F	G	G
E431	H	H	FH	H	FH	G	G	VG	E	E
E432	FL	FL	L	VL	M	E	VG	G	VG	E
E433	FL	M	FH	H	FH	G	F	E	E	VG
E434	M	FH	FH	H	VH	VP	P	F	VG	VG
E441	FH	VH	M	FH	VH	E	VG	VG	VG	G
E442	H	M	H	FH	H	G	F	VG	E	VP
E443	FL	H	VH	H	FH	G	F	VP	G	F
E444	VH	H	H	FH	VH	G	E	F	VG	F
E451	VH	VH	H	H	FH	W	VP	F	G	F
E452	VH	VH	H	VH	H	W	W	G	F	F
E453	FH	VH	FH	VH	M	VP	P	G	VG	F
E454	VH	VH	H	FH	FH	G	F	F	VP	P
E455	M	H	VH	H	H	E	VG	G	VG	F
E461	FH	FL	FL	FH	FH	G	G	F	F	F
E462	FH	M	FH	FH	H	VG	E	E	F	G
E471	H	H	FH	H	H	F	F	G	VG	VG
E472	FL	FH	FH	L	FH	E	E	E	VG	G
E473	H	H	FH	FH	H	G	E	VG	F	F
E474	VH	M	VH	H	VH	W	VP	F	G	E
E511	FH	FL	FL	M	FH	F	F	G	F	F
E512	FH	FH	FH	FH	H	E	VP	E	F	G
E513	H	H	M	H	H	E	E	E	E	E
E521	H	FH	VH	H	H	VP	G	E	G	G
E522	VH	M	VH	H	VH	VP	P	P	W	G
E523	H	VH	FH	FL	FL	E	VG	G	E	E
E531	M	M	FH	H	FL	G	E	F	G	F
E532	H	FL	H	M	H	F	F	VG	F	G
E533	M	M	M	M	M	F	F	F	G	G
E541	M	H	FL	M	M	VG	E	F	G	G
E542	H	H	FH	FH	H	G	F	F	F	VG
E543	M	H	M	VL	VL	G	F	F	F	VP

- Step 3: Approximation of the linguistic terms by the corresponding fuzzy intervals [32]: These linguistic variables were approximated by fuzzy intervals [5] chosen from literature [5, 25] and presented in Table 4.

Table 4: Linguistic variables and fuzzy numbers for weighting and rating of agility (Adapted from [25])

Importance Weight		Performance Rating	
Linguistic variable	Fuzzy number	Linguistic variable	Fuzzy number
Very Low (VL)	(0, 0.05, 0.15)	Worst (W)	(0, 0.5, 1.5)
Low (L)	(0.1, 0.2, 0.3)	Very Poor (VP)	(1, 2, 3)
Fairly Low (FL)	(0.2, 0.35, 0.5)	Poor (P)	(2, 3.5, 5)
Medium (M)	(0.3, 0.5, 0.7)	Fair (F)	(3, 5, 7)
Fairly High (FH)	(0.5, 0.65, 0.8)	Good (G)	(5, 6.5, 8)
High (H)	(0.7, 0.8, 0.9)	Very Good (VG)	(7, 8, 9)
Very High (VH)	(0.85, 0.95, 1.0)	Excellent (E)	(8.5, 9.5, 10)



To calculate the average fuzzy weight and performance rating of each attribute [5], the literature recommended using average operation method [5, 27].

Example: Average fuzzy weight of the attribute  $E_{111} = [H+FH+M+H+FH]/5 = (0.7, 0.8, 0.9)/5, (0.5, 0.65, 0.8)/5, (0.3, 0.5, 0.7)/5, (0.7, 0.8, 0.9)/5, (0.5, 0.65, 0.8)/5 = (0.54, 0.68, 0.82)$

Example: Average fuzzy performance rating of the attribute  $E_{111} = [E+VG+VG+G+F]/5 = (8.5, 9.5, 10)/5, (7, 8, 9)/5, (7, 8, 9)/5, (5, 6.5, 8)/5, (3, 5, 7)/5 = (6.1, 7.4, 8.6)$

The following step consists of calculating the rating of each criterion [5]. An example of this calculation for the criterion  $E_{11}$  is shown below.

Example: Rating of the criterion

$$E_{11} = \frac{\sum_{k=1}^{k=4} (\text{Average fuzzy performance rating} \otimes \text{Average fuzzy weight})}{\sum_{k=1}^{k=4} \text{Average fuzzy weight}}$$

$$= [(6.1, 7.4, 8.6) \otimes (0.54, 0.68, 0.82) \oplus (3.4, 5.0, 6.6) \otimes (0.54, 0.68, 0.82) \oplus (2.4, 3.8, 5.3) \otimes (0.58, 0.71, 0.84) \oplus (4.9, 6.5, 8.0) \otimes (0.54, 0.68, 0.82)] / [(0.54, 0.68, 0.82) \oplus (0.54, 0.68, 0.82) \oplus (0.58, 0.71, 0.84) \oplus (0.54, 0.68, 0.82)] = (4.17, 5.65, 7.11)$$

By using R language, fuzzy calculations are presented in Table 5.

Table 5: Fuzzy index of agile criteria rating

Agile criteria	Agile attributes	Average fuzzy performance rating	Average fuzzy weight	Criteria rating
E <sub>11</sub>	E <sub>111</sub>	(6.1, 7.4, 8.6)	(0.54, 0.68, 0.82)	(4.17, 5.65, 7.11)
	E <sub>112</sub>	(3.4, 5.0, 6.6)	(0.54, 0.68, 0.82)	
	E <sub>113</sub>	(2.4, 3.8, 5.3)	(0.58, 0.71, 0.84)	
	E <sub>114</sub>	(4.9, 6.5, 8.0)	(0.54, 0.68, 0.82)	
E <sub>12</sub>	E <sub>121</sub>	(5.0, 6.5, 8.0)	(0.46, 0.62, 0.78)	(3.52, 4.88, 6.28)
	E <sub>122</sub>	(3.0, 4.7, 6.4)	(0.54, 0.68, 0.82)	
	E <sub>123</sub>	(1.4, 2.6, 4.0)	(0.50, 0.65, 0.80)	
	E <sub>124</sub>	(1.4, 2.6, 4.0)	(0.58, 0.71, 0.84)	
E <sub>13</sub>	E <sub>125</sub>	(6.8, 8.0, 9.0)	(0.58, 0.71, 0.84)	(4.49, 5.99, 7.47)
	E <sub>131</sub>	(5.0, 6.5, 8.0)	(0.76, 0.86, 0.94)	
	E <sub>132</sub>	(6.7, 8.0, 9.0)	(0.58, 0.71, 0.84)	
	E <sub>133</sub>	(5.4, 6.8, 8.2)	(0.62, 0.74, 0.86)	
	E <sub>134</sub>	(3.8, 5.6, 7.4)	(0.42, 0.59, 0.76)	
	E <sub>135</sub>	(5.0, 6.5, 8.0)	(0.46, 0.62, 0.78)	
	E <sub>136</sub>	(2.6, 4.1, 5.6)	(0.79, 0.89, 0.96)	
E <sub>21</sub>	E <sub>137</sub>	(3.2, 4.7, 6.3)	(0.54, 0.68, 0.82)	(3.66, 4.93, 6.23)
	E <sub>211</sub>	(4.6, 6.2, 7.8)	(0.66, 0.77, 0.88)	
	E <sub>212</sub>	(1.8, 2.9, 4.2)	(0.42, 0.59, 0.76)	
	E <sub>213</sub>	(2.0, 3.5, 5.0)	(0.50, 0.65, 0.80)	
	E <sub>214</sub>	(6.4, 7.7, 8.8)	(0.76, 0.86, 0.94)	
	E <sub>215</sub>	(1.8, 3.2, 4.7)	(0.58, 0.71, 0.84)	
E <sub>22</sub>	E <sub>221</sub>	(5.3, 6.8, 8.2)	(0.24, 0.41, 0.58)	(4.45, 5.77, 7.04)
	E <sub>222</sub>	(3.9, 5.0, 6.1)	(0.54, 0.68, 0.82)	
	E <sub>223</sub>	(3.1, 4.4, 5.7)	(0.50, 0.65, 0.80)	
	E <sub>224</sub>	(5.7, 7.1, 8.4)	(0.62, 0.74, 0.86)	
E <sub>23</sub>	E <sub>224</sub>	(5.7, 7.1, 8.4)	(0.62, 0.74, 0.86)	(3.56, 5.05, 6.55)
	E <sub>231</sub>	(4.2, 5.9, 7.6)	(0.76, 0.86, 0.94)	
	E <sub>232</sub>	(2.2, 3.5, 4.9)	(0.54, 0.68, 0.82)	
	E <sub>233</sub>	(2.4, 3.8, 5.3)	(0.7, 0.8, 0.9)	
E <sub>24</sub>	E <sub>234</sub>	(5.7, 7.1, 8.4)	(0.50, 0.65, 0.80)	(3.4, 5.0, 6.6)
	E <sub>241</sub>	(3.4, 5.0, 6.6)	(0.50, 0.65, 0.80)	
	E <sub>251</sub>	(5.8, 7.1, 8.4)	(0.50, 0.65, 0.80)	
E <sub>25</sub>	E <sub>252</sub>	(3.8, 5.3, 6.8)	(0.50, 0.65, 0.80)	(3.68, 4.89, 6.18)
	E <sub>253</sub>	(1.0, 2.0, 3.2)	(0.42, 0.59, 0.76)	
	E <sub>261</sub>	(5.7, 7.1, 8.4)	(0.54, 0.68, 0.82)	

E <sub>31</sub>	E <sub>262</sub>	(2.6, 3.8, 5.2)	(0.58, 0.71, 0.84)	(3.29, 4.63, 6.03)
	E <sub>263</sub>	(1.6, 2.9, 4.3)	(0.50, 0.65, 0.80)	
	E <sub>264</sub>	(3.2, 4.7, 6.2)	(0.54, 0.68, 0.82)	
E <sub>32</sub>	E <sub>311</sub>	(2.4, 3.8, 5.3)	(0.58, 0.71, 0.84)	(3.08, 4.57, 6.06)
	E <sub>312</sub>	(1.6, 2.9, 4.4)	(0.7, 0.8, 0.9)	
	E <sub>313</sub>	(2.8, 4.4, 6.0)	(0.79, 0.89, 0.96)	
	E <sub>314</sub>	(6.4, 7.7, 8.8)	(0.50, 0.65, 0.80)	
E <sub>33</sub>	E <sub>321</sub>	(6.0, 7.4, 8.6)	(0.7, 0.8, 0.9)	(5.09, 6.55, 7.92)
	E <sub>322</sub>	(4.0, 5.6, 7.2)	(0.58, 0.71, 0.84)	
E <sub>34</sub>	E <sub>331</sub>	(5.7, 7.1, 8.4)	(0.58, 0.71, 0.84)	(4.69, 5.97, 7.21)
	E <sub>332</sub>	(4.5, 6.2, 7.8)	(0.7, 0.8, 0.9)	
	E <sub>333</sub>	(2.2, 3.5, 5.0)	(0.38, 0.56, 0.74)	
	E <sub>334</sub>	(2.2, 3.5, 4.9)	(0.58, 0.71, 0.84)	
	E <sub>335</sub>	(7.8, 8.9, 9.6)	(0.62, 0.74, 0.86)	
E <sub>35</sub>	E <sub>341</sub>	(2.8, 4.4, 6.0)	(0.50, 0.65, 0.80)	(4.21, 5.79, 7.32)
	E <sub>342</sub>	(3.6, 5.3, 7.0)	(0.50, 0.65, 0.80)	
	E <sub>343</sub>	(5.3, 6.8, 8.2)	(0.54, 0.68, 0.82)	
E <sub>36</sub>	E <sub>344</sub>	(4.9, 6.5, 8.0)	(0.62, 0.74, 0.86)	(4.20, 5.86, 7.45)
	E <sub>351</sub>	(2.6, 4.4, 6.2)	(0.58, 0.71, 0.84)	
	E <sub>352</sub>	(4.9, 6.5, 8.0)	(0.42, 0.59, 0.76)	
E <sub>41</sub>	E <sub>353</sub>	(5.3, 6.8, 8.2)	(0.58, 0.71, 0.84)	(5.01, 6.04, 7.03)
	E <sub>361</sub>	(1.0, 2.0, 3.3)	(0.38, 0.56, 0.74)	
	E <sub>362</sub>	(8.2, 9.2, 9.8)	(0.54, 0.68, 0.82)	
	E <sub>363</sub>	(4.7, 6.2, 7.6)	(0.66, 0.77, 0.88)	
E <sub>42</sub>	E <sub>411</sub>	(6.4, 7.7, 8.8)	(0.42, 0.59, 0.76)	(3.81, 5.30, 6.74)
	E <sub>412</sub>	(2.6, 4.1, 5.7)	(0.46, 0.62, 0.78)	
	E <sub>413</sub>	(4.9, 6.5, 8.0)	(0.42, 0.59, 0.76)	
	E <sub>414</sub>	(4.2, 5.9, 7.6)	(0.66, 0.77, 0.88)	
	E <sub>415</sub>	(3.9, 5.3, 6.6)	(0.58, 0.71, 0.84)	
	E <sub>416</sub>	(1.4, 2.6, 3.9)	(0.54, 0.68, 0.82)	
E <sub>43</sub>	E <sub>421</sub>	(6.4, 7.7, 8.8)	(0.73, 0.83, 0.92)	(6.21, 7.52, 8.67)
	E <sub>422</sub>	(6.4, 7.7, 8.8)	(0.50, 0.65, 0.80)	
	E <sub>423</sub>	(5.7, 7.1, 8.4)	(0.46, 0.62, 0.78)	
E <sub>44</sub>	E <sub>431</sub>	(6.8, 8.0, 9.0)	(0.62, 0.74, 0.86)	(5.84, 7.14, 8.27)
	E <sub>432</sub>	(7.2, 8.3, 9.2)	(0.16, 0.29, 0.43)	
	E <sub>433</sub>	(6.4, 7.7, 8.8)	(0.44, 0.59, 0.74)	
	E <sub>434</sub>	(4.0, 5.3, 6.6)	(0.57, 0.71, 0.84)	
E <sub>45</sub>	E <sub>441</sub>	(6.9, 8.0, 9.0)	(0.60, 0.74, 0.86)	(5.14, 6.53, 7.82)
	E <sub>442</sub>	(4.9, 6.2, 7.4)	(0.58, 0.71, 0.84)	
	E <sub>443</sub>	(3.4, 5.0, 6.6)	(0.59, 0.71, 0.82)	
	E <sub>444</sub>	(5.3, 6.8, 8.2)	(0.72, 0.83, 0.92)	
	E <sub>451</sub>	(2.4, 3.8, 5.3)	(0.72, 0.83, 0.92)	
E <sub>46</sub>	E <sub>452</sub>	(2.2, 3.5, 5.0)	(0.79, 0.89, 0.96)	(3.34, 4.76, 6.23)
	E <sub>453</sub>	(3.6, 5.0, 6.4)	(0.60, 0.74, 0.86)	
	E <sub>454</sub>	(2.8, 4.4, 6.0)	(0.68, 0.80, 0.90)	
	E <sub>455</sub>	(6.1, 7.4, 8.6)	(0.65, 0.77, 0.88)	
E <sub>47</sub>	E <sub>461</sub>	(3.8, 5.6, 7.4)	(0.38, 0.53, 0.68)	(5.28, 6.76, 8.16)
	E <sub>462</sub>	(6.4, 7.7, 8.8)	(0.50, 0.65, 0.80)	
	E <sub>471</sub>	(5.0, 6.5, 8.0)	(0.66, 0.77, 0.88)	
	E <sub>472</sub>	(7.5, 8.6, 9.4)	(0.36, 0.50, 0.64)	
E <sub>51</sub>	E <sub>473</sub>	(5.3, 6.8, 8.2)	(0.62, 0.74, 0.86)	(5.01, 6.42, 7.74)
	E <sub>474</sub>	(3.5, 4.7, 5.9)	(0.71, 0.83, 0.92)	
	E <sub>511</sub>	(3.4, 5.3, 7.2)	(0.34, 0.50, 0.66)	
E <sub>52</sub>	E <sub>512</sub>	(5.2, 6.5, 7.6)	(0.54, 0.68, 0.82)	(6.16, 7.34, 8.37)
	E <sub>513</sub>	(8.5, 9.5, 10.0)	(0.62, 0.74, 0.86)	
	E <sub>521</sub>	(4.9, 6.2, 7.4)	(0.69, 0.80, 0.90)	
	E <sub>522</sub>	(2.0, 3.2, 4.5)	(0.71, 0.83, 0.92)	
E <sub>53</sub>	E <sub>523</sub>	(7.5, 8.6, 9.4)	(0.49, 0.62, 0.74)	(4.48, 5.75, 6.93)
	E <sub>531</sub>	(4.9, 6.5, 8.0)	(0.40, 0.56, 0.72)	
	E <sub>532</sub>	(4.2, 5.9, 7.6)	(0.52, 0.65, 0.78)	
E <sub>54</sub>	E <sub>533</sub>	(3.8, 5.6, 7.4)	(0.3, 0.5, 0.7)	(4.33, 6.01, 7.67)
	E <sub>541</sub>	(5.7, 7.1, 8.4)	(0.36, 0.53, 0.70)	
	E <sub>542</sub>	(4.2, 5.9, 7.6)	(0.62, 0.74, 0.86)	
E <sub>54</sub>	E <sub>543</sub>	(3.0, 4.7, 6.4)	(0.26, 0.38, 0.52)	(4.38, 6.01, 7.57)

In order to calculate the rating of each enabler, we firstly aggregate the five experts' weights and ratings, by using median operation [25], and then we carry out the same calculation as that of the criteria rating (Table 6). An example of the rating of the enabler  $E_1$  is shown below.

Example: Rating of the enabler

$$E_1 = \frac{\sum_{j=1}^{j=3} (\text{Criteria rating} \otimes \text{Fuzzy importance weight of the agile criteria})}{\sum_{j=1}^{j=3} \text{Fuzzy importance weight of the agile criteria}}$$

$$= [(4.17, 5.65, 7.11) \otimes (0.5, 0.65, 0.8) \oplus (3.52, 4.88, 6.28) \otimes (0.5, 0.65, 0.8) \oplus (4.49, 5.99, 7.47) \otimes (0.5, 0.65, 0.8)] / [(0.5, 0.65, 0.8) \oplus (0.5, 0.65, 0.8) \oplus (0.5, 0.65, 0.8)] = (5.14, 6.55, 7.86)$$

Table 6: Fuzzy index of agile enabler rating

Agile enablers	Agile criteria	Criteria rating	Fuzzy importance weight of the agile criteria	Enabler rating	Fuzzy importance weight of the agile enablers
E <sub>1</sub>	E <sub>11</sub>	(4.17, 5.65, 7.11)	(0.5, 0.65, 0.8)	(5.14, 6.55, 7.86)	(0.5, 0.65, 0.8)
	E <sub>12</sub>	(3.52, 4.88, 6.28)	(0.5, 0.65, 0.8)		
	E <sub>13</sub>	(4.49, 5.99, 7.47)	(0.5, 0.65, 0.8)		
E <sub>2</sub>	E <sub>21</sub>	(3.66, 4.93, 6.23)	(0.5, 0.65, 0.8)	(3.67, 5.04, 6.44)	(0.5, 0.65, 0.8)
	E <sub>22</sub>	(4.45, 5.77, 7.04)	(0.5, 0.65, 0.8)		
	E <sub>23</sub>	(3.56, 5.05, 6.55)	(0.7, 0.8, 0.9)		
	E <sub>24</sub>	(3.4, 5.0, 6.6)	(0.5, 0.65, 0.8)		
	E <sub>25</sub>	(3.68, 4.89, 6.18)	(0.5, 0.65, 0.8)		
	E <sub>26</sub>	(3.29, 4.63, 6.03)	(0.5, 0.65, 0.8)		
E <sub>3</sub>	E <sub>31</sub>	(3.08, 4.57, 6.06)	(0.7, 0.8, 0.9)	(4.36, 5.79, 7.16)	(0.7, 0.8, 0.9)
	E <sub>32</sub>	(5.09, 6.55, 7.92)	(0.7, 0.8, 0.9)		
	E <sub>33</sub>	(4.69, 5.97, 7.21)	(0.7, 0.8, 0.9)		
	E <sub>34</sub>	(4.21, 5.79, 7.32)	(0.5, 0.65, 0.8)		
	E <sub>35</sub>	(4.20, 5.86, 7.45)	(0.5, 0.65, 0.8)		
	E <sub>36</sub>	(5.01, 6.04, 7.03)	(0.5, 0.65, 0.8)		
E <sub>4</sub>	E <sub>41</sub>	(3.81, 5.30, 6.74)	(0.5, 0.65, 0.8)	(4.88, 6.31, 7.64)	(0.5, 0.65, 0.8)
	E <sub>42</sub>	(6.21, 7.52, 8.67)	(0.5, 0.65, 0.8)		
	E <sub>43</sub>	(5.84, 7.14, 8.27)	(0.5, 0.65, 0.8)		
	E <sub>44</sub>	(5.14, 6.53, 7.82)	(0.7, 0.8, 0.9)		
	E <sub>45</sub>	(3.34, 4.76, 6.23)	(0.7, 0.8, 0.9)		
	E <sub>46</sub>	(5.28, 6.76, 8.16)	(0.5, 0.65, 0.8)		
	E <sub>47</sub>	(5.01, 6.42, 7.74)	(0.7, 0.8, 0.9)		
E <sub>5</sub>	E <sub>51</sub>	(6.16, 7.34, 8.37)	(0.5, 0.65, 0.8)	(4.90, 6.28, 7.61)	(0.5, 0.65, 0.8)
	E <sub>52</sub>	(4.48, 5.75, 6.93)	(0.7, 0.8, 0.9)		
	E <sub>53</sub>	(4.33, 6.01, 7.67)	(0.3, 0.5, 0.7)		
	E <sub>54</sub>	(4.38, 6.01, 7.57)	(0.3, 0.5, 0.7)		

- Step 4: Calculation of the FAI of HealthOrg: We carry out the same calculation as that of the enabler rating [32].

$$FAI = \frac{\sum_{i=1}^{i=5} (\text{Enabler rating} \otimes \text{Fuzzy importance weight of the agile enablers})}{\sum_{i=1}^{i=5} \text{Fuzzy importance weight of the agile enablers}}$$

$$= [(5.14, 6.55, 7.86) \otimes (0.5, 0.65, 0.8) \oplus (3.67, 5.04, 6.44) \otimes (0.5, 0.65, 0.8) \oplus (4.36, 5.79, 7.16) \otimes (0.7, 0.8, 0.9) \oplus (4.88, 6.31, 7.64) \otimes (0.5, 0.65, 0.8) \oplus (4.90, 6.28, 7.61) \otimes (0.5, 0.65, 0.8)] / [(0.5, 0.65, 0.8) \oplus (0.5, 0.65, 0.8) \oplus (0.7, 0.8, 0.9) \oplus (0.5, 0.65, 0.8) \oplus (0.5, 0.65, 0.8)] = (4.57, 5.98, 7.34)$$

The overall agility of HealthOrg is (4.57, 5.98, 7.34).

- Step 5: Matching the FAI with the appropriate linguistic level [32]: After determining the FAI of the organization, we converted it into linguistic terms. To do this, we used the Euclidean distance method in which we seek to obtain the minimum distance between FAI and the linguistic level (Table 8). Table 7 presents the linguistic terms of different agility levels and their fuzzy intervals [5].

Table 7: Fuzzy values of agility levels (Adapted from [25])

Level of agility	Fuzzy intervals
Slowly Agile	(0, 1.5, 3)
Fairly Agile	(1.5, 3, 4.5)
Agile	(3.5, 5, 6.5)
Very Agile	(5.5, 7, 8.5)
Extremely Agile	(7, 8.5, 10)

Table 8: Agility level of HealthOrg

FAI for HealthOrg	(4.57, 5.98, 7.34)
D (FAI, Slowly Agile)	$\{(4.57-0)^2 + (5.98-1.5)^2 + (7.34-3)^2\}^{1/2} = 7.78$
D (FAI, Fairly Agile)	$\{(4.57-1.5)^2 + (5.98-3)^2 + (7.34-4.5)^2\}^{1/2} = 5.13$
D (FAI, Agile)	$\{(4.57-3.5)^2 + (5.98-5.0)^2 + (7.34-6.5)^2\}^{1/2} = 1.67$
D (FAI, Very Agile)	$\{(4.57-5.5)^2 + (5.98-7)^2 + (7.34-8.5)^2\}^{1/2} = 1.80$
D (FAI, Extremely Agile)	$\{(4.57-7)^2 + (5.98-8.5)^2 + (7.34-10)^2\}^{1/2} = 4.40$

The minimum distance between FAI and the level of agility is that obtained with the “Agile” level. Then, HealthOrg is considered as an agile enterprise.

- Step 6: Fuzzy performance importance index (FPII) calculation: Although HealthOrg is agile; some attributes weakened its agility during COVID-19 era. In order to identify them, we calculate FPII and the ranking score for each agile attribute (Table 9) [5]. An example of it for E<sub>111</sub> is calculated as:

FPII<sub>111</sub> = [(1, 1, 1) – Average fuzzy weight of E<sub>111</sub>] ⊗ Average fuzzy performance rating of

$$E_{111} = [(1, 1, 1) - (0.54, 0.68, 0.82)] \otimes (6.1, 7.4, 8.6) = (2.81, 2.37, 1.55)$$

$$\text{Ranking score of } E_{111} = (2.81 + 4 \times 2.37 + 1.55) / 6 = 2.31$$

Table 9: FPII and ranking score of agile attributes

Agile Attributes	Average fuzzy weight	Fuzzy performance average rating	FPII	Ranking score
E <sub>111</sub>	(0.54, 0.68, 0.82)	(6.1, 7.4, 8.6)	(2.81, 2.37, 1.55)	2.31
E <sub>112</sub>	(0.54, 0.68, 0.82)	(3.4, 5.0, 6.6)	(1.56, 1.60, 1.19)	1.52
E <sub>113</sub>	(0.58, 0.71, 0.84)	(2.4, 3.8, 5.3)	(1.01, 1.10, 0.85)	<b>1.04</b>
E <sub>114</sub>	(0.54, 0.68, 0.82)	(4.9, 6.5, 8.0)	(2.25, 2.08, 1.44)	2.00
E <sub>121</sub>	(0.46, 0.62, 0.78)	(5.0, 6.5, 8.0)	(2.70, 2.47, 1.76)	2.39
E <sub>122</sub>	(0.54, 0.68, 0.82)	(3.0, 4.7, 6.4)	(1.38, 1.50, 1.15)	1.42
E <sub>123</sub>	(0.50, 0.65, 0.80)	(1.4, 2.6, 4.0)	(0.70, 0.91, 0.80)	<b>0.86</b>
E <sub>124</sub>	(0.58, 0.71, 0.84)	(1.4, 2.6, 4.0)	(0.59, 0.75, 0.64)	<b>0.70</b>
E <sub>125</sub>	(0.58, 0.71, 0.84)	(6.8, 8.0, 9.0)	(2.86, 2.32, 1.44)	2.26
E <sub>131</sub>	(0.76, 0.86, 0.94)	(5.0, 6.5, 8.0)	(1.20, 0.91, 0.48)	<b>0.89</b>
E <sub>132</sub>	(0.58, 0.71, 0.84)	(6.7, 8.0, 9.0)	(2.81, 2.32, 1.44)	2.25
E <sub>133</sub>	(0.62, 0.74, 0.86)	(5.4, 6.8, 8.2)	(2.05, 1.77, 1.15)	1.71
E <sub>134</sub>	(0.42, 0.59, 0.76)	(3.8, 5.6, 7.4)	(2.20, 2.30, 1.78)	2.20
E <sub>135</sub>	(0.46, 0.62, 0.78)	(5.0, 6.5, 8.0)	(2.70, 2.47, 1.76)	2.39
E <sub>136</sub>	(0.79, 0.89, 0.96)	(2.6, 4.1, 5.6)	(0.55, 0.45, 0.22)	<b>0.43</b>
E <sub>137</sub>	(0.54, 0.68, 0.82)	(3.2, 4.7, 6.3)	(1.47, 1.50, 1.13)	1.43
E <sub>211</sub>	(0.66, 0.77, 0.88)	(4.6, 6.2, 7.8)	(1.56, 1.43, 0.94)	1.37
E <sub>212</sub>	(0.42, 0.59, 0.76)	(1.8, 2.9, 4.2)	(1.04, 1.19, 1.01)	1.13
E <sub>213</sub>	(0.50, 0.65, 0.80)	(2.0, 3.5, 5.0)	(1.00, 1.22, 1.00)	1.15
E <sub>214</sub>	(0.76, 0.86, 0.94)	(6.4, 7.7, 8.8)	(1.54, 1.08, 0.53)	<b>1.06</b>
E <sub>215</sub>	(0.58, 0.71, 0.84)	(1.8, 3.2, 4.7)	(0.76, 0.93, 0.75)	<b>0.87</b>
E <sub>221</sub>	(0.24, 0.41, 0.58)	(5.3, 6.8, 8.2)	(4.03, 4.01, 3.44)	3.92
E <sub>222</sub>	(0.54, 0.68, 0.82)	(3.9, 5.0, 6.1)	(1.79, 1.60, 1.10)	1.55

E223	(0.50, 0.65, 0.80)	(3.1, 4.4, 5.7)	(1.55, 1.54, 1.14)	1.47
E224	(0.62, 0.74, 0.86)	(5.7, 7.1, 8.4)	(2.17, 1.85, 1.18)	1.79
E231	(0.76, 0.86, 0.94)	(4.2, 5.9, 7.6)	(1.01, 0.83, 0.46)	<b>0.80</b>
E232	(0.54, 0.68, 0.82)	(2.2, 3.5, 4.9)	(1.01, 1.12, 0.88)	<b>1.06</b>
E233	(0.7, 0.8, 0.9)	(2.4, 3.8, 5.3)	(0.72, 0.76, 0.53)	<b>0.71</b>
E234	(0.50, 0.65, 0.80)	(5.7, 7.1, 8.4)	(2.85, 2.48, 1.68)	2.41
E241	(0.50, 0.65, 0.80)	(3.4, 5.0, 6.6)	(1.70, 1.75, 1.32)	1.67
E251	(0.50, 0.65, 0.80)	(5.8, 7.1, 8.4)	(2.90, 2.48, 1.68)	2.42
E252	(0.50, 0.65, 0.80)	(3.8, 5.3, 6.8)	(1.90, 1.85, 1.36)	1.78
E253	(0.42, 0.59, 0.76)	(1.0, 2.0, 3.2)	(0.58, 0.82, 0.77)	<b>0.77</b>
E261	(0.54, 0.68, 0.82)	(5.7, 7.1, 8.4)	(2.62, 2.27, 1.51)	2.20
E262	(0.58, 0.71, 0.84)	(2.6, 3.8, 5.2)	(1.09, 1.10, 0.83)	<b>1.05</b>
E263	(0.50, 0.65, 0.80)	(1.6, 2.9, 4.3)	(0.80, 1.01, 0.86)	0.95
E264	(0.54, 0.68, 0.82)	(3.2, 4.7, 6.2)	(1.47, 1.50, 1.12)	1.43
E311	(0.58, 0.71, 0.84)	(2.4, 3.8, 5.3)	(1.01, 1.10, 0.85)	<b>1.04</b>
E312	(0.7, 0.8, 0.9)	(1.6, 2.9, 4.4)	(0.48, 0.58, 0.44)	<b>0.54</b>
E313	(0.79, 0.89, 0.96)	(2.8, 4.4, 6.0)	(0.59, 0.48, 0.24)	<b>0.46</b>
E314	(0.50, 0.65, 0.80)	(6.4, 7.7, 8.8)	(3.20, 2.69, 1.76)	2.62
E321	(0.7, 0.8, 0.9)	(6.0, 7.4, 8.6)	(1.80, 1.48, 0.86)	1.43
E322	(0.58, 0.71, 0.84)	(4.0, 5.6, 7.2)	(1.68, 1.62, 1.15)	1.55
E331	(0.58, 0.71, 0.84)	(5.7, 7.1, 8.4)	(2.39, 2.06, 1.34)	1.99
E332	(0.7, 0.8, 0.9)	(4.5, 6.2, 7.8)	(1.35, 1.24, 0.78)	1.18
E333	(0.38, 0.56, 0.74)	(2.2, 3.5, 5.0)	(1.36, 1.54, 1.30)	1.47
E334	(0.58, 0.71, 0.84)	(2.2, 3.5, 4.9)	(0.92, 1.01, 0.78)	<b>0.96</b>
E335	(0.62, 0.74, 0.86)	(7.8, 8.9, 9.6)	(2.96, 2.31, 1.34)	2.26
E341	(0.50, 0.65, 0.80)	(2.8, 4.4, 6.0)	(1.40, 1.54, 1.20)	1.46
E342	(0.50, 0.65, 0.80)	(3.6, 5.3, 7.0)	(1.80, 1.85, 1.40)	1.77
E343	(0.54, 0.68, 0.82)	(5.3, 6.8, 8.2)	(2.44, 2.18, 1.48)	2.11
E344	(0.62, 0.74, 0.86)	(4.9, 6.5, 8.0)	(1.86, 1.69, 1.12)	1.62
E351	(0.58, 0.71, 0.84)	(2.6, 4.4, 6.2)	(1.09, 1.28, 0.99)	1.20
E352	(0.42, 0.59, 0.76)	(4.9, 6.5, 8.0)	(2.84, 2.66, 1.92)	2.57
E353	(0.58, 0.71, 0.84)	(5.3, 6.8, 8.2)	(2.23, 1.97, 1.31)	1.90
E361	(0.38, 0.56, 0.74)	(1.0, 2.0, 3.3)	(0.62, 0.88, 0.86)	<b>0.83</b>
E362	(0.54, 0.68, 0.82)	(8.2, 9.2, 9.8)	(3.77, 2.94, 1.76)	2.88
E363	(0.66, 0.77, 0.88)	(4.7, 6.2, 7.6)	(1.60, 1.43, 0.91)	1.37
E411	(0.42, 0.59, 0.76)	(6.4, 7.7, 8.8)	(3.71, 3.16, 2.11)	3.08
E412	(0.46, 0.62, 0.78)	(2.6, 4.1, 5.7)	(1.40, 1.56, 1.25)	1.48
E413	(0.42, 0.59, 0.76)	(4.9, 6.5, 8.0)	(2.84, 2.66, 1.92)	2.57
E414	(0.66, 0.77, 0.88)	(4.2, 5.9, 7.6)	(1.43, 1.36, 0.91)	1.30
E415	(0.58, 0.71, 0.84)	(3.9, 5.3, 6.6)	(1.64, 1.54, 1.06)	1.48
E416	(0.54, 0.68, 0.82)	(1.4, 2.6, 3.9)	(0.64, 0.83, 0.70)	<b>0.78</b>
E421	(0.73, 0.83, 0.92)	(6.4, 7.7, 8.8)	(1.73, 1.31, 0.70)	1.28
E422	(0.50, 0.65, 0.80)	(6.4, 7.7, 8.8)	(3.20, 2.69, 1.76)	2.62
E423	(0.46, 0.62, 0.78)	(5.7, 7.1, 8.4)	(3.08, 2.70, 1.85)	2.62
E431	(0.62, 0.74, 0.86)	(6.8, 8.0, 9.0)	(2.58, 2.08, 1.26)	2.03
E432	(0.16, 0.29, 0.43)	(7.2, 8.3, 9.2)	(6.05, 5.89, 5.24)	5.81
E433	(0.44, 0.59, 0.74)	(6.4, 7.7, 8.8)	(3.58, 3.16, 2.29)	3.08
E434	(0.57, 0.71, 0.84)	(4.0, 5.3, 6.6)	(1.72, 1.54, 1.06)	1.49
E441	(0.60, 0.74, 0.86)	(6.9, 8.0, 9.0)	(2.76, 2.08, 1.26)	2.06
E442	(0.58, 0.71, 0.84)	(4.9, 6.2, 7.4)	(2.06, 1.80, 1.18)	1.74
E443	(0.59, 0.71, 0.82)	(3.4, 5.0, 6.6)	(1.39, 1.45, 1.19)	1.40
E444	(0.72, 0.83, 0.92)	(5.3, 6.8, 8.2)	(1.48, 1.16, 0.66)	1.13
E451	(0.72, 0.83, 0.92)	(2.4, 3.8, 5.3)	(0.67, 0.65, 0.42)	<b>0.61</b>
E452	(0.79, 0.89, 0.96)	(2.2, 3.5, 5.0)	(0.46, 0.38, 0.20)	<b>0.36</b>
E453	(0.60, 0.74, 0.86)	(3.6, 5.0, 6.4)	(1.44, 1.30, 0.90)	1.26
E454	(0.68, 0.80, 0.90)	(2.8, 4.4, 6.0)	(0.90, 0.88, 0.60)	<b>0.84</b>
E455	(0.65, 0.77, 0.88)	(6.1, 7.4, 8.6)	(2.13, 1.70, 1.03)	1.66
E461	(0.38, 0.53, 0.68)	(3.8, 5.6, 7.4)	(2.36, 2.63, 2.37)	2.54
E462	(0.50, 0.65, 0.80)	(6.4, 7.7, 8.8)	(3.20, 2.69, 1.76)	2.62
E471	(0.66, 0.77, 0.88)	(5.0, 6.5, 8.0)	(1.70, 1.49, 0.96)	1.44
E472	(0.36, 0.50, 0.64)	(7.5, 8.6, 9.4)	(4.80, 4.30, 3.38)	4.23
E473	(0.62, 0.74, 0.86)	(5.3, 6.8, 8.2)	(2.01, 1.77, 1.15)	1.71
E474	(0.71, 0.83, 0.92)	(3.5, 4.7, 5.9)	(1.01, 0.80, 0.47)	<b>0.78</b>
E511	(0.34, 0.50, 0.66)	(3.4, 5.3, 7.2)	(2.24, 2.65, 2.45)	2.55
E512	(0.54, 0.68, 0.82)	(5.2, 6.5, 7.6)	(2.39, 2.08, 1.37)	2.01
E513	(0.62, 0.74, 0.86)	(8.5, 9.5, 10.0)	(3.23, 2.47, 1.40)	2.42
E521	(0.69, 0.80, 0.90)	(4.9, 6.2, 7.4)	(1.52, 1.24, 0.74)	1.20
E522	(0.71, 0.83, 0.92)	(2.0, 3.2, 4.5)	(0.58, 0.54, 0.36)	<b>0.52</b>
E523	(0.49, 0.62, 0.74)	(7.5, 8.6, 9.4)	(3.82, 3.27, 2.44)	3.22
E531	(0.40, 0.56, 0.72)	(4.9, 6.5, 8.0)	(2.94, 2.86, 2.24)	2.77
E532	(0.52, 0.65, 0.78)	(4.2, 5.9, 7.6)	(2.02, 2.06, 1.67)	1.99

E533	(0.3, 0.5, 0.7)	(3.8, 5.6, 7.4)	(2.66, 2.80, 2.22)	2.68
E541	(0.36, 0.53, 0.70)	(5.7, 7.1, 8.4)	(3.65, 3.34, 2.52)	3.25
E542	(0.62, 0.74, 0.86)	(4.2, 5.9, 7.6)	(1.60, 1.53, 1.06)	1.46
E543	(0.26, 0.38, 0.52)	(3.0, 4.7, 6.4)	(2.22, 2.91, 3.07)	2.82

Based on the five experts' experience, scale 1.1 was considered as the threshold which distinguishes the weaker attributes than the other ones. Table 10 showed these attributes and some suggestions to improve them [5].

Table 10: Weaker agile attributes and improvement proposals

Weak agile attribute	References	Improvement proposals
<ul style="list-style-type: none"> <li>Staff interchangeability</li> <li>Multi-skilled and flexible staff</li> <li>Implementation of job rotation system</li> </ul>	[15, 33]	Prepare employees to participate in the implementation of job rotation system
<ul style="list-style-type: none"> <li>Flexible employees to accept the adoption of new technologies</li> <li>Multi-functional, developed and trained employees</li> </ul>	[15, 29]	Develop a flexible working environment for employees
<ul style="list-style-type: none"> <li>Decentralized decision-making, knowledge and control</li> <li>Knowledge and skills management systems</li> <li>Staff empowerment to resolve patient issues</li> </ul>	[15, 29, 30, 31, 33]	Give authority to different level employees which contributes to improved their knowledge
<ul style="list-style-type: none"> <li>Loyalty and commitment to a project or a group</li> <li>Participative management style</li> <li>Quick evaluation and implementation of employee suggestions</li> <li>Involvement of suppliers and different agents in product/service development</li> </ul>	[15, 29, 30]	Remove barriers to facilitate the participation of different employees and suppliers
<ul style="list-style-type: none"> <li>Efficient information system and technology</li> <li>Exploitation of information technology (IT) in supply chain management</li> </ul>	[15, 27, 29, 30, 32, 33]	Link information systems to technology
<ul style="list-style-type: none"> <li>Adopting supply chain management concepts to improve the efficiency of outsourcing</li> <li>Simple process to implement</li> <li>Intelligent Engineering Design Support System</li> <li>Active policy to keep work areas clean and tidy</li> <li>Execution of short range planning</li> <li>Company's procurement policy based on time schedule</li> <li>Improved manufacturing technology</li> <li>First-time correct design</li> <li>Reduction of non value-adding costs</li> </ul>		Use advanced technologies and production methods



## 5. Conclusion

This study evaluated organization agility of a public hospital in Morocco in times of COVID-19. The enablers influencing agility were studied, as were the agile criteria and attributes. After a literature review, an assessment model was presented and tested via the fuzzy logic approach. Empirical results showed that HealthOrg is agile. The COVID-19 outbreak has revealed how different enablers can influence the hospital agility. It has also shown how some agile enablers need to be enhanced in order to increase the healthcare organization agility.

This article offers initial empirical exploration on how Moroccan healthcare organizations cope with the COVID-19 crisis. It allows identifying the required changes to improve the agility of the organization. There will be increasing improvement for hospitals in technology and human resources departments; COVID-19 has demonstrated their importance in making the healthcare organization extremely agile.

Despite the above benefits for using the assessment model, there is some limitation: this model does not take into account the different agile drivers and capacities which must be aligned with the agile enablers. Also, the organizational agility assessment has been done for a single healthcare organization; however future research should replicate the assessment model in other organizations, in public and private sector. Also, it is highly recommended to compare the results obtained in times of COVID-19 with those provided by previous studies. Moreover, further practical suggestions for healthcare sector through COVID 19 outbreaks should be provided.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

The authors acknowledge the financial support of the National Centre for Scientific and Technical Research (CNRST) under the Excellence Research Scholarships Program.

## References

- [1] K. C. Ng Kee Kwong, P. R. Mehta, G. Shukla, A. R. Mehta, "COVID-19, SARS and MERS: a neurological perspective" *Journal of Clinical Neuroscience*, 2020. <https://doi.org/10.1016/j.jocn.2020.04.124>
- [2] I. Chakraborty, P. Maity, "COVID-19 outbreak: Migration, effects on society, global environment and prevention" *Science of the Total Environment*, **728**, 2020. <https://doi.org/10.1016/j.scitotenv.2020.138882>
- [3] K. Barro, A. Malone, A. Mokede, C. Chevance, "Management of the COVID-19 epidemic by public health establishments—analysis by the fédération hospitalière de France" *Journal of Visceral Surgery*, 2020. <https://doi.org/10.1016/j.jviscsurg.2020.04.011>
- [4] Y. Y. Yusuf, M. Sarhadi, A. Gunasekaran, "Agile manufacturing: the drivers, concepts and attributes" *International Journal of Production Economics*, **62**(1–2), 33–43, 1999. [https://doi.org/10.1016/S0925-5273\(98\)00219-9](https://doi.org/10.1016/S0925-5273(98)00219-9)
- [5] M. Suresh, R. Patri, "Agility assessment using fuzzy logic approach: a case of healthcare dispensary" *BMC Health Services Research*, **17**(1), 2017. <https://doi.org/10.1186/s12913-017-2332-y>
- [6] M. Christopher, "The agile supply chain: competing in volatile markets" *Industrial Marketing Management*, **29**(1), 37–44, 2000. [https://doi.org/10.1016/S0019-8501\(99\)00110-8](https://doi.org/10.1016/S0019-8501(99)00110-8)
- [7] R. Kitzmiller, E. Hunt, S. B. Sproat, "Adopting best practices: 'agility' moves from software development to healthcare project management" *Computers Informatics Nursing*, **24**(2), 75–82, 2006. <https://doi.org/10.1097/00024665-200603000-00005>
- [8] S. Brown, J. Bessant, "The manufacturing strategy-capabilities links in mass customisation and agile manufacturing – an exploratory study" *International*

- Journal of Operations & Production Management*, **23**(7), 707–730, 2003. <https://doi.org/10.1108/01443570310481522>
- [9] H. Sharifi, Z. Zhang, "Agile manufacturing in practice - application of a methodology" *International Journal of Operations & Production Management*, **21**(5/6), 772–794, 2001. <https://doi.org/10.1108/01443570110390462>
- [10] G. Fliedner, R. J. Vokurka, "Agility: competitive weapon of the 1990s and beyond?" *Production and inventory management journal*, **38**(3), 19–24, 1997.
- [11] E. Bottani, "A fuzzy QFD approach to achieve agility" *International Journal of Production Economics*, **119**(2), 380–391, 2009. <https://doi.org/10.1016/j.ijpe.2009.02.013>
- [12] J. Prince, J. M. Kay, "Combining lean and agile characteristics: creation of virtual groups by enhanced production flow analysis" *International Journal of Production Economics*, **85**(3), 305–318, 2003. [https://doi.org/10.1016/S0925-5273\(03\)00118-X](https://doi.org/10.1016/S0925-5273(03)00118-X)
- [13] P. g. Saleeshya, A. S. Babu, A. s. Vishnu, "A model to assess the agility of manufacturing organisations: systems approach and application" *International Journal of Productivity and Quality Management*, **8**(3), 265–295, 2011. <https://doi.org/10.1504/IJPM.2011.042509>
- [14] H. Dong, C. L. Li, "Agile supply chain performance evaluation of automobile manufacturing based on regression analysis" *Applied Mechanics and Materials*, 2662-2666, 2013. <https://doi.org/10.4028/www.scientific.net/AMM.397-400.2662>
- [15] S. Vinodh, S. R. Devadasan, B. Vasudeva Reddy, K. Ravichand, "Agility index measurement using multi-grade fuzzy approach integrated in a 20 criteria agile model" *International Journal of Production Research*, **48**(23), 7159–7176, 2010. <https://doi.org/10.1080/00207540903354419>
- [16] P. G. Saleeshya, A. S. Babu, "A combined AHP- and DEA-based approach to measure agility of manufacturing systems" *International Journal of Business and Systems Research*, **6**(4), 431–455, 2012. <https://doi.org/10.1504/IJBSR.2012.049472>
- [17] S. G. Azevedo, H. Carvalho, V. C. Machado, "Agile index: automotive supply chain" *International Journal of Industrial and Manufacturing Engineering*, **5**(7), 1433–1439, 2011. <https://doi.org/10.5281/zenodo.1078048>
- [18] S. A. Raj, A. Sudheer, S. Vinodh, G. Anand, "A mathematical model to evaluate the role of agility enablers and criteria in a manufacturing environment" *International Journal of Production Research*, **51**(19), 5971–5984, 2013. <https://doi.org/10.1080/00207543.2013.825381>
- [19] K. Khalili-Damghani, M. T. Taghavifard, "A three-stage fuzzy DEA approach to measure performance of a serial process including JIT practices, agility indices, and goals in supply chains" *International Journal of Services and Operations Management*, **13**(2), 147–188, 2012. <https://doi.org/10.1504/ijom.2012.048828>
- [20] R. Zitkiene, M. Deksnys, "Organizational agility conceptual model" *Montenegrin Journal of Economics*, **14**(2), 115–129, 2018. <https://doi.org/10.14254/1800-5845/2018.14-2.7>
- [21] A. Gunasekaran, "Agile manufacturing: enablers and an implementation framework" *International Journal of Production Research*, **36**(5), 1223–1247, 1998. <https://doi.org/10.1080/002075498193291>
- [22] M. Nejatian, M. H. Zarei, "Moving towards organizational agility: are we improving in the right direction?" *Global Journal of Flexible Systems Management*, **14**(4), 241–253, 2013. <https://doi.org/10.1007/s40171-013-0048-3>
- [23] Z. Zhang, H. Sharifi, "A methodology for achieving agility in manufacturing organisations" *International Journal of Operations & Production Management*, **20**(4), 496–513, 2000. <https://doi.org/10.1108/01443570010314818>
- [24] S. Tolf, M. E. Nyström, C. Tishelman, M. Brommels, J. Hansson, "Agile, a guiding principle for health care improvement?" *International Journal of Health Care Quality Assurance*, **28**(5), 468–493, 2015. <https://doi.org/10.1108/IJHCQA-04-2014-0044>
- [25] C.-T. Lin, H. Chiu, Y.-H. Tseng, "Agility evaluation using fuzzy logic" *International Journal of Production Economics*, **101**(2), 353–368, 2006. <https://doi.org/10.1016/j.ijpe.2005.01.011>
- [26] W. Triaa, L. Gzara, H. Verjus, "Organizational agility key factors for dynamic business process management" in 2016 IEEE 18th Conference on Business Informatics (CBI), Paris, France, 2016. <https://doi.org/10.1109/CBI.2016.16>
- [27] A. T. Eshlaghy, A. N. Mashayekhi, A. Rajabzadeh, M. M. Razavian, "Applying path analysis method in defining effective factors in organisation agility" *International Journal of Production Research*, **48**(6), 1765–1786, 2010. <https://doi.org/10.1080/00207540802566410>
- [28] C.-S. Tsai, C.-W. Chen, C.-T. Li, "Align agile drivers, capabilities and providers to achieve agility: a fuzzy-logic QFD approach" *Supply Chain*, 2008.
- [29] D. Vázquez-Bustelo, L. Avella, E. Fernández, "Agility drivers, enablers and outcomes: empirical test of an integrated agile manufacturing model" *International Journal of Production Research*, **51**(19), 5971–5984, 2013. <https://doi.org/10.1080/00207543.2013.825381>



International journal of operations & production management, **27**(12), 1303–1332, 2007. <https://doi.org/10.1108/01443570710835633>

- [30] B. Sherehiy, W. Karwowski, J. K. Layer, “A review of enterprise agility: concepts, frameworks, and attributes,” *International Journal of Industrial Ergonomics*, **37**(5), 445–460, 2007. <https://doi.org/10.1016/j.ergon.2007.01.007>
- [31] R. Patri, M. Suresh, “Modelling the Enablers of Agile Performance in Healthcare Organization: A TISM Approach” *Global Journal of Flexible Systems Management*, **18**(3), 251–272, 2017. <https://doi.org/10.1007/s40171-017-0160-x>
- [32] M. Deksnys, “Organizational agility in high growth companies” Ph.D Thesis, Mykolas Romeris University, 2018.
- [33] Yuh-Chuan Shih, Ching-Tomg Lin, “Agility index of manufacturing firm-a fuzzy-logic-based approach” in 2002 IEEE International Engineering Management Conference, Cambridge, UK, 2002. <https://doi.org/10.1109/IEMC.2002.1038476>
- [34] S. Vinodh, S. R. Devadasan, “Twenty criteria based agility assessment using fuzzy logic approach” *International Journal of Advanced Manufacturing Technology*, **54**(9), 1219–1231, 2011. <https://doi.org/10.1007/s00170-010-3015-6>
- [35] V. Vaishnavi, M. Suresh, “Assessing the readiness level of healthcare for implementing agility using fuzzy logic approach,” *Global Journal of Flexible Systems Management*, **21**(2), 163–189, 2020. <https://doi.org/10.1007/s40171-020-00237-7>

## **A Support Vector Machine Based Technique for Fault Detection in A Power Distribution Integrated System with Renewable Energy Distributed Generation**

Katleho Moloi\*, Yskandar Hamam, Jacobus Andries Jordaan

*Faculty of Built Environment and Engineering, Dept. of Electrical Engineering, Tshwane University of Technology, Pretoria, 1035, South Africa*

### **ARTICLE INFO**

*Article history:*

*Received: 18 May, 2020*

*Accepted: 17 July, 2020*

*Online: 19 August, 2020*

*Keywords:*

*Fault Detection*

*Feature Extraction*

*Packet Wavelet Transform*

*Power System Integration*

*Renewable Energy Sources*

*Support Vector Machine*

### **ABSTRACT**

*The integration of renewable energy distributed generation (REDG) into the energized distribution power grid has become more popular in recent years. This has been escalated by the general global energy shortages. The REDG has proven to be effective for energy sustainability and reliability. However, there are technical challenges which arise from integrating REDG into the energized power grid. These challenges include the effectiveness of power grid protection against faults. In this paper, a fault diagnostic algorithm is proposed to detect faults in a power system integrated with REDGs. The algorithm utilizes wavelet packet transform (WPT) for signal filtering, support vector machine (SVM) for fault classification and detection. The proposed algorithm is validated using the Eskom 90 bus electrical system and the results obtained show that faults can be detected with a high accuracy of 99%. The Eskom 90 bus system is modelled using DigSilent platform and the algorithm is tested on the WEKA software.*

## **1. Introduction**

The reliable and sustainable source of energy plays a critical role for the potential growth in a state [1]. The global energy sector has been faced with many challenges over the years, such as energy shortages, high levels of air pollution from burning fossil fuel to generate electricity, and the high cost of coal. These problems have led to finding an alternative source of electricity supply to meet the required demands. Renewable energy distribution generators (REDG) for instance, photovoltaic (PV), wind turbine (WT), hydropower, biomass, etc..., have been effective for energy supply with minimum environmental impact. The flexible introduction of REDGs enables their application to be necessary for formulating a power mix framework for energy sustainability.

The integration of REDGs into the existing distribution power grid has numerous practical aids such as voltage improvement, dependability increase, network performance increase, and power loss minimization [2, 3]. However, integrating REDGs into the power distribution grid changes the traditional trajectory of the energy supply. These changes in the power flow may influence the power balance of the entire power grid [4]. Furthermore, the meteorological and geological dependency of REDGs affects the

expansion planning of the power grid, which results in escalated operational costs [5]. Coupled with these problems are there technical effects that arise from integrating REDGs into the power grid such as frequency variation [6], voltage fluctuations [7], reliable and secure power flow [8, 9]. Generally, the performance of the power grid with or without REDGs highly depends on the reliability of the protection system. The high penetration level of REDGs into the grid affects the traditional topology of the protection system. This may result in a catastrophic incident if a fault is not cleared timeously and with great effect. Integrating REDGs into the existing power grid has operational challenges that may affect the technical performance of the system. These challenges include voltage variations, power supply forecasting, frequency fluctuation and load demand management.

A power system grid is prone to faults, whether internally or externally. Electrical protection schemes are essential to timeously eradicate the presence of a fault in the power system [10, 11]. In recent years, numerous methods for power system fault diagnostic have been proposed. In [12], a method base on voltage imbalances and third harmonic distortion (THD) was used to determine the faulty section for an integrated power system network. The passive methods including frequency variations [13, 14], degree of alteration of power frequency [15], power and voltage variation

\*Corresponding Author: Katleho Moloi, moloikt023@gmail.com

[16], energy mismatch and harmonic content recognition [17] and the degree of variation of reactive power [18], these techniques were applied in power grid integrated network for fault diagnostic. Moreover, undetected faults have an undesirable effect on the overall performance of the power grid, the active algorithms used for fault diagnostic includes frequency signal injection [19], singular and dual harmonic current injection [20, 21], output power variation [22], reactive power control [23, 24], and the impedance measurement method [25, 26], these methods were used to detect faults in power grid integrated network with REDGs.

To improve the fault detection algorithms, artificial intelligence (AI) and machine learning (ML) have been adopted to solve the problem. The AI and ML algorithms have the advantage of discriminating and classifying different types of faults. Fault classification is an important element for reliability improvement and network security. The authors in [27], proposed a fault classification technique based on the decision tree algorithm. The technique was implemented in a PV plant system connected with the distribution network. An artificial neural network (ANN) based method was implemented for fault diagnostic in a system interconnected with WT energy system [28]. Another method based on ANN for islanding and fault detection in microgrids integrated with the power system was proposed [29]. Other AI and ML algorithms used for fault diagnostic in the power system includes fuzzy logic [30, 31], adaptive neuro-fuzzy interface system (ANFIS) [32], and Bayesian classifier [33]. Signal processing plays a major role in improving the fault detection algorithms, signal decomposition techniques which includes wavelet transforms (WT) [34], Hilbert-Huang transform (HHT), multi-resolution analysis (MRA) and Intrinsic Mode Function (IMF) [35] were used for signal tracking to improve the fault detection algorithm. Statistical data analysis is a critical aspect of the decision-making process. The support vector machine (SVM) has been successfully used in power engineering applications for fault diagnostics. The authors in [36], applied SVM for fault detection in high voltage systems. The quadratic function was implemented as an optimal kernel function for effective fault diagnostics. A technique based on SVM for fault diagnostic in active distribution integrated network with the PV plant was proposed [37]. The technique employed SVM for fault classification and isolation. A technique based on a modified multi-class support vector machine (MMC-SVM) was proposed to detect and classify different faults on a power system [38]. A maximal overlap discrete wavelet transform (MODWT) based technique was proposed for fault detection in a power system [39]. In [40], a method based on the multi-agent system (MAS) technique was proposed for fault diagnostic and location in a power system integrated network with DGs. Probabilistic Boolean network (PBN) technique was proposed for fault diagnostic in smart power system grid for performance and technical improvement [41]. A novel technique based on an unknown input observer (UIO) method was proposed for fault detection and isolation on microgrids [42].

### *1.1. Manuscript Organization*

In this work, a hybrid fault diagnostic method based on packet wavelet transform (PWT) and support vector machine (SVM) is proposed. The PWT is used for signal decomposition and feature extraction, this is done to reduce computation burden and improve

processing time. The particle swarm optimization (PSO) algorithm is implemented for determining the ideal parameters of SVM for maximum classification and detection. The paper is summarized as follow:

- The SVM method is used for a dual purpose (fault detection and classification).
- Feature extraction and selection technique using WPT is used to minimize the data size to effectively reduce the computational burden and time.

The remaining sections are organized as signal processing, feature extraction and selection are discussed in section 2, In section 3, the SVM method discussed. Section 4. Discussed the implementation of the proposed fault diagnostic method. The results are discussed in section 5, and lastly, a conclusion is drawn in section 6.

## **2. Signal Processing**

Signal tracking and spectrum analysis is an essential element of signal decomposition and have been utilized for numerous signal analysis applications. The application of signal tracking is mostly used to distinguish the signal of interest from a range of present signals. Most power utilities use spectral analysis to analyze the signals recorded from a power system during the fault conditions to determine the nature of the fault. In this present work, WPT is used to track the signals of interest to enhance the fault detection algorithm.

### *2.1. Wavelet Packet Transform*

The WPT technique is a mathematical tool used for tracking and analyzing signals. The WPT may be observed as the simplification of discrete wavelet transform (DWT) that produces more efficient signal analysis results [43]. The WPT can be utilized for numerous expansions of a signal at different levels. The DWT has generally been used for many power systems applications and has proven to be effective. However, DWT doesn't produce good results for small values. This has led to the implementation of WPT in the present work. When using the WPT, the signal ( $S$ ) is passed through numerous filters containing both the low and high pass filtering process. The low frequency is represented by the approximation ( $a$ ) coefficient and the high frequency is represented by the detail ( $d$ ) coefficient. The ( $a$ ) and ( $d$ ) coefficients are supplementary decomposed repetitively to a particular point  $K$ . The main advantage of WPT over DWT is that more features are obtained and the frequency resolution of WPT is greater than that of DWT.

In the present work, 30 kHz is used as the sampling frequency, and level 4 the decomposition is considered. Shannon's entropy criterion is selected. The criterion is used to calculate the entropy at each decomposition level [44]. The best decomposition results are obtained when the parent entropy is greater than the total entropy of the decomposed level. The WPT decomposition tree of the present work is presented in Figure 1.

### *2.2. Extraction of features*

The extraction of certain features from a signal is a process of minimizing huge data matrix by selecting significant feature which represents a specific pattern to improve computational processing

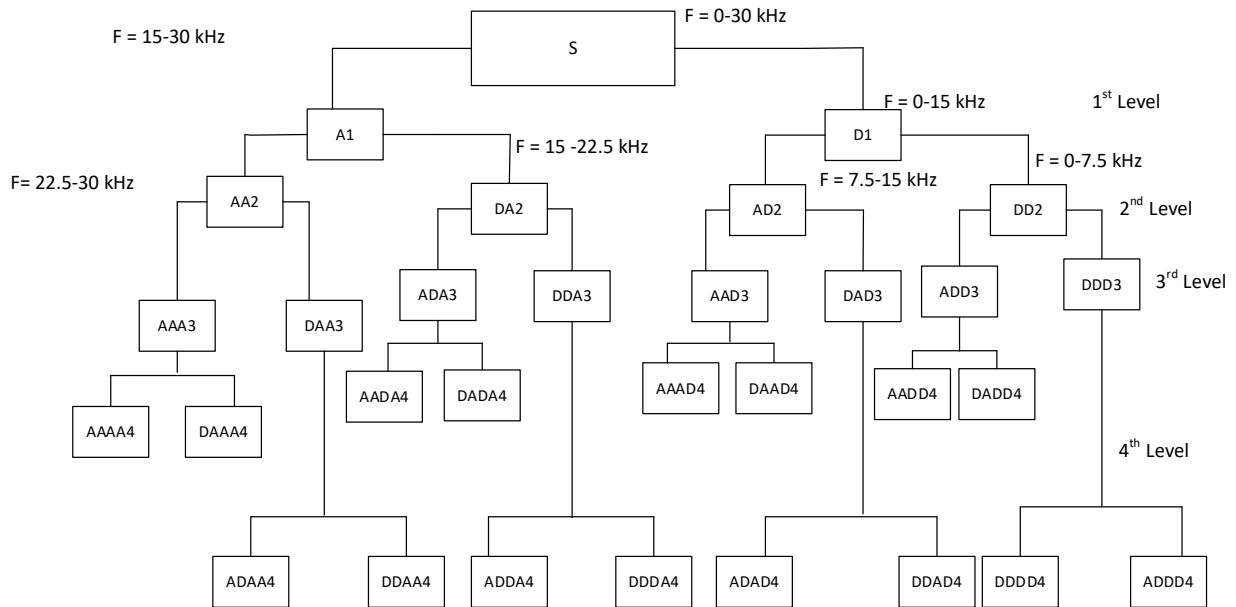


Figure 1: WPT decomposition tree

time. The data matrix is then converted to a feature matrix with fewer data points. The entropy and the energy features are extracted from the fault current signal. Energy is mathematically defined as [45, 46]:

$$E(t_1, t_2) = \int_{t_1}^{t_2} (|x|)^2 dt \quad (1)$$

where,  $E$  is the energy signal between the time range of  $(t_1, t_2)$  and  $x(t)$  is the signal. The value of the energy signal is greater than the value of the normal signal. The entropy feature is used to measure the signal information content [45]. The measured information includes the cost function of the signal  $x(t)$  defined by the entropy 'EN' such that the energy signal at zero  $E(0) = 0$ , the entropy is mathematically defined in eq. (2).

$$EN(x) = \sum_i EN(x_i) \quad (2)$$

where,  $(x_i)$  is the decomposed coefficient of the signal  $x(t)$ . The value of the entropy 'EN' is larger for transient signals and small for normal signals.

### 2.3. Feature Selection

Feature selection is a process of selecting the best features from the developed data matrix. The best features are ranked based on the correlation with the target output, the redundant features are then rejected. The feature selection process is imperative as large redundant data increases computational processing time and gives erroneous results which affect the efficiency of the classification algorithm [47]. The general concept of the feature selection technique is to select the feature which best represents the target. The forward feature selection technique is employed for selecting the best features. The technique is used to calculate the best features at each step iterative and a sub-input features matrix is developed, thus removing the redundant features [47]. The mean square error (MSE) function is used to evaluate the ranking of features

determined by the K-nearest neighbor (KNN) technique [48]. The features are ranked from high to low depending on the margin of error.

### 3. Pattern Recognition and Classification

Pattern recognition and classification have been a subject of interest for many researches in the past decades. The interest has arisen because of the many application ranging from speech recognition, image identification, power system fault, and optical character recognition. It is therefore important to build intelligent machines that can reliably and accurately be used to solve classification problems. Generally, classification is defined as a process of categorization, in which data, objects, and ideas are recognized and understood to produce an accurate response.

#### 3.1. Support Vector Machine

Support vector machine (SVM) was originally established to resolve statistical problems in empirical data modelling. When using SVM, the input data is plotted into a high dimensional space to determine the separating margin between two classes of data. The hyperplane is the separating index between the two classes of data [49]. The hyperplane is optimal when the distance between the class of data sets is maximized. The hyperplane can be calculated using the quadratic programming method defined mathematically as:

$$\min \frac{1}{2} |w|^2 + C \left( \sum_{i=1}^l \zeta_i \right), \quad (3)$$

$$\text{subject to } y_i(w \cdot x_i + b) \geq 1 - \zeta_i, \quad \zeta_i \geq 0 \forall i$$

where  $x_i$  is the  $i$ th example and the class label which is either +1 or -1 is represented by  $y_i$ . The problem is solved using its dual form

$$\max L_D = \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j (x_i^T x_j) \quad (4)$$



subject to  $0 \leq \alpha_i \leq C \quad \forall i, \quad \sum_i \alpha_i y_i = 0.$

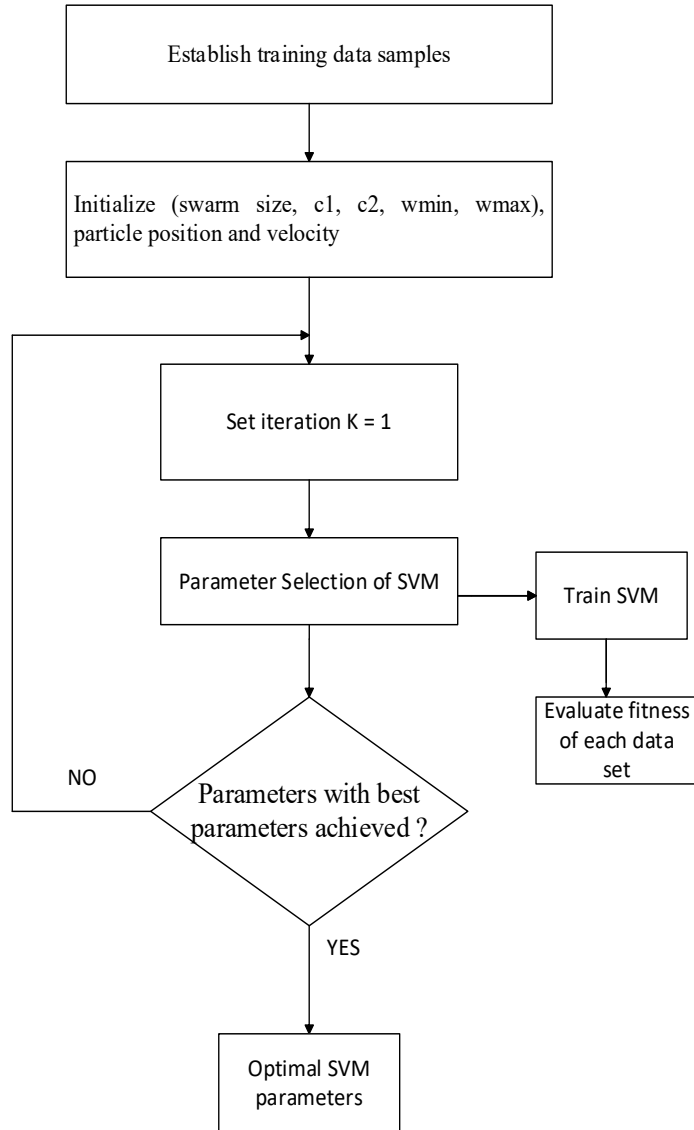


Figure 2: SVM parameter selection process using PSO

Kernel function can be used to solve the problem of non-linearity in statistical analysis. The linear, quadratic, radial bias function (RBF) and sigmoid are the most commonly used kernel functions. The advantage of SVM is that it gives a global solution, it is inclined to overfitting and it converges to local minima. The selection of SVM parameters is a significant and critical task for accurate fault diagnostic in a power distribution system [50]. To improve the performance of SVM, a particle swarm optimization (PSO) scheme is used to select the best parameters of the SVM.

The PSO technique was initially developed by Eberhart and Kennedy to solve optimization in 1995 [51]. The PSO technique is founded on the ordinary conduct of birds during a flight in space. The PSO relies on updating the initial position and velocity of each particle at every iterative until an optimal solution is determined. The process of using the PSO method begins with generating the random particles given by  $N$ . The best optimization solution is determined by calculating the fitness value of each

particle. and finally, the velocity of each particle ( $v_i$ ) is modernized by the mathematical representation defined as:

$$v_i(t+1) = w \times v_i(t) + C_1 \times rand(X_i^{best} - X_i(t)) + C_2 \times rand(X^{gbest} - X_i(t)) \quad (5)$$

where,  $X^{gbest}$  is the best global solution,  $X_i^{best}$  is the solution at the current position,  $C_1$  and  $C_2$  are the non-negative constants representing the best local and global position weight respectively, and  $w$  is the inertia coefficient. The position of the particle  $X_i$  will be updated using the following expression:

$$X_i(t+1) = X_i(t) + v_i \times (t+1) \quad (6)$$

To estimate the fitness value of the SVM by utilizing the PSO technique the fitness function is represented as:

$$f = \sqrt{\frac{1}{N} \sum_{k=1}^N [y^k - y(k)]^2} \quad (7)$$

where,  $N$  denotes the discrete sample number,  $y(k)$  is the discrete signal, and  $y^k$  is the SVM output. The process of obtaining optimal parameters for SVM application is presented by a flow chart in Figure 2.

#### 4. Proposed Fault Detection Technique

Section 4, discussed the proposed fault diagnostic technique which is applied in a power grid network. The fault detection taxonomy proposed in this work is presented in Figure 3. The fault current signal with one cycle is analyzed to detect the fault type that occurred in the power grid network. The fault current signal is decomposed into large frequency sub-bands using WPT. From the disintegrated fault current signal, the statistical signal features (energy and entropy) are extracted. The total set of features to build a matrix is 32 (16 coefficients  $\times$  2 statistical features). The fault current signal data is then generated considering simulation conditions. The generated data is then divided into the training and testing data.

From the total extracted features using WPT, other features do not forecast the desired results accurately and thus reducing the efficiency of the scheme. A feature selection technique is employed to eliminate features that do not present desired results. The selection technique uses a ranking algorithm to eliminate the features which are redundant and compromises the accuracy of the proposed technique. The best selected features are then fed into a classifier for fault identification. The PSO method is utilized to determine the optimal parameters of the SVM classifier. The fault classification scheme using the SVM technique is presented in Figure 4. From the presented scheme, each phase of the power system has a classification scheme to identify faults occurring in each phase, another SVM scheme is placed to detect ground faults. To accurately classify different fault conditions, the SVM output is either '+1' or '0', where '+1'

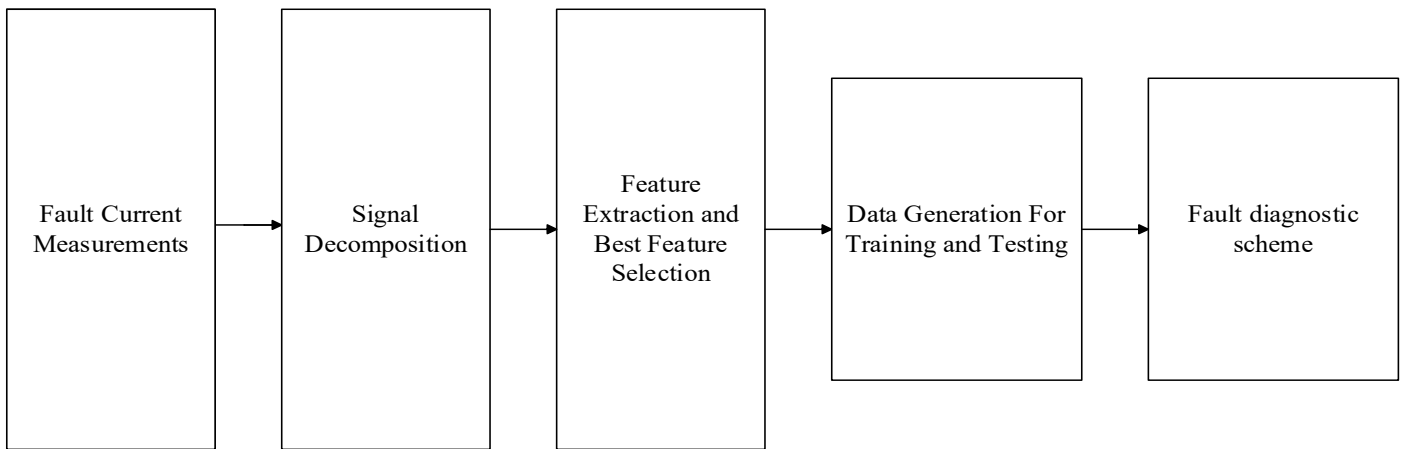


Figure 3: Fault detection taxonomy

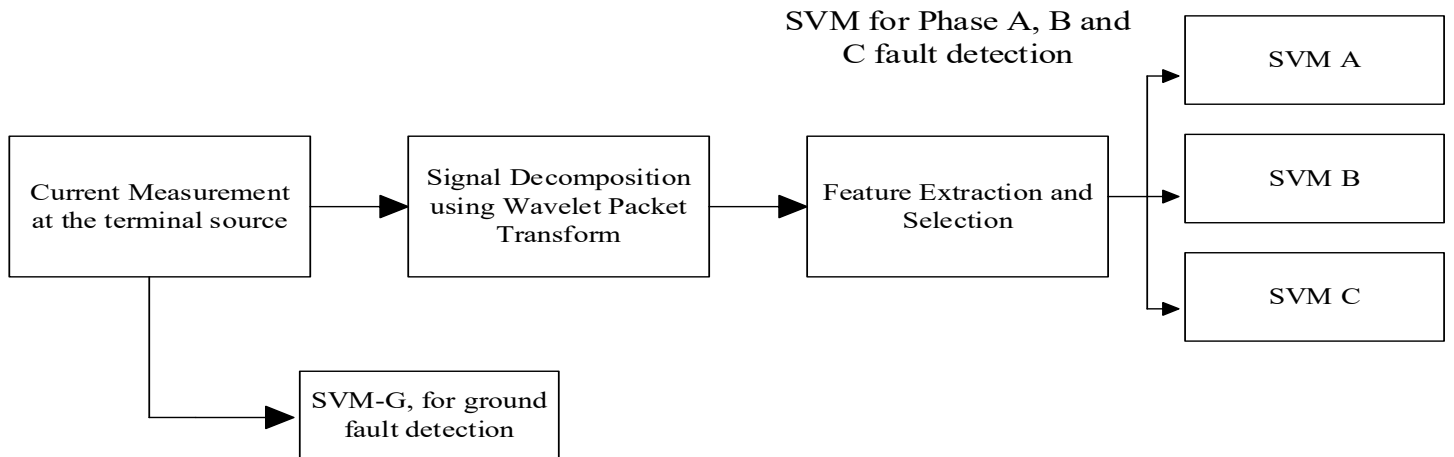


Figure 4: SVM Fault classification scheme

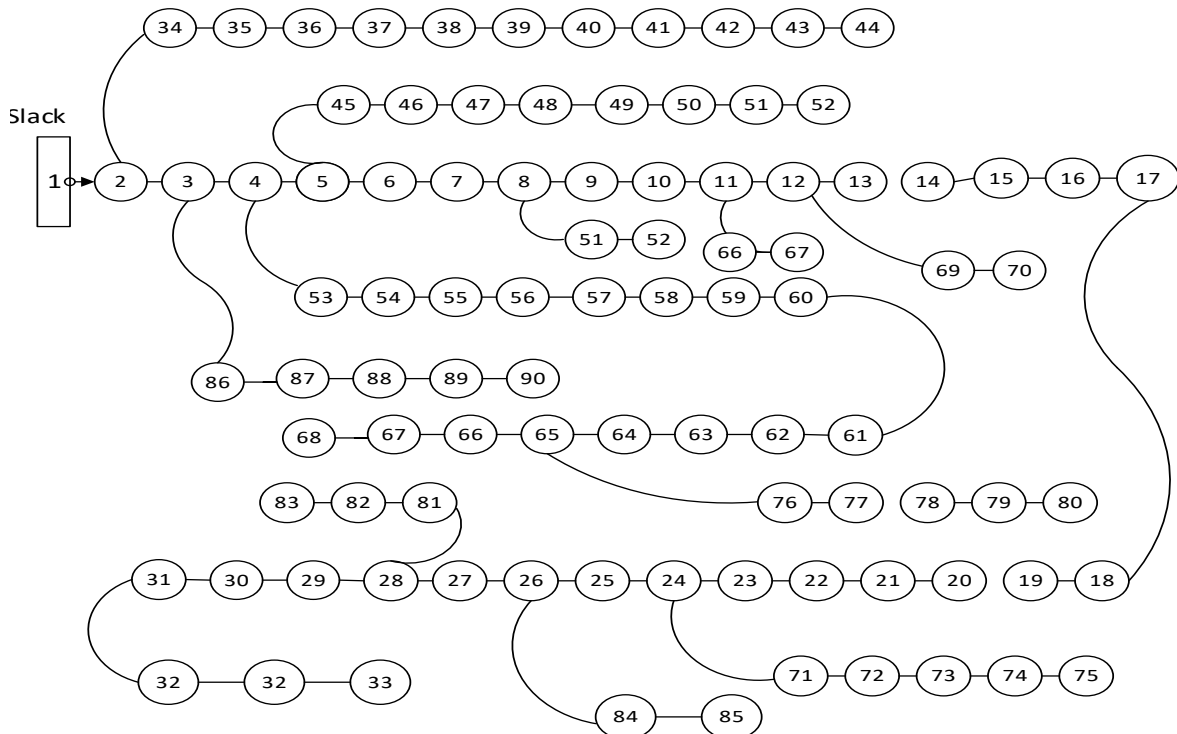


Figure 5: Eskom Network

shows that there is a fault and ‘0’ shows that there is no fault in the power system. In practice the line to line fault is usually misclassified as a line to line to ground, this problem may affect the restoration time. To solve this problem, a separate SVM scheme is positioned between the phase and ground where a zero-phase sequence current pointer is utilized as a directory value as presented in Figure 4. The index threshold value is determined using a trial and error method, in the present work, the value is set at 0.03. The ground fault detection is detected if the directory value is bigger than the set minimum value. The current index can be defined mathematically as:

$$Index = \frac{|I_a + I_b + I_c|}{mean(I_a, I_b, I_c)} \quad (8)$$

where,  $I_a, I_b,$  and  $I_c$  are the instantaneous current values. The classification accuracy (CA) is determined by:

$$CA = \frac{AFC}{No. of tested samples} \times 100 \quad (9)$$

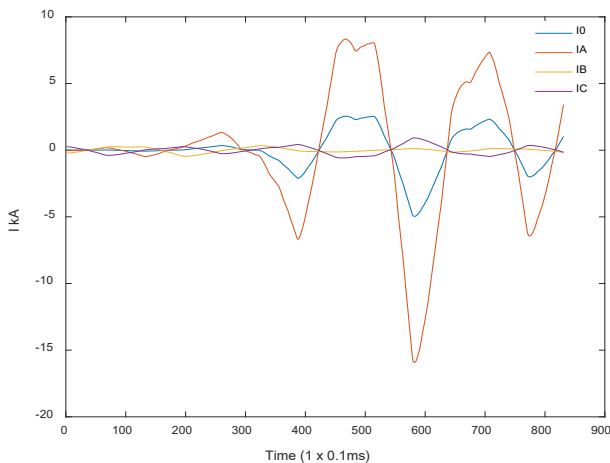
where, AFC represents the accurate fault classification.

### 5. Power System Case-Study

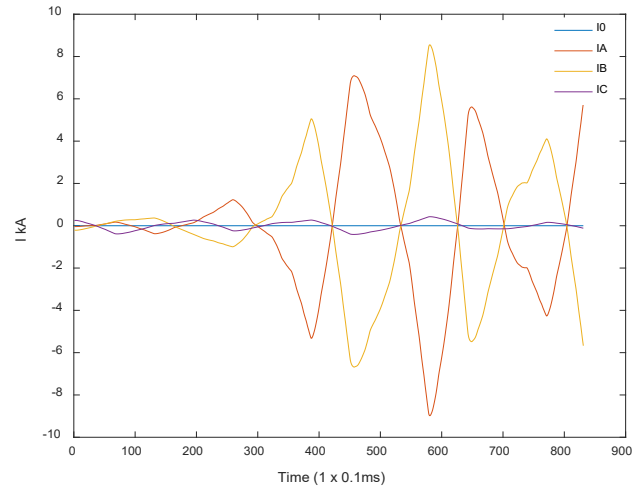
In this work, an Eskom 90 bus 22 kV system is considered. The power system network is modelled using Digsilent Power Factory platform. The PV and WT sources are connected into the network at the location optimally determined to satisfy technical consideration (power loss, voltage stability, fault levels, and power quality, etc...). The proposed Eskom network is presented in Figure 5. The base voltage of the network is 22 kV and the base apparent power is 100 MVA. The total load connected to the system is 115.5MVA. the maximum fault level at the substation is 15 kA. The PV connected into the system is rated at 50kW and the WT is rated at 120 kVA.

### 6. Results and Discussion

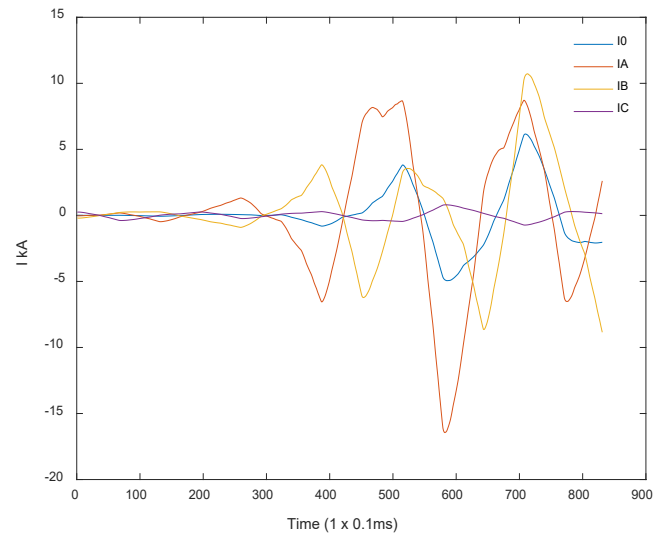
An integrated power mix energy distribution system is considered in this work. The simulated fault current signals are presented in (a)



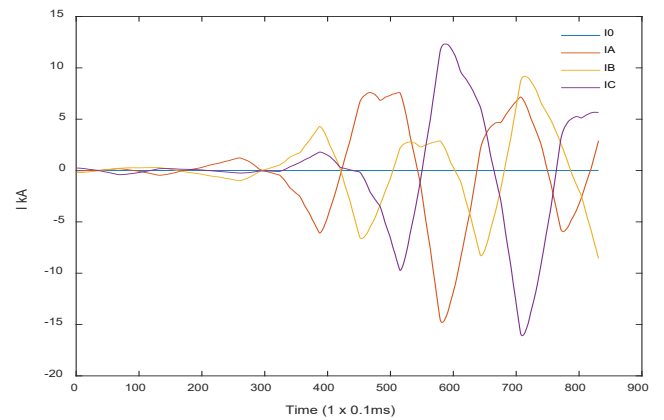
(a)



(b)



(c)



(d)

Figure 6: (a) Line-to-ground fault signal, (b) Line-to-line fault current signal, (c) Line-to-line-to-ground fault current signal, (d) Three phase fault current signal

The sampling frequency considered in the present work is taken to be 30 kHz. The WPT scheme is used to decompose the signal at level 4, and the entropy and energy features are extracted. The choice of a mother wavelet is vital for analyzing the signals. In the present work, Daubechies 4 is selected as the best for transient signal analysis. Some of the WPT decomposed features are presented in (d)

Figure 6. Furthermore, the best features are obtained from the extracted features. The best features using WPT to accurately predict the target is presented in Table 1. After the acquisition of best features, the features are subdivided into training and testing data sets. The training and testing parameters with various conditions are presented in Table 2.

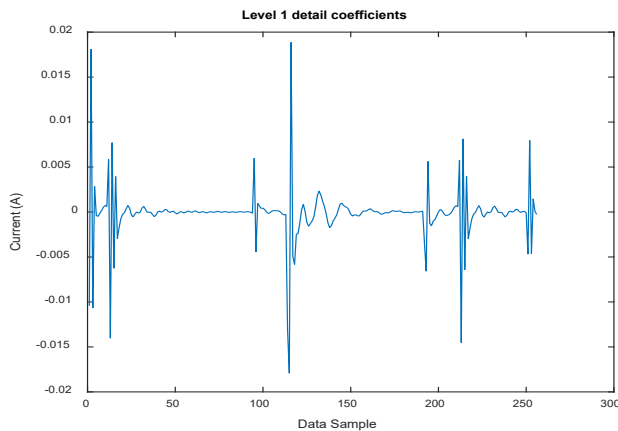
Different fault current cases are simulated and the data is subsequently fed into the SVM for classification. A fault training matrix using SVM is detailed in Table 3. In the present work, SVM output is either a '+1 or 0', where +1 shows that there is a fault and 0 is the output of a non-faulty section in the corresponding phase.

Table 1: Best Features

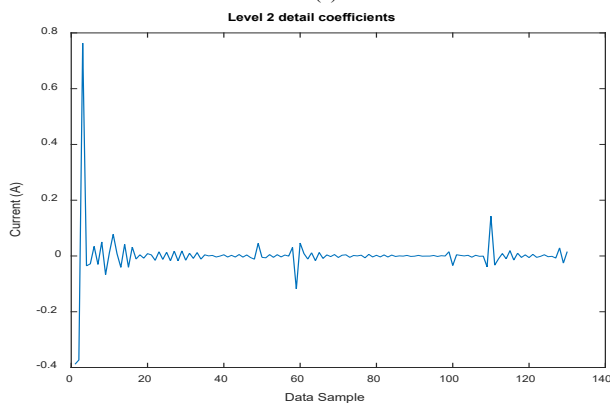
Signal type	Features	Coefficients
Current	Entropy Energy	ADAD4 DDDD4

Table 2: Training and testing parameters

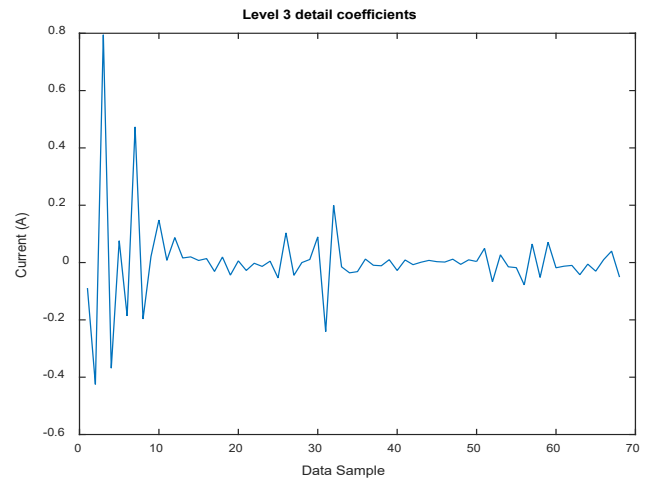
Data set	Fault resistance ( $\Omega$ )	Incipient Angle ( $\theta$ )
Test and Training	0, 1, 5, 15, 35, 80, 105, 150 2, 16, 28, 55, 68, 140, 190	$0^\circ, 30^\circ, 40^\circ, 60^\circ, 70^\circ$ $15^\circ, 48^\circ, 76^\circ, 88^\circ, 10^\circ$



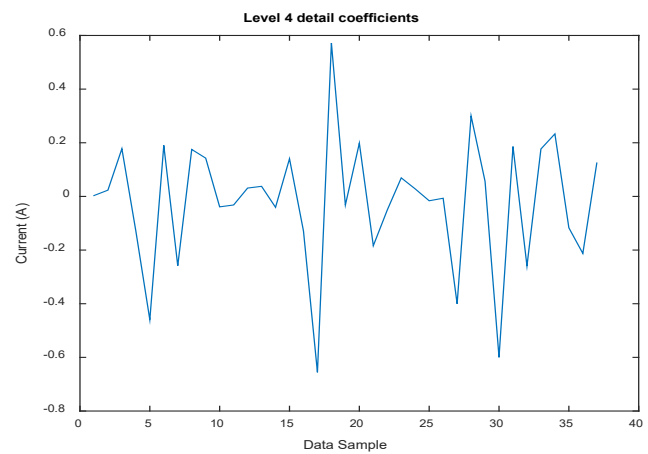
(a)



(b)



(c)



(d)

Figure 6: (a) Level 1 Detail Wavelet coefficients, (b) Level 2 Detail Wavelet coefficients, (c) Level 3 Detail Wavelet coefficients, (d) Level 4 Detail Wavelet coefficients

Table 3: SVM classification matrix

Fault Type	SVM <sub>A</sub>	SVM <sub>B</sub>	SVM <sub>C</sub>	SVM <sub>G</sub>
A-G	+1	0	0	+1
B-G	0	+1	0	+1
C-G	0	0	+1	+1
A-B	+1	+1	0	0
A-C	+1	0	+1	0
B-C	0	+1	+1	0
A-B-G	+1	+1	0	+1
A-C-G	+1	0	+1	+1
B-C-G	0	+1	+1	+1
A-B-C	+1	+1	+1	0



An evaluation fault classification process of selecting the optimal mother wavelet is determined and the results are shown in Table 4. In this present work, we further investigated other signal tracking analysis techniques such as the wavelet transform (WT) and the Fourier transform (FT). Based on the analysis done, the WPT signal analysis technique performed better than both the FT and WT. This is because when using the WPT, both the low and the high frequencies are measured and this improves the data analysis of a signal. From the results, dB4 has the highest accuracy level and thus it is selected for the scheme application. The PSO technique is used to obtain optimal parameters of the SVM. In Table 5, the best SVM parameters are presented. The PSO parameters are presented in Table 6.

Table 4: Mother wavelet performance evaluation

Mother wavelet	WPT fault diagnostic accuracy (%)	WT fault diagnostic accuracy (%)	FT fault diagnostic accuracy (%)
dB1	86.3	88.1	86.1
dB2	72.1	68.2	63.2
dB3	88.0	90	85.4
<b>dB4</b>	<b>99.8</b>	<b>86</b>	<b>81.2</b>
dB7	96.4	89.3	88.6
dB14	93.7	91.4	90.1

Table 5: SVM best parameters

SVM parameters	Fault classification
Kernel function	Radial bias function
Gamma ( $\gamma$ )	0.35
Cost ( $c$ )	Not used
Nu ( $\nu$ )	0.5

Table 6: PSO parameters

PSO parameters	Value
$c1 = c2$	5
Particle size	70
No. of iteration	1000
$wmin$	0.7
$wmax$	0.8

In Table 7, the classification results of different fault cases are presented. From the presented results it may be seen that the accuracy of classification is 1321. In the present work, we tested the Naïve Bayes (NB), Neural Network (NN), Decision tree (DT), and K-Nearest Neighbor. These techniques are implemented in a free machine learning platform Waikato Environment for Knowledge Analysis (WEKA). The classification process is carried by developing a classifier based on the training and testing of different class labels. Subsequently, the test data set is applied to the classifier to predict the accuracy of the classification.

The description of the different classifiers is discussed below:

- Naïve Bayes (NB): The NB algorithm is recognized as a fast learning technique. It is a simplified version of the Bayesian classifier and functions and under certain assumptions, (i) Attributes are conditionally independent for the class label, (ii) The prediction process is not affected by the latent attributes [52].
- Decision Tree (DT): The DT is a well-known efficient data mining algorithm for solving difficult problems by formulating computer graphic illustrations. The DT algorithm has been used proficiently to solve real-world problems [53].
- Neural Network (NN): The NN technique was developed by using the biological analysis of a human brain. This algorithm performs better with big data analysis and has proven to be efficient for classification and prediction purposes [54].

The performance of the different classifiers is evaluated using the confusion matrix. The correlation of the predicted instances values of the NB, DT and NN classifiers are presented in Table 8, Table 9 and Table 10 respectively.

Table 7: SVM accuracy

Fault type	Tested samples	Samples correctly classified	Samples incorrectly classified	Classification accuracy (%)
L-G (a-g, b-g, c-g)	60 400	60 075	325	99.4
L-L (a-b, a-c, b-c)	60 400	60 100	300	99.5
L-L-G (ab-g, ac-g, bc-g)	60 400	59 905	495	99.2
LLL (abc)	30 200	16 270	530	98.2
Total	211 400	196 350	1 650	99

Table 8: Naive Bayes fault classification performance

NB classifier	Line-Ground	Line-Line	Line-Line-Ground	Line-Line-Line
No. of cases	50 400	50 400	50 400	16 800
Correct classified	50 010	50 000	50 130	16 080
Incorrectly classified	390	400	270	720
Correct classified %	99.2	99.2	99.5	95.6
Mean absolute error	0.0085	0.0811	0.0712	0.0625
Root mean	0.0862	0.092	0.881	0.755
Kappa statistic	0.967	0.938	0.935	0.933
Confidence level (95%)	96.32	95.661	96.225	95.213

Table 9: Decision Tree fault classification performance

NB classifier	Line-Ground	Line-Line	Line-Line-Ground	Line-Line-Line
No. of cases	50 400	50 400	50 400	16 800
Correct classified	49 820	50 110	48 930	16 580
Incorrectly classified	580	290	1 470	220
Correct classified %	98.8	99.4	97.1	98.6
Mean absolute error	0.0093	0.0098	0.0397	0.0052
Root mean	0.0185	0.095	0.0891	0.0712
Kappa statistic	0.889	0.959	0.966	0.9221
Confidence level (95%)	95.101	96.261	97.181	97.43

Table 10: Neural Network fault classification performance

NB classifier	Line-Ground	Line-Line	Line-Line-Ground	Line-Line-Line
No. of cases	50 400	50 400	50 400	16 800
Correct classified	50 120	50 325	47 635	15 990
Incorrectly classified	280	75	2 765	810
Correct classified %	99.4	99.8	94.5	95.2
Mean absolute error	0.0792	0.0093	0.0235	0.0125
Root mean	0.0193	0.0822	0.0822	0.0961
Kappa statistic	0.722	0.823	0.917	0.9882
Confidence level (95%)	96.123	98.225	95.921	95.182

The performance comparison of the SVM, NB, DT, and NN techniques applied for classification is presented from

Table 11 to Table 14 respectively. To improve the computational time analysis, the fault current data set is subdivided and a quarter of the data sample is used for fault identification. From the presented results the average precision of SVM, NB, DT, and NN is 99.9%, 83%, 99.7%, and 94% respectively. For this application, the SVM classifier performed better than other tested classifiers in the present work.

Table 11: SVM performance with ¼ data set

Fault Type	Line-Ground	Line-Line	Line-Line-Ground	Line-Line-Line	Average
TP	0.992	0.999	0.997	0.999	0.997
FP	0.000	0.005	0.008	0.001	0.004
Precision	1.000	1.000	0.998	1.000	0.999
Recall	0.998	0.994	0.995	0.991	0.995
F-Measure	0.999	0.993	0.991	0.998	0.995
ROC	0.999	0.994	0.989	0.998	0.995

Table 12: NB Performance with ¼ data set

Fault Type	Line-Ground	Line-Line	Line-Line-Ground	Line-Line-Line	Average
TP	0.931	0.920	0.951	0.972	0.944
FP	0.024	0.004	0.006	0.001	0.009
Precision	0.851	0.561	0.945	0.922	0.830
Recall	0.957	0.883	0.992	0.856	0.922
F-Measure	0.858	0.864	0.840	0.951	0.878
ROC	0.967	0.960	0.910	0.951	0.947

Table 13: DT Performance with ¼ data set

Fault Type	Line-Ground	Line-Line	Line-Line-Ground	Line-Line-Line	Average
TP	0.989	0.962	0.840	0.856	0.912
FP	0.032	0.001	0.003	0.001	0.009
Precision	0.997	0.998	1.000	0.991	0.997
Recall	0.991	0.999	0.985	0.993	0.992
F-Measure	0.991	0.980	0.972	0.966	0.977
ROC	0.918	0.986	0.958	0.980	0.961

Table 14: NN performance with ¼ data set

Fault Type	Line-Ground	Line-Line	Line-Line-Ground	Line-Line-Line	Average
TP	0.992	0.942	0.966	0.981	0.971
FP	0.001	0.000	0.000	0.002	0.001
Precision	0.999	0.957	0.922	0.887	0.941
Recall	0.965	0.921	0.989	0.968	0.961
F-Measure	0.965	0.982	0.972	0.951	0.968
ROC	0.976	0.991	0.987	0.992	0.987

## 7. Conclusion

An increase in electricity demand has enlarged the technical variations of the load-demand phenomenon. To solve this problem external electricity sources have been considered. The environmental attributes of using REDG improves the quality of the air and thus contributing positively to the health of the people. However, integrating REDGs has technical challenges that must be addressed. In the present work, we focus on the fault diagnostic mechanism when the REDGs are integrated into the distribution system. An Eskom power system is modelled, and various fault studies are carried out. In the present work, a fault diagnostic technique is proposed. The method consists of the signal processing scheme, a feature extraction section, feature selection section, and a fault diagnostic section. The WPT is used to decompose the signal into frequency sub-bands, subsequently, the entropy and energy features are selected from the decomposed signal. From the selected features a feature selection scheme is used to select the best features to improve the computational time and reduce burden. The selected features are then fed into the SVM classifier to determine the fault occurrence in the network. The PSO algorithm is used to determine the best parameters of the SVM. We further investigated the effectiveness of other classification algorithms. From the results obtained the SVM classifier performed better with the accuracy of 99%. The future work will entail a fault location scheme in an integrated system.

## References

- [1] K. Kusakana, H. J. Vermaak and G. P. Yuma, "Optimization of hybrid standalone renewable energy systems by linear programming," *Advanced Science Letters*, **19**(4), 2501-2504, 2013. DOI: 10.1166/asl.2013.4948
- [2] T. Dragicovic, "Model predictive control of power converters for robust and fast operation of AC microgrids," *IEEE Transaction on Power Electronics*, **33**(7), 6304-6317, 2018. DOI: 10.1109/TPEL.2017.2744986
- [3] T. Dragicovic and M. Novak, "Weighting factor design in model predictive control of power electronic converters: An artificial neural network approach," *IEEE Transaction on Industrial Electronics*, 1-12, 2018. DOI: 10.1109/TIE.2018.2875660
- [4] P. M. S. Carvalho, P. F. Correia and L. Ferreira, "Distributed reactive power generation control for voltage rise mitigation in distribution networks," *IEEE Transaction on Power Systems*, **23**(2), 766-772, 2008. DOI: 10.1109/TPWRS.2008.919203
- [5] W. E. Khttam and M. M. A. Salama, "Distributed generation technologies, definition and benefits," *IEEE Transaction on Power Systems*, **71**(2), 119-128, 2004 DOI: 10.1016/j.epr.2004.01.006.
- [6] A. M. Foley, P. G. Leahy, A. Marvuglia and E. J. McKeogh, "Current methods and advances in forecasting of wind power generation," *Renewable Energy*, **37**(1), 1-8, 2012. DOI: 10.1016/j.renene.2011.05.033
- [7] K. Meng, H. Yang, Z. Y. Dong, W. Guo, F. Wen and X. Xu, "Flexible operational planning framework considering multiple wind energy forecasting service providers," *IEEE Transaction on Sustainable Energy*, **7**(2), 708-717, 2016. DOI: 10.1109/TSTE.2015.2497698
- [8] M. N. Kabir, Y. Mishra, G. Ledwich, Z. Y. Dong and K. P. Wong, "Coordinated control of control-connected photovoltaic reactive power and battery energy storage systems to improve the voltage profile of a residential distribution feeder," *IEEE Transaction on Industrial Informatics*, **10**(2), 967-977, 2014. DOI: 10.1109/TII.2014.2299336
- [9] Z. Y. Dong, M. Lu, Z. Lu and K. P. Wong, "A differential evolution based method for power system planning," in *IEEE International Conference on Evolutionary Computation*, 2006 DOI: 10.1109/CEC.2006.1688646.
- [10] D. Mlakic, H. R. Baghaee and S. Nikolovski, "Gibbs phenomenonbased based hybrid islanding detection strategy for VSC-based microgrids using frequency shift, THDU and RMSU," *IEEE Transactions on Smart Grids*, 1-12, 2018. DOI: 10.1109/TSG.2018.2883595
- [11] "IEEE Standard on Interconnecting Distributed Resources with Electric Power Systems," IEEE Std. 1547, 2003.
- [12] S. I. Jang and K. H. Kim, "An islanding detection method for distributed generations using voltage unbalance and total harmonic distortion of current," *IEEE Transactions on Power Delivery*, **19**(2), 745-752, 2014. DOI: 10.1109/TPWRD.2003.822964
- [13] M. Khodaparastan, H. Vahedi, F. Khazaeli and H. Oraee, "A novel hybrid islanding detection method for inverter-based DGs using SFS and ROCOF," *IEEE Transactions on Power Delivery*, **32**(5), 2162-2170, 2017. DOI: 10.1109/TPWRS.2008.2007002
- [14] X. Wang, W. Freitas, V. Dinavahi and W. Xu, "Investigation of positive feedback anti-islanding control for multiple inverter-based distributed generators," *IEEE Transactions on Power Systems*, **24**(2), 785-795, 2009. DOI: 10.1109/TPWRS.2008.2007002

- [15] S. J. Huang and F. S. Pai, "A new approach to islanding detection of dispersed generators with self-commutated static power converters," *IEEE Transactions on Power Delivery*, **15**(2), 500-507, 2000. DOI: 10.1109/61.852975
- [16] S. Salman, "New loss of mains detection algorithm for embedded generation using rate of change of voltage and changes in power factors," in *7th International Conference on Developments in Power Systems Protection*, 2001. DOI: 10.1049/cp:20010105
- [17] A. Samui and S. R. Samantaray, "Assessment of ROCPAD relay for islanding detection in distributed generation," *IEEE Transactions on Smart Grid*, **2**(2), 391-398, 2011. DOI: 10.1109/TSG.2011.2125804
- [18] S. Nikolovski, H. R. Baghaee and D. Mlakic, "Islanding detection of synchronous generator-based DGs using rate of change of reactive power," *IEEE Systems Journal*, **9**, 1-11, 2019. DOI: 10.1109/JSYST.2018.2889981
- [19] A. S. Subhadra, P. Reddy and S. B. Modi, "Islanding detection in a distribution system with modified DG interface controller," *International Journal of Applied Power Engineering (IJAPE)*, **6**(3), 135-143, 2017. DOI: 10.11591/ijape
- [20] S. Cavalcante, P. R. Araujo, R. D. Keuton, Alvesa, F. B. Costa and T. O. A. Rocha, "A wavelet-based hybrid islanding detection system applied for distributed generators interconnected to AC microgrids," *Electrical Power and Energy Systems*, **121**(2), 1-10, 2020. DOI: 10.1016/j.ijepes.2020.106032
- [21] D. D. Reigosa, F. Briz, C. B. Charro and J. M. Guerrero, "Passive islanding detection using inverter nonlinear effects," *IEEE Transactions on Power Electronics*, **32**(11), 8434-8445, 2017. DOI: 10.1109/TPEL.2016.2646751
- [22] M. Inci, "Active/reactive energy control scheme for grid-connected fuel cell system with local inductive loads," *Energy*, **197**(3), 1-15, 2004. DOI: 10.1016/j.energy.2020.117191
- [23] A. T. Kollu and N. Ghaffarzadeh, "A novel phaselet-based approach for islanding detection in inverter-based distributed generation systems," *Electric Power Systems Research*, **182**(2), 1-9, 2020. DOI: 10.1016/j.epr.2020.106226
- [24] G. Wang, F. Gao, J. Liu, Q. Li and Y. Zhao, "Design consideration and performance analysis of a hybrid islanding detection method combining voltage unbalance/total harmonic distortion and bilateral reactive power variation," *IEEE Transactions on Power Electronics and Applications*, **5**(1), 86-100, 2020. 10.24295/CPSSTPEA.2020.00008
- [25] M. C. Wrinch, "Negative sequence impedance measurement for distributed generator islanding detection," University of British Columbia, Vancouver, BC, Canada, 2008. DOI: 10.14288/1.0065504
- [26] N. Liu, C. Diduch, L. Chang and J. Su, "A reference impedance based passive islanding detection method for inverter-based distributed generation system," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, **3**(4), 1205-1217, 2015. DOI: 10.1109/JESTPE.2015.2457671
- [27] S. S. Madani, Abbaspour, M. Beiraghi, P. Z. Dehkordi and A. M. Ranjbar, "Islanding detection for PV and DFIGN using decision tree and AdaBoost algorithm," in *3rd IEEE PES Innovative Smart Grid Technologies Europe*, Berlin, Germany, 2012. DOI: 10.1109/ISGTEurope.2012.6465818
- [28] F. Bignucolo, A. Raciti and R. Caldori, "Coordinating active and reactive energy balances in islanded networks supported by renewables and BESS," in *3rd Renewable Power Generation Conference*, Naples, Italy, 2014. DOI: 10.1049/cp.2014.0824
- [29] I. Phafula, E. D. M. Koch and K. Nixon, "Preliminary Study of Fault Detection on an Islanded Microgrid Using Artificial Neural Networks," in *IEEE 2020 International SAUPEC/RobMech/PRASA Conference*, Cape Town, South Africa, 2020. DOI: 10.1109/SAUPEC/RobMech
- [30] M. Khodaparastan, H. Vahedi, F. Khazaeli and H. Oraee, "A novel hybrid islanding detection method for inverter-based DGs using SFS and ROCOF," *IEEE Transactions on Power Delivery*, **32**(5), 2162-2170, 2017. DOI: 10.1109/TPWRD.2015.2406577
- [31] S. R. Samantaray, K. El-Arroudi, G. Joos and I. Kamwa, "A fuzzy rule-based approach for islanding detection in distributed generation," *IEEE Transactions on Power Delivery*, **25**(3), 1427-1433, 2010. DOI: 10.1109/TPWRD.2010.2042625
- [32] D. Mlakic, H. R. Baghaee and S. Nikolovski, "A novel ANFIS-based islanding detection for inverter-interfaced microgrids," *IEEE Transactions on Smart Grid*, **9**, 1-13, 2018. DOI: 10.1109/TSG.2018.2859360
- [33] P. Kumar, V. Kumar and B. Tyagi, "A Novel Islanding Detection Technique Based on Event Index Value for Reconfigurable Microgrid," in *IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy*, Cochin, India, 2020. DOI: 10.1109/PESGRE45664.2020.9070479
- [34] A. Garg, A. Sinha, M. J. B. Reddy and D. K. Mohanta, "Detection of islanding in microgrid using wavelet-MRA," in *IEEE Conference on Power, Control, Communication and Computational Technologies for Sustainable Growth*, Kurnool, India, 2015. DOI: 10.1109/PCCCTSG.2015.7503917
- [35] M. Mishra and P. K. Rout, "Detection and classification of micro-grid faults based on HHT and machine learning techniques," *IET Generation, Transmission & Distribution*, **12**(2), 388-397, 2018. DOI: 10.1049/iet-gtd.2017.0502
- [36] J. Li and Y. Guo, "Support vector machine based fault detection and diagnosis for HVAC systems," *International Journal of Intelligent Systems Technologies and Applications*, **18**(2), 204-222, 2019. DOI: 10.1504/IJISTA.2019.097752
- [37] H. R. Baghaee, D. Mlakic, S. Nikolovski and T. Dragičević, "Support Vector Machine-based Islanding and Grid Fault Detection in Active Distribution Networks," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, **3**(15), 1-19, 2019. DOI: 10.1109/JESTPE.2019.2916621
- [38] F. Mohammadi, G. A. Nazri and M. Saif, "A Fast Fault Detection and Identification Approach in Power Distribution Systems," in *IEEE 2019 International Conference on Power Generation Systems and Renewable Energy Technologies (PGSRET)*, Istanbul, Turkey, 2019. DOI: 10.1109/PGSRET.2019.8882676
- [39] V. Ashok, A. Yadav and A. Y. Abdelaziz, "MODWT-based fault detection and classification scheme for cross-country and evolving faults," *Electric Power Systems Research*, **175**, 1-20, 2019. DOI: 10.1016/j.epr.2019.105897
- [40] A. Mohamed, B. Younes, T. Lamhamdi, H. E. Moussaoui and H. E. Markhi, "Fault Location and Isolation Technique in Smart Distribution Systems with Distributed Generation," in *IEEE 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, Meknes, Morocco, 2020. DOI: 10.1109/IRASET48871.2020.9092095
- [41] P. J. Rivera-Torres and O. L. Santiago, "Fault Detection and Isolation in Smart Grid Devices Using Probabilistic Boolean Networks," *Computational Intelligence in Emerging Technologies for Engineering Applications*, **872**, 165-185, 2020. DOI: 10.1007/978-3-030-34409-2\_10
- [42] H. H. Alhelou, M. E. H. Golshan, N. D. Hatziaargyriou and M. P. Moghaddam, "Novel Unknown Input Observer-based Measurement Fault Detection and Isolation scheme for Micro-Grid Systems," *IEEE Transactions on Industrial Informatics*, **15**(3), 1-10, 2020. DOI: 10.1109/TII.2020.2970211
- [43] P. Ray and D. P. Mishra, "Support vector machine based fault classification and location of a long transmission line," *Engineering Science and Technology, an International Journal*, **10**, 1368-1380, 2016. DOI: 10.1016/j.jestch.2016.04.001
- [44] K. Thirumala, A. Kanjolia, T. Jain and A. C. Umarikar, "Empirical wavelet transform and dual feed-forward neural network for classification of power quality disturbances," *International Journal of Power and Energy Conversion*, **11**(1), 1-19, 2009. DOI: 10.1504/IJPEC.2020.104805
- [45] M. Dashtdar, M. Esmailbeig, M. Najafi and M. E. N. Bushehri, "Fault Location in the Transmission Network Using Artificial Neural Network," *Automatic Control and Computer Sciences*, **54**(1), 39-51, 2020. 10.3103/S0146411620010022



- [46] D. Mourada and E. H. Shehab\_Eldinb, "An enhanced ground fault selection and distance measurement approach for single and double circuit fixed series compensated transmission line," *Electric Power Systems Research*, **187**(3), 1-11, 2020. DOI: 10.1016/j.epsr.2020.106495
- [47] S. Luo, W. Yang and H. Tang, "A novel feature selection method to boost variable predictive model-based class discrimination performance and its application to intelligent multi-fault diagnosis," *Expert Systems Application*, **53**(2), 104-118, 2020. DOI: 10.1177/0020294019877497
- [48] T. T. Le, R. J. Urbanowicz, J. H. Moore and B. A. McKinney, "Statistical Inference Relief (STIR) feature selection," *Journal of Bioinformatics*, **35**(8), 1358-1365, 2020. DOI: 10.1093/bioinformatics/bty788
- [49] P. D. Raval and A. S. Pandya, "A Novel Fault Classification Technique in Series Compensated Transmission Line using Ensemble Method," *International Journal of Pattern Recognition and Artificial Intelligence*, **34**(4), 1-18, 2020. /DOI: 10.1142/S0218001420500093
- [50] H. A. FayedabAmir and F. Atiyac, "Speed up grid-search for parameter selection of support vector machines," *Journal of Applied Soft Computing*, **80**(4), 202-210, 2019. DOI: 10.1016/j.asoc.2019.03.037
- [51] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Sixth International Symposium on Micro Machine and Human Science*, 1995 DOI: 10.1109/MHS.1995.494215.
- [52] A. H. Jahromi and M. Taheri, "A non-parametric mixture of Gaussian Naive Bayes classifiers based on local independent features," in *Artificial Intelligence and Signal Processing Conference (AISP)*, Shieaz, 2017. DOI: 10.1109/AISP.2017.8324083
- [53] A. Trabelsi, Z. Elouedi and E. Lefevre, Decision tree classifiers for evidential attributes values and class labels, *Fuzzy Sets Systems*, 2018. oi.org/DOI: 10.1016/j.fss.2018.11.006
- [54] J. Gui, T. Liu, D. Tao, Z. Sun and T. Tan, "Representative vector machines: a unified framework for classical classifiers," *IEEE Trans. Cybern.*, **46**(8), 1877-1888, 2016. DOI: 10.1109/TCYB.2015.2457234

## Overview of Solar Radiation Estimation Techniques with Development of Solar Radiation Model Using Artificial Neural Network

Amar Choudhary<sup>1,\*</sup>, Deependra Pandey<sup>1</sup>, Saurabh Bhardwaj<sup>2</sup>

<sup>1</sup>Electronics and Communication Engineering Department, Amity School of Engineering and Technology, Amity University, Lucknow, 226010, India

<sup>2</sup>Electronics and Instrumentation Engineering Department, Thapar Institute of Engineering and Technology, Thapar University, Patiala, 147001, India

### ARTICLE INFO

Article history:

Received: 08 May, 2020

Accepted: 07 August, 2020

Online: 19 August, 2020

Keywords:

Renewable Energy

Artificial Neural Network

Machine Learning

Solar Radiation Estimation

Fuzzy Logic

Support Vector Machine

### ABSTRACT

Estimation Solar radiation is the most significant part of the optimization of solar power. This may be achieved if the solar radiation is predicted well in advance. Meteorological stations have radiation measuring devices like pyranometer, pyr heliometer, radiometer, solarimeter, etc. however, it may not be available at the location of interest for researchers. Due to this limitation solar radiation estimation models are devised based on location details like Altitude, Latitude, Longitude, and metrological details like Wind Speed, Ambient Temperature, Relative Humidity, Day Temperature, etc. These radiation models provide Global Solar Radiation (GSR) as output. These models are statistically tested based on error calculation like Mean Bias Error, Mean Absolute Error, Root Mean Square Error, etc. This paper is framed to briefly provide the idea behind different solar radiation estimation models with the methodology used. Soft computing-based models are mainly analyzed here. ANN-based Global Solar Irradiance Estimation Model has been developed using geographical parameters like Elevation, Latitude, Longitude, Longitude, and meteorological parameters like Months of a year, Days of a month, Temperature, Atmospheric Pressure, Relative Humidity, and Wind Speed. The data are downloaded from the National Solar Radiation Database (NSRDB) for 2014 (latest available). From this paper, the reader will come to know about various techniques used in solar radiation estimation. The developed ANN-based model has better results for training, testing, validation, and all with Regression value of 0.94304, 0.9488, 0.94766, 0.94556 respectively. The MSE is found to be 0.0089147 at epoch 0. The obtained values of R and MSE indicates the suitability of the developed model.

### 1. Introduction

Optimization of renewable sources of energy is one of the thrust areas for the researchers, now a days. There are many advantages of renewable sources over non-renewable sources like abundancy, non-pollutant, etc. in comparison to other sources of renewable energy like biomass, wind, hydropower, geothermal, solar energy is most preferable due to its profound abundancy over the earth. The sun radiates about 1,20,000 TW of radiation per hour which is more than sufficient to fulfill the energy need of the world for a year [1-2]. Solar energy with huge potential can meet the need of earth's energy requirements [3]. Solar energy has a great lead as

per the application point of view over other renewable sources of energy [4]. The solar radiation is available as extraterrestrial and global solar radiation. The first one is found above the atmosphere while the second one is under the atmosphere. Global solar radiation is measured by the measuring instruments like Pyr heliometer in case of direct beam solar radiation and by Solarimeter, Pyranometer, Radiometer in case of diffused solar radiation [5]. These measuring devices are of very high cost and they are installed at a few meteorological stations. This means, measuring devices may not be available at the locations of interest for the researchers [6]. Due to this limitation, solar radiation is estimated based on the location and meteorological parameters

\*Corresponding Author: Amar Choudhary, amar.giet.ece@gmail.com

such as latitude, longitude, altitude, sunshine hour, air temperature, wind speed, cloud cover, humidity, days of a month, etc.

As per the nature of the forecasting models, it may be categorized as Mathematical/Classical/Statistical Models, Machine Learning Models, Cloud Motion Models, Numerical Weather Prediction Models, and Hybrid Models. Persistence Models provide a standard forecast in comparison to the realization of other models.

Empirical models mostly depend on the following factors [7]:

- Astronomical elements such as hour angle, solar declination, earth-sun distance, etc.
- Meteorological elements like humidity, air temperature, sunshine duration, precipitation, evaporation, etc.
- Geographical elements like altitude, longitude, latitude, etc.
- Physical elements like water vapor content scattering of dust, scattering of air molecules, etc.
- Geometrical elements such as tilt angle, sun elevation, sun azimuth, etc.

Further, based on types of input meteorological parameters, empirical models may be categorized as Sunshine, Cloud, Temperature, and other meteorological parameter-based models.

This paper is oriented as follows. Section 2 and section 3, provide a brief survey of the classical solar radiation estimation model and development of a new global solar irradiation model. Results have been quoted in section 4 and conclusion of the work has drawn section 5.

## 2. Summary of Empirical/ Classical/ Statistical Model of Solar Radiation Estimation

In earlier days Solar Radiation estimation was carried out using various mathematical relations [8-9], which were widely tested and evaluated across the globe. Later on, its revised versions like quadratic, cubic, exponential, and logarithmic, were also advised by various researchers [10-13]. A comparative study reveals that some revised models have better results than that of the A-P Model [14-17]. Another author found similar results while evaluating linear, quadratic, cubic, and exponential models in Iran [15]. Insignificant difference between these models were reported after testing and evaluating [17]. After this, several researchers modified A-P Model by incorporating other parameters like atmospheric pressure, precipitation, air temperature, relative humidity, etc. [18-24].

As sunshine models are subjected to availability of sunshine hour [25-26], so to overcome with this, a model was devised based on maximum and minimum temperatures [27]. Later on, this model was also improved by several researchers [28]. By using precipitation, atmospheric pressure, and relative humidity data, model developed by [27] and was modified by [29]. Many researchers found that the accuracies of the model of [27] and other models [30] highly varies for various geographical locations and local climate of the location of interest [31].

In addition to the above two categorizations, some authors have used other parameters like precipitation, atmospheric pressure, and relative humidity to estimate solar radiation. However, due to the complex radiation process, it is a challenging task nowadays also

to develop the perfect empirical model [32]. Several empirical models were developed, evaluated, and reviewed by researchers [33-37].

## 3. Development of Global Horizontal Solar Radiation Estimation Model based on ANN

In Section 2, maximum types of classical as well as machine learning models are briefed and found that Artificial Neural Network-based estimation models have better performance in comparison to others [38-39]. The development of the ANN-based model is detailed below.

### 3.1. Geographical and Meteorological data Collection and Processing

The present analysis is carried out for New-Delhi (National Capital of India). The measured geographical parameters such as Latitude, Longitude, Elevation and meteorological parameters such as Months of a year, Days of a month, Temperature, Atmospheric Pressure, Humidity, and Wind Speed are downloaded from National Solar Radiation Database (NSRDB) for 2014 (latest available). As, data of different downloaded parameters were having different ranges and units, so max-min normalization of data was performed ranging between 0-1 by equation (1);

$$v'_i = \left[ \left( \frac{v_i - B}{A - B} \right) \cdot (M - N) \right] + N \quad (1)$$

where  $v'_i$  is the Normalized value of the variable,  $v_i$  is downloaded value of the variable,  $A$  is Maximum value,  $B$  is the minimum value,  $M$  is the new maximum value, and  $N$  is the new minimum value.

### 3.2. Methodology

A computer program has been performed under MATLAB R2016a using Neural Network/Data Manager Tool: nntool. Its configuration details are listed in Table 1 below.

Table 1: NN Tool Customization

Sl. No.	Particulars	Configuration Details
1	Network Type	Feed Forward Back Propagation
2	Training Function	TRAINLM
3	Adaptation Learning Function	LEANGDM
4	Error Function	MSE
5	Number of Hidden Layers	02
6	Properties for Layer-1	Transfer Function: TANSIG, No. of Neurons: 10
7	Properties of Layer-2	Transfer Function: TANSIG
8	Training Info	Input and Output
9	Training Parameters	Epochs: 1000, max fail: 1000
10	Data Division	Random (dividerand)
11	Training	Levenberg-Marquardt (trainlm)
12	Performance	Mean Squared Error (MSE)
13	Calculation	MEX
14	Plot Interval	1 Epochs

The Mean Square Error (Equation-2) is used for error calculation and evaluation of the developed model.

$$MSE = \left(\frac{1}{n}\right) \sum_{j=1}^n (SR_{i(ANN)} - SR_{i(actual)})^2 \quad (2)$$

where n is the number of inputs,  $SR_{i(ANN)}$  is estimated Global Solar Irradiance,  $SR_{i(actual)}$  estimated Global Solar Irradiance.

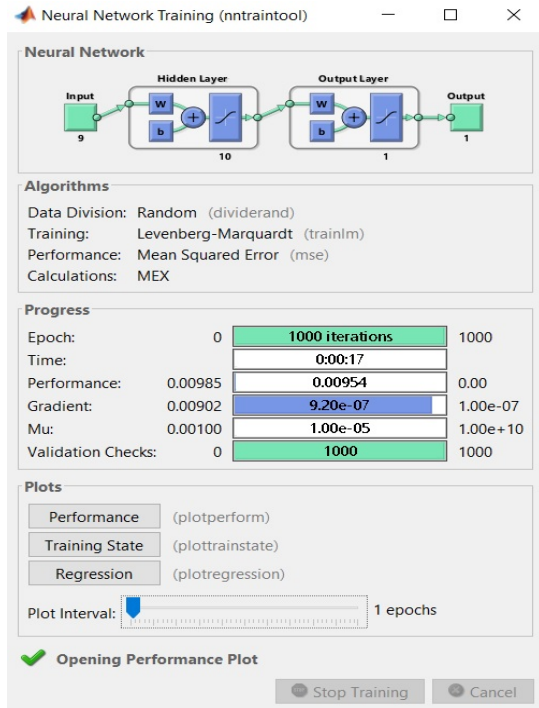


Figure 1. Neural Network Training Environment

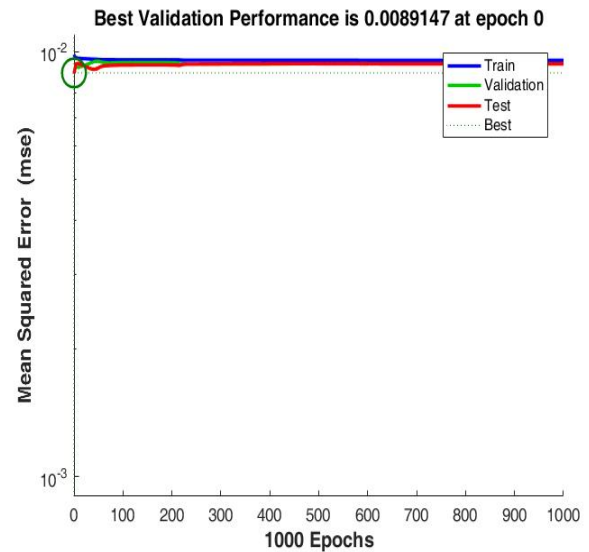


Figure 2. Performance Plot

#### 4. Results

The training environment of the Neural Network Train tool is in Figure 1. Where architecture of the applied neural network is represented along with detail of algorithms, plot interval, and progress of training.

Figure 2 is the performance plot after the training, testing, and validation. It is a plot between epochs and Mean Squared Error. The best validation performance is 0.0089147 at epoch 0. Also, training, validation, testing, and best performance curve is shown by blue, green, red, and dotted lines respectively.

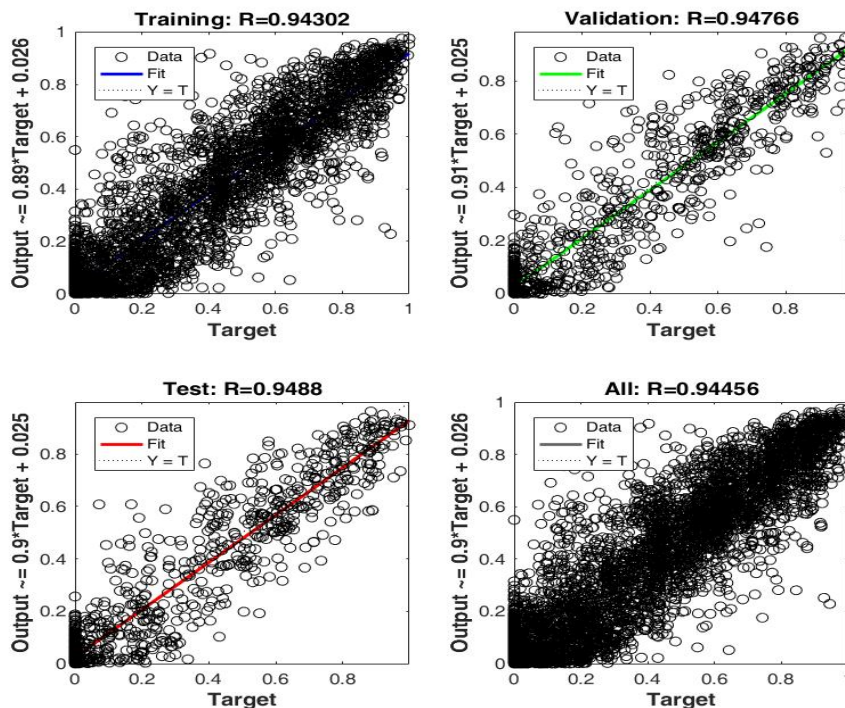




Figure 3 is the plot between Target and Output. This graph is plotted for training, testing, validation, and all. The R-value of Training is 0.94304, for Testing is 0.9488, for validation is 0.94766 and for all it is 0.94556. These values are listed in Table 2. R-value closer to 1 and MSE value closer to 0 are assumed to be a better one.

Table 2: R Value Analysis

Sl. No.	Particulars	R
1	Training	0.94302
2	Validation	0.94766
3	Testing	0.9488
4	All	0.94456

## 5. Conclusion

The implantation of Artificial Neural Network in the modeling of Global Solar Radiation is reported. The developed model shows that selection of ANN model has lesser MSE and considerably good values of R. The model is developed by using meteorological parameters like Months of a year, Days of a month, Temperature, Pressure, Humidity, Wind Speed and Latitude, Longitude, Elevation of New Delhi, India for 2014. This model may be used for the estimation of Global Solar Irradiance for other stations also.

## Conflict of Interest

We declare that there is no conflict of interest of this article.

## Acknowledgment

We are highly thankful to National Solar Radiation Database (NSRDB) for enabling the researchers to access the Solar Radiation data freely from its website.

## References

[1] K. Kapoor, K. K. Pandey, A. K. Jain, & A. Nandan, "Evolution of solar energy in India: A review", *Renewable and Sustainable Energy Reviews*, **40**, 475–487, 2014. <https://doi.org/10.1016/j.rser.2014.07.118>

[2] A. B. Jemaa, S. Rafa, N. Essounbouli, A. Hamzaoui, F. Hnaïen, & F. Yalaoui, "Estimation of Global Solar Radiation Using Three Simple Methods", *Energy Procedia*, **42**, 406–415, 2013. <https://doi.org/10.1016/j.egypro.2013.11.041>

[3] S. A. Kalogirou, "Artificial Neural Networks and Genetic Algorithms for the Modeling, Simulation, and Performance Prediction of Solar Energy Systems", *Green Energy and Technology*, 225–245, 2013. [https://doi.org/10.1007/978-1-4471-5143-2\\_11](https://doi.org/10.1007/978-1-4471-5143-2_11)

[4] D. Pandey, A. Choudhary, "A Review of Semiconductor Solar PV Cell and Development of Solar Radiation Estimation Models", *Journal of Semiconductor Devices and Circuits*, **5**(1):20-6, 2018.

[5] T. Khatib, A. Mohamed, & K. Sopian, "A review of solar energy modeling techniques", *Renewable and Sustainable Energy Reviews*, **16**(5), 2864–2869, 2012. <https://doi.org/10.1016/j.rser.2012.01.064>

[6] J. Zhang, L. Zhao, S. Deng, W. Xu, & Y. Zhang, "A critical review of the models used to estimate solar radiation", *Renewable and Sustainable Energy Reviews*, **70**, 314–329, 2017. <https://doi.org/10.1016/j.rser.2016.11.124>

[7] F. Besharat, A. A. Dehghan, & A. R. Faghih, "Empirical models for estimating global solar radiation: A review and case study", *Renewable and Sustainable Energy Reviews*, **21**, 798–821, 2013. <https://doi.org/10.1016/j.rser.2012.12.043>

[8] A. Angstrom, "Solar and terrestrial radiation. Report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation", *Quarterly Journal of the Royal Meteorological Society*, **50**(210), 121–126, 2007. <https://doi.org/10.1002/qj.49705021008>

[9] J. A. Prescott, "Evaporation from a water surface in relation to solar radiation", *Trans. Roy. Soc. S. Aust.*, **46**:114-8, 1940.

[10] H. Ögelman, A. Ecevit, & E. Tasdemiroğlu, "A new method for estimating solar radiation from bright sunshine data", *Solar Energy*, **33**(6), 619–625, 1984. [https://doi.org/10.1016/0038-092x\(84\)90018-5](https://doi.org/10.1016/0038-092x(84)90018-5)

[11] V. Bahel, H. Bakhsh, & R. Srinivasan, "A correlation for estimation of global solar radiation", *Energy*, **12**(2), 131–135, 1987. [https://doi.org/10.1016/0360-5442\(87\)90117-4](https://doi.org/10.1016/0360-5442(87)90117-4)

[12] J. Almorox, & C. Hontoria, "Global solar radiation estimation using sunshine duration in Spain", *Energy Conversion and Management*, **45**(9-10), 1529–1535, 2004. <https://doi.org/10.1016/j.enconman.2003.08.022>

[13] D. B. Ampratwum, & A. S. S. Dorvlo, "Estimation of solar radiation from the number of sunshine hours", *Applied Energy*, **63**(3), 161–167, 1999. [https://doi.org/10.1016/s0306-2619\(99\)00025-2](https://doi.org/10.1016/s0306-2619(99)00025-2)

[14] A. Manzano, M. L. Martín, F. Valero, C. Armenta, "A single method to estimate the daily global solar radiation from monthly data", *Atmospheric Research*, **166**, 70–82, 2015. <https://doi.org/10.1016/j.atmosres.2015.06.017>

[15] K. Mohammadi, S. Shamshirband, M. H. Anisi, K. A. Alam, & D. Petković, "Support vector regression based prediction of global solar radiation on a horizontal surface", *Energy Conversion and Management*, **91**, 433–441, 2015. <https://doi.org/10.1016/j.enconman.2014.12.015>

[16] A. Teke, & H. B. Yildirim, "Estimating the monthly global solar radiation for Eastern Mediterranean Region", *Energy Conversion and Management*, **87**, 628–635, 2014. <https://doi.org/10.1016/j.enconman.2014.07.052>

[17] M. Chelbi, Y. Gagnon, & J. Waewsak, J. "Solar radiation mapping using sunshine duration-based models and interpolation techniques: Application to Tunisia", *Energy Conversion and Management*, **101**, 203–215, 2015. <https://doi.org/10.1016/j.enconman.2015.04.052>

[18] K. H. Lee, "Improving the correlation between incoming solar radiation and sunshine hour using DTR", *International Journal of Climatology*, **35**(3), 361–374, 2014. <https://doi.org/10.1002/joc.3983>

[19] M. H. Saffaripour, M. A. Mehrabian, & H. Bazargan, "Predicting solar radiation fluxes for solar energy system applications", *International Journal of Environmental Science and Technology*, **10**(4), 761–768, 2013. <https://doi.org/10.1007/s13762-013-0179-2>

[20] K. Bakirci, "Prediction of global solar radiation and comparison with satellite data", *Journal of Atmospheric and Solar-Terrestrial Physics*, 152-153, 41–49, 2017. <https://doi.org/10.1016/j.jastp.2016.12.002>

[21] J. Liu, J. Liu, H. W. Linderholm, D. Chen, Q. Yu, D. Wu, & S. Haginoya, "Observation and calculation of the solar radiation on the Tibetan Plateau", *Energy Conversion and Management*, **57**, 23–32, 2012. <https://doi.org/10.1016/j.enconman.2011.12.007>

[22] J. L. Chen, & G. S. Li, "Estimation of monthly average daily solar radiation from measured meteorological data in Yangtze River Basin in China", *International Journal of Climatology*, **33**(2), 487–498, 2012. <https://doi.org/10.1002/joc.3442>

[23] J. Fan, X. Wang, L. Wu, F. Zhang, H. Bai, X. Lu, & Y. Xiang, Y., "New combined models for estimating daily global solar radiation based on sunshine duration in humid regions: A case study in South China", *Energy Conversion and Management*, **156**, 618–625, 2018. <https://doi.org/10.1016/j.enconman.2017.11.085>

[24] G. N. Okonkwo, A. O. Nwokoye, "Estimating global solar radiation from temperature data in Minna location", *European Scientific Journal*, **10**(15), 2014.

[25] H. Khorasanizadeh, K. Mohammadi, & M. Jalilvand, "A statistical comparative study to demonstrate the merit of day of the year-based models for estimation of horizontal global solar radiation", *Energy Conversion and Management*, **87**, 37–47, 2014. <https://doi.org/10.1016/j.enconman.2014.06.086>

[26] S. S. Sharifi, V. Rezaverdinejad, & V. Nourani, "Estimation of daily global solar radiation using wavelet regression, ANN, GEP and empirical models: A comparative study of selected temperature-based approaches", *Journal of Atmospheric and Solar-Terrestrial Physics*, **149**, 131–145, 2016. <https://doi.org/10.1016/j.jastp.2016.10.008>

[27] H. Hargreaves George, & A. Samani Zohrab, "Reference Crop Evapotranspiration from Temperature", *Applied Engineering in Agriculture*, **1**(2), 96–99, 1985. <https://doi.org/10.13031/2013.26773>

[28] M. Benganem, & A. Mellit, "A simplified calibrated model for estimating daily global solar radiation in Madinah, Saudi Arabia", *Theoretical and Applied Climatology*, **115**(1-2), 197–205, 2013. <https://doi.org/10.1007/s00704-013-0884-2>

[29] R. Chen, E. Kang, S. Lu, J. Ji. X. Yang, Z. Zhang, & J. Zhang, "New methods to estimate global radiation based on meteorological data in China", *Energy Conversion and Management*, **47**(18-19), 2991–2998, 2006. <https://doi.org/10.1016/j.enconman.2006.03.025>

[30] K. L. Bristow, & G. S. Campbell, "On the relationship between incoming solar radiation and daily maximum and minimum temperature", *Agricultural and Forest Meteorology*, **31**(2), 159–166, 1984. [https://doi.org/10.1016/0168-1923\(84\)90017-0](https://doi.org/10.1016/0168-1923(84)90017-0)

[31] J. Fan, B. Chen, L. Wu, F. Zhang, X. Lu, & Y. Xiang, Y., "Evaluation and development of temperature-based empirical models for estimating daily global solar radiation in humid regions", *Energy*, **144**, 903–914, 2018. <https://doi.org/10.1016/j.energy.2017.12.091>

- [32] J. L. Chen, L. He, Q. Chen, M. Q. Lv, H. L. Zhu, Z. F. Wen, & S. J. Wu, "Study of monthly mean daily diffuse and direct beam radiation estimation with MODIS atmospheric product", *Renewable Energy*, **132**, 221–232, 2019. <https://doi.org/10.1016/j.renene.2018.07.151>
- [33] M. Despotovic, V. Nedic, D. Despotovic, & S. Cvetanovic, "Review and statistical analysis of different global solar radiation sunshine models", *Renewable and Sustainable Energy Reviews*, **52**, 1869–1880, 2015. <https://doi.org/10.1016/j.rser.2015.08.035>
- [34] F. Besharat, A. A. Dehghan, & A. R. Faghieh, "Empirical models for estimating global solar radiation: A review and case study", *Renewable and Sustainable Energy Reviews*, **21**, 798–821, 2013. <https://doi.org/10.1016/j.rser.2012.12.043>
- [35] N. Chukwujindu Samuel, "A comprehensive review of empirical models for estimating global solar radiation in Africa", *Renewable and Sustainable Energy Reviews*, **78**, 955–995, 2017. <https://doi.org/10.1016/j.rser.2017.04.101>
- [36] W. Yao, Z. Li, Y. Wang, F. Jiang, & L. Hu, "Evaluation of global solar radiation models for Shanghai, China", *Energy Conversion and Management*, **84**, 597–612, 2014. <https://doi.org/10.1016/j.enconman.2014.04.017>
- [37] M. H. Sonmete, C. Ertekin, H. O. Menges, H. Haciseferoğullari, & F. Evrendilek, "Assessing monthly average solar radiation models: a comparative case study in Turkey", *Environmental Monitoring and Assessment*, **175**(1-4), 251–277, 2010. <https://doi.org/10.1007/s10661-010-1510-8>
- [38] A. Choudhary, D. Pandey, and A. Kumar, "A Review of Various Techniques for Solar Radiation Estimation," 3<sup>rd</sup> International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE), NOIDA, India, 169-174, 2019. <https://doi.org/10.1109/RDCAPE47089.2019.8979001>
- [39] A. Choudhary, D. Pandey, and S. Bhardwaj, "Artificial Neural Network based Solar Radiation Estimation: A Case Study of Indian Cities", *International Journal on Emerging Technologies*, **11**(4): 257–262, 2020.

## Customer Satisfaction Recognition Based on Facial Expression and Machine Learning Techniques

Moulay Smail Bouzakraoui<sup>\*1</sup>, Abdelalim Sadiq<sup>1</sup>, Abdessamad Youssfi Alaoui<sup>2</sup>

<sup>1</sup>Department of Informatics, Faculty of sciences, Ibn Tofail University, Kenitra, 14000, Morocco

<sup>2</sup>Laboratory of Advanced Digital Enterprise Modeling and Information Retrieval, National School of Computer Science and Systems Analysis, Mohammed V University, Rabat, 10000, Morocco

### ARTICLE INFO

Article history:

Received: 14 June, 2020

Accepted: 21 July, 2020

Online: 19 August, 2020

Keywords:

Facial expression

customer emotion

Emotion recognition

### ABSTRACT

Nowadays, Customer satisfaction is important for businesses and organizations. Manual methods exist, namely surveys and distributing questionnaires to customers. However, marketers and businesses are looking for quick ways to get effective and efficient feedback results for their potential customers. In this paper, we propose a new method for facial emotion detection to recognize customer's satisfaction using machine-learning techniques. We used a facial landmark point; we extract geometric features from customer's emotional faces using distances from landmarks points. Indeed, we used distances between the neutral side and the negative or positive feedback. After that, we classified these distances by using different classifier, namely Support Vector Machine (SVM), KNN, Random Forest, Adaboost, and Decision Tree. To assess our method, we verified our algorithm by using JAFFE datasets. The proposed method reveals 98,66% as accuracy for the most performance SVM classifier.

## 1. Introduction

This paper is an extension of work originally presented in world conference on complex systems (WCCS'19) [1]. We decide to develop this work because, we observe in the last decade, business primarily has an important impact based on the way, how customers observe their products and services [2]. Customer satisfaction can be measured using manual methods such as the satisfaction survey, the interview, and the focus group. These methods are not efficient and effective to the cost, time and reliability of the data. Facial expressions are used to communicate non-verbally. They are a special way to express our emotions and appreciations. In the context of customer satisfaction, a negative feedback emotion is often related to a lower perceived quality of service. The facial expression contributes 55% to communicate with a speech. More than that, 70-95% of negative feedback can be understood verbally [3]. Companies have long been interested in understanding the purchasing decision-making process of consumers [4]. In this work, we aim at detecting the positive or negative emotions customers from the analysis of facial expressions. In this way, this type of information is useful. For example, we can calculate the statistics about the products as well

as the replacement of their exposures. Appreciated products must be brought to the fore, while negatively appreciated products must be replaced or otherwise requested. For this reason, we are so existed to propose a new method for facial emotion detection to recognize customers' satisfaction using facial expression and machine learning techniques.

The remainder of this paper is structured as follows: Firstly, we offer an outline of the related work. Secondly, we introduce our approach that contains a system for customer satisfaction using facial expression and machine learning techniques. We reserved a separate section to the experimental results. In the last section, we sum up our work.

## 2. Related work

Currently in the field of marketing and advertising, facial recognition is used to study consumers' emotions in two forms: positive and negative emotions. Charles Darwin was the first to provide robust basis for emotions, representing their significance, their usefulness and communication. Already, many of facial expressions of emotions have an adaptive significance that serve to communicate something, most emotions are expressed alike on the human face nerveless of culture or race [5], his work "Expression of emotions in humans and animals", remains a

<sup>\*</sup>Corresponding Author: Moulay Smail Bouzakraoui, Department of Informatics, faculty of sciences, university Ibn Tofail, Kenitra, Morocco, bouzakraoui.smail14@gmail.com

reference for many scientists. In [6], authors proposed method to classify muscle movement to code the facial expression. The facial movements are determined through action units (AU). Each AU is based on the affinity of the muscles. This method, used mostly in classifying facial movement named Facial Action Coding

System (FACS). We can define six universal basic emotions: happy, sadness, anger, surprise, fear and disgust. Various researchers have supported the universality of the expression of emotions. Generally, in reaction to similar stimuli, people prompt similar expressions with local variations [7]. Frequently, we used two methods: motion-based method and deformation of face. The change of face is taken into account for the motion-based method [8, 9]. On the contrary, we take into account a neutral image and another image for the deformation-based method [10, 11]. In [10], authors use FACS (Face Action Code System) as features to classify emotions. Their model consists to count images, which classified correctly and weighted by the system. They found that, 2% of images (6 image) are failed completely in tracking. They also reported 91% as average correct recognition ratio. Otherwise, in [11] a model which able to use also speech content is developed. They create a system to compute the Human Computer Interaction (HCI). In order to test their model, they utilize 38 subjects of affect recognition approach with 11 HCI-related affect states. Their model a gave 90% for the bimodal fusion as an average recognition accuracy.

In [12] feature extraction contains geometric and appearance-based models. The first model involves feature extraction about eye, mouth, nose, eyebrow... However; the second model covers the specific part of the face. They also evaluate their model by employing SVM, ANN, KNN, HMM, Bayesian network, and sparse representation-based classification. The experiment results report that, sparse representation-based classification is the more performant classifier for their model.

Deep learning algorithms are also used to extract pattern from facial expressions. More specifically, we develop a convolutional neural network to detect features from faces, and to classify these features to different emotions. This type of network contains two parts. The first part contains convolutional layers, which are based on applying mathematic convolution operation. The result of this part is features extracted from faces. The second part contains feed-forward network. This part of network is able to classify features extracted to different facial expression. In [13, 14] the based results researches of this algorithm is proposed. For example, authors, in [13], apply KNN to classify emotions. After that, they find 76.7442% and 80.303% as accuracy in the JAFFE and CK+, respectively, which illustrate viability and success of their model. In [14], authors employ also Decision Tree and Multi-Layer Perceptron to classify emotions. However, they find that, Convolutional Neural Network report the greatest recognition accuracy. An Affdex by Affectiva [15] is one of the most implements for visual emotional analysis. Which permits giving the emotional trend of a subject, through detection emotions [16], and the Microsoft Cognitive Services based on the Azure platform. These two studies are able to perceive age, sex

and ethnicity and depend on Convolutional and/or Recurrent Neural Network [17]. In addition, they find that, the proposed model is yields performance increases for facial expression recognition using CK+ dataset. A multimodal affect recognition system was developed to determine if the customer exhibits negative affect such as being unhappy, disgusted, frustrated, angry or positive affect such as happy, satisfied and content with the product being offered [18]. Furthermore, authors apply Support Vector Machine an emotion template to evaluate their algorithm. They test the real-time performance, which evaluate the feedback assessment in order to compute accuracy and viability for multimodal recognition. To appreciate customer satisfaction special facial emotion namely "happy", "surprised" and "neutral" is proposed in [19]. They evaluate their model by using datasets as Radboud Faces and Cohn-Kanade (CK+). Output results report that their proposed algorithm a gave an important recognition accuracy utilizing Action Unit (AU) features, Support Vector Machine (SVM) and K-Nearest Neighbors (KNN). In the last years, eye-tracking system has been developed to analyze the research of customer [20]. Companies invest intensely within the advancement and publicizing of products, the return on investment (ROI) must be justified by the organization. Hence, the specific observing of the fascination of a product by the customer amid its promotion, and how to develop their promotion campaigns. Eye tracking is very promising in navigation advertising. In [21], authors developed an algorithm aims to compute sales assistant, which depends specially on emotion detection. Otherwise, [22] present a study, which contains products features impact the costumer's emotional. Furthermore, they used 21 emotional categories of products.

### 3. Our Approach

In this section, we present an algorithm to recognize the emotions of customers towards a given product. In order to analyze the facial expressions of a client to derive satisfaction from the product offered, our algorithm is based on three essential steps. The first is to acquire the image of the face and the second allows extracting the expressions in the form of geometric features. To obtain these features, our algorithm transforms the input image into geometric primitives such as points and curves. These primitives are located via distinctive patterns such as the eyes, mouth, nose and chin by measuring their relative position, their width and other parameters such as the distance between the nose, eyes, mouth. Our contribution is therefore to propose a method for selecting the optimal distances that can make the difference between facial expression linked to emotion and customer satisfaction. In the last step, these different distances are used to classify a client's emotions towards a product, into three types; positive, negative or neutral appreciation. These results will be analyzed and provided to the company to understand how customers perceive the product.

Otherwise, the pipeline of our algorithm begin by detecting the face into an input image. After that, we compute the key point of the face by applying a mask, in order to compute the distances between key points. We aim to measure the variation of face's shape in order to classify appreciation of customer's toward a product. For this raison, our algorithm is based on computing



face's key points to calculate distances between those landmarks. Finally, we employ a set of classifier to evaluate our algorithm

3.1. Features extraction (Geometric features)

In order to represent the emotions of a face, this step consists in locating the facial landmarks (for example the mouth, the eyes, the eyebrows, the nose and the chin). The model presented in figure (1) allows us to locate the face marks. This model has a drawback; the deformation of these marks modifies the facial expressions.

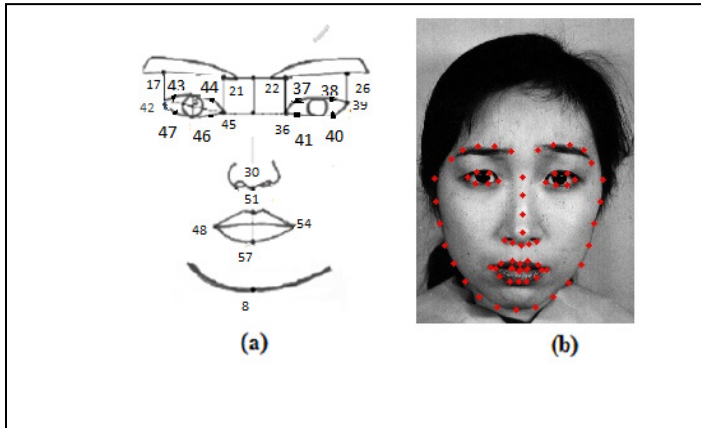


Figure1: (a) Facial model depicting the position of the chosen landmark point, (b) Landmark point on the test subject face

However, we used twenty-two facial landmarks points to calculate eighteen distances. Therefore, a vector of eighteen values will represent each face. The selected distance are:

**For the eyes:** six features

Left eye width, Right eye width, Right eye height, left eye height, Right eye height, left eye height

**For the mouth:** two features

Mouth Width, mouth height

**For the chin:** two features

Chin to nose distance, Chin to mouth distance

**For the eyebrow:** six features

Left eyebrow width  
 Right eyebrow width  
 Outer left eyebrow distance to eye  
 Inner left eyebrow distance to eye  
 Outer right eyebrow distance to eye  
 Inner right eyebrow distance to eye

**For the nose:** two features

Nose distance to left eye  
 Nose distance to right eye

Table1: Definitions of Geometric features

Distance	Description	Formula	Distance	Description	Formula
D1	Mouth Width	d(X48,X54)	D12	Right eye height	d(X38,X40)
D2	Mouth height	d(X51,X57)	D13	Outer left eyebrow distance to eye	d(X17,X42)
D3	Chin to mouth distance	d(X8,X57)	D14	Outer right eyebrow distance to eye	d(X26,X39)
D4	Chin to nose distance	d(X8,X30)	D15	Inner right eyebrow distance to eye	d(X22,X36)
D5	Left eyebrow width	d(X17,X21)	D16	Inner left eyebrow distance to eye	d(x21,x45)
D6	Right eyebrow width	d(X22,X26)	D17	Nose distance to left eye	d(X45,X30)
D7	Left eye width	d(X42,X45)	D18	Nose distance to right eye	d(X36,X30)
D8	Right eye width	d(X36,X39)			
D9	Left eye height	d(X43,X47)			
D10	Left eye height	d(X44,X46)			
D11	Right eye height	d(X37,X41)			

### 3.2. Training and classification

Using the most used algorithms in machine learning, namely: SVM, RANDOM FOREST, Decision Tree KNN, and Adaboost classifiers, the geometric features presented in the previous sections will be classified into three classes according to customer’s satisfaction.

- **SUPPORT VECTOR MACHINE (SVM):** SVM is a discriminative algorithm, which used to develop a supervised model. It used to classify or to predict regression [23]. In this algorithm, we aim to find a separator, which can classify neural, positive, and negative emotion class. This separator is baptized hyperplane. To find a very performant hyperplane, we tested kernel methods. We found that RBF Kernel is the best kernel for our model. We utilized xi, as a vector support, and  $\sigma$ , as positive float, which is, specify a priori, to compute RBF Kernel. Alternatively,  $RBFK(x_i, x) = \exp\left(\frac{-\|x_i-x\|^2}{2\sigma^2}\right)$ . We used also S as a set of vectors support, which contains xi,b as bias, and Lagrangian coefficient  $\alpha_i$ , is a vector support, to compute hyperplane equation:

$$H(x) = \sum_{i \in S} \alpha_i y_i K(x_i, x) + b. \tag{1}$$

- **RANDOM FOREST:** Random forest [24] is an algorithm constructed by combination between trees. This algorithm used decision trees concept to classify. In decision Tree, we calculate Gini index or Gain ratio to construct the tree. In Random Forest, we compute just Gini index to find the most popular attributes founded by Decision Trees [25]. To compute Gini index, we used the probability, which link between class k and selected case  $p(q/p)$ . Or, Gini index can written as follow:

$$Gini(p) = 1 - \sum_{q=1}^c p(q/p)^2. \tag{2}$$

- **K NEAREST NEIGHBORS (KNN):** KNN [26] used k closest neighbors to classify. It can be used for classification or clustering. The concept of this algorithm is finding the class, which have the majority of k closest neighbors. Precisely, we compute distances between examples and we take the examples, which have the smaller distances.
- **ADABOOST (ADAPTIVE BOOSTING):** This classier is baptized also Adaptive Boosting. It is one of the classifiers, which based on the use of boosting concept. It combines between multiple classifiers. It is an iterative algorithm. Adaboost consists to set up a solid classifier by combining a various inadequately performing classifiers in order to product high performant classifier. Any classifier accepts weight can be used as base classifier of Adaboost classifier. Adaboost should meet two conditions:
  1. The training should be done by using examples weighted.
  2. In every iteration, it attempts to give a very good fit to these examples by minimizing the value of error.

### 3.3. Experimental Results

In this work, the Jaffe dataset is used to test the performance of our method. The database contains 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) posed by 10 Japanese female models. 60 Japanese subjects have rated each

image on 6 emotion adjectives. The database was planned and assembled by Michael Lyons, Miyuki Kamachi, and Jiro Gyoba. To use this dataset for our algorithm; we have classified its images into three classes. The first one contains images for satisfied persons, the second is for non-satisfied persons and the third class is for neural persons. Indeed, each face is represented in the form of a vector of 68 elements. Each element is the coordinate of a key point. All key points are calculated using a mask. These operations are applied to all Jaffe images. Next, the distances, shown by Table 1, are calculated. Therefore, the input of our classifiers SVM, RANDOM FOREST, KNN, decision tree and Adaboost classifiers will be vectors, which contain eighteen distances. Alternatively, a distance vector will represent each face. On the other hand, we computed confusion matrix of each classifier. We had also calculated accuracy, sensitivity, and specificity of each classifier. The following equations present the formula of specificity, sensitivity, and accuracy:

- Specificity is the ratio between non-satisfied and neutral faces, which are not classified as a satisfied face and the number total of non-satisfied and neutral faces.

$$Specificity = \frac{\text{non - satisfied \& neutral faces correctly classified}}{\text{All non - satisfied \& neutral faces}}. \tag{3}$$

- Sensitivity is the ratio between the number of satisfied faces, which are classified correctly, and the number of all satisfied faces (i.e. the satisfied faces, which are classified correctly, satisfied faces, which are classified neutral faces, and satisfied faces, which are classified as non-satisfied faces).

$$Sensitivity = \frac{\text{satisfied faces correctly classified}}{\text{Total of satisfied faces}}. \tag{4}$$

- Accuracy is the ratio between the number of satisfied, neutral, and non-satisfied faces are correctly classified and the number of all faces.

$$Accuracy = \frac{\text{Faces are correctly classified}}{\text{Total of all faces}}. \tag{5}$$

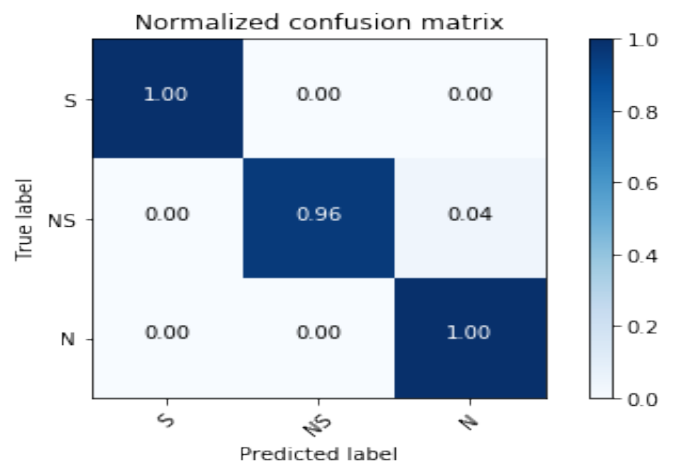


Figure. 2. Confusion matrices of our algorithm by applying SVM a JAFFE dataset

We also calculated the ROC curve (Receiver Operating Characteristics curve). ROC curve determinate the probability of distinguishing capability of our algorithm, i.e. how much our algorithm is able to separate between satisfied, neutral, and non-satisfied faces.

Table- 2: The sensitivity, specificity, and accuracy of applying our algorithm by using SVM, KNN, Decision Tree, Random Forest, and AdaBoost, and by utilizing JAFFE datasets

Dataset		Accuracy	Sensitivity	Specificity
JAFFE	SVM(RBF Kernel)	98.66%	100%	98%
	KNN	88.87%	90%	75.75%
	Decision Tree	98.66%	100%	98%
	Random Forest	77.76%	97%	62.77%
	AdaBoost	98.66%	100%	98%

We found that our algorithm detected 97% of satisfied faces correctly. In addition, our method extracted 91% non-satisfied faces correctly. Finally, it detected 100% of neutral faces correctly.

After computing confusion matrices of our algorithm, we computed the sensitivity, specificity, and sensitivity of our algorithm. We found the results showing into table 2. We calculate these statistic criteria for all classifier. We calculate the sensitivity of applying our algorithm by using SVM with RBF kernel, KNN, AdaBOOST, and Random Forest. We found that the most performant classifier is SVM with RBF Kernel. Table 2 shows the results of each classifier by using JAFFE datasets. Our algorithm report that, geometrics features a gave a great model to measure satisfaction customers. Our algorithm consists to measure the variation of face’s shape, by computing distances between face landmarks. Since, we based our classification on using the different variation between satisfied, neutral, and non-satisfied faces. For this reason, our algorithm a gave a great model to detect the satisfaction of customers by using emotions of customers faces.

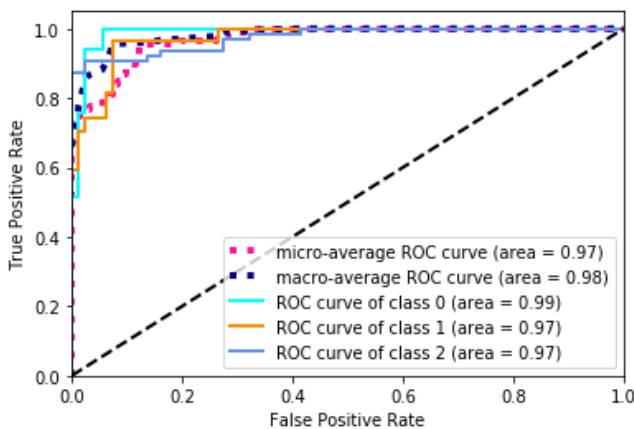


Figure. 3.ROC curve computed by applying our algorithm using SVM and JAFFE dataset.

#### 4. Conclusion

One of the greatest challenges for marketing era is to ensure customer satisfaction. We have proposed a new method based on the classification of various facial expressions, to assess customer satisfaction using different classifiers: SVM, RANDOM FOREST, KNN and AdaBOOST. Distances between facial landmarks points represent the geometrics features. A vector of eighteen values represents customer’s face expression. We tested our method on the JAFFE database and we obtained a performants result. In the future work, we will develop a multimodal algorithm, which contains speech and motion recognition.

#### References

- [1] M.S. Bouzakraoui, S.Abdelalim, Y.A. Abdssamed, "Appreciation of Customer Satisfaction Through Analysis Facial Expressions and Emotions Recognition" in 2019 4th Edition of World conference on Complex Systems, Ouarzazate, Morocco, 2019. <https://doi.org/10.1109/ICoCS.2019.8930761>.
- [2] F. Conejo, "The Hidden Wealth of Customers: Realizing the Untapped Value of your Most Important Asset", Journal of Consumer Marketing, **30**(5), 463-464, 2013. <https://doi.org/10.1108/JCM-05-2013-0570>.
- [3] A. Mehrabian, Communication without words, Psychology today, 1968.
- [4] S. M. C. Loureiro, "Consumer-Brand Relationship: Foundation and State-of-the-Art", Customer-Centric Marketing Strategies: Tools for Building Organizational Performance, 21, 2013, <https://doi.org/10.4018/978-1-4666-2524-2.ch020>
- [5] C. Darwin, The expression of the emotions in man and animals, D. Appleton and Company, London, 1913
- [6] J. Hamm, C.G. Kohler, R.C. Gur, R. Verma, "Automated Facial Action Coding System for dynamic analysis of facial expressions in neuropsychiatric disorders", Journal of neuroscience methods, **200**(2), 237-256, 2011. <https://doi.org/10.1016/j.jneumeth.2011.06.023>
- [7] A. A. Marsh, H. A. Effenbein, N. Ambady - "Nonverbal "accents" cultural differences in facial expressions of emotion", Psychological Science, **14**(4), 373-376, 2003. <https://doi.org/10.1111/1467-9280.24461>
- [8] M. Nishiyama, H. Kawashima, T. Hirayama, "Facial Expression representation based on timing structures in faces", in 2005 workshop on Analysis Modeling of faces and Gestures, Beijing, China, 2005. [https://doi.org/10.1007/11564386\\_12](https://doi.org/10.1007/11564386_12)
- [9] K.K. Lee, Human expression and intention via motion analysis: Learning, recognition and system implementation, The Chinese University of Hong Kong Hong, 2004.
- [10] M. Pantic, L.J.M. Rothkrantz, "An expert system for multiple emotional classification of facial expressions" in 1999 IEEE International Conference on Tools with Artificial Intelligence, Chicago, IL, USA, 1999. <https://doi.org/10.1109/TAL.1999.809775>
- [11] Z. Zeng, J.T., M. Liu, T. Zhang, N. Rizzoto, Z. Zhang, "Bimodal HCI-related Affect Recognition" in 2004 International Conference Multimodal Interfaces, State College, PA, USA, 2004. <https://doi.org/10.1145/1027933.1027958>
- [12] X. Zhao, S. Zhang, "A review on facial expression recognition: feature extraction and classification" IETE Technical Review, **33**(5), 505-517, 2016. <https://doi.org/10.1080/02564602.2015.1117403>
- [13] K. Shan, J. Guo, W. You, D. Lu, R. Bie, "Automatic Facial Expression Recognition Based on a Deep Convolutional-Neural Network Structure" in 2017 IEEE 15th Int. Conf. Softw. Eng. Res. Manag. Appl. 123–128, London, United Kingdom, 2017. <https://doi.org/10.1109/SERA.2017.7965717>
- [14] T.A. Rashid, "Convolutional neural networks based method for improving facial expression recognition" in 2016 Intelligent Systems Technologies and Applications, Jaipur, India, 2016. <https://doi.org/10.1007/978-3-319-47952-1>
- [15] S. Ceccacci, A. Generosi, L. Giraldi, M. Mengoni, "Tool to Make Shopping Experience Responsive to Customer Emotions" International Journal of Automation Technology, **12**(3), 319-326, 2018. <https://doi.org/10.20965/ijat.2018.p0319>
- [16] D. Matsumoto, M. Assar, "The effects of language on judgments of universal facial expressions of emotion" Journal of Nonverbal Behavior, **16**, 85-99, 1992. <https://doi.org/10.1007/BF00990324>
- [17] R. Cui, M. Liu, M. Liu, "Facial expression recognition based on ensemble of multiple CNNs" in 2016 Chinese Conf. of Biometric Recognition, Chengdu, China, 2016. [https://doi.org/10.1007/978-3-319-46654-5\\_56](https://doi.org/10.1007/978-3-319-46654-5_56)
- [18] L. Huang, T. Chuan-Hoo, W. Ke, K. Wei, "Comprehension and Assessment of Product Reviews: A Review-Product Congruity Proposition" Journal of

Management Information Systems, **30**(3), 311-343, 2014.  
<https://doi.org/10.2753/MIS0742-1222300311>

- [19] M. Slim, R. Kachouri, A. Atitallah, "Customer satisfaction measuring based on the most significant facial emotion", in 2018 15th IEEE International Multi-Conference on Systems, Signals & Devices, Hammamet, Tunisia, 2018. <https://doi.org/10.1109/SSD.2018.8570636>
- [20] M. Wedel, R. Pieters - Review of marketing research, 2008.
- [21] G. S. Shergill, A. Sarrafzadeh, O. Diegel, A. Shekar, "Computerized sales assistants: the application of computer technology to measure consumer interest – a conceptual framework", Journal of Electronic Commerce Research, 9(2), 176 – 191, 2008. DOI:10.1.1.491.5130
- [22] P. M. A. Desmet, P. Hekkert, "The basis of product emotions", Pleasure with products: Beyond usability. ed. WS Green; PW Jordan, 61-67, 2002. DOI:10.1201/9780203302279.ch4
- [23] C. Cortes, V. Vapnik, "Support-Vector Networks", Machine Learning, 20, 273-297, 1995. <https://doi.org/10.1023/A:1022627411411>
- [24] B.Leo, "Random Forests", Machine Learning, 45, 5-32, 2001. <https://doi.org/10.1023/A:1010933404324>
- [25] Raileanu, L. Elena, K. Stoffel, "Theoretical Comparison between the Gini Index and Information Gain Criteria," Annals of Mathematics and Artificial Intelligence 41, 77-93, 2004. <https://doi.org/10.1023/B:AMAI.0000018580.96245.c6>
- [26] G. Laencina, J. Pedro, "K nearest neighbours with mutual information for simultaneous classification and missing data imputation" Neurocomputing, 72(7-9), 1483-1493, 2009. <https://doi.org/10.1016/j.neucom.2008.11.026>



# Real-Time Traffic Sign Detection and Recognition System for Assistive Driving

Adonis Santos<sup>\*,1,2</sup>, Patricia Angela Abu<sup>3</sup>, Carlos Oppus<sup>1</sup>, Rosula Reyes<sup>1</sup>

<sup>1</sup>Electronics, Computer, and Communications Engineering Dept., Ateneo de Manila University, Quezon City, 1108, Philippines

<sup>2</sup>College of Engineering, First Asia Institute of Technology and Humanities, Tanauan City, 4232, Philippines

<sup>3</sup>Department of Information Systems and Computer Science, Ateneo de Manila University, Quezon City, 1108, Philippines

---

## ARTICLE INFO

Article history:

Received: 12 June, 2020

Accepted: 04 August, 2020

Online: 25 August, 2020

---

Keywords:

traffic sign detection and recognition

embedded system

computer vision

machine learning

convolutional neural network

---

## ABSTRACT

Road traffic accidents are primarily caused by drivers error. Safer roads infrastructure and facilities like traffic signs and signals are built to aid drivers on the road. But several factors affect the awareness of drivers to traffic signs including visual complexity, environmental condition, and poor drivers education. This led to the development of different ADAs like TSDR that enhances vehicle system. More complex algorithms are implemented for improvement but this affects the performance of a real-time system. This study implements a real-time traffic sign detection and recognition system with voice alert using Python. It aims to establish the proper trade-off between accuracy and speed in the design of the system. Four pre-processing and object detection methods in different color spaces are evaluated for efficient, accurate, and fast segmentation of the region of interest. In the recognition phase, ten classification algorithms are implemented and evaluated to determine which will provide the best performance in both accuracy and processing speed for traffic sign recognition. This study has determined that Shadow and Highlight Invariant Method for the pre-processing and color segmentation stage provided the best trade-off between detection success rate (77.05%) and processing speed (31.2ms). Convolutional Neural Network for the recognition stage not only provided the best trade-off between classification accuracy (92.97%) and processing speed (7.81ms) but also has the best performance even with lesser number of training data. Embedded system implementation utilized Nvidia Jetson Nano with interface Waveshare IMX219-77 camera, Nvidia 7" LCD and generic speaker and programmed in Python with OpenCV, sci-kit learn and Pytorch libraries. It is capable of running at an adaptive frame rate from 8-12 frames per second with no detection and down to approximately 1 frame per second when there is traffic sign detected.

---

## 1 Introduction

The World Health Organization (WHO) in 2013 reported that road traffic accidents that result to loss of lives and damages to properties will continue to become a global challenge due to rapid motorization and insufficient action of national governments [1]. In the Philippines, traffic accidents are mainly due to driver's error [2] or violation [3] and this prompts government agencies to focus on regulation of driver by means of licensing and provide safer roads by building infrastructures like traffic lights and signs. Another solution includes the development of advanced driver-assistance systems (ADAs) that enhance vehicle systems to aid drivers on the

road. But this system faces its own challenge, for instance, traffic sign detection and recognition (TSDR) system is often exposed to different weather phenomenon and road conditions [4] that degrades their effectivity [5]. This can be solved by employing more complex algorithms but this has an adverse effect to processing speed that is a must for smart vehicle systems. This trade-off between system accuracy and speed is still the biggest challenge in the implementation of ATSDR systems.

This paper "Real-Time Traffic Sign Detection and Recognition System for Assistive Driving" is an extension of work originally presented by the authors in International Symposium on Multimedia and Communication Technology (ISMAT) 2019 entitled "Traffic

---

\*Corresponding Author: Adonis Santos, Ateneo de Manila University, adonis.santos@obf.ateneo.edu

Sign Detection and Recognition for Assistive Driving” [6]. The goal of this study is to help solve the problem of drivers neglect and lack of road education by implementing an ATSDR system that automatically detects and recognizes road traffic signs then provides information to the driver about the meaning of the signs. This work is the full implementation of a study that the previous work [6] is a part of. The differences between this work and the previous work include (1) gathering and utilization of new traffic road image dataset and road sign image datasets for training and testing the system, (2) evaluation of four pre-processing and detection methods for the detection stage, (3) addition of a Convolutional Neural Network (CNN) deep learning algorithm in the evaluation for the recognition stage aside from the original nine methods, (4) the detection and recognition methods are evaluated not only using the segmentation and classification success rate (accuracy) metric but also processing speed, (5) system-level optimization is done to improve accuracy by fine-tuning sub-processes and a negative class is added to reduce false detection, (6) the models are fully implemented not only in a personal computer but in an embedded system using microcomputer with interface camera, display and speaker, and (7) the system installed in a sedan car is tested in an actual road setup.

Traffic signs defined by the Department of Public Works and Highways to be recognized must be strategically positioned, clear and fully visible [7] and captured in good weather condition during daytime and will be limited only to six regulatory signs corresponding to speed limit, no overtaking and no turning shown in Figure 1. These are generally characterized by their red circular boundaries. The detection and recognition will only be limited to a single traffic sign in a single video frame. The capturing device will be installed to a sedan car with a target maximum speed of at least sixty (60) kilometers per hour. The system will be able to provide voice alert to the driver before the car completely passed by the traffic sign. Implementation of the system will make use of methods from previous studies using Python and open-source libraries.



Figure 1: Traffic signs to be recognized

Several implementations of traffic sign detection and recognition systems used different algorithms in pre-processing, detection and recognition phases. Pre-processing stage is necessary to suppress noise and improve image performance [8]. A study [6] showed the detection improvements in traffic sign recognition provided by the use of pre-processing and filtering methods by means of contrast

stretching, color normalization and image enhancement before getting the regions of interest. Other techniques are applied depending on the choice of segmentation method in the detection stage which can be color-based, shape-based or combination of both, though color-based is widely implemented in the detection of road signs [9]. Shape based segmentation may have problem when there are similar objects to the traffic signs. It can be quite computationally expensive, and by nature, sensitive to noise. Many shape detectors are slow in computing over large and complex images [10] which will become an issue in a real-time system. The complexity of color based lies in the usage of three intensity values depending on the color space instead of just working in gray level image as in shape based. But still, color segmentation is used more often as color usually provides enough information on the meaning of the traffic sign to be recognized than shape segmentation. Difficulties in color-based detection include lack of standard colors among countries, and various phenomenons and different imaging conditions that cause color variations. In these cases, traffic sign images may appear with varying illumination, blurred, faded, highlighted, noisy, and physically damaged [5] affecting the performance of the color segmentation process. A study [11] solved this problem by instead of working on absolute RGB values, relative RGB values are used for they are almost unchanged with various illumination circumstances. Another study [8] combined four color spaces namely RGB, HSV, YIQ, and XYZ based on histogram to provide more accurate segmentation. Color is segmented using K-means clustering and effective robust kernel based Fuzzy C-means clustering resulting to a more accurate segmentation but has increased computational complexity. In [10], color segmentation in RGB color space followed by shape segmentation using joint transform correlator (JTC) template matching had been implemented for a more efficient system. A study [12] solved the unreliability issue of RGB due to sensitivity to illumination variation by working on HSV color space along with shape based filtering through template matching. HSV is a non-linear transform of RGB where any color can be represented by hue and the depth or purity of color by saturation [13] whose values are not affected by light intensity [12]. In [14], three methods of segmentation based on improved HSL color space were implemented that provided independence between chromatic and achromatic components resulting to robustness in changes in external light conditions. Based on success rate in segmentation, thresholding using global mean of luminance is the best method followed by segmentation using region growing then by improved HSL normalization. Another study [15] solved the effect of poor lighting by RGB channel enhancement using histogram equalization followed by color constancy algorithm before segmentation in HSV color space. Color constancy is the ability to correctly determine the color of objects in view irrespective of the illuminant [16].

Several studies evaluated different segmentation algorithms based on segmentation success rate and processing time. A study [5] evaluated four color segmentation methods namely, Dynamic Threshold Algorithm, Modified de la Escaleras Algorithm, Fuzzy Color Segmentation Algorithm, and Shadow and Highlight Invariant Algorithm, in red color segmentation of traffic sign in different road conditions and distances. Fuzzy Color Segmentation has the best segmentation success rate followed by Shadow and Highlight Invariant. In processing time, Shadow and Highlight Invariant has

the best average time, Fuzzy Algorithm has the best standard deviation and the most stable. Using the dataset with traffic image taken at different distances and conditions, Shadow and Highlight Invariant has the best performance. Another study [17] evaluated shape-based and color-based segmentation methods for traffic sign recognition using Matlab. Metrics in evaluation include number of recognized signs, lost signs, falsely recognized signs and speed of recognition. Among the evaluated methods are RGB Normalized Thresholding, Hue and Saturation Thresholding, Hue and Saturation Color Enhancement Thresholding, Ohta Space Thresholding, Grey-Scale Edge Removal, Canny Edge Removal, and Color Edge Removal. Results showed that color space thresholding methods are the best methods in segmentation success rate and speed. RGB normalized method provided best recognition score among the algorithms evaluated. Edge detection methods had the worst execution speed and excessive false detection percentage. Findings showed that edge-detection methods may be used as a complement to other color-segmentation methods, but they cannot be used alone, and normalization, as in the RGB Normalized method or Ohta Space Thresholding, improves performance and represents a low-cost operation.

In recognition stage, the use of neural networks in classification of road signs provides considerable saving in computational requirement as compared to template matching [9]. Several studies implemented deep learning algorithms for traffic sign recognition due to their high accuracy with faster processing speed [18]-[20]. In [21], detection and classification of circular traffic signs used CNN after pre-processing and Hough Transform using TensorFlow using accuracy metric attaining 98.2%. In [22], 24 traffic sign classes were used in recognition using deep CNN that achieved 100% accuracy. Another study [23] used masked R-CNN in the detection (Region Proposal Network) and recognition (Fast R-CNN) of over 200 traffic sign categories where data augmentation is implemented and resulted to less than 3% error rate. The disadvantage of neural network lies on the training overhead and architecture of multi-layer networks cannot be adapted for online application. In order to distinguish among different classes for classification, specific features of the input data must be chosen. In [4], histogram of oriented gradient is used as feature descriptor for classification of traffic signs as it captures color and shape as one feature.

## 2 Methodology

Shown in Figure 2 is the methodology for the proposed real-time traffic sign detection and recognition system starting with image acquisition for training the detection and recognition models followed by testing. The system will be designed in Python using the OpenCV, sci-kit learn, and Pytorch libraries and will be implemented in a microcomputer with interfaced camera and speaker.

### 2.1 Data Acquisition

Input traffic data will be acquired by the system using a camera interfaced to the microcomputer. Videos taken will be processed per frame by the system. Some testing process will make use of traffic images to be stored in the microcomputer.

Four sets of data will be used in the implementation and testing of the system. For the detection phase, a set of local traffic images from Google Images containing actual road and traffic images with different traffic signs to be detected will be used to test the models. Another set of data is composed of road traffic images from dataset of Napier University containing actual road with traffic signs to be detected captured in different distances. The recognition phase model will utilize one set of data containing images cropped to fully view the standard traffic sign from actual Philippine roads. The last set of data is actual video footage from local road to test the full performance of the traffic sign detection and recognition system.

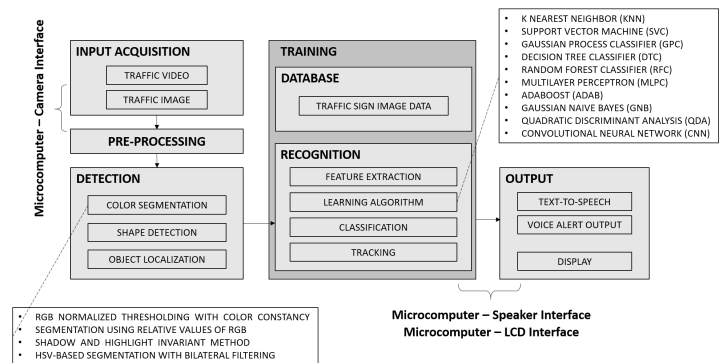


Figure 2: Overview of methodology

### 2.2 Pre-processing and Detection

The detection phase is composed of pre-processing, color-based segmentation, shape-based detection and object localization. This study will evaluate four different pre-processing and color-based segmentation methods that solved the problem of lighting variations affecting the detection phase. Methods to be evaluated that use different filtering techniques in different color spaces are RGB Normalized Thresholding with Color Constancy Algorithm, Segmentation using Relative Values of RGB, Shadow and Highlight Invariant Method, and HSV-based Segmentation with Bilateral Filtering.

#### 2.2.1 RGB Normalized Thresholding with Color Constancy (CCM)

Histogram equalization of separate RGB channels is done followed by color constancy algorithm to extract the true color of the image [15]. The algorithm computes local space average color using a parallel grid of processing elements used to shift the color of the input pixel in the direction of the gray vector. Target color segmentation is done in HSV color space [16].

#### 2.2.2 Segmentation using Relative Values of RGB (RRGB)

Segmentation that is based in three rules using the RGB color space. The pixel will be regarded as the target red color when (i) red component is greater than  $\alpha$ , (ii) difference between red and green components is between  $\beta_1$  and  $\beta_2$ , and (iii) the difference between the red and blue components is between  $\gamma_1$  and  $\gamma_2$ . The values of  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ ,  $\gamma_1$ , and  $\gamma_2$  are experimentally found to represent the red color perceived in different lighting conditions [11].

### 2.2.3 Shadow and Highlight Invariant Method (SHI)

Utilized the normalized HSV color space values. Segmentation happens in the hue image where hue is 255 if the pixel's hue is between 240 and 255, or hue is between 0 and 10. Hue will be set to 0 if pixel's saturation is less than 40, or pixel's value does not fall between 30 and 230. The hue image and seed image from the new hue is used, applying region growing algorithm to find target proper region [5].

### 2.2.4 HSV-based Segmentation with Bilateral Filtering (BFM)

Traffic images undergo pre-processing in the RGB color space before the detection stage. Bilateral filtering smoothens the image while preserving the edges by means of a nonlinear combination of nearby pixel values [24]. The weight assigned to each neighbor decreases with both the distance in the image plane and the distance on the intensity axis.

One of the four methods will be used for color-based segmentation to improve the efficiency and speed of detection of the next stage by reducing search areas that do not contain any possible traffic sign [9, 25] defined by pixels falling outside the target red color.

The next phase is the shape detection using Hough transform to detect the circular shape of the traffic sign. Hough transform has been recognized as a robust technique in curve detection that can detect objects even when polluted by noise [26]. It transforms a set of feature points in the image space into a set of accumulated votes in a parameter space. For each feature point, votes are accumulated in an accumulator array for all parameter combinations. Only one circle, by means of highest number of votes, will be detected as increasing the number of detection will also increase the time required to process multiple regions for the succeeding phases. The last stage in the detection phase is the object localization or the cropping of the candidate traffic sign from the traffic image to be used for the classification stage.

## 2.3 Feature Extraction

Histogram of Oriented Gradients (HOG) features are extracted from the candidate traffic sign images by the Python program to become the key feature for classification for the nine machine learning algorithms. The image from the dataset undergoes RGB to grayscale conversion and edge detection. The resulting image is then resized before using the built-in HOG extractor library in Python [6].

The steps above are important in ensuring optimum accuracy of the features when training the machine learning classifiers. HOG feature provides excellent performance relative to other existing feature sets [27].

## 2.4 Classifier

To classify the candidate traffic sign images, learning algorithms will be used. For this study, ten learning algorithms are implemented using sci-kit learn library [28] and Pytorch [29] in Python. The classifiers [30] to be evaluated based on accuracy and processing speed for traffic sign recognition are as follows.

### 2.4.1 K Nearest Neighbor (KNN)

KNN is a method where classification is computed from a simple majority vote of the nearest neighbors of each point [31]. Existing Python implementation of KNN uses model `KNeighborsClassifier(n_neighbors=1)` where a single neighbor is obtained for prediction.

### 2.4.2 Support Vector Machine (SVC)

An SVM is a discriminative classifier formally defined by a separating hyperplane. The algorithm outputs an optimal hyperplane using labeled training data which categorizes where the new examples are [32]. Implementation of the SVC will make use of the model `SVC(kernel="linear", C=0.025)` for linear kernel and with regularization parameter  $C$  of 0.025.

### 2.4.3 Gaussian Process Classifier (GPC)

Gaussian Processes predict by computing for empirical confidence interval using probabilistic Gaussian distribution [33]. For GPC implementation, the model `GaussianProcessClassifier(1.0 * RBF(1.0))` will be used utilizing the default squared-exponential kernel as radial basis function.

### 2.4.4 Decision Tree Classifier (DTC)

Decision Trees are non-parametric supervised learning method used for classification and regression with goal of creating a model that predicts the value of a target variable by learning simple decision rules inferred from the data features [34]. The model to be used will be `DecisionTreeClassifier(max_depth=5)` with maximum depth of the decision tree equal to 5. The deeper the tree, the more complex the if-then-else decision rules and the fitter the model.

### 2.4.5 Random Forest Classifier (RFC)

Random Forest algorithm is a supervised classification algorithm which creates a forest with a number of trees. The higher the number of trees in the forest gives the high accuracy results [35]. The model to be used is `RandomForestClassifier(max_depth=5, n_estimators=10, max_features=1)` utilizing 10 decision trees, each with maximum depth of 5 and considering only 1 feature when looking for the best split.

### 2.4.6 Multilayer Perceptron (MLPC)

MLPC can be viewed as a logistic regression classifier where the input is first transformed using learnt non-linear transformation that projects the input data into a space where it becomes linearly separable [36]. Existing implementation model `MLPClassifier(alpha=1e-5)` will be used with L2 penalty parameter or regularization term equal to 0.00001. Smaller value of alpha may fix high bias which is a sign of underfitting by encouraging larger weights, potentially resulting in a more complicated decision boundary, while larger alpha value helps in avoiding overfitting by penalizing weights with large magnitudes [30]



### 2.4.7 AdaBoost (AdaB)

An AdaBoost classifier is a meta-estimator that begins by fitting a classifier on the original dataset and then fits additional copies of the classifier on the same dataset but where the weights of incorrectly classified instances are adjusted such that subsequent classifiers focus more on difficult cases [37]. The model `AdaBoostClassifier()` will be used utilizing default hyper-parameter values where the base estimator is a decision tree classifier with a maximum depth of 1 and the maximum number of estimators is 50.

### 2.4.8 Gaussian Naive Bayes (GNB)

GNB is a special type of NB algorithm used when the features have continuous values and assumes that all the features are following a Gaussian distribution i.e, normal distribution. It scales cubically with the size of the dataset and might be considerably faster [38]. The model to be utilized calls `GaussianNB()`.

### 2.4.9 Quadratic Discriminant Analysis (QDA)

QDA is closely related to LDA but uses quadratic decision surface. It can be derived from simple probabilistic model where prediction is obtained by using Bayes' rule. The model fits a Gaussian density to each class [39]. In QDA, individual covariance matrix is estimated for every class of observations. QDA becomes useful when there is prior knowledge that individual classes exhibit distinct covariances. The model that will be utilized is `QuadraticDiscriminantAnalysis()`.

### 2.4.10 Convolutional Neural Network (CNN)

It is a neural network mathematical model based on connected artificial neurons similar to biological neural networks. Neurons are organized in layers and connections are established between neurons of adjacent layers. The output layer neurons amount is equal to the number of classifying classes with probabilities showing the possibility that the input vector belongs to a corresponding class [18]. The training process is to minimize the cost function with minimization methods based on the gradient decent also known as backpropagation [19]. The CNN model to be utilized is based on LeNet-5 architecture [40] composed of seven layer combination of convolution, pooling with rectified linear unit (ReLU) activation function and fully connected layers shown in Figure 3.

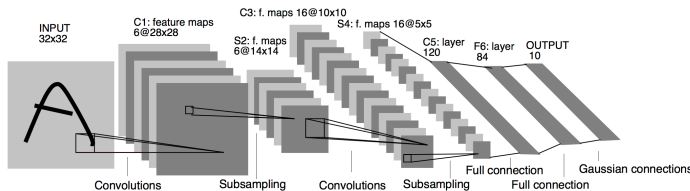


Figure 3: CNN model based on LeNet-5 architecture [40]

Table 1 summarizes the modification to hyper-parameter values done for the implementation of all the learning algorithms using scikit learn and Pytorch libraries in Python. All other hyper-parameters not modified are set to their default values.

Table 1: Summary of modified hyper-parameters for learning algorithms implementation

Learning Algorithm	Modified Hyper-parameter
KNN	n_neighbors=1
SVC	kernel="linear", C=0.025
GPC	None
DTC	max_depth=5
RFC	max_depth=5, n_estimators=10, max_features=1
MLPC	alpha=1e-5
AdaB	None
GNB	None
QDA	None
CNN	6 classes output

## 2.5 Evaluation

The evaluation phase of the designed model will be divided into three: detection, recognition, and system level testing. For this study, the detection and recognition models will be tested individually based on accuracy and speed using Lenovo U410 ultrabook before a system level testing is performed.

### 2.5.1 Detection Model Testing

After implementation of the four pre-processing and segmentation models, traffic images will be fed. Accuracy will be computed using Equation 1. *TP* is defined by correct detection on image with traffic sign, *TN* when there is no detection for traffic image without traffic sign, *FP* for detection when the image has no traffic sign or detection of other part of the image with traffic sign, and *FN* for non-detection of the traffic image with traffic sign. Processing speed of the sub-processes will also be measured for the entire detection process.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

### 2.5.2 Recognition Model Testing

For the testing of the implemented recognition model, cross validation method will be used wherein portion of the dataset will be used for training while the remaining will be for testing. A confusion matrix will then be generated. Evaluation metric to be used is accuracy that determines whether a label predicted for a sample exactly matches the corresponding true label of the sample. Equation 1 will be used to compute for accuracy. Processing time of the entire recognition model and its sub-processes will also be recorded.

### 2.5.3 System Level Testing

For the system level testing, the best detection and recognition models based on the evaluation will be integrated to create the entire model of the system using Nvidia Jetson Nano with interface Wave-share 7" LCD, IMX219-77 camera and generic speaker. Accuracy and speed of detection will be tested using an actual video taken in local road. System level accuracy will be computed to be the product of detection accuracy and recognition accuracy using Equation 1.

A successful recognition means that the system has provided a voice alert about the detected traffic sign. A text-to-speech python module will be used for this function and the corresponding voice alert for each traffic sign detected is shown in Table 2. A text annotation will also be generated to the display.

Table 2: Voice alert and text annotation for target traffic signs

Traffic Sign	Voice Alert	Text Annotation
Left Turn Prohibition	No Left Turn	No Left Turn
Right Turn Prohibition	No Right Turn	No Right Turn
U-Turn Prohibition	No U-Turn	No U-Turn
Overtaking Prohibition	No Overtaking	No Overtaking
30kph Speed Limit	Thirty kph	30 kph
60kph Speed Limit	Sixty kph	60 kph

### 3 Results and Analysis

#### 3.1 Dataset

A total of 2,194 images are used to test the detection phase of the models. Of these, 2,170 are traffic images from online sources with traffic sign having different viewing angle and position on image. These are evenly distributed to the six traffic sign classes and an additional class for images without traffic sign with 1366x768 pixels resolution. The remaining 24 images with 1936x1296 pixels resolution are foreign traffic images captured at different distances from 10 meters to 120 meters with 10-meter increment. For the recognition models, 12,120 images will be used to evaluate speed and accuracy. This dataset is augmented using different cropping methods including zooming, north-west cropping and south-east cropping to act as learning compensation when segmentation of the traffic sign is not centered. The dataset will be divided into train data (10%, 5%, 1%) and test data (90%, 95%, 99%) to determine the effect of the number of test data to models' accuracy. Another set contains six video footages with 360 frames of traffic images, 308 frames of which contain traffic sign. These are recorded using the actual system camera setup attached in a sedan car while approaching the traffic sign to be detected. Figure 4 shows a sample frame image from a video of a local road while Figure 5 shows sample images from the Napier University traffic dataset.



Figure 4: Sample traffic image from local road

Figure 6 shows the sample traffic signs from the dataset to be classified. This dataset contains a total of 2,020 images per class for the training and testing phases and is composed of different signs ranging from high resolution images up to 699x693 pixels down to low-resolution images of 12x12 pixels.



Figure 5: Sample traffic images from Napier University dataset



Figure 6: Sample traffic sign dataset

#### 3.2 Traffic Sign Detection

The detection process is summarized by Figure 7 where the high-lighted blocks are to be replaced by each of the four detection methods to be evaluated. From the original traffic image, red pixels are segmented using color thresholding. The resulting mask undergoes edge detection before Hough transform is applied to detect circular shape. The center and radius information is used to crop the candidate traffic sign for the next phase of classification. A traffic sign is considered detected if the cropped image fully shows the entirety of the traffic sign including the red circular boundary.



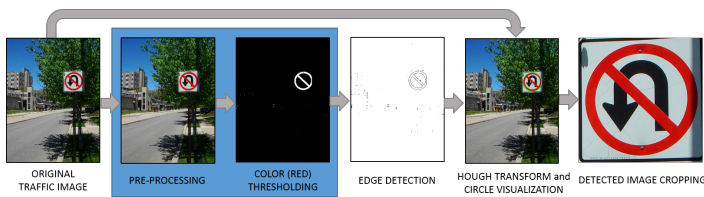


Figure 7: Traffic sign detection process

Experimentation for segmenting the red boundary of the traffic sign using HSV and relative RGB color spaces have resulted to the following threshold range summarized in Table 3 and Table 4. Figure 8 shows masks of segmented red color from an original traffic image using Relative RGB and HSV color segmentation.

Table 3: Range values for HSV color segmentation

Color Component/Channel	Range
Hue	$0 \leq h \leq 10$ or $160 \leq h \leq 179$
Saturation	$80 \leq s \leq 255$
Value	$60 \leq v \leq 255$

Table 4: Range values for relative RGB color segmentation

Color Component/Channel	Range
Red	$60 \leq r \leq 255$
Red minus Green	$35 \leq r-g \leq 135$
Red minus Blue	$8 \leq r-b \leq 204$



Figure 8: Sample relative RGB and HSV segmented red mask

Table 5: Accuracy of pre-processing and color segmentation methods

Method	Total Images	TP	TN	FP	FN	Accuracy
RRGB	2,170	535	299	11	1,462	32.12%
SHI	2,170	576	292	18	1,284	40.00%
BFM	2,170	503	304	6	1,357	37.19%
CCM	2,170	586	294	16	1,274	40.55%

Table 5 summarized the evaluation of the four pre-processing and color segmentation methods. In terms of accuracy, Color Constancy method provided the best performance successfully detecting 586 traffic signs from 2,170 traffic images. It is closely followed

by the Shadow and Highlight Invariant method with 576 detection. Bilateral Filtering method and Color Constancy method have Shadow and Highlight Invariant method for color detection but results show improvement with Color Constancy method while performance degradation for Bilateral Filtering method due to (1) color improvement done by the Color Constancy method to the traffic image, and (2) images from the dataset are generally blurry due to vehicle vibration and environmental parameters that affect the performance of the Bilateral Filtering method. For Relative RGB implementation, color segmentation resulted to less false detection. Figure 9 shows the result of successful detection of traffic signs for six traffic images taken from local road using Shadow and Highlight Invariant method.



Figure 9: Sample detection of traffic sign

Table 6: Number of successful segmentation of pre-processing and color segmentation methods vs. traffic sign capturing distance

Distance	RRGB	SHI	BFM	CCM
120m	0	0	0	0
110m	0	0	0	0
100m	0	0	0	1
90m	0	0	0	0
80m	0	0	0	1
70m	0	2	1	1
60m	0	2	2	2
50m	0	2	2	2
40m	0	2	2	2
30m	0	2	2	2
20m	0	2	2	2
10m	2	2	2	2

Table 6 shows the evaluation of segmentation success of traffic sign with varying capturing distances. The traffic sign is captured twice per given distance. Color Constancy algorithm correctly segmented 15 out of 24 traffic images in the dataset with detection

distance up to 100 meters. Shadow and Highlight Invariant Method correctly segmented 14 out of 24 images with detection distance up to 70 meters. Bilateral Filtering Method correctly detected 13 out of 24 images with detection distance up to 70 meters. Relative RGB Method has the worse performance detecting only 2 out of 24 images with detection distance up to 10 meters and false detecting 13 images. This result validates the accuracy evaluation of the four methods previously discussed in Table 5 and suggests that traffic sign distance affects the image feature clarity of the sign on the image and causes significant decrease in detection success rate.

Table 7 summarizes the processing speed of the four pre-processing and color segmentation methods including execution speed of their sub-processes. This timing does not include shape detection and ROI localization as these processes are common to all methods. Relative RGB segmentation has the fastest processing time closely followed by the Shadow & Highlight Invariant method. The slight difference is due to color space conversion from RGB to HSV. Different processor runtime caused variable timing result for color segmentation process for Shadow & Highlight Invariant, Bilateral Filtering, and Color Constancy. These two employed simple pre-processing while Color Constancy and Bilateral Filtering methods are pre-processing methods themselves and perform multiple and complex image filtering using kernels.

Considering accuracy and processing speed, the pre-processing and segmentation method that provided the best trade-off is Shadow and Highlight Invariant method with 40.00% segmentation success rate and 31.2 ms processing speed.

Table 7: Processing speed of pre-processing and color segmentation methods and sub-processes

Method	Pre-processing	Color Segmentation	Total
RRGB	10.4 ms	15.6 ms	26.0 ms
SHI	5.2 ms	26.0 ms	31.2 ms
BFM	270.8 ms	18.2 ms	289.0 ms
CCM	494.8 ms	20.8 ms	515.6 ms

### 3.3 Feature Extraction

HOG feature extraction of a traffic image is shown in Figure 10. It starts from grayscale conversion followed by edge detection then by HOG feature extraction. HOG feature represents the general structure of the image using the dominant orientation of pixel groups. This process transforms the 2D image into 1D vector histogram representing the gradient direction of the image.

### 3.4 Classifier Evaluation

Table 8 shows the result of implementation of traffic sign classification using Multi-layer Perceptron Classifier with variable image input resolution. This shows that more feature pixels will not always provide the best accuracy for a given classifier and the optimum image resolution considering accuracy and processing speed is 50x50 pixels.

Table 9 summarizes the result of evaluating the 10 classification algorithms in terms of accuracy with the number training data varied. Most of the model showed downward trend in accuracy when

training data is decreased. GPC, though has one of the best accuracy for higher number of training data, provided the worse performance for decreasing training data since optimization is being done using squared-exponential kernel to best fit few training data leading to misclassification. Convolutional Neural Network implementation obtained the best accuracy among the models. This is due to learning optimization using backpropagation decreasing the loss function resulting to higher classification success rate. With lesser training data, CNN still has the best accuracy. This is important to consider to ensure that the model will still provide better accuracy even when trained with fewer data. Figure 11 shows the normalized confusion matrix generated for CNN classifier.

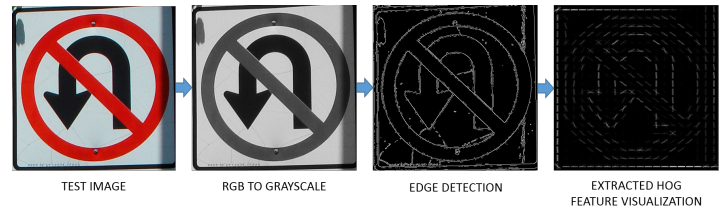


Figure 10: Test image feature extraction

Table 8: Effect of input image size to accuracy and processing speed for MLPC

Input Image Resolution	Accuracy	Processing Time
200x200 pixels	88.89%	454.5 ms
100x100 pixels	88.49%	331.2 ms
50x50 pixels	91.23%	283.2 ms
32x32pixels	85.32%	321.0 ms

Table 9: Effect of number of training data to accuracy

Classifier	10% Train	5% Train	1% Train	Average
KNN	94.63%	92.06%	86.70%	91.13%
SVM	89.72%	87.37%	57.21%	78.10%
GPC	93.87%	80.95%	16.63%	63.82%
DTC	72.45%	68.35%	57.90%	66.23%
RFC	76.59%	73.27%	60.17%	70.01%
MLPC	93.54%	91.19%	86.36%	90.36%
ADAB	62.50%	54.95%	35.64%	51.03%
GNB	85.44%	85.46%	77.50%	82.80%
QDA	54.07%	30.80%	26.28%	37.05%
CNN	96.64%	94.63%	87.64%	92.97%

K-Nearest Neighbor performed the second best in terms of accuracy closely followed by Multilayer Perceptron. This can be accounted to the characteristics of dataset. Each class of the dataset used is primarily composed of very similar images, and results suggest that KNN works better for data with close similarity.

The result of the evaluation of processing speed of the classification algorithms and their sub-processes for 90% of dataset is summarized in Table 10. CNN got the longest training time due to its nature of performing loss function optimization. Total recognition time is measured from color conversion until prediction and CNN obtained the fastest processing time. This is due to lesser



pre-processing being implemented by the method unlike the other algorithms that still perform color conversion and extract HOG features to every input traffic sign image to classify. Recognition time for all classifiers is below 20 ms except for GPC with 750.6 ms due to the use of more complex squared-exponential kernel.



Figure 11: Normalized confusion matrix for CNN classifier

Table 10: Processing speed of classifiers and sub-processes

Classifier	Feature Extract	Train Time	Color Convert	Image Resize	Predict	TOTAL TIME
KNN	16.04s	0.05s	1.56ms	10.42ms	5.21ms	17.19ms
SVC	16.04s	0.64s	1.56ms	10.42ms	0.00ms	11.98ms
GPC	16.04s	272.8s	1.56ms	10.42ms	738.6ms	750.6ms
DTC	16.04s	0.47s	1.56ms	10.42ms	0.00ms	11.98ms
RFC	16.04s	0.02s	1.56ms	10.42ms	2.08ms	14.06ms
MLPC	16.04s	6.62s	1.56ms	10.42ms	0.00ms	11.98ms
AdaB	16.04s	5.32s	1.56ms	10.42ms	6.77ms	18.75ms
GNB	16.04s	0.03s	1.56ms	10.42ms	2.60ms	14.58ms
QDA	16.04s	0.43s	1.56ms	10.42ms	4.16ms	16.15ms
CNN	-	581.4s	-	5.21ms	2.60ms	7.81ms

Table 11 shows result of optimization of the best detection and classification models based on the evaluation. These models are Shadow and Highlight Invariant method and Convolutional Neural Network classifier.

Table 11: Optimization of detection and recognition models

Parameter	Original Model	First Optimization	Second Optimization
Total Images	2,170	2,170	2,170
TP <sub>detection</sub>	576	1,372	1,372
TN <sub>detection</sub>	292	300	300
FP <sub>detection</sub>	18	40	40
FN <sub>detection</sub>	1,284	458	458
TP <sub>recognition</sub>	547	1,240	1,257
TN <sub>recognition</sub>	NA	NA	4
FP <sub>recognition</sub>	47	172	147
FN <sub>recognition</sub>	NA	NA	4

The first optimization resulted to improvement of correct detection and significant decrease in false detection. The process involved modification of threshold parameters of Hough circle transform for it to detect circle less strictly since traffic sign boundaries most often are elliptical in shape in the traffic image due to variable viewing

angle. Hough transform is also adjusted to only detect circles with diameter less than one fifth of the image height. The limiting of circular detection is important since detecting larger circles will result to more pixel features thus will require longer processing time. Also in actual application, traffic signs do not often occupy the entire traffic image.

For the second optimization where a seventh class is added, results in Table 11 show further increase in correct recognition and decrease in false recognition. The additional class filters signs that are similar to any of the six classes, and once recognized as false class, no output will be generated by the model.

Results show that the optimized integrated model has a detection performance of 77.05% and recognition performance of 89.31%. The relatively low detection performance is due to the failure of the Hough transform to detect the circular boundary of the traffic sign since the appearance of the traffic signs are affected by the viewing angle and effect of environmental setting to image quality. Although adjusting the detector threshold to lower value resulted to detection improvement, further decrease in threshold resulted to more false detection. The overall accuracy of the model is 68.81%.

### 3.5 System Implementation

After implementing the system using Nvidia Jetson Nano platform, a timing comparison between the performance of the system using Lenovo U410 Ultrabook is presented in Table 12. Training time and accuracy of both implementation are relatively the same. The huge difference lies in processing time with Nvidia Jetson Nano almost doubled that of the Lenovo U410. This is due to the clocking difference of the processors of the two platforms.

Table 12: Performance comparison between Lenovo U410 and Nvidia Jetson Nano

Parameter	Lenovo U410	Nvidia Jetson Nano
Training Time	678.94 s	679.10 s
Processing Speed	131.3 ms	238.0 ms
Accuracy	95.58%	97.24%

Figure 12 shows some of the successful detection and recognition where traffic signs are annotated to each frame of actual local road video. Table 13 summarizes the timing analysis of the entire TSDR system with text annotation and voice alert output. The system has a total processing time of 1.0953 seconds. It takes about 0.73 ms to display the text annotation of the detected and recognized traffic sign while it takes 938.21 ms for the system to utter the entire phrase for voice alert. For voice alert system, the next frame can only be processed after it has finished all the required processing for current frame including uttering the meaning of the sign.

Given this timing values, the system without voice alert and only using text annotation as output can run at a maximum frame rate of 1/116.7ms or approximately 8 frames per second. For the system with voice alert output, the system requires a frame rate of 1/1.0953 seconds or less than 1 frame per second for it to run in real-time. Since voice alert must be implemented, having a frame rate of less than 1 fps will result to poor sampling causing traffic signs to not be detected especially for fast moving vehicles. A solution is to implement adaptive frame rate such that during non-detection,

the system will run at approximately 8 to 12 fps and when there is a traffic sign detected, frame rate shifts to approximately 1 fps to cater for the voice alert output. Figure 13 shows the hardware configuration block diagram of the system while Figure 14 shows the actual setup of the TSDR system for road testing. Figure 15 shows the testing of the TSDR system using actual footage from local road.

fied ( $TP_{recognition}$ ), and 33 were incorrectly recognized ( $FP_{recognition}$ ). The system has a detection accuracy of 32.50% and a recognition accuracy of 60.24%. The overall accuracy of the system is 19.58%. The low accuracy is due to the poor image quality caused by several factors including (1) poor camera, (2) vehicle motion and vibration, and (3) environmental setting.



Figure 12: Sample detection and recognition of traffic frames

Table 13: Timing of sub-processes of the TSDR system

Process	Processing Time
Input Acquisition	13.69 ms
BGR to HSV Conversion	2.49 ms
Red Color Segmentation	5.22 ms
Gaussian Blurring	1.52 ms
Median Blurring	7.64 ms
Binary Thresholding	11.44 ms
Erosion	0.75 ms
Dilation	0.63 ms
Circle Detection	70.96 ms
ROI Cropping	7.09 ms
Image Transformation	4.79 ms
Prediction	30.13 ms
Text Annotation	0.73 ms
Voice Alert Output	938.21 ms
TOTAL TIME	1.0953 s

Table 14 shows the result of testing the TSDR system using a total of 360 frames from six different actual footage (F1-F6) from local roads. The footage contains 308 frames where traffic sign is fully visible with approaching distance and 52 frames where the traffic sign is not visible. The system had detection on 83 traffic images ( $TP_{detection} + FP_{detection}$ ), of which 30 were classified correctly to not contain traffic signs ( $TN_{recognition}$ ), 20 were correctly classi-

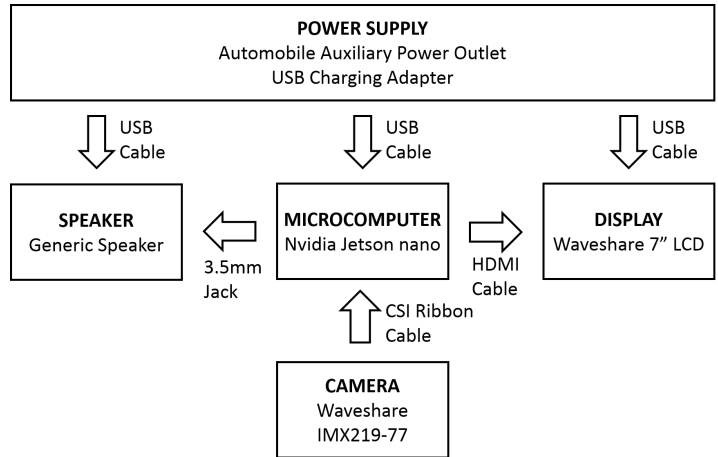


Figure 13: TSDR hardware configuration diagram

Detection and recognition of traffic sign in every frame is ideal for a ATSDR system but in actual application, one traffic sign correctly detected and recognized is enough to provide road information to the driver. Out of 6 sets of image sequences, 3 sets or 50% have at least one of the frames correctly detected and recognized.

Table 14: Test using frames from actual local road footage

Parameter	F1	F2	F3	F4	F5	F6	TOTAL
No. of Frames	87	73	64	32	48	56	360
$TP_{detection}$	12	29	2	0	0	23	66
$TN_{detection}$	12	7	8	8	8	8	51
$FP_{detection}$	3	8	0	0	0	6	17
$FN_{detection}$	60	29	54	24	40	19	226
$TP_{recognition}$	0	7	2	0	0	11	20
$TN_{recognition}$	9	15	0	0	0	6	30
$FP_{recognition}$	6	15	0	0	0	12	33
$FN_{recognition}$	0	0	0	0	0	0	0
Sign Detected?	No	Yes	Yes	No	No	Yes	50%



Figure 14: TSDR car setup for actual road testing



Figure 15: Sample frames from actual TSDR footage

## 4 Conclusion

Among the four evaluated pre-processing and detection methods, Shadow and Highlight Invariant method provided the best trade-off between segmentation success rate and processing time due to the less complex image filtering approach and use of HSV color space. Convolutional Neural Network classifier has the best performance in both accuracy and processing speed among the ten evaluated algorithms for traffic sign recognition due to the optimization method of the loss function using back propagation and the in-system feature extraction method of the algorithm. In classification, additional class to contain the not targeted traffic signs provides improvement of false detection performance.

An embedded system using Nvidia Jetson Nano platform with interface IMX219-77 camera, Waveshare 7" LCD and generic speaker using Python and OpenCV, sci-kit learn and Pytorch libraries has been developed. Implementation of adaptive frame rate improves sampling and provides an approach to handle variable frame processing time for a real-time system.

This study has not only provided comparative analysis on detection and recognition performance of several algorithms based on accuracy and processing speed but also able to implement a real-time embedded system. Further improvement to the system includes the use of better high definition camera with better image sensor and optical image stabilization. This can also be improved by the use of better detection and recognition algorithm like masked R-CNN. Also, further study can be done on the extension of application of the study to include other traffic signs and signals, and road markings.

**Conflict of Interest** The authors declare no conflict of interest.

**Acknowledgment** The authors would like to acknowledge Ateño de Manila University, FAITH Colleges, and Fastech Synergy Philippines for supporting this research through their engineering scholarship program.

## References

- [1] World Health Organization, "Global status report on road safety", 2015, Retrieved from [https://www.who.int/violence\\_injury\\_prevention/road\\_safety\\_status/2015/en/](https://www.who.int/violence_injury_prevention/road_safety_status/2015/en/)
- [2] M. R. De Leon, "Towards Safer Roads: Road safety initiatives of DOTr", ASEAN Automobile Safety Forum, May 2017.
- [3] A. C. Mendoza, "Applications of Psychology to the understanding and modification of road user behavior", University of the Philippines National Center for Transportation Studies, August 2006.
- [4] Y. MaryReeza, T. Latha, A. MaryAnsalinShalini, "Traffic sign detection and recognition in driver support system for safety precaution", ARPN J. Eng. Appl. Sci., **10**, 2015, 2209-2213.
- [5] S. Feng, "Evaluation of colour segmentation algorithms in red colour of traffic sign detection", M.S. Thesis, Department of Computer Engineering, Hogskolan Dalarna, 2010.
- [6] A. Santos, P. A. Abu, C. Oppus, R. Reyes, "Traffic sign detection and recognition for assistive driving", in 2019 International Symposium on Multimedia and Communication Technology (ISMAC), Quezon City, Philippines, 2019, pp. 1-6, doi: 10.1109/ISMAC.2019.8836161.
- [7] Department of Public Works and Highways, "Highway safety design standards, road safety design manual", May 2012.
- [8] C. Mythili, V. Kavitha, "Color image segmentation using ERKFCM", Int. J. Comput. Appl., **41**(20), 21-28, March 2012, doi: 10.5120/5809-8074.
- [9] H. Fleyeh, M. Dougherty, "Road and traffic sign detection and recognition", Advances OR and AI Methods in Transportation.
- [10] V. Deshmukh, G. Patnaik, and M. Patil, "Real-time traffic sign recognition system based on colour image segmentation", 2013, Int. J. Comput. Appl., **83**(3), 30-35, doi: 10.5120/14430-2575 10.5120/14430-2575.
- [11] C. Lin, C. Su, H. Huang, K. Fan, "Colour image segmentation in various illumination circumstances" in Proceedings of the 9th WSEAS International Conference on Circuits, Systems, Electronics, Control & Signal Processing, World Scientific and Engineering Academy and Society (WSEAS), Stevens Point, Wisconsin, USA, 2010, 179184, doi: 10.13140/2.1.1045.3446.
- [12] H. Rashmi, M. Shashidhar, G. Prashanth Kumar, "Automatic tracking of traffic signs based on HSV" International Journal of Engineering Research & Technology, **3**(5), May 2014.
- [13] S. Sural, G. Qian, S. Pramanik, "Segmentation and histogram generation using the HSV color space for image retrieval" in Proceedings of the International Conference on Image Processing, Rochester, NY, USA, 2002, pp. II-II, doi: 10.1109/ICIP.2002.1040019.
- [14] H. Fleyeh, "Color detection and segmentation for road and traffic signs," IEEE Conference on Cybernetics and Intelligent Systems, 2004, Singapore, 2004, pp. 809-814, doi: 10.1109/ICCIS.2004.1460692.
- [15] H. Fleyeh, "Traffic signs color detection and segmentation in poor light conditions", MVA2005 IAPR Conference on Machine Vision Applications, May 16-18, 2005 Tsukuba Science City, Japan.
- [16] M. Ebner, "Color constancy using local color shifts", The 8th European Conf. on Computer Vision, Prague, Czech Republic, 2004, 276-287, doi: 10.1007/978-3-540-24672-5\_22.
- [17] H. Gomez-Moreno, S. Maldonado-Bascon, P. Gil-Jimenez, S. Lafuente-Arroyo, "Goal evaluation of segmentation algorithms for traffic sign recognition" in IEEE T. Intel. Transp., **11**(4), pp. 917-930, Dec. 2010, doi: 10.1109/TITS.2010.2054084.
- [18] A. Shustanov, P. Yakimov, "CNN design for real-time traffic sign recognition", Procedia Eng., **201**, 718-725, 2017, doi: 10.1016/j.proeng.2017.09.594.
- [19] S. Prasomphan, T. Tathong, P. Charoenprateepkit, "Traffic sign detection for panoramic images using convolution neural network technique", HPCCT 2019: Proceedings of the 2019 3rd High Performance Computing and Cluster Technologies Conference, 2019, 128-133, doi: 10.1145/3341069.3341090.
- [20] Y. Wu, Y. Liu, J. Li, H. Liu, X. Hu, "Traffic sign detection based on convolutional neural networks", Proceedings of the International Joint Conference on Neural Networks, 2013, 1-7, doi: 10.1109/IJCNN.2013.6706811.
- [21] Y. Sun, P. Ge, D. Liu, "Traffic sign detection and recognition based on convolutional neural network," 2019 Chinese Automation Congress (CAC), Hangzhou, China, 2019, pp. 2851-2854, doi: 10.1109/CAC48633.2019.8997240.
- [22] D. Alghmgham, G. Latif, J. Alghazo, L. Alzubaidi, "Autonomous traffic sign (ATSR) detection and recognition using deep CNN", Procedia Comput. Sci., **163**, 2019, Pages 266-274, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2019.12.108>.



- [23] D. Tabernik, D. Skoaj, "Deep learning for large-scale traffic-sign detection and recognition" *IEEE T. Intel. Transp.*, **21**(4), pp. 1427-1440, April 2020, doi: 10.1109/TITS.2019.2913588.
- [24] C. Tomasi, R. Manduchi, "Bilateral filtering for gray and color images," *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, Bombay, India, 1998, pp. 839-846, doi: 10.1109/ICCV.1998.710815.
- [25] N. Kehtarnavaz, A. Ahmad, "Traffic sign recognition in noisy outdoor scenes," *Proceedings of the Intelligent Vehicles '95. Symposium*, Detroit, MI, USA, 1995, pp. 460-465, doi: 10.1109/IVS.1995.528325.
- [26] M. Rizon, H. Yazid, P. Saad, A. Shakaff, A. Saad, "Object detection using circular hough transform", *American Journal of Applied Sciences*, **2**, 2005, 10.3844/ajassp.2005.1606.1609.
- [27] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, **1**, San Diego, CA, USA, 2005, pp. 886-893, doi: 10.1109/CVPR.2005.177.
- [28] J. Lim, C. Oppus, "Identifying car logos using different classifiers in Python with HOG features".
- [29] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, "PyTorch: An imperative style, high-performance deep learning library", *Adv. Neur. In.*, **32**, 2019.
- [30] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, "Scikit-learn: machine learning in python", *J. Mach. Learn. Res.* **12**, pp. 2825-2830, 2011.
- [31] Nearest Neighbors Classification. Retrieved from <https://scikit-learn.org/stable/modules/neighbors.html>
- [32] Introduction to Support Vector Machines. Retrieved from [https://docs.opencv.org/3.4/d1/d73/tutorial\\_introduction\\_to\\_svm.html](https://docs.opencv.org/3.4/d1/d73/tutorial_introduction_to_svm.html)
- [33] Gaussian Processes. Retrieved from [https://scikit-learn.org/stable/modules/gaussian\\_process.html](https://scikit-learn.org/stable/modules/gaussian_process.html)
- [34] Decision Trees. Retrieved from <https://scikit-learn.org/stable/modules/tree.html>
- [35] How the Random Forest Algorithm Works in Machine Learning. Retrieved from <https://dataaspirant.com/2017/05/22/random-forest-algorithm-machine-learning/>
- [36] Multilayer Perceptron. Retrieved from <http://deeplearning.net/tutorial/mlp.html>
- [37] An AdaBoost Classifier. Retrieved from <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.AdaBoostClassifier.html>
- [38] Gaussian Nave Bayes Classifier Implementation in Python. Retrieved from <https://dataaspirant.com/2017/02/20/gaussian-naive-bayes-classifier-implementation-python/>
- [39] Linear and Quadratic Discriminant Analysis. Retrieved from [https://scikit-learn.org/stable/modules/lda\\_qda.html](https://scikit-learn.org/stable/modules/lda_qda.html)
- [40] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition," in *P. IEEE*, **86**(11), pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.



# Performance Analysis of Go-Back-N ARQ Protocol Used in Data Transmission Over Noisy Channels

Fayza Ahmed Nada\*

Department of Information Technology, Faculty of Computers and Information, Suez University, Suez 43533, Egypt

## ARTICLE INFO

Article history:

Received: 22 June, 2020

Accepted: 07 August, 2020

Online: 25 August, 2020

Keywords:

Performance Analysis

Go-Back-N ARQ protocols

Sliding Window Protocols

System Occupancy

Waiting Time

## ABSTRACT

Automatic Repeat reQuest (ARQ) protocols are used to control transmission errors in data communications. They are used at Data Link Control (DLC) sublayer of Data Link Layer (DLL) of OSI (Open Systems Interconnection) network model to achieve flow and error control and provide smooth and reliable transmission between network nodes. ARQ protocols use acknowledgments and timeouts to satisfy reliable data transmission over noisy channels. Types of ARQ protocols include Stop-and-wait ARQ, Go-Back-N ARQ, and Selective Repeat ARQ. This paper presents a new mathematical model to analyze performance of Go-Back-N ARQ protocol over noisy channels. Many performance measures of Go-Back-N ARQ protocols are discussed. The analysis uses Stochastic Process and a two-dimensional Markov Chain to present the proposed model. We study the distributions of system occupancy and waiting time when implementing Go-Back-N ARQ protocol in data transmission. Results include: Probability Generating Functions (PGF) of system occupancy and waiting time. Moreover, mean values of previous system parameters are derived. ARQ protocols are essential for peer-to-peer protocols that provide reliable data transmission. The obtained results can be adopted and implemented in simulations of similar communication systems or approximating relevant systems.

## 1 Introduction

Controlling transmission errors is essential for data traffic between network nodes. Automatic Repeat reQuest (ARQ) protocols are implemented at Data Link Control (DLC) sublayer of Data Link Layer (DLL) of Open System Interconnection (OSI) network model to provide error control mechanism and enable reliability of data transmission between different network nodes. ARQ protocols are essential for peer-to-peer protocols that provide reliable data transmission. There are two categories of ARQ protocols: protocols for noisy channels and protocols for noiseless Channel. However, noiseless channels are not available in real life. Noiseless channels are ideal channels which assumes successful delivery for all data. So, there is more concentration on noisy channel protocols, which emphasis flow control and error control mechanisms.

Go-Back-N ARQ (GBN ARQ), Selective Repeat ARQ (SR ARQ) and Stop and Wait ARQ (SW ARQ) are different types of ARQ protocol. Each protocol has a different strategy and methodology for handling data transmission over unreliable channels. The basic algorithm modules used in ARQ protocols are: ACKnowledgments (ACK is a small frame that the receiver sends to the sender to

notify successful delivery of a data frame), Negative AcKnowledge-ment (NAK is a small frame that the receiver sends to the sender to indicate unsuccessful transmission. It tells the sender that some type of error transmission has occurred and that retransmission is required) and time-outs (a period of time that a sender waits before retransmitting the frame). As long as the ACK is not received and time-out expired, the sender will retransmit the frame. This process is repeated until the sender receives the acknowledgment or number of retransmissions exceeds a predefined number. All ARQ protocols use sliding window concept (explained in next section) to tell the sender which (if any) frames need to be retransmitted. Many research articles handled and focused on studying different types of ARQ protocols.

In [1] "analysis and modelling of Automatic Retransmission re-Quest (ARQ) protocols as part of a wireless communications system is presented. It considers systems in increasing complexity and is divided into two parts. The first discusses a single link communications and the second proposes point-to-multipoint systems assuming wireless communications. The first part presents a proper discussion and analysis of ARQ protocols communications. The second part investigates power control policies to maximize the amount

\*Fayza A. Nada, Department of Information Technology, Faculty of Computers and Information, Suez University, Suez 43533, Egypt, +201023135051 & f.nada@suezuni.edu.eg

of data being transferred at the downlink of a cellular communications system as well as at a network with per hop retransmissions”. While [2]: “proposes to analyze the performance of ARQ protocols with applying a Higher Order Logic theorem prover (HOL). The delay of ARQ protocols has been modelled formally as a function of geometric random variable in higher order logic. Actually, this reference develops the behavior of delay of the main ARQ protocols, i.e., SW ARQ, GBN ARQ and SR ARQ. A type of verification of the expected value of the delay formulas has been used in HOL.” “A simulation link for studying and analyzing GBN ARQ protocol” is applied in [3]. However, [4] “analyzed both GBN ARQ protocol together with SR ARQ with intra-block (GBN-SR) and Revised GBN-SR (Rev-GBN-SR) schemes. The presented (Rev-GBN-SR) mechanism has shown better performance with applying bigger range and when the traffic is moving from a low error rate regions to high error rate regions”.

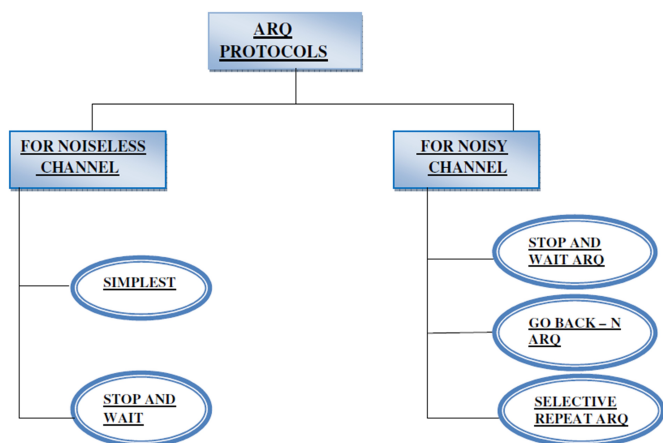


Figure 1: ARQ Protocols

In [5]: “GBN ARQ is discussed with both reliable and unreliable feedback with focus on throughput. Probability measurements have been used to present the results in both cases. The obtained results can be considered as a basic building stone on which larger systems may be built and established.” While [6] used “a two dimensional Markov Chain to study the transmission delay of the GBN ARQ protocol on unreliable correlated channels. The Probability Generating Function (PGF) of the transmission delay of a frame is obtained considering decay factor of a Markov chain. Moreover, an expression of the throughput is obtained. It is shown how the decay factor has affected the protocol performance and the transmission delay.”

On the other hand, [7] studied: “intra block of GBN ARQ and SR ARQ protocols but assuming limited retransmissions”. However, [8] ”proposed a strategy for controlling error transmissions with multiple parallel channels. It assumes that receiver has a buffer of fixed size (called re-sequencing buffer). This buffer stores out of order packets that are successfully received. Actually, the proposed model can be considered as a combination (hybrid ARQ) GBN ARQ and SR ARQ protocols”.

It is clear from previous review that the main performance measures (specially the system occupancy and the waiting time) of a communication system, when using ARQ protocols, have not been studied properly before. The basic contribution of this article is

studying such systems. A new mathematical model for measuring and analyzing the performance of Go-Back-N ARQ protocol is presented. Moreover, the protocol is proposed to be used over noisy channels. Many performance measures of Go-Back-N ARQ protocol are focused on. Complete network evaluation metrics are analyzed when applying Go-Back-N ARQ protocol in data transmission. This work can be considered as a reference for most evaluation results related to Go-Back-N ARQ protocol. Add to this that the analysis results can be adopted and implemented in simulations of relevant communication systems or approximating other systems.

We study the distributions of the waiting time and system occupancy when implementing Go-Back-N ARQ protocol in data traffics. Results include: Probability Generating Functions (PGF) of system occupancy and waiting time. Also, mean values of previous system parameters are derived. Such analysis and results is important in networks analysis, communication systems, and simulation models of larger systems. The paper flows as follows, Next section explains and presents the main concept of sliding window protocols and sheds the light on different types of ARQ protocols. Section 3 introduces the assumptions of the mathematical model under study. In section 4 the PGF of system occupancy and the mean system occupancy are presented. PGF of waiting time and mean waiting time are derived in section 5. Conclusion is the last section.

## 2 Sliding Window Protocols

Sliding window protocols are implemented at the data link layer (OSI layer 2) and at transport layer (OSI layer 4) for applying error control mechanisms on transmitted traffics. They ensure the reliability and sequential delivery of data transmitted between network nodes. According to the sliding window concept, the sender will be allowed to send multiple frames before waiting for any ACK from the receiver. So, sliding window is trying to enhance the network throughput and reduce traffic delay [9].

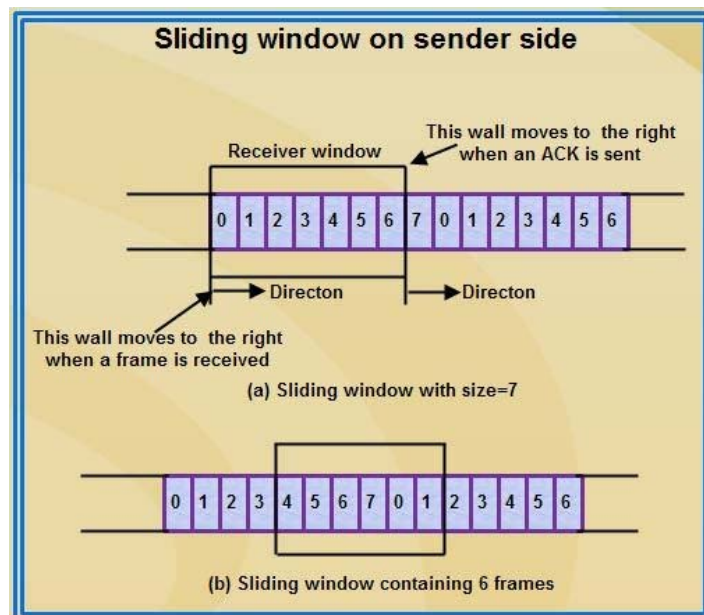


Figure 2: Sliding Window Protocol

Sliding window protocols assume that both the sender and the receiver are provided with buffers called sending and receiving window respectively. However, there is a type of dependency or correlation between the maximum size of sending window and the length of the frame sequence field at the frame header. If the length of the sequence number field at the frame header is  $n$ , then the frame number will range from 0 to  $2^n - 1$ . Consequently, the maximum size of the sending window can not exceed  $2^n - 1$ . Thus, in order to accommodate a sending window with maximum size of  $2^n - 1$ , the sequence number field should be of  $n$ -bit length. In such case the frames sequence numbers are numbered as modulo- $n$ . For example, if the sending window size is 4, then the frame sequence number field should be of 2-bits length to represent the sequence numbers 00, 01, 10, 11 which is interpreted as 0, 1, 2, 3.

On the other hand the size of the receiving window specifies the maximum number of frames that the receiver can accept at a time. Although both Go-Back-N protocol and Selective Repeat protocol are sliding window protocols, there are many differences between them. In case of transmission errors, GBN ARQ protocol will retransmit all the frames after the damaged or lost frame (even if they are correctly received) while in SR ARQ, only damaged or lost frames will be retransmitted.

### 2.1 Go-Back-N ARQ Protocol

The following points show the GBN ARQ strategy:

- The sender starts by sending multiple frames according to the size of sending window which is  $2^n - 1$ , where  $n$  is the size of sequence number field in the frame header.
- The receiver sends an ACK for each correctly received frame.
- The receiver sends a NAK for each damaged or lost frame showing the number of expected frame.

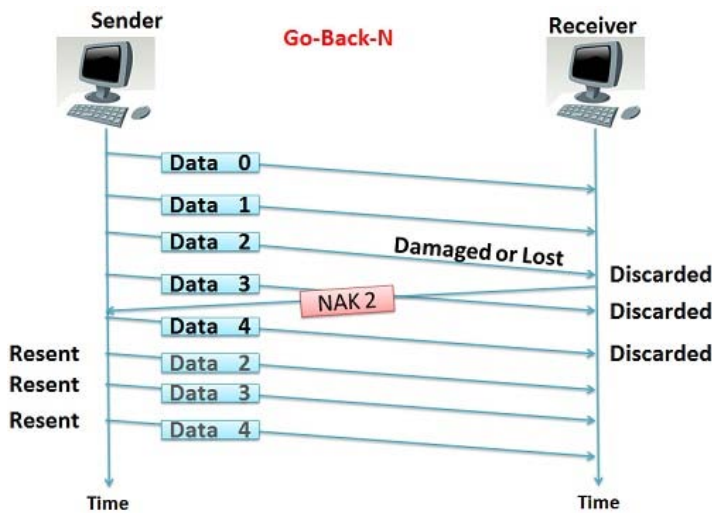


Figure 3: Go-Back-N ARQ Protocol

- If the frame is out of sequence, The receiver discards that frame even if it is correct and sends a NAK for the expected frame.

- After sending a NAK, the receiver discards all the received frames even if they are correctly received (No ACKs are sent for those frames).
- The sender sets a timer for receiving an ACK. If the timer expired (ACK is lost or damaged), the sender will retransmit all non-acknowledged frames.
- After receiving a NAK, the sender will retransmit all the frames sent after the frame referred by the NAK.
- Both ACK number and NAK number indicate the number of the frame expected by the receiver.
- Applying GBN ARQ may lead to wasting bandwidth if the error rate is high.

### 2.2 Selective Repeat ARQ Protocol

The following points show the SR ARQ strategy:

- The sender starts by sending multiple frames according to the size of sending window which is  $2^n - 1$ , where  $n$  is the size of sequence number field in the frame header.
- The receiver sends an ACK for each correctly received frame.
- The receiver sends a NAK for each damaged or lost frame showing the number of expected frame.
- If the frame is out of sequence, the receiver accepts it (sends ACK). The new point is that the receiver is adapted to be able to sort the frames in a proper sequence before submitting to the higher OSI model (Network layer). However, the receiver must hold a buffer to save all the out of sequence received frames till the retransmission process is completed and frames are placed in the correct order.

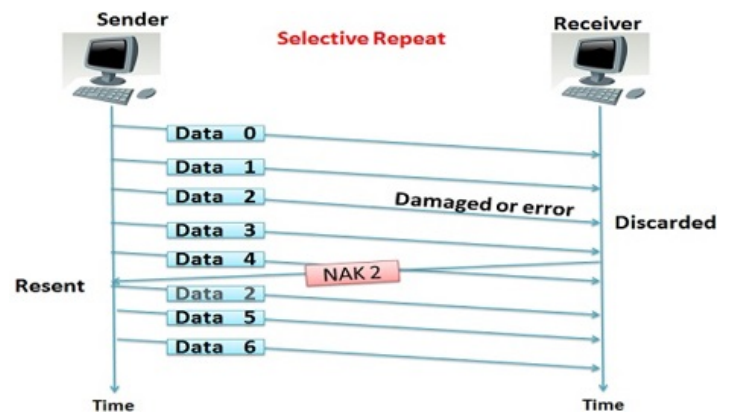


Figure 4: Selective Repeat ARQ Protocol

- When receiving out of sequence frame, the receiver sends a NAK for the frame it expects.
- After sending a NAK, the receiver accepts all subsequent received frames (if they are correct) and send an ACK for each correct frame.

- The sender sets a timer for receiving an ACK. If the timer expired (ACK is lost or damaged), the sender will retransmit the non-acknowledged frames only.
- After receiving a NAK, the sender will retransmit the NAKed frame only.
- Both ACK number and NAK number indicate the number of the frame expected by the receiver.
- Compared to GBN ARQ, SR ARQ has better performance and needs less window size. Performance evaluation of selective repeat ARQ protocol has been presented in [10].

- The receiver uses a return channel for sending ACK or NAK to the sender (to specify the situation of the received packet).
- The RV  $R$  represents how many packets can be sent during the time of a round trip propagation delay.
- No transmission errors are considered on the return channel.

### 3 New Mathematical Model Assumptions

The system is modelled as a stochastic process with existence of a sender and a receiver under the following assumptions:

- Time axis is discretized and divided to equal units called slots.
- Messages are divided into fixed length packets which fits exactly one slot and are transmitted through a channel.
- Message transmission starts only after transmission of all packets of the previous message.
- The receiver is considered with a buffer with infinite capacity.
- At messages level the service is FCFS but GBN protocol is considered when transmitting packets of the same message.
- Messages arrive to the system just before the slot boundaries.
- Let the Random Variables (RV)  $A^k$  represent the sizes of the messages batches arriving in slot  $k$ . The RVs  $A^k$  are independent and have the same distribution, with common PGF denoted by  $A^k(z)$ ,

$$a_n^k = \Pr[A^k = n] \tag{1}$$

$$A^k(z) = \sum_{n=0}^{\infty} a_n^k z^n, \tag{2}$$

$$= \sum_{n=0}^{\infty} a_n z^n, \text{ (at the steady state).} \tag{3}$$

- Let the RV  $B$  represents the number of packets in each message,

$$b_n = \Pr[B = n] \tag{4}$$

$$B(z) = \sum_{n=1}^{\infty} b_n z^n. \tag{5}$$

- Bernoulli process is applied to the random channel transmission errors. If  $p$  is the probability of error occurrence in a packet and  $\varepsilon$  denotes the channel bit error rate, then  $p$  can be written as:

$$p = 1 - (1 - \varepsilon)^N, \tag{6}$$

where  $N$  denotes the number of bits in a packet.

### 4 System Occupancy and Mean System Occupancy

The discussed system is considered as a discrete time queueing system with one server. Let the RV  $S$  denotes the transmission (service) time of any transmitted message, with PGF  $S(z)$ . And the RV  $S_i$  represents the transmission time of a message where  $i$  represents the number of packets in the message, with PGF  $S_i(z)$ ,

$$\begin{aligned} S_i(z) &= \sum_{n=1}^{\infty} \Pr[S_i = n] z^n \\ &= S(z|B = i). \end{aligned} \tag{7}$$

And

$$\begin{aligned} S(z) &= \sum_{i=1}^{\infty} \Pr[B = i] S(z|B = i) \\ &= \sum_{i=1}^{\infty} b_i S_i(z) \\ &= b_1 (\Pr[S_1 = 1]z + \Pr[S_1 = 2]z^2 + \dots) \\ &\quad + b_2 (\Pr[S_2 = 1]z + \Pr[S_2 = 2]z^2 + \dots). \end{aligned} \tag{8}$$

For stability condition of the discussed discrete time queueing system we should have

Arrival rate < departure rate,

$$E[A] < E[D], \frac{E[A]}{E[D]} < 1.$$

But if  $E[D]$  is the mean departure rate, then  $1/E[D]$  represents the mean service time  $E[S]$ . So we get the stability condition

$$E[A]E[S] < 1. \tag{9}$$

Let  $P$  be a RV represents the queue length (system occupancy), with the PGF  $P(z)$  We have proved in [11] that

$$P(z) = [1 - A'(1)S'(1)] \left\{ \frac{[z - 1]S(A(z))}{[z - S(A(z))]} \right\}, \tag{10}$$

which represents the queue length of the system at random slot marks. Let

$$E[A] = A'(1), E[S] = S'(1). \tag{11}$$

Substituting from (11) in (10), then

$$P(z) = [1 - E[A]E[S]] \left\{ \frac{[z - 1]S(A(z))}{[z - S(A(z))]} \right\}. \tag{12}$$



To determine the average queue length, we have

$$E[P] = P^{(1)}(1) = \frac{d}{dz} P(z) \Big|_{z=1}. \tag{13}$$

Let

$$u(z) = [1 - E[A]E[S]][z - 1]S(A(z)), \tag{14}$$

$$v(z) = [z - S(A(z))]. \tag{15}$$

Substituting from (14) and (15) in (12), hence

$$P(z) = \frac{u(z)}{v(z)}. \tag{16}$$

Therefore

$$P^{(1)}(z) = \frac{v(z)u^{(1)}(z) - u(z)v^{(1)}(z)}{v(z)^2}$$

$$P^{(1)}(1) = \frac{0}{0}.$$

Applying l'hospital rule twice, we get

$$P^{(1)}(1) = \frac{v^{(1)}(1)u^{(2)}(1) - u^{(1)}(1)v^{(2)}(1)}{2v^{(1)}(1)^2}. \tag{17}$$

We have

$$u^{(1)}(1) = [1 - E[A]E[S]]. \tag{18}$$

And

$$u^{(2)}(1) = [1 - E[A]E[S]] [2E(A)E(S)]. \tag{19}$$

We also have

$$v^{(1)}(1) = [1 - E[A]E[S]]. \tag{20}$$

And

$$v^{(2)}(1) = - [E[S]A^{(2)}(1) + E[A]S^{(2)}(1)E[A]]. \tag{21}$$

Substituting from (18), (19), (20), and (21) in (17) we get

$$P^{(1)}(1) = \frac{[1 - E[A]E[S]] [2E[A]E[S]] + [E[S]A^{(2)}(1) + E[A]^2S^{(2)}(1)]}{2[1 - E[A]E[S]]} = E[A]E[S] + \frac{[E[S]A^{(2)}(1) + E[A]^2S^{(2)}(1)]}{2[1 - E[A]E[S]]}. \tag{22}$$

But we have the relations

$$A^{(2)}(1) = E[A^2] - E[A], \tag{23}$$

$$S^{(2)}(1) = E[S^2] - E[S]. \tag{24}$$

Substituting from (23) and (24) in (22), we get

$$P^{(1)}(1) = E[A]E[S] + \frac{[E[S][E[A^2] - E[A]] + E[A]^2[E[S^2] - E[S]]]}{2[1 - E[A]E[S]]}.$$

From which we get

$$E[P] = P^{(1)}(1) = E[A]E[S] + \frac{[E[A]^2E[S^2] + (E[A^2] - E[A] - E[A]^2)E[S]]}{2[1 - E[A]E[S]]}, \tag{25}$$

which represents the mean system occupancy.

## 5 Waiting Time and Mean Waiting Time

The waiting time is defined as the time between arrival point and the beginning of message transmission. Waiting time distribution has been studied in [12] and we have proved that the PGF of the waiting time (including the service time) is given by

$$W(z) = [1 - A'(1)S'(1)] \times \left\{ \frac{[z - 1]S(z)[1 - A(S(z))]}{[z - A(S(z))][1 - S(z)]A'(1)} \right\}. \tag{26}$$

Letting

$$E[A] = A'(1), E[S] = S'(1),$$

and excluding the service time, we can write

$$W(z) = [1 - E[A]E[S]] \times \left\{ \frac{[z - 1][1 - A(S(z))]}{[z - A(S(z))][1 - S(z)]E[A]} \right\}. \tag{27}$$

From Little's rule have

$$E[W] = \frac{E[P]}{E[A]} - E[S]. \tag{28}$$

Substituting from (25) in (28), we find

$$E[W] = \frac{E[A]E[S]}{E[A]} + \frac{[E[A]^2E[S^2] + (E[A^2] - E[A] - E[A]^2)E[S]]}{2E[A][1 - E[A]E[S]]} - E[S] = \frac{[E[A]^2E[S^2] + (E[A^2] - E[A] - E[A]^2)E[S]]}{2E[A][1 - E[A]E[S]]}.$$

From which we get

$$E[W] = \frac{E[A]E[S^2]}{2[1 - E[A]E[S]]} + \frac{(E[A^2] - E[A] - E[A]^2)E[S]}{2E[A][1 - E[A]E[S]]}, \tag{29}$$

which represents the mean waiting time.

## 6 Conclusion

GBN ARQ protocol is a sliding window protocol applied at the DLC sublayer of DLL layer of OSI model to provide error and flow control for data transmission at this level. The main work in this article is analyzing and studying the performance of GBN ARQ protocol. Stochastic process has been used to model the discussed system. Many system parameters have been analyzed and calculated. The main results obtained in this paper are: PGF of system occupancy and waiting time. The mean values of previous system parameters are also derived. The analysis results can be used and implemented in simulations of similar communication systems or approximating other related systems.

## References

- [1] A. Giovanidis, *ARQ Protocols in Wireless Communications: Modeling, Analysis and Optimization* (German Edition), ISBN-13: 978-3838120553.
- [2] O. Hasan, S. Tahar, "Performance Analysis of ARQ Protocols using a Theorem Prover" in 2008 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), 20-22 April 2008, Austin, TX, USA, DOI: 10.1109/ISPASS.2008.4510741.
- [3] H. Saleh, M. Elfituri, "Packet communication within a Go-Back-N ARQ system using Simulink" in 2016 4th International Conference on Control Engineering & Information Technology (CEIT), 16-18 Dec. 2016, Hammamet, Tunisia, DOI: 10.1109/CEIT.2016.7929028.
- [4] Y. Hayashida, A. Maeda, N. Sugimachi, S. Fujii, "Performance analysis of Go-Back-N ARQ scheme with selective repeat in intra-block" *IEEE Transactions on Communications*, **50**(3), 391 - 395, 2002, DOI: 10.1109/26.9909006.
- [5] K. Ausavapattanakun, A. Nosratinia, "Analysis of Go-Back-N ARQ in block fading channels" *IEEE Transactions on Wireless Communications*, **6**(8), 2793 - 2797, 2007, DOI: 10.1109/TWC.2007.06062.
- [6] Y. Hayashida, M. Komatsu, "Delay performance of Go-back-N ARQ scheme with Markovian error channel" in 1993 2nd IEEE International Conference on Universal Personal Communications, 12-15 Oct. 1993, Ottawa, Ontario, Canada, DOI: 10.1109/ICUPC.1993.528425.
- [7] S. Fujii, T. Kaneko, Y. Hayashida, "Effectiveness of copy-transmission in Go-Back-N ARQ system with selective repeat in intra-block and with limited retransmissions" in 2010 2nd International Conference on COMMunication Systems and NETworks (COMSNETS 2010), 5-9 Jan. 2010, Bangalore, India, DOI: 10.1109/COMSNETS.2010.5432007.
- [8] Q. Li, C. Chen, "A Hybrid ARQ Protocol for the Communication System with Multiple Channels" in 2018 International Conference on Information and Communication Technology Convergence (ICTC), 17-19 Oct. 2018, Jeju, South Korea, DOI: 10.1109/ICTC.2018.8539714.
- [9] H. D. Le, V. Mai, C. T. Nguyen, A. T. Pham, "Design and analysis of sliding window ARQ protocols with rate adaptation for burst transmission over FSO turbulence channels" *IEEE/OSA Journal of Optical Communications and Networking*, **11**(5), 151 - 163, 2019, DOI: 10.1364/JOCN.11.000151.
- [10] F. A. Nada, "Service Time Distribution of Selective Repeat ARQ Protocol Used In Transmitting Short Messages Over Noisy Channels" in 2020 12th IEEE International Conference On Electrical Engineering (ICEENG-12), 7-9 Jul. 2020, Cairo, Egypt, .
- [11] F. A. Nada, "Steady State Analysis of Buffer Contents in a General Communication System" in 2019 9th International Conference on Intelligent Computing and Information Systems (ICICIS), 8-10 Dec. 2019, Cairo, Egypt. DOI: 10.1109/ICICIS46948.2019.9014737.
- [12] F. A. Nada, "Unfinished Work & Waiting Time of General Discrete-Time Communication System" in 2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE), 8-9 Feb. 2020, Aswan, Egypt. DOI: 10.1109/ITCE48509.2020.9047804.

## Modeling and Transformation from Temporal Object Relational Database into Mongoddb: Rules

Soumiya Ain El Hayat\* , Mohamed Bahaj

Department of Mathematics & Computer, Faculty of Science and Technologies, Settat, 26000, Morocco

### ARTICLE INFO

Article history:

Received: 13 June, 2020

Accepted: 22 July, 2020

Online: 25 August, 2020

Keywords:

Temporal database

Mongoddb

Nosql

Temporal object relational database

Migration

### ABSTRACT

With such a big volume of data growing tremendously every day, and the storage of important volume of data becoming increasingly more flexible, NoSQL (Not only SQL or Non-Relational) database are designed to store a large amount of information and are growing for big data systems in web analytics. It an approach does not require any specific schema and avoid the use of the joins to store or retrieve the information. To ensure the availability and scalability many industries are now replacing their object relational database in many systems by adopting NoSQL database technology for e-business applications. It is document-oriented databases which help in grouping data more logically. This paper describes a disciplined approach of migration and proposes a model transformation from temporal Object relational database (TORDB) Into Mongo db. Also, this paper approach deals with a novel data integration methodology in order to manipulate and the store the varying time data in Json (Java Script Object Notation) documents.

### 1. Introduction

In the recent years, a varying Time management database is one of the most interest systems of the information technology to support temporal features associated with data. It models the information in the real world. The time is an important property to characterize a data on the web. The temporal database has appeared since the 1980s, there were a several authors produced articles and books about varying-time, but the commercial adoption has been slow [1]. During a long time, developers and researchers developed several applications that support new data type concepts bases on the time. The temporal data processing is very important in dynamically evolving systems, industry, communication systems and also in systems processing sensitive data, which incorrect change would cause a great harm [2]. The literature on varying time management offers three Periods type for temporal data support: valid time period, transaction time period and Bitemporal data which combine both to make a complete description of data. Hence, the need to retain trace and audit the change made to a data and the ability to plan based on the past or future assumptions are important uses cases for temporal [3].

\*Corresponding Author: Soumiya Ain El Hayat ,FST , Settat, Morocco ,+212658521182, soumya.ainelhayat@gmail.com

Today, we notice the advent of big data. There is a revolution going on in databases system management. With the development of data acquisition technologies, the information to be stored expands strikingly in volume and velocity. NoSQL database have evolved intensely in the last years due to their flexible structure, less constrained than relational ones and offering faster access to information. Nosql is now used in many fields of industries and companies to support applications and systems not well served by relational and object relational database. NOSQL is released and widely used in many domains. Nosql database provides a mechanism for storage and retrieval of instructed data other than tabular or object relation used in relational and object relational database respectively [4]. The Nosql model fulfils the scalability problems. Nosql databases are mostly open source , non-relational , distributed and designed for large volumes of data across many clusters supporting replication and partitioning , parallel processing and what is usually called horizontal scaling [5].

One of the post popular and leading Nosql system management is Mongoddb. Mongo db is an open source database based on distributed document released in 2009. It stores data as JSON-Like document with dynamic schemas (The format is called BSON) [6]. Mongoddb is a document-oriented database which

holds a set of collection that are similar to relational database Tables where each collection contains a set of documents. One collection can hold different document with number of fields, content and size are not similar. It is a schema less that means the mongodb can be a distributed database and does not have a predefined schema so that allows providing additional data type and inserting new fields. Mongodb document is a set of key value pairs. Key values databases provide a hash table where the unique key and a pointer to a specific set of values are stored and data can be retrieved using the key. Mongodb is suitable solution for distributing data and managing balance between instances while using replication to increase the level of availability [7].

With the information industry dependent on the time, developing rapidly in recent years, the dataset using in different system are becoming extremely large in volume with a high variety of data. For this reason, Mongodb has been invented to overcome the limitation of relational and Object Relational Database, and provides new mechanisms for managing huge amount of data that are different from the typical relational and object relational Model. In addition, the mongodb is adopted to handle the huge evolution of temporal data in distributed environments in which continue to rise in complex applications and social network.

The main goal of this approach is to provide a reliable, reasonable and efficient method to convert the schema and migrate the temporal data from the implemented temporal object relational database into Mongodb system. Our proposed approach provides a new model transformation from object relational tables including Bitemporal data features towards documents-oriented databases based on Json files. Several rules are defined to facilitate the migration process.

## 2. Related Works

A research work in [8] proposed a generic standard-based architecture that allows Nosql systems focusing on mongo db to be manipulated using SQL query and seamlessly interact with any software supporting JDBC. Ajit Singh demonstrated data conversion to mongo db , in order to make data more interactive and innovative using the data stored in cloud [9]. The model transformation and data migration from relational database into Mongodb taken into consideration the query characteristics of each model, in addition, an algorithm to automate the migration are discussed in [10]. Another approach presented Algorithm for automatic mapping of relational database to mongodb using entity relationship (ER) Model to provide the conceptual schema and modelling relationship between the different entities [11]. A framework for mapping MySQL database to mongodb by developing an algorithm that uses the metadata stored in relational system as input is discussed in [12]. The work in [13] described a migration process from Object relational database to Nosql document database end provided a review of different proposed approach. An overview of Nosql to evaluate the scalability and efficiency in storage of data in oriented document database case study in order to show the representational format and querying management process of Mongodb [14]. The work in [15]introduced a disciplined approach called Jschem(Temporal

Json Schema) for the temporal management Json documents by creating a temporal json documents from conventional document that can vary over time, the generated document uses such features of temporal management data. The mapping Process of spatio temporal disaster data into Mongodb database using the data represented by the aspects of space and time is shown in [16].Boicea, Radulesu and agapin discussed the difference between oracle and Mongodb this comparison dealt with several criteria including theoretical , Concept , restrictions and query processing of SQL database get a document oriented database Management [17].

During our criticized analysis, we noticed that many studies are devoted to the analysis and the migration from relational database into Mongodb. Most of the previous approaches don't deal with the use of temporal data in order to make a complete description and efficient correctness to historical data. The temporal management features in Mongodb concepts was overlooked in many research, which reflects that the purpose for those researchers was only to provide methods to manipulate the data rather than on gain from the offered features of schema less of Nosql to handle a huge amount of historical data. In addition, works about the conversion from temporal database based on object relational into MongoDB are not as frequent.

From the all overhead works, we covered the transformation and the evaluation of mongodb to relational database. We conclude that mongodb has overcome many limitations of the conventional Relational and object relational database. The contribution of our work focuses on temporal aspect in emerging of temporal oriented-documents database. Our approach is based on semantic enriched mechanism which simplify the development of schema translation which enhanced by additional varying time data features. And formalize the rules to simplify the transformation process from temporal Object relational database into Mongodb which is based on temporal json file enriched with bitemporal data dimension. Our study examines the possibility to maintain and understand the structure of varying time attributes defining in TORDB tables and also to promote the storage of data into json documents using some semantics concepts.

## 3. Modelling and transformation Rules from TORDB into Mongodb: Phases

In this section, we present the most necessary phases for translation process. We are going to propose several rules that allow developing the determined schema. In the first, we present a comparison between temporal object relational database and Mongodb. After that we will explain the different elements composed TORDB design, and then we will define the mongodb schema including bitemporal data. In this paper, we will not explain the rules for translation of TORDB model into their equivalent in the TJson\_schema.

### 3.1. Comparison between Object relational database and Mongodb features:

Oracle Database is an Object-relational database management system produced and marketed by oracle Corporation [18], which



combine the relational database concept and object-oriented database capabilities. In object relational database, the user has the ability to create his own predefined type called UDT or User Defined Type, which is used to specify the structure and the behaviour of the data, based on Object mechanism in order to define and create complex data type. Therefore, ORDBs with time-varying features has addressed to enhance ORDB efficiency and makes a complete description of recorded data. For this reason, we create User-defined time approach. It is a time representation created by object relational techniques to satisfy specific needs of users. The User-defined time uses a bitempral data dimension which support both valid time and transaction time. The bitemporal data presents the changing knowledge of the changing world, hence, links data values with facts and also determines when the facts were valid, in order to provide history of data values and their changes over time. The following statement show the Bitemporal period Object to store valid and transaction Time attribute:

**Query 1: Bitemporal Object**

```
Create type Bit_period as object
(vt_Start date,
vt_end date,TT_start date,
TT_End date)/
CREATE TYPE bitemporal_period IS TABLE OF
bit_period;
```

Mongodb is an open source Nosql database based on oriented document structure. It was developed during 2007 by software called 10gen Company. Mongodb documents are stored in binary form that are similar to Json document model called BSON format, that supports such primitives data type (String, integer, date, Boolean, float and binary).the main features in mongodb are collections and documents. Mongodb documents have a flexible schema in which the collection dos not impose the necessary document schema. However, the temporal object relational databases require a table schema to be declared and created before inserting the data. Any temporal object table has a certain design that shows the relationships between them. Although mongodb does not support join operations as SQL databases, the relationship between documents can be represented using either the referenced or embedded concepts. The relationship in mongodb define how various documents logically dependant to each other. The relationship in mongodb can be expressed via embedded or referenced concepts. Where, embedded documents maintain all the related data in one document. These renormalized data representations allow applications handle and retrieve data from a single document.

In the following subsection, we focus on comparing mongodb to Temporal Object relational database. The table presents the main techniques of TORDB methodology and Mongodb.

On the other hand, the referenced document stores the relationships separately between document by adding id-Field that references or links from one document to another. This approach designs the normalized relationship. Actually, the

difficult part of transformation process is how we can convert the relationships of the existing temporal object relational database into Mongodb document.

Table 1 the Differences between TORDB and Mongodb features

TORDB	Mongodb
Database	Database
Temporal Object\Temporal Table	Temporal Collection
Row	Document
Column	Field
Data Type	Data Type
Primary Key	Id_Field
Simple UDT  REF   Nested table   Array   nested table (REF)	Referenced Document \ Embedded Document

In our work, each User Defined Time or UDT is converted to a MongoDB collection, in this example the collection name is customer contains Bitemporal period object. The customer\_table holds data rows of customer\_type objects. Also, the customer collection will store customer objects with the same attribute as documents in BSON (Binary encoded JSON) format including varying time attributes. Then statements below shows the creation of customer collection and how we can generate documents with bitemporal Object in mongodb :

**Document1: Example of Json file with bitemporal data**

```
db.createCollection("Customer") \ Creation of Customer
Collection
db.Customer.insert(
{
"Customer_Id":23
"Name": "Soumiya"
"Bitemporal_data":
{
"Vt_start": "2020-02-03" \ Valid time start
"Vt_End": "2028-02-28" \ Valid Time End
"TT_start": new Date() \ Transaction Time Start takes
the sysdate as Default value .
"TT_End": "9999-12-31" }} \ Transaction Time End sets
the value of insertion of transaction time to the highest
value ("31/12/9999")
```

3.1 Definition of the data Model

3.1.1 Temporal Object Relational Database Model (TORDB\_Model):

Our approach in [19] consists in structuring a generic model of semantic enrichment. We defined the component of a TORDB Model, which constructs a comprehensive schema of the temporal object-relational model. The TORDB model is expressed as a set of temporal typed table based on structured type ST and temporal structured type TST which include the user defined Time (Bitemporal Period) for specifying temporal period interval. Each

TST is declared as a set of simple attributes and varying time attributes. In this obtained Model, the temporal attributes are based on a bitemporal object that actually defined in a nested table as a collection type.

Generally, The TORDB model is denoted as three-tuples:

$$\text{TORDBModel} = \{ \text{TT} | \text{TT} = (\text{TTs}, \text{STs}, \text{Tm}) \} \quad (1)$$

where:

- TT s is set of temporal and non temporal typed table, STs is a set of temporal structured type or simple structured type , and Tm is a time-varying Period.

The sets TTs, STs and Tm are defined as follows:

$$\text{STs} = \{ \text{Sn} , \text{S} , \text{AT} \} \quad (2)$$

where Sn is the name of a structured type , S is the super type of ST, and AT is a set of structured type's attributes:

$$\text{AT} = \{ \text{A} | \text{A} := \{ \text{N}, \text{T}, \text{D}, \text{NL}, \text{BitT}, \text{M} \} \}, \quad (3)$$

where

N: is the name of attribute, T: data type and can be primitive, UDT or reference. NL: if the attribute allows Null value or not. D: default value. M: means if the AT is a single valued or collection valued.

BitT: denotes if the attribute contains a bitemporal object is defined:

$$\text{BitT} = \{ (\text{VT\_Start}, \text{VT\_End}, \text{TT\_start}, \text{TT\_End}) \} \quad (4)$$

$$\bullet \text{TTs} = \{ \text{typedtable} | \text{Ttable} = \{ \text{TTn}, \text{STn}, \text{PK}, \text{Tp} \} \} \quad (5)$$

where: TTn is the name of typed table, STn is the name of the structured type based upon which TT is defined, PK : primary key, TP: means if the TT is temporal or not.

**Association:** For each relationship Rel where RelType= “associate with” is translated into:

$$\text{TT\_Association} = \{ \text{TT} | \text{TT} = (\text{TTn}, \text{STn}, \text{AT} \cup (\text{Ref}(\text{ST}')), \text{PK}, \text{Bit\_P}) \} \quad (6)$$

In association relationship, we define a method for storing the reference type of the structured type ST' (REF(ST')) referencing to the class which is related as a Nested Table collection with Bitemporal data dimension.

**Aggregation:** each relationship Rel where RelType= “Aggregation” is expressed as a collection of UDT (User Defined Type) representing the C' class that interacts with class C.

$$\text{TT\_Aggregation} = \{ \text{TT} | \text{TT} = (\text{TTn}, \text{STn}, \text{AT} \cup (\text{NT}(\text{UDT}(\text{ST}')), \text{PK}, \text{Bit\_P}) \} \quad (7)$$

**Composition:** each relationship rel where RelType = “composition” is mapped into an attribute typed as a nested table maintains the attributes of the ST' corresponding to a class part C'.

$$\text{TT\_Composition} = \{ \text{TT} | \text{TT} = (\text{TTn}, \text{STn}, \text{AT} \cup \text{NT}(\text{ST}'), \text{PK}, \text{Bit\_P}) \} \quad (8)$$

**Inheritance:** each relationship rel where RelType = “inheritance”, the class C' inherits all the properties of its super class that are matching to ST'. For the defining the inheritance relationship in Object relational model, we add the keyword UNDER during the creation of ST that represents the child or sub class C.

$$\text{TT\_ihneritance} = \{ \text{TT} | \text{TT} = (\text{STn}, \text{ST.ATUSuper\_T.AT}, \text{PK}, \text{Bit\_P}) \} \quad (9)$$

### 3.1.2. Temporal Json schema (Tjson-schema):

Tjson-schema is a representation of Json document with historical data that is enhanced with additional semantic data to offer a new description of mongodb document. We have chosen the most commonly used in the oriented documents that able to identify the relevant collections, their documents and their relationships. The model constructs a data reference design in order to facilitate the understandability of metadata stored in Json document integrating Bitemporal Data. Also, it overcomes the complications that occur during the transformation process.

Tjson-schema can be defined as follows:

$$\text{Tjson-schema} = \{ \text{TJ} | \text{TJ} = \{ \text{Col\_name} , \text{Tdoc} , \text{RELCOL} \} \} \quad (11)$$

where:

- Col\_name : each collection has a name , where the collection can be defined as a set of temporal and simple documents.
- Tdoc: denotes temporal Json documents including varying Time fields. Generally, the document in mongo db is a set of a key value pairs:

$$\text{TDOC} = \{ \text{doc\_ID}, \text{Fields}, \text{Bitemporal\_Period}, \text{Primary Key} \}$$

- The Json document uses Object Id as a default id which is generated during the creation of mongodb document.
- Fields = means a set of fields and can be identified by the following elements:

$$\text{Field} = \{ \text{Field\_Name} , \text{Field\_Name} \} \quad (12)$$

where:

Field\_Name= is the name of the field.

Field\_Type= means data type (integer, string, date)

- Bitemporal\_Period= means the embedded document in Json document which composed by:

$$\text{Bitemporal\_Period} = \{\text{VT\_start}, \text{VT\_end}, \text{TT\_start}, \text{TT\_end}\} \quad (13)$$

- Relcol=Relations Tjson-schema. Each collection has a set of relationship between documents and can be identified as follow:

$$\text{Relcol} = \{\text{RelType}, \text{RelName}, \text{Dircol}\} \quad (14)$$

where:

Reltype= each collection or Document has a relationship Type with other documents. The Reltype offers 2 types of relations: Embedded document of referencing document.

Dircol= is name of doc' that is related to the document Doc.

**Document 2: Account** Document extraction with Mongoddb

```
{
  "_id":1345672
  "Num_Account": 1
  "Type_Account": "Saving_Account"
  "account_h":
    {
      "VT_start": ISODATE("2010-01-01")
      "VT_End": ISODATE("2012-11-30")
    }
  "TT_Start": ISODATE ("2010-01-03")
  "TT_End": ISODATE("2012-12-04")
  "Customer":{
    "_id": "145673" }}
```

3.2. Transformation Rules from TORDB into mongoddb:

3.2.1. Association

R1: For the two UDTs Namely Customer\_Type and Account\_Type contain Bitemporal data and are related with association (1..N) relationship, using reference mechanism which allow retrieving data rapidly without using join between tables. The transformation of association in Mongoddb consists on generating a new document where the Customer\_Type will be referenced in Account\_type document and the object type will be embedded in the both document.

$$\text{Col1} = \{\text{NameCol1}, \text{NameDoc}, \text{Doc(Fields)}, \text{Embedded(Bitemporal\_Period)}\} \quad (14)$$

$$\text{Col2} = \{\text{NameCol2}, \text{NameDoc}, \text{Doc(Fields)}, \text{Object\_ID(Doc)}, \text{Embedded(Bitemporal\_Period)}\} \quad (15)$$

The example below shows the structure of customer and account document, also the TORDB query Statement:

The creation will proceed according to the following syntax:

$$\text{Col1} = \{\text{Customer}, \text{Customer\_Json Document}, (\_id, \text{Id\_Cust}, \text{Name}), \text{Customer\_h} (\text{VT\_start}, \text{VT\_End}, \text{TT\_Start}, \text{TT\_End})\}$$



Figure1. An example of Association 1 to N relationship between Customer and Account

$$\text{Col2} = \{\text{Account}, \text{Account\_Json Document}, (\_id, \text{Num\_Account}, \text{Type\_Account}), \text{Customer} (\_id), \text{account\_h} (\text{VT\_start}, \text{VT\_End}, \text{TT\_Start}, \text{TT\_End})\}$$

The defined TORDB Query for the corresponding example:

**Query2:** creation statement for account table

```
CREATE TYPE account_t AS OBJECT
(acc_no NUMBER,
acc_type varchar(20),
Customer REF customer_T,
account_h bitemporal_period);
CREATE TABLE account_table OF account_t NESTED
TABLE account_h STORE AS account_h_tab ;
```

On the other hand, in the case of association many to many (N,N), the both document representing account and Branch\_bank (see the example below), will be mapped in composition between the both document where each document integrates referenced document of another. The transformation result is:

$$\text{Col1} = \{\text{NameCol1}, \text{NameDoc}, \text{Doc(Fields)}, \text{Object\_ID(Doc)}, \text{Embedded(Bitemporal\_Period)}\} \quad (16)$$

$$\text{Col2} = \{\text{NameCol2}, \text{NameDoc}, \text{Doc(Fields)}, \text{Object\_ID(Doc)}, \text{Embedded(Bitemporal\_Period)}\} \quad (17)$$

The example shows the structure of branch\_bank and customer Json document representing N to N Relationship:

**Document 3:** Extraction of branch bank Json file

Banch\_bank document:

```
{
  "_id":45679
  "branch_pk":1
  "city": "Casablanca"
  "phone": 06453278
  "Bitemporal_Period":
    {
      "VT_start": DATE("2013-03-01")
      "VT_End": DATE("9999-12-31")
      "TT_Start": DATE ("2013-03-03")
      "TT_End":DATE("9999-12-31")
    }
  "Customer":{
    "Customer_id": " 2334" }}
Customer_Json Document :
{
  "_id":2334
```

```

"Num_Cust":
"Name_cust":
"Bitemporal_Period":
{
  "VT_start": DATE("2010-01-01")
  "VT_End": DATE("2012-11-30")
  "TT_Start": DATE("2010-01-03")
  "TT_End": DATE("2012-12-04")}
"Branch_bank":{
  "_id": "45679" }}
    
```

3.2.2. Composition:

In ORDB, the composition relationship is represented by declaring Nested table in the whole class which stores all attributes of the whole part. This relationship will be converted in a strong composition in mongo db between the account and balance documents. The transformation model generates one document account contains embedded collection of balance documents. The result of composition relationship:

Col={NameCol, NameDoc, Doc (Fields), Embedded (Bitemporal\_Data), Embedded( Collection (Part\_Fields))} (18)

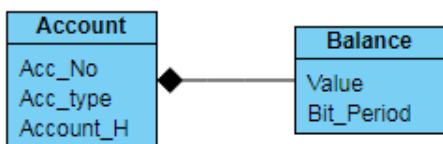


Figure2. An example of Composition Class diagram

The corresponding TJson document of the example above:

Col={ Account, Account \_Json Document , ( \_id, Num\_Account, Type\_Account), account\_h (VT\_start, VT\_End,TT\_Start, TT\_End),Balance{Value,Balance\_h(VT\_start,VT\_End,TT\_Start, TT\_End))}

Account document embedded the balance document:

**Document4:** Extraction of balance Json File

```

{
  "_id":65789
  "Num_Account": 2
  "Type_Account": "saving_account"
  "Bitemporal_Period":
  {
    "VT_start": "2015-01-01"
    "VT_End": "9999-12-31"
    "TT_Start": "2015-01-01"
    "TT_End": "9999-12-31"
  }
  "Balance":[ {
    "value":324561.098

    "Blance_h":
    "VT_start": "2015-01-01"
    "VT_End": "9999-12-31"
  }
    ]
}
    
```

```

"TT_Start": "2015-01-01"
"TT_End": "9999-12-31" }
{
  " value": 4356.098

  " Blance_h ":
  {
    "VT_start": "2019-07-24"
    "VT_End": "2019-09-23"
    "TT_Start": "2019-07-25"
    "TT_End": "2019-09-27"
  }
}}
    
```

3.2.3. Aggregation:

An aggregation represents a binary relationship. It is a weak form of composition where the part is shareable and independent from the whole, and its properties can be linked with more than one whole class component. For example, if all the composites (whole) are deleted, the sheared part can be existed.

For the aggregation Relationship, the relation is identified a collection of (UDT) in addition of bitemporal data. Then, the branch bank can be composed by one or more than one shareable and existence independent collection. In mongoddb, the Aggregation relationship will be considered as a referencing collection of \_ID document represented the account document:

Col1={NameCol1,NameDoc,doc(Fields),Embedded(Bitemporal\_Period)} (19)

Col2={NameCol2,NameDoc2,Doc(Fields),Referencing(collection(Doc1)+BitemporalPeriod),Embedded(Bitemporal\_Period)} (20)

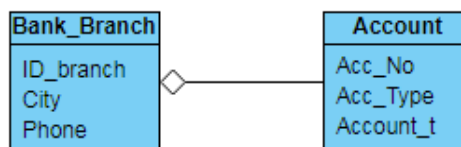


Figure3. Aggregation relationship Class

The transformation will be defined as follow:

Col1={ Account, Account \_Json Document , ( \_id, Num\_Account, Type\_Account), account\_h (VT\_start, VT\_End,TT\_Start, TT\_End))

Col2={Branch\_Bank, Banch\_bankdocument, { \_id, branch\_id, city,phone},Account( collection( \_ID , account\_h (VT\_start, VT\_End,TT\_Start, TT\_End)) , Banch\_bank\_h(VT\_start, VT\_End,TT\_Start, TT\_End))}

The following TJson document presents the example above:

**Document5:** Temporal Json File for Balance Bank

```

{
  "_id": 678
  "branch_id": 9
  "city": "Casablanca"
}
    
```



```

"phone":074467543

Accounts :[ {
  "Account":56432
  "Account_h":
    {
      "VT_start": "2019-08-22"
      "VT_End": "9999-12-31"
      "TT_Start": "2019-08-25"
      "TT_End": "9999-12-31" }}
    {
      "Account":980654
      "Account_h":
        {
          "VT_start": "2018-07-24"
          "VT_End": "2019-09-23"
          "TT_Start": "2018-07-25"
          "TT_End": "2019-09-27" }}
    }}

```

```

Col2={NameCol2,NameDoc,doc(Fields+doc1(Fields)), (22)
Embedded(Bitemporal_Period)}}

```

The details are illustrated in the following example:

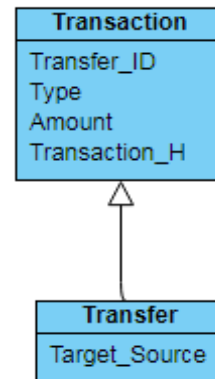


Figure4. Inheritance Relationship Example

The example below presents the TORDB Query Creation for Branch Bank and Account Tables:

**Query 3:** creation statement of the aggregation relationship

```

Create type Account_NT as object
(Account Account_T,
Account_h bitemporal_period) /

CREATE TYPE Account IS TABLE OF Account_NT ;
Create type branch_bank_type as object (
Branch_Num number,
Address varchar(40),
City varchar(20),
Accounts Account,
Branch_h bit_period)

CREATE TABLE branch_bank_table of
branch_bank_type NESTED TABLE accouns STORE
AS accounth_tab, NESTED TABLE branch_h STORE
AS branch_tab ;

```

3.2.4. Inheritance:

Is called also a generalization is a relationship between two classes or more, where one entity represent a parent or super class and the other one is considered as a child or sub class. The child inherits the behavior and all the properties of the parent.

The inheritance is a very important In ORDB. For the creation statement of UDT that represents the inheritance, we add the keyword under for the sub class. This model will be transformed in Mongo db by generating two documents separately modeling transaction and transfers Types. The transfer document maintains the same structure of transaction type with Additional properties of subtype is defined in the usual way with time varying features.

```

Col1={NameCol1,NameDoc,Doc(Fields), (21)
Embedded(Bitemporal_Period)}}

```

```

Col1={Transaction, Transaction_Json Document doc(_id,
trans_ID, Type_Trans, Amount), Transaction_h h (VT_start,
VT_End,TT_Start, TT_End)}

```

```

Col2={Transfer, Transfer_Json Document ,( _id, trans_ID,
Type_Trans, Amount, target_source), Transaction_h h (VT_start,
VT_End,TT_Start, TT_End)}

```

The following TJson document presents inheritance relationship:

**Document6:** Temporal Json File for Transfer class

```

{
  "_id":87654
  "trans_ID": 34
  "Type_Trans": "transfer"
  "Account":9876
  "Amount": 23450
  "Transaction_h":
    {
      "VT_start": "2018-04-02"
      "VT_End": "2018-04-05"
      "TT_Start": "2018-04-02"
      "TT_End": "2018-04-04" }}
  "target_source": "soumiya}

```

Inheritance Creation Query in TORB using Under Keyword:

**Query 4:** creation statement for inheritance relationship

```

Create or Replace type Transaction_Type as object
(trans_ID number,
Type_Trans varchar(20),
Amount number,
Account REF account_type,
Transaction_h bitemporal_period) NOT FINAL ;
Create table transaction_table of Transaction_type
NESTED TABLE Transaction_h STORE AS
transactionh_tab;

Create or Replace type Transfer_T UNDER
Transaction_Type (target_source varchar(20)) ;

```

#### 4. Conclusion

This work provides a new approach to the problem of transforming and migrating of massive data based on temporal object Relational database into Temporal Json Documents using MongoDB. We presented the basics phases of modeling and converting Temporal Object relational including User defined time with Bitemporal data into oriented document database. To do that, we formalized the rules by specifying the basics steps involved in the temporal object-relational database design, in order to capture the relationship's type between objects. This solution is done by providing a TORDB Model from an Object relational database that contains user defined Time, and we use it as an input enriched with additional semantic data, which is translatable into any of the target database schemas. Furthermore, Temporal Json document design has defined including the varying time data by exploiting the range of powerful concepts provided by Nosql database. Currently, any paper deals with the transformation process and its functionalities from TORDB into MongoDB. However, this article proposes an analyst part for our future work, where we will implement a complete framework for automating the migration mechanism of the schema, structure and the data into MongoDB without any human interference.

#### Conflict of Interest

The authors declare no conflict of interests regarding the publication of this paper.

#### References

- [1] K. Kulkarni., J.E. Michels, "Temporal features in SQL: 2011", *ACM SIGMOD Record*, **41**(3), 34--43, 2012. , <https://doi.org/10.1145/2206869.2206883>
- [2] M. Kvet, K. Matiasco, "Transaction management in the fully temporal system", in 2014 16th International Conference on Computer Modelling and Simulation (UKSim), Cambridge, UK, 2014. <https://doi.org/10.1109/UKSim.2014.26>
- [3] M. Kaufmann, P. M. Fischer, N. May, D. Kossmann, "Benchmarking Bitemporal Database Systems: Ready for the Future or Stuck in the Past?" *EDBT*, 738-749, 2014. DOI: 10.5441/002/edbt.2014.80
- [4] Z.Gansen, L.Qiaoying, L.Libao, L.Zijing. "Schema Conversion Model of SQL Database to NoSQL". In 2014 Ninth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing. Guangdong, China, 2014. DOI: 10.1109/3PGCIC.2014.137
- [5] Byrne, B, Nelson, David and Jayakumar, "Big Data Technology - Can We Abandon the Teaching of Normalisation?", in 2017 19th annual International Conference on Education and New Learning Technologies, Barcelona, Spain, 2017. DOI: 10.21125/edulearn.2017.1113
- [6] A. Boicea, F. Radulescu and L. I. Agapin, "MongoDB vs Oracle -- Database Comparison", In 2012 Third International Conference on Emerging Intelligent Data and Web Technologies, Bucharest, 2012. DOI: 10.1109/EIDWT.2012.32
- [7] T. Fouad, M. Bahaj. "Model Transformation From Object Relational Database to NoSQL Document database". In 2019 International Conference on Networking, Information Systems & Security, Rabat, Morocco , 2019 . <https://doi.org/10.1145/3320326.3320381>
- [8] R. Lawrence, "Integration and Virtualization of Relational SQL and NoSQL Systems Including MySQL and MongoDB", in 2014 The International Conference on Computational Science and Computational Intelligence , Las Vegas, NV, USA, 2014 , DOI: 10.1109/CSCI.2014.56
- [9] A. Singh, "Data Migration from Relational Database to MongoDB". *Global Journal of Computer Science and Technology: C Software & Data Engineering* ,Vol 19, Issue 2 ,2019. <http://dx.doi.org/10.2139/ssrn.3372802>
- [10] T. Jia, X. Zhao, Z. Wang, D. Gong and G. Ding, , "Model Transformation and Data Migration from Relational Database to MongoDB", in 2016 IEEE International Congress on Big Data (BigData Congress), San Francisco, CA, 2016. DOI: 10.1109/BigDataCongress.2016.16
- [11] L. Stanescu, M. Brezovan, D.D. Burdescu , "An Algorithm for Mapping the Relational Databases To MongoDB--A Case Study", *International Journal of Computer Science & Applications*, **14**(1), 2017.
- [12] L.Stanescu, M.Brezovan , CS. Spahui, DD. Burdescu , "A framework for mapping the mysql Databases to MongoDB-- Algorithm, Implementation and experiments. *International Journal of Computer Science and Applications*, **15**(1), 65 – 82, 2018
- [13] D. Chauhan, K.L. Bansal, "Using the Advantages of NoSQL: A case study on MongoDB", *International Journal on Recent and Innovation Trends in Computing and Communication*, **5**(2), ISSN 232/-8169.2017
- [14] S. Brahmia, Z. Brahmia, F. Grandi, R. Bouaziz, "A Disciplined Approach to Temporal Evolution and Versioning Support in JSON Data Stores," In *Emerging Technologies and Applications in Data Processing and Management*", IGI Global, 114-133, 2019. DOI: 10.4018/978-1-5225-8446-9.ch006
- [15] Y. Widyani, H. Laksmiwati and E. D. Bangun, "Mapping spatio-temporal disaster data into MongoDB", In 2016 International Conference on Data and Software Engineering (ICoDSE), Denpasar, Indonesia, 2016. DOI: 10.1109/ICODSE.2016.7936157
- [16] S .Ain El Hayat,F. Toufik,M. Bahaj, "UML/OCL based design and the transition towards temporal object relational database with bitemporal data". *Journal of King Saud University - Computer and Information Sciences*, 2019. <https://doi.org/10.1016/j.jksuci.2019.08.012>

## Modelling and Simulation of 3-DOF Lower Limb Rehabilitation Robot using Force Feed Forward Control

Pham Van Bach Ngoc<sup>1,\*</sup>, Le Thi Hong Gam<sup>2,3</sup>, Dam Hai Quan<sup>3,4</sup>, Bui Trung Thanh<sup>3</sup>, Nguyen Luong Thien<sup>1</sup>

<sup>1</sup>Space Technology Institute, Vietnam Academy of Science and Technology, Ha Noi, 100000, Vietnam

<sup>2</sup>Faculty of Physics, Thai Nguyen University of Education, Thai Nguyen, 250000, Vietnam

<sup>3</sup>Mechanical Engineering, Hung Yen University of Technology and Education, Hung Yen, 160000, Vietnam

<sup>4</sup>Electronics Telecommunications, Thai Nguyen College of Economics and Finance, Thai Nguyen, 250000, Vietnam

### ARTICLE INFO

Article history:

Received: 13 June, 2020

Accepted: 22 July, 2020

Online: 25 August, 2020

Keywords:

Lower Limb Rehabilitation robot

Perceiving reality

Force control

3-DOF

Feed forward control

### ABSTRACT

*This research the authors design the intelligent control for 3-DOFs lower limb rehabilitation robot base on the complex dynamics equation. The Force Feed-Forward Method (FFM) is promote to control the of 3-DOFs lower limb rehabilitation robot including dynamics characteristics. The robot can sense the force of the therapist which exerted on the robot and patient's leg, then produces necessary forces through joints at the hip, knee, and ankle. The force feedforward controller is used to compensate the force generated by the therapist to perform patient-active exercises. In this paper, firstly authors briefly introduce 3-DOFs lower limb rehabilitation robot, next the kinematics and dynamics equation of 3-DOFs lower limb rehabilitation robot established base on Lagrange-Euler method are presented, and then the control method is introduced. Last, the performance of the proposed control methods has been confirmed by numerical simulations of the robot in all three joints: hip, knee, and ankle.*

## 1. Introduction

The rehabilitation for patients after stroke using rehabilitation robot has three stages, early rehabilitation stage, intermediate training stage, and advanced training stage. In order to get a satisfying control effect, various control methods have been used for Robot which is suitable for each stage of the rehabilitation process. In the early rehabilitation stage, the position control strategy with the trajectory tracking control method is appropriate to continue repetitive passive exercises. After the early rehabilitation stage with continued repetitive passive exercises following the desired trajectory, the patient's movement ability and muscle strength could be partial recovery. From this time, the fixed exercises can cause negative reactions, limiting creativity, proactive and active training of patients. Thus, the force control method aims to create the interaction between patient and robot is a very important requirement in this phase of training. In [1], the authors proposed a new method to control the force when the patient performs lower limb extensions. The aim of this control method is increasing real-time resistance to lower limb force asymmetry. The force-position hybrid control method in high rate

is applied in Rutgers Ankle system [2] to resistive force on the foot of the patient to perform virtual reality-based exercises. In [3] authors use the combination of impedance controller and PD controller to control hip, knee, ankle joint of a 3-DOF rehabilitation robot. However, with the impedance controller, its parameters directly affect the effectiveness of the patient's training. If parameters of the impedance controller are low, the patient's foot motion may be out of range of physiological motion. Conversely, if parameters of the impedance controller are high, it can reduce the patient's proactive training [4]. A force feedback system with sensors/observers and encoders can measure the angle of rotation of the joints and can transfer feedback forces to the controller to help the robot perceiving reality and supporting for patients [5]. Force feedback control has many advantages and has been studied worldwide for upper limb exoskeleton [6]. For the lower limb rehabilitation robot, several studies using a force feedback controller have been researched and developed. In [7], the exoskeleton has six flexible force sensors are placed in its shoe and two load cells are mounted between the end of the piston rod and the lower leg joint. Force feedback control is realized by comparing ground reaction force and applied force of the hydraulic cylinder. A force control strategy is developed in [8] to exert user-

\*Pham Van Bach Ngoc, 18 Hoang Quoc Viet Street, pbnoc@imech.vast.vn

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050474>

denied assistive force. The force controller involves two sets of PD plus feed-forward controllers corresponding to low- and high-force ranges, in which the interaction force between the subject and the robot (the feedback force) is calculated based on the Hooke's law. In [9], the authors introduced a rehabilitation robot which uses a force-position hybrid fuzzy controller to limit movement in the desired direction, while maintaining a constant force along the moving direction. The position and direction of movement of the patient's foot are collected, after which a resistance force is used to rehabilitation for stroke patients. The selection of impedance parameters is not rooting to make. These parameters need to be re-selected during the training process to suit the patient's ability and progress [10]. Therefore, it is really necessary to have adaptive control strategies to improve the dynamic control performance of the rehabilitation robot. In [11], "robot-in-charge" and "patient-in-charge" control strategies were applied for LOPES, in which an impedance controller is used to improve the proactive training of the patient. This study's result shows that the assistive force from the robot is effectively delivered to lower-limbs and the force feed-forward controller can be used to improve the force control performance of the lower limb rehabilitation robot. In [12], authors presented the structure of a novel force-position hybrid controller to encourage patient activeness to gait rehabilitation using driven gait orthotics. In this research, the proportional derivative (PD) close loop position controller and forced controller are applied. Impedance control strategy is a great approach for encouraging the patient's active performing and real-time adjust the manipulator's position and interactive force. In [13], the authors proposed a position-force controller, applied to gait rehabilitation robot, ALEX. With this controller, ALEX can guide patients to move in physiological trajectories.

This paper focus on develop a new control method Force Feed-forward for 3-DOFs lower limb rehabilitation robot for perceiving reality. The force feed-forward controller is designed for robot system can track the position, magnitude, and direction of the contact force between the patient's lower limb and the robot. Moreover, it can produce necessary force through joints at the hip, knee, and ankle. To accomplish this task, we considered a problem, which consists of two steps. Firstly, the therapist exerts a force at links of the robot, the force is measured by force sensors placed at there. Meanwhile, the angular joint trajectories are determined by encoders located at joints. Secondly, a force feed forward controller is used to control DC motors located at joints to compensate the force generated by the therapist to perform patient-active exercises. Furthermore, force sensors also measured the force exerted by the patient's lower limb as the input signal of the FFM controller. Finally, numerical simulation results of the robot system with FFM controller will indicate whether the controller has achieved the desired control performance. Moreover, the simulation result is used to estimate the maximum torque of a DC motor for the next design steps of robot control.

**2. Mathematical model**

*2.1. Robot's mechanism*

The mechanical structure of the robot provides supports for rehabilitation training. It is a serial robot that has three main joints of the lower limb, including a joint at the hip, a joint at the knee, and a joint at the ankle with three links: base link, thigh, leg, and a

foot platform. Movement of each joint is driven by a geared DC motor. The mechanism of Robot is shown in Fig. 1. The linkage diagram of 3 DOFs rehabilitation Robot is shown in Figure 2. A coordinate system is attached to each link of the Robot. The coordinate system is attached to the base link with its origin located at the hip joint and axis pointing to the right. This coordinate system coincides with the base coordinate system . Because all of the joints axes are parallel to each other, so all the twist angles and translational distances are zero. To perform force control for the robot using FFM control method, three DC motors are installed at joints with encoders are used to measure rotation angle, and force sensors are used to measure the force exerted by physiotherapists in training stages or force exerted by the user's foot in practice stages.

*2.2. Kinematic model*

Figure 2 show the linkage diagram of 3-DOFs rehabilitation Robot, the Denavit-Hartenberg (D-H) parameters of Robot are given in Table 1, where the parameters  $a_i$  with  $(i = 1, 2, 3)$  are constant, the variables  $\theta_i$  vary by rotation around the  $z_i$  axis with  $(i = 1, 2, 3)$  .

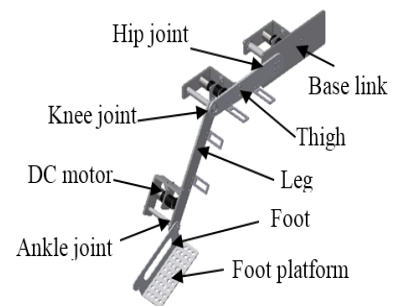


Figure 1: Mechanism of 3-DOF Rehabilitation Robot [1]

Table 1: D-H Parameters of 3-DOF Rehabilitation robot

Joint $i$	$\alpha_i$	$a_i$	$d_i$	$\theta_i$
1	0	$l_1$	0	$\theta_1$
2	0	$l_2$	0	$\theta_2$
3	0	$l_3$	0	$\theta_3$



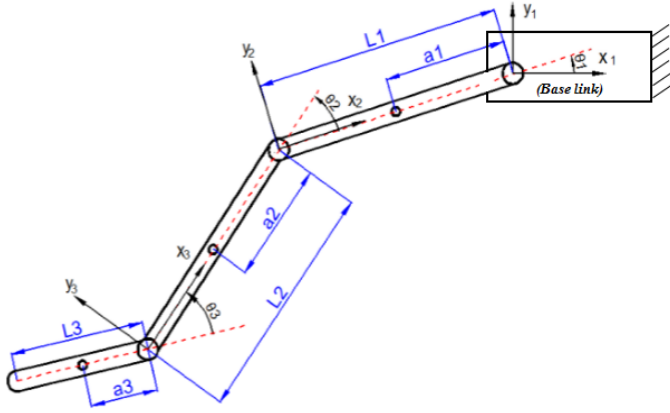


Figure 2: Linkage diagram of 3 DOF Rehabilitation Robot

The Denavit-Hartenberg transformation matrices are derived by substituting the D-H link parameters into equations (1) and (2), [14]:

$${}^{i-1}A_i = \begin{bmatrix} \cos \theta_i & -\sin \theta_i \cos \alpha_i & \sin \theta_i \sin \alpha_i & a_i \cos \theta_i \\ \sin \theta_i & \cos \theta_i \cos \alpha_i & -\cos \theta_i \sin \alpha_i & a_i \sin \theta_i \\ 0 & \sin \alpha_i & \cos \alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

$${}^0A_i = {}^0A_1 {}^1A_2 \dots {}^{i-1}A_i \quad (2)$$

The transformation matrix of 3DOF rehabilitation robot is:

$${}^0A_3 = \begin{bmatrix} \cos \theta_{123} & -\sin \theta_{123} & 0 & l_1 \cos \theta_1 + l_2 \cos \theta_{12} + l_3 \cos \theta_{123} \\ \sin \theta_{123} & \cos \theta_{123} & 0 & l_1 \sin \theta_1 + l_2 \sin \theta_{12} + l_3 \sin \theta_{123} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

where  $\theta_{12} = \theta_1 + \theta_2$  and  $\theta_{123} = \theta_1 + \theta_2 + \theta_3$

From equation (3), we can get the direct kinematic model and inverse kinematic model of Robot.

### 2.3. Dynamic model

The dynamic model of 3DOF Robot is written in matrix form as [14]:

$$\mathbf{M}(q)\ddot{q} + \mathbf{V}(q, \dot{q}) + \mathbf{G}(q) = \boldsymbol{\tau} \quad (4)$$

where

$$\mathbf{M}(q) = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$$

is the manipulator inertia matrix, with

the elements  $m_{ij}$  of  $M(q)$  have been calculated as follow:

$$m_{11} = a_1^2 m_1 + (a_2^2 + l_2^2 + 2a_2 l_1 \cos \theta_1) m_2 + (a_3^2 + l_1^2 + l_2^2 + 2l_2 l_1 \cos \theta_2 + 2a_3 l_1 \cos \theta_{23} + 2a_3 l_1 \cos \theta_3) m_3$$

$$m_{12} = m_{21} = (a_2^2 + a_2 l_1 \cos \theta_2) m_2 + (a_3^2 + l_2^2 + 2a_3 l_2 \cos \theta_3 + a_3 l_1 \cos \theta_{23} + l_1 l_2 \cos \theta_2) m_3$$

$$m_{13} = m_{31} = (a_3^2 + a_3 l_1 \cos \theta_{23} + a_3 l_2 \cos \theta_3) m_3$$

$$m_{22} = a_2^2 m_2 + (a_3^2 + l_2^2 + 2a_3 l_2 \cos \theta_3) m_3$$

$$m_{23} = m_{32} = (a_3^2 + a_3 l_2 \cos \theta_3) m_3$$

$$m_{23} = m_{32} = (a_3^2 + a_3 l_2 \cos \theta_3) m_3$$

$$m_{33} = (a_3^2 + \frac{1}{12} l_3^2) m_3$$

where the parameters  $m_i$ ,  $l_i$  are the mass and the length of link  $i$ , respectively;  $a_i = \frac{l_i}{2}$

$\mathbf{V}(q, \dot{q}) = [V_1 \ V_2 \ V_3]^T$  is the velocity coupling vector, the elements  $V_i$  of vector  $V(q, \dot{q})$  have been calculated as follow:

$$V_1 = -a_2 l_1 m_2 \sin \theta_1 \dot{\theta}_1^2 - 2(l_1 l_2 m_3 \sin \theta_2 + a_3 l_1 m_3 \sin \theta_{23}) \dot{\theta}_1 \dot{\theta}_2 - (l_1 l_2 m_3 \sin \theta_2 + a_3 l_1 m_3 \sin \theta_{23}) \dot{\theta}_2^2 - 2(a_3 l_1 m_2 \sin \theta_3 + a_3 l_1 m_2 \sin \theta_{23}) \dot{\theta}_1 \dot{\theta}_3 - 2(a_3 l_2 m_3 \sin \theta_3 + a_3 l_1 m_3 \sin \theta_{23}) \dot{\theta}_2 \dot{\theta}_3 - (a_3 l_2 m_3 \sin \theta_3 + a_3 l_1 m_3 \sin \theta_{23}) \dot{\theta}_3^2$$

$$V_2 = (l_1 l_2 m_2 \sin \theta_2 + a_3 l_1 m_2 \sin \theta_{23}) \dot{\theta}_1^2 - 2a_3 l_2 m_3 \sin \theta_3 \dot{\theta}_1 \dot{\theta}_3 - 2a_3 l_2 m_3 \sin \theta_3 \dot{\theta}_2 \dot{\theta}_3 - a_3 l_2 m_3 \sin \theta_3 \dot{\theta}_3^2$$

$$V_3 = (a_3 l_1 m_2 \sin \theta_3 + a_3 l_1 m_2 \sin \theta_{23}) \dot{\theta}_1^2 + 2a_3 l_2 m_3 \sin \theta_3 \dot{\theta}_1 \dot{\theta}_2 + a_3 l_2 m_3 \sin \theta_3 \dot{\theta}_2^2$$

$$\mathbf{G}(q) = [G_1 \ G_2 \ G_3]^T$$
 is the gravitational forces vector, with the elements  $G_i$  of vector  $\mathbf{G}(q)$  are determined:

$$G_1 = m_1 g a_1 \cos \theta_1 + m_2 g (l_1 \cos \theta_1 + a_2 \cos \theta_{12}) + m_3 g (l_1 \cos \theta_1 + l_2 \cos \theta_{12} + a_3 \cos \theta_{123})$$

$$G_2 = m_2 g a_2 \cos \theta_{12} + m_3 g (l_2 \cos \theta_{12} + a_3 \cos \theta_{123})$$

$$G_3 = m_3 g a_3 \cos \theta_{123}$$

$$\text{with } g = 9.8 \text{ (m/sec}^2\text{) is the gravity.}$$

$$\mathbf{q} = [\theta_1 \ \theta_2 \ \theta_3]^T$$
 is vector of generalized Lagrange coordinates,

$$\mathbf{F} = [f_1 \ f_2 \ f_3]^T$$
 is the vector of generalized forces.

Substituting matrix  $\mathbf{M}(q)$ , vector  $\mathbf{V}(q, \dot{q})$  and vector  $\mathbf{G}(q)$  into equation (4), the torques  $f_1, f_2, f_3$  are calculated as follow:

$$\begin{aligned} \tau_1 = & \begin{bmatrix} a_1^2 m_1 + (a_2^2 + l_2^2 + 2a_2 l_1 \cos \theta_1) m_2 + \\ \left( a_3^2 + l_1^2 + l_2^2 + 2l_2 l_1 \cos \theta_2 + 2a_3 l_1 \cos \theta_{23} \right) m_2 \\ + 2a_3 l_1 \cos \theta_3 \end{bmatrix} \ddot{\theta}_1 \\ & + \begin{bmatrix} (a_2^2 + a_2 l_1 \cos \theta_2) m_2 + \\ \left( a_3^2 + l_2^2 + 2a_3 l_2 \cos \theta_3 + a_3 l_1 \cos \theta_{23} + l_2 l_1 \cos \theta_2 \right) m_3 \end{bmatrix} \ddot{\theta}_2 \\ & + \left[ (a_3^2 + a_3 l_1 \cos \theta_{23} + a_3 l_2 \cos \theta_3) m_3 \right] \ddot{\theta}_3 \\ & - a_2 l_1 m_2 \sin \theta_1 \dot{\theta}_1^2 - 2(l_1 l_2 m_3 \sin \theta_2 + a_3 l_1 m_3 \sin \theta_{23}) \dot{\theta}_1 \dot{\theta}_2 \\ & - (l_1 l_2 m_3 \sin \theta_2 + a_3 l_1 m_3 \sin \theta_{23}) \dot{\theta}_2^2 \\ & - 2(a_3 l_1 m_2 \sin \theta_3 + a_3 l_1 m_2 \sin \theta_{23}) \dot{\theta}_1 \dot{\theta}_3 \\ & - 2(a_3 l_2 m_3 \sin \theta_3 + a_3 l_1 m_3 \sin \theta_{23}) \dot{\theta}_2 \dot{\theta}_3 \\ & - (a_3 l_2 m_3 \sin \theta_3 + a_3 l_1 m_3 \sin \theta_{23}) \dot{\theta}_3^2 + m_1 g a_1 \cos \theta_1 \\ & + m_2 g (l_1 \cos \theta_1 + a_2 \cos \theta_{12}) \\ & + m_3 g (l_1 \cos \theta_1 + l_2 \cos \theta_{12} + a_3 \cos \theta_{123}) \end{aligned} \quad (5)$$

$$\begin{aligned} \tau_2 = & \begin{bmatrix} (a_2^2 + a_2 l_1 \cos \theta_2) m_2 + \left( a_3^2 + l_2^2 + 2a_3 l_2 \cos \theta_3 \right. \\ \left. + a_3 l_1 \cos \theta_{23} + l_1 l_2 \cos \theta_2 \right) m_3 \end{bmatrix} \ddot{\theta}_1 \\ & + \left[ a_2^2 m_2 + (a_3^2 + l_2^2 + 2a_3 l_2 \cos \theta_3) m_3 \right] \ddot{\theta}_2 \\ & + \left[ (a_3^2 + a_3 l_2 \cos \theta_3) m_3 \right] \ddot{\theta}_3 (l_1 l_2 m_2 \sin \theta_2 + a_3 l_1 m_2 \sin \theta_{23}) \dot{\theta}_1^2 \\ & - 2a_3 l_2 m_3 \sin \theta_3 \dot{\theta}_1 \dot{\theta}_3 \\ & - 2a_3 l_2 m_3 \sin \theta_2 \dot{\theta}_3 - a_3 l_2 m_3 \sin \theta_3 \dot{\theta}_3^2 \\ & + m_2 g a_2 \cos \theta_{12} + m_3 g (l_2 \cos \theta_{12} + a_3 \cos \theta_{123}) \end{aligned} \quad (6)$$

$$\begin{aligned} \tau_3 = & (a_3^2 m_3 + a_2 l_1 m_3 \cos \theta_{23} + a_3 l_2 m_3 \cos \theta_3) \ddot{\theta}_1 \\ & + (a_3^2 m_3 + a_3 l_2 m_3 \cos \theta_3) \ddot{\theta}_2 + \left( a_3^2 m_3 + \frac{1}{12} l_3^2 \right) \ddot{\theta}_3 \\ & + (a_3 l_1 m_2 \sin \theta_3 + a_3 l_1 m_2 \sin \theta_{23}) \dot{\theta}_1^2 \\ & + 2a_3 l_2 m_3 \sin \theta_3 \dot{\theta}_1 \dot{\theta}_2 + a_3 l_2 m_3 \sin \theta_3 \dot{\theta}_2^2 + m_3 g a_3 \cos \theta_{123} \end{aligned} \quad (7)$$

### 3. Design the force feed-forward controller

Rewrite general equation of motion of 3-DOF Rehabilitation Robot as:

$$\mathbf{M}(q)\ddot{q} + \mathbf{V}(q, \dot{q}) + \mathbf{G}(q) = \mathbf{u}_{dk} + \mathbf{J}(q)^T \mathbf{F}_{ff} \quad (8)$$

Solving equation (8) yield

$$\ddot{q} = \left[ \mathbf{u}_{dk} + \mathbf{J}(q)^T \mathbf{F}_{ff} - \mathbf{V}(q, \dot{q}) - \mathbf{G}(q) \right] \mathbf{M}(q)^{-1} \quad (9)$$

where  $\mathbf{u}_{dk}$  is control torques vector,  $\mathbf{J}(q)^T$  is Jacobian sub-matrix, which is used to transfer force value from the task space to the joint space,  $\mathbf{F}_{ff}$  is force exerted vector.

In the intermediate training stage, the task of the controller is to produce a reasonable interaction between the patient and the Robot. To accomplish this task, in the same way, we built a FFM controller for an upper limb exoskeleton in our previous research [14], we considered a problem, which consists of two steps:

**Step 1:** In the beginning, the therapist moves Robot by rotating its joints following the desired loading motion. The force of the therapist  $F_h$  is measured by force sensors placed there and it is generated to compensate for weight of the robot and the force of loading. At the same time, the angular joint trajectories are determined by encoders located at joints.

**Step 2:** When the robot becomes active, control torques  $u_r \neq 0$ , the force, which is measured by force sensors, is used as the input of the force feedforward controller to compensate the force generated by the therapist to perform patient-active exercises. Moreover, force sensors also measured the force exerted by the patient's lower limb and sent these forces to the input of the FFM controller. In this step, the force exerted by the therapist is desired to decrease to zero and the force exerted by the patient's lower limb is considered as a disturbance. The target of the control is to cancel the disturbance.

From their problem, the feedforward control approach with the formula yield:

$$\mathbf{u}_{dk} = w_p \left[ \mathbf{J}(q_{(i)})^T \mathbf{F}_{ff(i)} \right] + w_l \left[ \sum_{j=1}^{i-1} \mathbf{J}(q_{(i)})^T \mathbf{F}_{ff(j)} \right] \quad (10)$$

where  $\mathbf{q}_{(i)}$  ( $i = 1, 2, 3$ ) joints angles are measured by encoders and  $\mathbf{F}_{ff(i)}$  ( $i = 1, 2, 3$ ) force of the therapist or the patient's lower limb are measured by force sensors and  $w_p, w_l$  are constant.

The block diagram of the FFM controller shown in Figure 3.

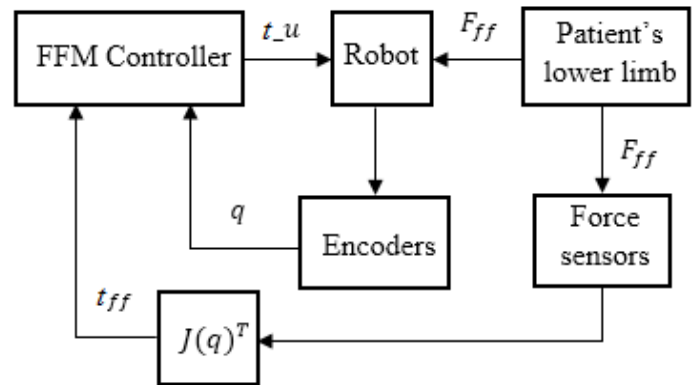


Figure 3: Block diagram of the FFM controller

The FFM controller block is designed to perform two tasks: trajectory tracking and force control. There are two loops in the block diagram above. The inner loop is used to feedback data acquiring from encoders to implement trajectory tracking control. The outer loop is used to apply force control. The  $\mathbf{J}(q)^T$  block is used to transfer force value from the task space to the joint space.

#### 4. Numerical simulation of the force feed-forward controller

Based on the controller proposed in section III, the numerical simulation of robot kinematic, dynamic, and controller are conducted on the MATLAB/Simulink software to evaluate the performances of the proposed force controller. In the simulation program, valuable system parameters are listed in Table 2, where the parameters  $m_i$  is the total mass of link, the motor, and the joint of stitch  $i$ ;  $l_i$  the length of link  $i$ ;  $a_i = \frac{1}{2}l_i$ ;  $K_g$  is the gear of motor.

Table 2: Simulation parameters of 3-dof rehabilitation robot

Parameter	$i = 1$	$i = 2$	$i = 3$
$m_i$ (kg)	3.2	3.1	0.25
$l_i$ (m)	0.4	0.48	0.33
$a_i$ (m)	0.2	0.24	0.165
$K_g$	19	19	19

**Step 1:** In this step, based on the desired loading motion, variables  $\mathbf{q} = [\theta_1 \ \theta_2 \ \theta_3]^T$ ,  $w_p, w_l$  are given as follows:

$$\theta_{1(i)} = \omega_1 t_i ; \theta_{2(i)} = \omega_2 t_i ; \theta_{3(i)} = \omega_3 t_i$$

where  $t_i$  ( $i = 1, 2, 3$ ) is time,  $\omega_1, \omega_2, \omega_3$  are angular velocities of joints, in this simulation their values are chosen as  $\omega_1 = \omega_2 = \omega_3 = 0.025 \left(\frac{rad}{s}\right)$  and the gains  $w_p = 1; w_l = 1$ .

The force  $F_{ff}$  is an unknown parameter and measures by force sensors. In reality, this force may include force exerted by the therapist, force exerted by the patient, and the weight of the patient's foot. In this work  $F_{ff}$  is desired to decrease with time as expressed as:

$$\begin{aligned} \mathbf{F}_{ff} &= \left[ \mathbf{F}_{ff(i)x} \quad \mathbf{F}_{ff(i)y} \quad \mathbf{F}_{ff(i)z} \right]^T \\ &= 50 \times e^{-t_i} \times [1 \quad 1 \quad 1]^T (N) \end{aligned} \tag{11}$$

Simulation results are shown in figure 4 to figure 9, in case of applying force to foot platform link to rotate the ankle joint, the force exerted and the respective motion of the link are shown in figures following. In this simulation study, the motion of the foot platform is investigated at its center of gravity.

Fig.4 to Fig.6 shows that when the therapist applies training activity, the foot platform of the robot moves according to the trajectory in the XY plane. Fig.7 to Fig.9 shows that at the beginning of motion, the force exert by the therapist is maximum because of the foot platform is in the static state and start to move. This force decrease to zero when the platform moves to the end of the trajectory. In this step, the trajectory is sensed by the

encoder, and the force exerted by the therapist is sensed by the force sensor.

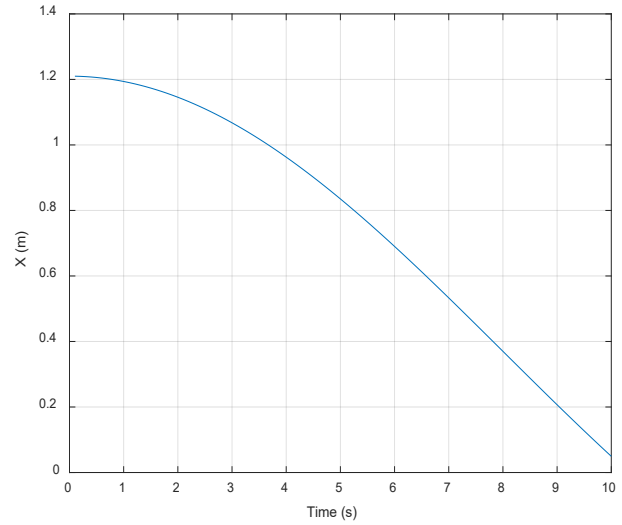


Figure 4: Motion of the footplatform in X axis

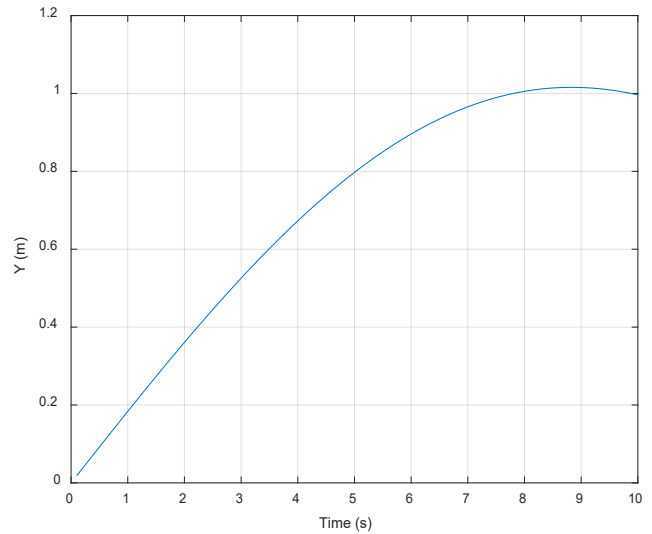


Figure 5: Motion of the footplatform in Y axis

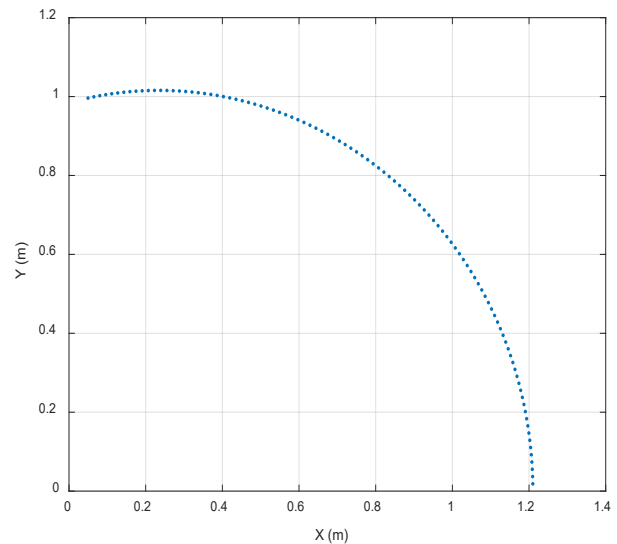


Figure 6: Motion of the footplatform in XY plane

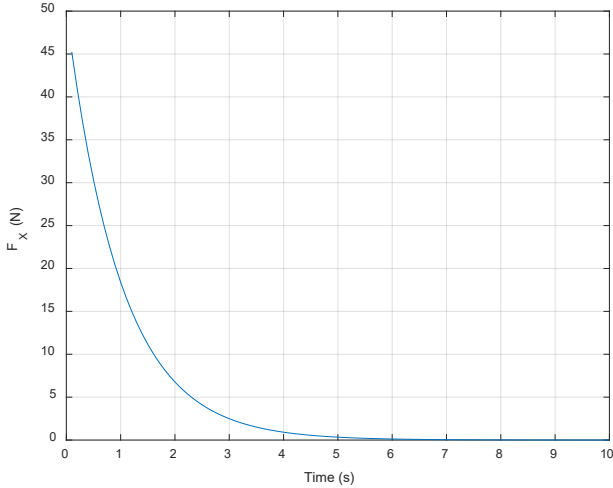


Figure 7: Force in X direction

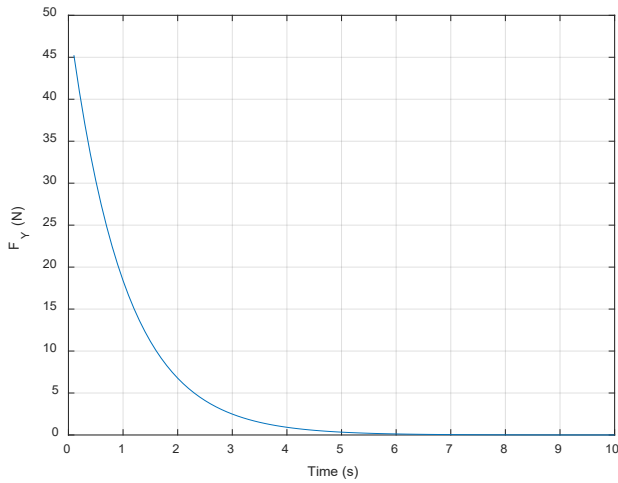


Figure 8: Force in Y direction

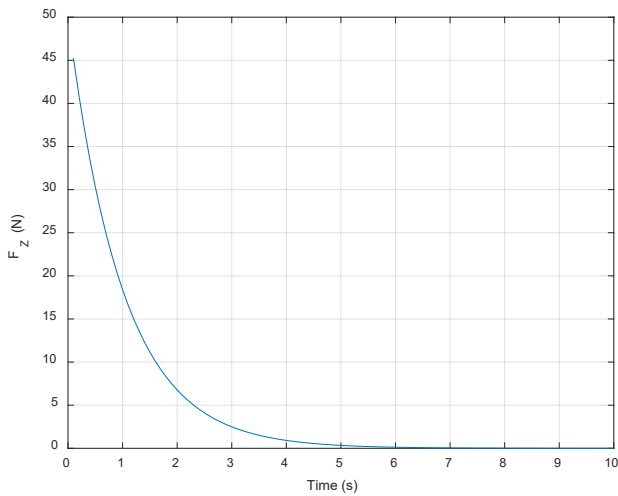


Figure 9: Force in Z direction

**Step 2:** In this step, when the force  $F_{ff}$  decrease to zero, the force feedforward controller computed the control torques to compensate  $F_{ff}$  force. By using the equation (15), the control torques at the ankle joint in  $X, Y, Z$  directions of the base coordinate system are shown in figures 10, 11 and 12, respectively.

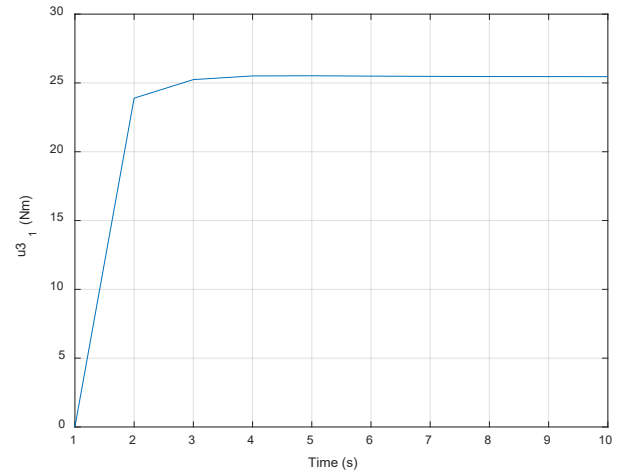


Figure 10: Control torque in X direction

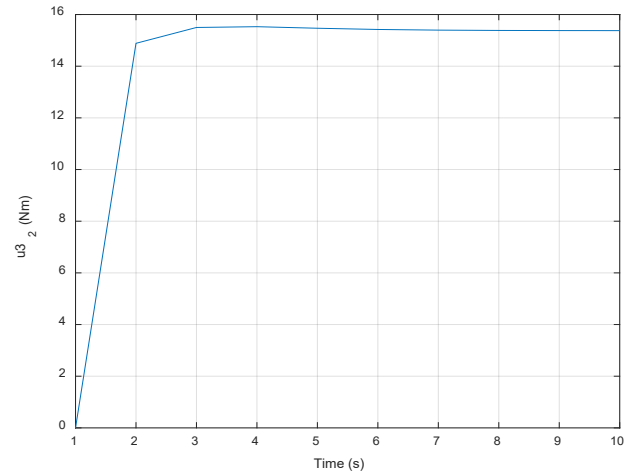


Figure 11: Control torque in Y direction

There are initial conditions for the calculation of control torque in step 2.



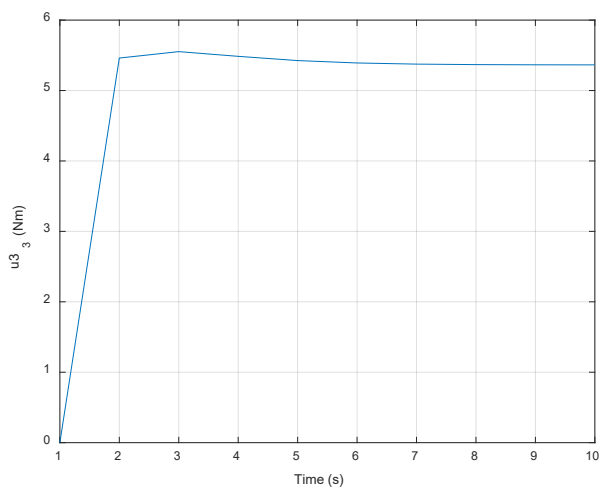


Figure 12: Control torque in Z direction

The above results show that when the  $F_{ff}$  force decreases, the controller has calculated and created an increasing control torque to compensate for the reduced force. In the beginning,  $t_0 = 0(s)$ , the force exerted is maximum  $F_{ff} = 45(N)$ , the control torque is minimum, it is zero. In the period of 0 to 3 seconds, the  $F_{ff}$  force decreases, the control torque increases. When the  $F_{ff} = 0(N)$  the control torque is maximum with an approximate value of  $25(Nm)$  in the  $X$  direction, equal to  $15.6(Nm)$  in the  $Y$  direction and equal to  $5.5(Nm)$  in the  $Z$  direction. The maximum value of control torque will be kept stable during the operation of the robot. During this process, if the patient's lower limb actively exerts a force on the robot, this force is measured by the force sensors and sensed to the controller input. The force exerted by the patient's lower limb is considered as a disturbance, which is canceled by the FFM controller. However, this feedforward control still has some disadvantages and need to be evaluated in a real model to adjust the parameters towards a more realistic control.

## 5. Conclusions

In this paper, an advanced method for force control of a 3-DOF lower limb rehabilitation robot for perceiving reality using Force Feed-forward Model was presented. The robot system can sense the force feedback from reality and apply torque through joints hip, knee, and ankle. When the therapist or patient exerts a force at the robot's links, the exerted force is measured by force sensors placed at there and used for computing torque generated at the joints based on the configuration of the robot. When the exerted force decreases, the FFM controller was used to compute and compensate the force generated by the therapist to perform patient-active exercises. Numerical simulation results showed that the computed force control and corresponding loading motions. The results could be used to estimate the maximum moment of motors for the design steps of the robot control. In future work, we intend to present some experimental test results to compare with the numerical simulation results, which have just been presented in this paper.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

This work is funded by MOET project under grant number B2018-SKH-04

## References

- [1] A.M. Simon, R.G. Brent, D.P. Ferris, "Symmetry-based resistance as a novel means of lower limb rehabilitation," *Journal of Biomechanics*, **40** 1286–1292, 2007, doi:10.1016/j.jbiomech.2006.05.021.
- [2] E.D. Judith, J. Latonio, C. Grigore and R. Boian, "Post-Stroke Rehabilitation with the Rutgers Ankle System: A Case Study," *Virtual and Augmented Reality*, **10**(4), 416-430, 2001, doi:10.1162/1054746011470262 .
- [3] H.D. Trung, T.D. Vu. "A Simple Control Method for Exoskeleton for Rehabilitation," *SSRG International Journal of Electrical and Electronics Engineering (SSRG-IJEEE)*, **4**(8), 7-12, 2017, doi:10.14445/23488379/IJEEE-V4I8P102.
- [4] L.L. Cai, A.J. Fong, C.K. Ootoshi, Y. Liang, J.W. Burdick, R.R. Roy, "Implications of assist-as-needed robotic step training after a complete spinal cord injury on intrinsic strategies of motor learning," **26**, *J Neurosci*, 2006, doi:10.1523/JNEUROSCI.2266-06.2006.
- [5] C. Perry et al, "Upper arm power exoskeleton design," *IEEE/ASME Transactions on Mechatronics*, **12**(4), 408-417, 2007, doi:10.1109/TMECH.2007.901934.
- [6] H. Lo and S. Xie, "Exoskeleton robots for upper-limb rehabilitation: State of the art and future prospects," *Med Eng Phys*, **34**, 261-268, 2012, doi:10.1016/j.medengphy.2011.10.004.
- [7] S. Yusuf et al., "Force Feedback Control of Lower Extremity Exoskeleton Assisting of Load Carrying Human," *Applied Mechanics and Materials*, **598**, 546-550, 2014, doi:10.4028/www.scientific.net/AMM.598.546.
- [8] G. Chen, P. Qi, Z. Guo, Y. Haoyong, "Mechanical design and evaluation of a compact portable knee-ankle-foot robot for gait rehabilitation," *Mechanism and Machine Theory*, 51–64, 2016, doi:10.1016/j.mechmachtheory.2016.04.012.
- [9] B. Koopman et al, "Selective control of gait subtasks in robotic gait training: foot clearance support in stroke survivors with a powered exoskeleton," *Journal of NeuroEngineering and Rehabilitation*, **10**, 2013, doi:10.1186/1743-0003-10-3.
- [10] M.S. Ju et al., "A rehabilitation robot with force-position hybrid fuzzy controller: hybrid fuzzy control of rehabilitation robot," *IEEE Trans Neural Syst Rehabil Eng*, 349–358, 2005, doi:10.1109/TNSRE.2005.847354.
- [11] H. Valleryet al., "Reference trajectory generation for rehabilitation robots: complementary limb motion estimation," *IEEE Transaction on Neural Systems Rehabilitation Engineering*, **17**(1), 23–30, 2009, doi:10.1109/TNSRE.2008.2008278.
- [12] M. Bernhardt et al., "Hybrid force-position control yields cooperative behaviour of the rehabilitation robot LOKOMAT," *Proceedings of the IEEE 9th international conference on rehabilitation robotics*, 536–9, 2005, doi:10.1109/ICORR.2005.1501159.
- [13] S.K. Banala, S.K. Agrawal, J.P. Scholz, "Active Leg Exoskeleton (ALEX) for gait rehabilitation of motor-impaired patients," *Proceedings of the IEEE 10th International Conference on Rehabilitation Robotics (ICORR)*, 401-407, 2007, doi:10.1109/ICORR.2007.4428456.
- [14] T. Lung-Wen. *The Mechanics of Serial and Parallel Manipulators*, Wiley-Interscience Publication, 1999.

## Earthquake Response of Multi-Storey Infilled Reinforced Concrete Buildings

Miloud Mouzzoun\*, Abdelkader Cherrabi

Civil engineering department, Hassania School of Public Works, Casablanca, BP8108, Morocco

### ARTICLE INFO

Article history:

Received: 22 June, 2020

Accepted: 20 July, 2020

Online: 25 August, 2020

Keywords:

Frame

Infill

Seismic

Time history

Pushover

Strut

### ABSTRACT

The aim of this work is to study the effect of infill brick panels on the response of multi-storey buildings under seismic loading. An eight storey building is investigated. The building is analysed under gravity and seismic loads. Infill panel is replaced by two struts according to FEMA306. Time history and pushover analyses are performed to assess seismic strength of the building. Simulations are performed by SAP2000. Numerical results show that behavior of bare and infill frames under lateral loading are too distinct. There is a change in the manner in which the infill frame carries the lateral loads.

## 1. Introduction

Concrete frames infilled by bricks are generally used as structural system for multi-storey buildings to carry gravity and lateral seismic loads. Past earthquakes, have indicate that infill panels change the structural response of buildings under seismic loads. Figures 1 and 2 show collapses caused by soft storey and short column effects. Structural behavior of infilled frames was studied by many authors [1]-[7]. The studies were concentrated on the response of infilled frames by masonry under horizontal loading. Experimental results indicate that masonry infills modify the structural behavior of the frame. There is a modification in the manner in which the frame carries the lateral loads. In this article the seismic effect of infill on structural response of buildings is investigated. For this purpose, a multi-storey framed building is studied. Infill panel is modeled by two concentric diagonal struts according to FEMA306 [8]. Numerical simulations are performed using the software analysis SAP2000.

## 2. Material Nonlinearity

### 2.1. Concrete members

Beams and columns are idealized as frame elements with two concentrated plastic hinges at the extremities. For pushover analysis, P-M3 hinge is assigned to the columns and M3 hinge is

assigned to the beams. The hinge parameters are calculated according to FEMA356 [10].



Figure 2: Short column mechanism [9]

\*Corresponding Author: Miloud Mouzzoun, Hassania School of public works, Morocco, mouzzoun.mouloud@gmail.com

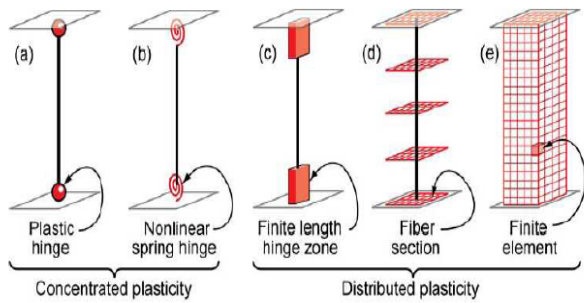


Figure 3: Introduction of material nonlinearity in concrete members [11]

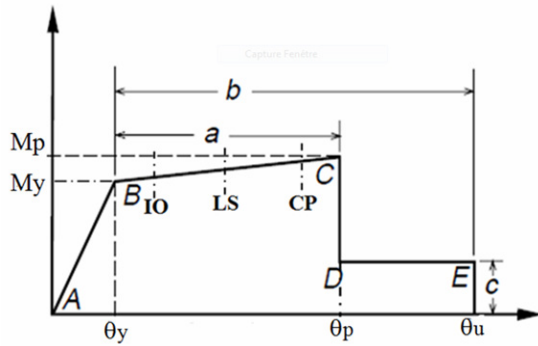


Figure 4: behavior of plastic hinge

## 2.2. Masonry infill wall

When lateral seismic forces are applied, infilled frame is replaced by two struts as shown in figure 5. The strut doesn't support any tension, it carries only compression loads. Strut width  $W$  is calculated by the equation (1).

$$w = 0.175d(\lambda H_c)^{-0.4} \quad (1)$$

$$\lambda = \sqrt[4]{\frac{E_m t \sin 2\theta}{4E_c I_c H_m}} \quad (2)$$

$$\theta = \tan^{-1} \frac{H_m}{L_m} \quad (3)$$

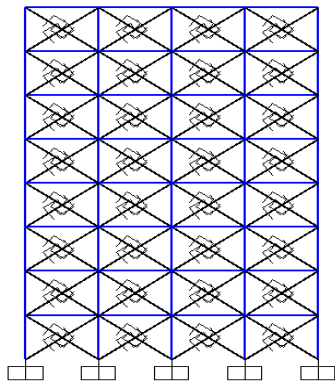


Figure 5: Frame with equivalent diagonal struts

## 3. Numerical investigations

The building studied here is an eight storey framed building, having 20 meters in the transversal direction and 25 meters in the longitudinal direction. Each storey height is 3m. The beams sections are 25cmx40cm; slabs thicknesses are 14cm. The columns sections vary from 30cmx30cm to 45cmx45cm. Building is analyzed under gravity and lateral seismic loads. Time history and Pushover analyses are performed by SAP2000 [12].

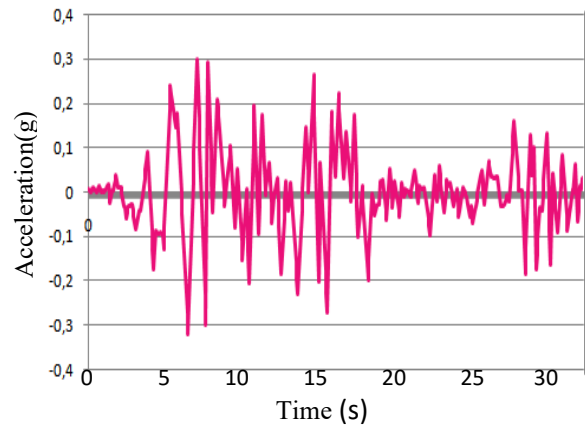


Figure 6: Elcentro Earthquake ground motion

## 4. Results and discussion

### 4.1. Periods of vibration

In the table 1, the periods of vibration of infill and bare frames are presented. It's found that introduction of infill walls reduces significantly the vibration periods.

Table 1: Natural periods of vibration

Mode Number	Infill frame	Bare frame
1	0.265	0.498
2	0.092	0.166
3	0.054	0.089
4	0.037	0.055
5	0.028	0.039
6	0.027	0.029

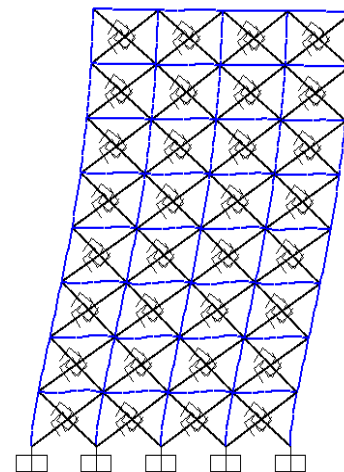


Figure 7: First mode of vibration of infill frame. T=0.265s

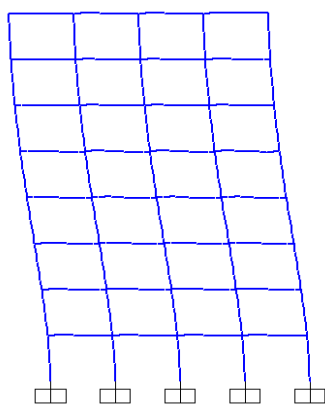


Figure 8: First mode of vibration of bare frame. T=0.498s

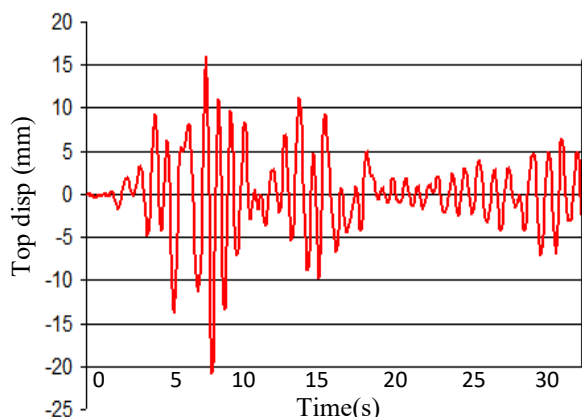


Figure 9: Top displacement time history. Infill frame

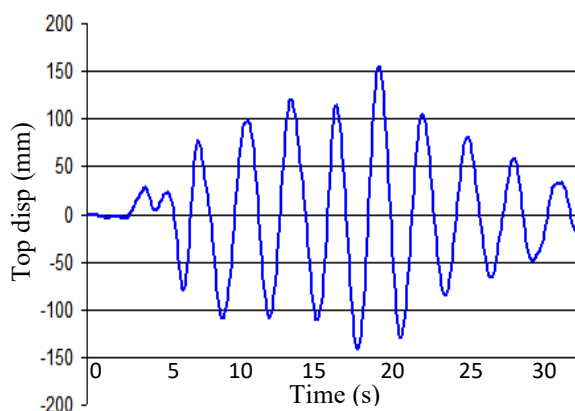


Figure 10: Top displacement time history. Bare frame

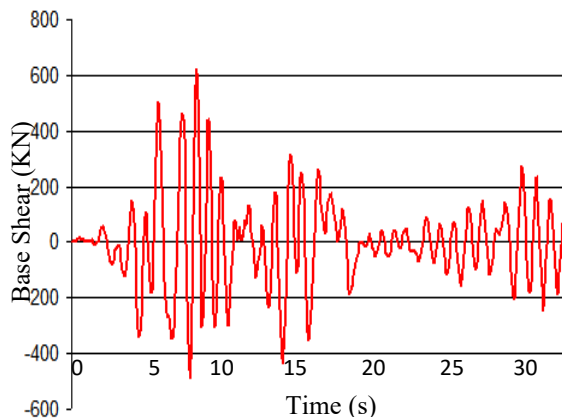


Figure 11: Base shear time history. Infill frame

#### 4.2. Top displacement and base shear

Base shear and roof displacement time histories are shown in the figures 9, 10, 11 and 12. The results show that top displacement decreases and base shear increases when infill panels were introduced in the model by diagonal strut concept.

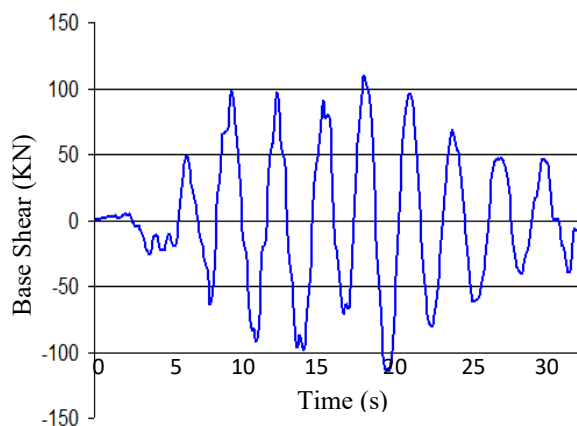


Figure 12: Base shear time history. Bare frame

#### 4.3. Bending moments and shear forces

In the tables 2 to 5, we present the results of internal forces. The obtained results show that there is a great reduction in the values of bending moments and shear forces when infill wall is introduced in the structural model.

Table 2: Bending moments in the columns

Storey	Bare frame	Infill frame
	M(KN.m)	M(KN.m)
8°	68,02	23,81
7°	103,71	37,57
6°	153,61	63,17
5°	168,90	70,58
4°	171,49	75,95
3°	215,12	82,65
2°	298,60	132,40
1°	450,97	280,97

Table 3: Shear forces in the columns

Storey	Bare frame	Infill frame
	V(KN)	V(KN)
8°	33,15	11,42
7°	54,81	18,86
6°	81,87	32,33
5°	97,47	39,07
4°	109,03	43,81
3°	122,19	55,08
2°	122,45	62,65
1°	117,48	86,27



Table 4: Shear forces in the beams

Beam	Bare frame	Infill frame
	V (KN)	V (KN)
B1	104,71	60,50
B2	107,74	61,64
B3	104,68	61,59
B4	105,48	61,44

Table 5: Bending moments in the beams

Beam	Bare frame	Infill frame
	M (KN.m)	M (KN.m)
B1	166,02	78,63
B2	168,16	81,90
B3	168,04	81,86
B4	172,27	82,07

#### 4.4. Seismic performance

Pushover analyses are carried out for infill and bare frames. The obtained pushover curves are superposed with the seismic demand curves to obtain the performance point. Two seismic demands are considered; moderate shaking with design acceleration  $A=0.16g$  and high shaking with design acceleration  $A=0.20g$ .

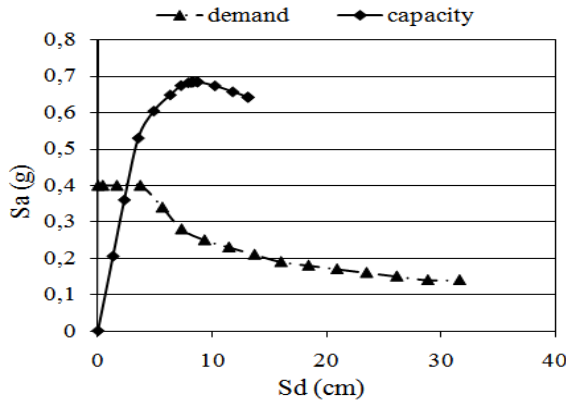


Figure 13: Seismic performance of infill frame under moderate shaking

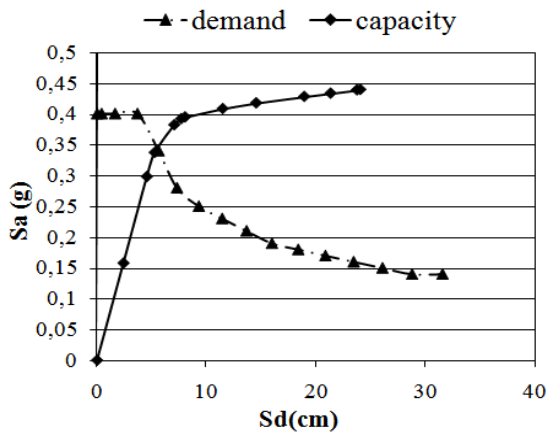


Figure 14: Seismic performance of bare frame under moderate shaking

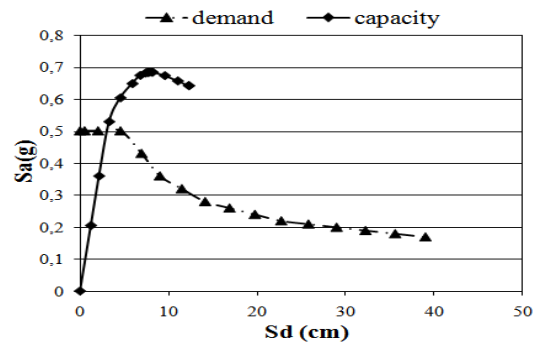


Figure 15: Seismic performance of infill frame under high shaking

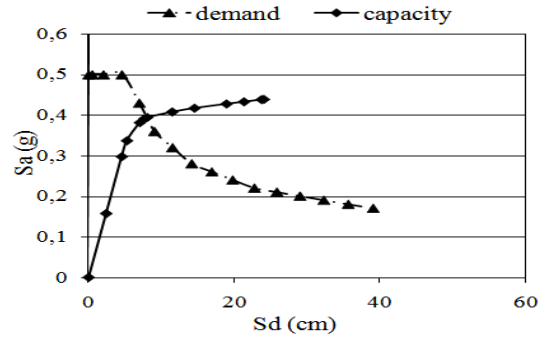


Figure 16: Seismic performance of bare frame under high shaking

#### 4.5. Lateral storey displacements

The seismic demands for infill and bare frames under moderate and high shaking are presented in the figures 17 to 20. The results show that infill frame performs well compared with bare frame.

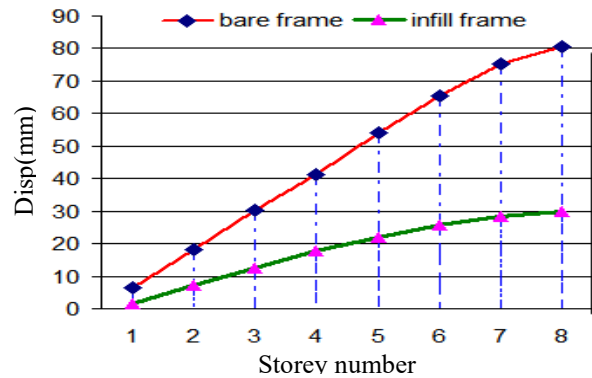


Figure 17: Lateral story displacements under moderate shaking

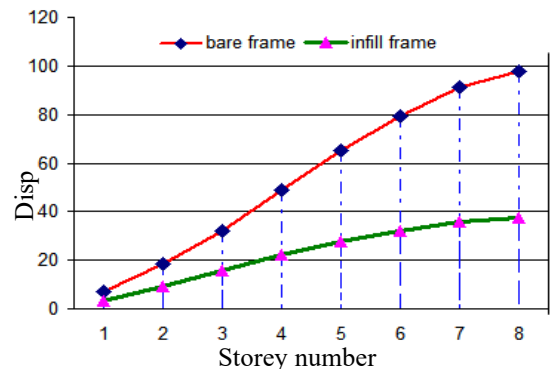


Figure 18: Lateral story displacements under high shaking

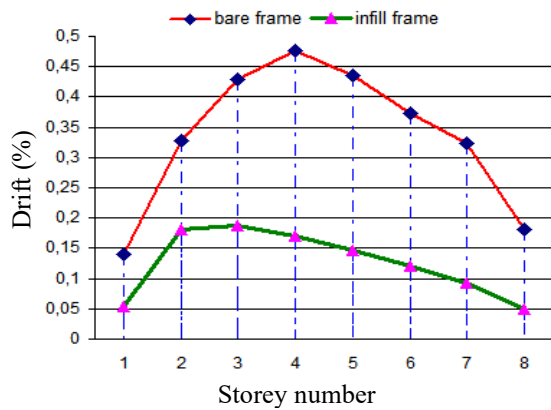


Figure 19: Inter-storey drifts under moderate shaking

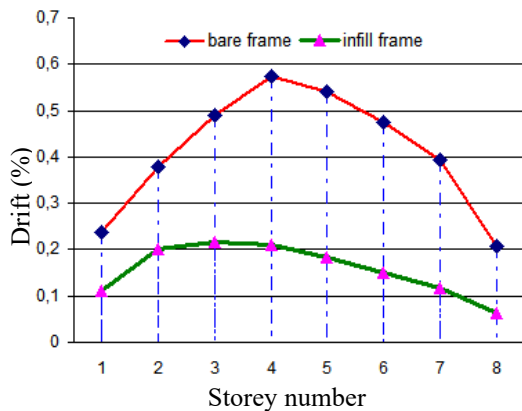


Figure 20: Inter-storey drifts under high shaking

## 5. Conclusion

The structural behavior of multi-storey buildings with infill panels under earthquake motion was investigated. The essential results are:

- Results show that natural periods of bare frame are greater than those of infill frame. Therefore, frame analysis without infill idealization underestimates the lateral shear design.
- Numerical simulations indicate that infill walls change considerably the internal forces. Maximum bending moments and shear forces of frame without infill are greater than those of infill frame.
- Pushover analysis shows that infill walls increase both lateral strength and stiffness of the building. Therefore, the seismic vulnerability of infilled building will be significantly reduced.
- Pushover analysis indicates that seismic performance of infill frame is largely superior to that of bare frame under different seismic intensities.
- Investigation results show that there is a change in the manner in which the infill frame carries the lateral loads. Diagonal strut concept of infill transforms the flexible frame into braced frame and then the values of internal forces will be changed.

Finally, this study has highlighted the influence of uniform infills on the earthquake response. Infill walls have a significant influence under seismic loads. Neglecting infills in the structural

analysis is not a safe way. In perspective, it seems important to study the seismic effect of irregular configurations of infill panels, in order to prevent their unfavorable effects as in the case of soft storey and short column failures.

## Conflict of Interest

The authors declare no conflict of interest.

## References

- [1] R.J. Mainstone, "on the stiffness and strength of Infilled frames" Proceedings of the Institution of the Civil Engineers, **49**(2). 1971. <https://doi.org/10.1680/iicep.1971.6267>.
- [2] D.V. Malick, RT Severn, "dynamic characteristics of infilled frames" Proceedings of the Institution of the Civil Engineers, **39**(2). 1968. <https://doi.org/10.1680/iicep.1968.8091>.
- [3] B. Stafford, C. Carter, "A method for analysis for infilled frames" Proceedings of Institution of Civil Engineers, No. 7218, 31-48, 1969. <https://doi.org/10.1680/iicep.1969.7290>.
- [4] A.B. Mehrabi, P.B. Shing, M.P. Schuller, J.L. Noland, "Experimental Evaluation of Masonry-Infilled RC Frames" ASCE Journal of Structural Engineering, **122**(3), 228-237, 1996. [https://doi.org/10.1061/\(ASCE\)0733-9445\(1997\)123:5\(604\)](https://doi.org/10.1061/(ASCE)0733-9445(1997)123:5(604)).
- [5] P. Negro, A. Colombo, "Irregularities induced by non structural masonry panels in framed buildings" Engineering Structures. **19**(7), 576-585. 1997. [https://doi.org/10.1016/S0141-0296\(96\)00115-0](https://doi.org/10.1016/S0141-0296(96)00115-0).
- [6] G. Al-Chaar, Evaluating strength and stiffness of unreinforced masonry structures. US Army Corps of Engineers. Construction Engineering Research Laboratories, 2002.
- [7] A. Hashemi, K.M. Mosalam, "Shake-Table Experiment on Reinforced Concrete Structure Containing Masonry Infill Wall" Journal of Earthquake Engineering and Structural Dynamics. **14**(35), 1827-1852, 2006. <https://doi.org/10.1002/eqe.612>.
- [8] FEMA306, Evaluation of Earthquake Damaged Concrete and Masonry Wall Buildings, Basic Procedures Manual, Federal Emergency Management Agency, 1999.
- [9] M. Mouzzoun, "Seismic behaviour of reinforced concrete frame buildings with masonry infill", PhD thesis, Mohammadia School of engineers, Morocco, 2015.
- [10] FEMA356, Federal Emergency Management Agency, NEHRP recommended Provisions for Seismic Regulations for New Buildings and Other Structures, 2000.
- [11] ATC40, Applied Technology Council, Partnership for Response, and United States. Federal Emergency Management Agency. Evaluation of Earthquake Damaged Concrete and Masonry Wall Buildings: Basic Procedures Manual, 1999.
- [12] CSI, Analysis Reference Manual for SAP2000, ETABS and SAFE, Computers and Structures, Inc. Berkeley, California, USA, 2005.

## Machine Learning for Network Intrusion Detection Based on SVM Binary Classification Model

Anouar Bachar<sup>\*1,2</sup>, Noureddine El Makhfi<sup>1</sup>, Omar EL Bannay<sup>2</sup>

<sup>1</sup>FST, Abdelmalek Essaadi University, PMIC Lab, Al-Hoceima, 32003, Morocco

<sup>2</sup>ENSA, Sultan Moulay Slimane University, ASTI Lab, Khouribga, 25000, Morocco

### ARTICLE INFO

Article history:

Received: 23 June, 2020

Accepted: 20 July, 2020

Online: 25 August, 2020

Keywords:

IDS

Intrusions

Machine Learning

SVM classification model

UNSWNB-15 dataset

### ABSTRACT

Recently, the number of connected machines around the worldwide has become very large, generating a huge amount of data either to be stored or to be communicated. Data protection is a concern for all institutions, it is difficult to manage the masses of data that are susceptible to multiple threats. In this work, we present a novel method of Intrusion Detection System (IDS) based on the detection of anomalies in computer networks. The aim is to use artificial intelligence techniques in the form of Machine Learning (ML) for intrusion detection. For this purpose, we have proposed a Support Vector Machine (SVM) classification model with two kernels, one Polynomial and the other Gaussian. This model is trained and tested with the recent UNSWNB-15 dataset. Regarding the results obtained, we have evaluated our model with six metrics capable of offering all potential threats. As a result, we have achieved a percentage of 94% for the detection rate (DR).

## 1. Introduction

The application of artificial intelligence in the field of computer network security has become an inescapable reality. In this context, we propose an extended version of an oral communication [1] which deals with network intrusion detection based on the SVM model.

The use of computer network services has become essential in several areas, particularly those related to daily life. The exchange of data can be very costly in the case of sometimes fatal threats, especially when it is a large-scale exchange at the level of governments or businesses [2]. These threats affect the integrity, confidentiality, and availability of these data. Possible solutions for the protection of computer networks are divided into two categories, those related to hardware or those related to software. We can cite firewalls, antivirus, cryptography, etc. as solutions. However, these solutions are insufficient to protect all networks [3]. In order to strengthen network security, another higher level of security such as IDSs is needed.

An IDS is a medium that allows us to protect our computer network from any unwanted activity as described in [4]. It filters incoming and outgoing network traffic to detect possible threats in the network. We distinguish between two types of IDS, the classical ones based on signatures as in the open-source software SNORT. They have a database of intrusion signatures. The

signature of each incoming entity is compared with the signatures present in this database. Thus, the signature-based-IDS is very efficient for the detection of common threats and generates a very low FAR (False Alarm Rate). Despite this, it is unable to detect modern attacks. The second type of IDS is based on artificial intelligence techniques using ML for the detection of new attacks (0 Day attack) as described in the work [5]. However, the only problem is the very high false alarm rate generated by this type of IDS. This is the reason why several scientific research works are interested in this topic.

In this article, we present an IDS based on the technique of ML. We have chosen to use an SVM binary classification model associated with a UNSWNB-15 dataset for training and testing the proposed model. This article is structured into seven sections: In section 2, we present the related work. In the next section, we discuss the basic principle of an IDS. Then, section 4 concerns the existing data sources and the description of the UNSWNB-15 data set that we have used in this paper. In Section 5, we present our methods. In section 6, we show the performance of our proposed model with a comparison and interpretation of the results obtained. In the last section, we have a general conclusion and the perspectives of this work.

## 2. Related Work

In recent years, protecting computer networks against various threats is a priority for most companies. The evolution of

\*Corresponding Author: Anouar Bachar, Email: [anouar.bachar@gmail.com](mailto:anouar.bachar@gmail.com)

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050476>

cyberattacks requires the use of highly developed ways for the detection of these threats. Among the first works, we can cite the work of James Canady [6] who used the Artificial Neural Network (ANN) classification model in 1998. Netssan P, et al [7] obtained a detection rate of 85.87% with the KDD-CUP 99 dataset by combining two naive Bayesian models and decision trees. Other work using the same KDD-CUP 99 dataset has been proposed by O. Deepen et al [8]. They obtained a DR of 99.9% so the FAR is too low in the order of 1.25%. Also, Divyatmika et al [9] achieved a detection rate close to 99% with the Multilayer Perceptron (MLP) classification model algorithm using KDD-CUP 99 dataset. Finally, G-H Zhou [10] proposed a very efficient model (Detection rate = 99.54%) using a KDD-CUP 99 dataset for an IDS.

In parallel to the above, several works have used different ML algorithms for the detection of threats in computer networks. We cite the work of G. Meena et al [11] who recently published a paper using some Machine Learning algorithms associated with the two datasets NSL-KDD and KDD-CUP 99. Also, L. Dhanabal et al [12] published a study evaluating the efficiency of ML algorithms with the NSL-KDD dataset.

The two datasets KDD-CUP 99 and NSL-KDD are used in most of the previous work. They remain ineffective for the detection of recent attacks, even though the detection rate is very high for the different classification algorithms [13]. Therefore, it is necessary to use a newly updated dataset containing the behavior of modern network traffic. In Section 4, we will develop a study of the different existing datasets to show the effectiveness of a modern intrusion detection system.

### 3. Intrusion Detection System

An IDS allows us to protect our computer network infrastructure from any malicious activity. These activities are generally targeting the integrity, confidentiality, and availability of data in our network.

We classify IDS according to several criteria such as the type of IDS and the associated classification method. Regarding the type, we have the (N-IDS) and the (H-IDS) (Figure 1):

- N-IDS (Network Intrusion Detection System): It secures the entire network. Its location in the network affects the false alarm rate. For example, if the NIDS is located upstream of a Firewall (Figure 1), then it generates fewer false alarms. Since the traffic is already filtered by the Firewall. Otherwise, if it is placed downstream of a Firewall (location 2). It generates more false alarms.
- H-IDS (Host Intrusion Detection System): It allows the machine to be secured on its own without the use of other systems.

The IDSs are classified according to the detection method as in the work [14]. The first is signature-based-IDS. In this case, the IDS have a signature base compared with the signature of each incoming entity. This allows the IDS to be very favorable for the detection of already known intrusions. However, this first class of IDS only detects modern attacks (0 Day attack) i.e. threats with unknown signatures [15]. The second detection class is anomaly-based-IDS as in [16]. In this work, we have used this last class of threat detection. We have focused on new machine learning

techniques. Knowing that an IDS builds a user profile based on the normal behavior of the system. It then monitors the network to identify any abnormal activity in the network. The anomaly-based-IDS can detect modern attacks, but it generates a high FAR in contrary to the signature-based-IDS.

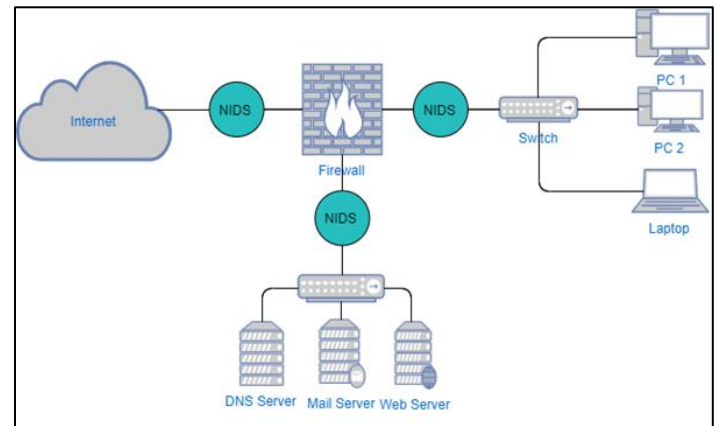


Figure1: Architecture of an IDS

## 4. Datasets

### 4.1. IDS-Dataset

Since the end of the 1990s, research work has been developing datasets for IDSs. These data sources are collected from network traffic simulators or by real systems, although the latter is very difficult to achieve because it may contain private data. One of the most widely used data sources is DRAPPA98 [17] published publicly in 1998 by MIT. Also, the dataset KDD-CUP 99 [18] published in 1999. This dataset contains the features of four categories of attacks (DOS, Probe, R2L, U2R). The major problem encountered is the considerable amount of redundant data which makes the use of this dataset difficult. We can also mention NSL-KDD dataset [19]. It is an updated version of the KDD-CUP 99. The NSL-KDD dataset is widely used for the network intrusion classification. Finally, we have other but less popular datasets such as CAIDA, DARPA, ISCX dataset, CICIDS.

### 4.2. Selecting-dataset

Previous researches are based on the two datasets KDD-CUP 99 and NSL-KDD for the classification of threats in computer networks. The authors of the works [13,20,21] have shown that these datasets encounter significant problems and their use is unreliable. These datasets suffer from a lack of diversity in traffic volumes. They contain duplicate data which does not cover real and recent attacks. To overcome these dataset issues, we have adopted the UNSWNB-15 [22] data source for training and testing our proposed model. Table 1 presents a comparison between the three datasets KDD-CUP 99, NSL-KDD, and UNSWNB-15.

Table 1: Comparison between the three datasets

Dataset	RNC	RNT	LO	TIC	FPC	MMC
KDD Cup 99	✓		✓	✓	✓	✓
NSL-KDD	✓		✓	✓	✓	✓
UNSW-NB15 (used in our method)	✓	✓	✓	✓	✓	✓



RNC: Realistic Network Configuration  
 RNT: Realistic Network Traffic  
 LO: Labelled Observations  
 TIC: Total Interaction Capture  
 FPC: Full Packet Capture  
 MMS: Many Malicious Scenarios

4.3. UNSWNB-15 dataset description

In this work, we have used a UNSWNB-15 dataset. This dataset was developed by ACCS (Australian Centre for Cyber Security) using IXIA Perfect storm tools. It is published in 2015 by authors Mustapha & Slay [23]. The UNSWNB-15 dataset contains 49 features categorized into six data associations (See Table 2 below).

Table 2: UNSWNB-15 dataset features

	N°	Features	Type
Flow features	1	srcip	N
	2	sport	I
	3	dstip	N
	4	dsport	I
	5	proto	N
Basic features	6	state	N
	7	dur	F
	8	sbytes	I
	9	dbytes	I
	10	sttl	I
	11	dttl	I
	12	sloss	I
	13	dloss	I
	14	service	N
	15	sload	F
	16	dload	F
	17	spkts	I
	18	dpkts	I
Content features	19	swin	I
	20	dwin	I
	21	stcpb	I
	22	dtcpb	I
	23	smeansz	I
	24	dmeansz	I
	25	Trans_depth	I
	26	Res_bdy_len	I
Time features	27	sjit	F
	28	djit	F
	29	stime	T
	30	ltime	T
	31	sintpkt	F
	32	dintpkt	F
	33	tcprrt	F
	34	synack	F
	35	ackdat	F
Additional features: general purpose	36	is_sm_ips_ports	B
	37	ct_state_ttl	I
	38	ct_flw_http_mthd	I
	39	is_ftp_login	B

Additional features: Connection features	40	ct_ftp_cmd	I
	41	ct_srv_src	I
	42	ct_srv_dst	I
	43	ct_dst_ltm	I
	44	ct_src_ltm	I
	45	ct_src_dport_ltm	I
	46	ct_dst_sport_ltm	I
Labelled features	47	ct_dst_src_ltm	I
	48	attack_cat	N
	49	Label	B

Type:

N: Nominal, I: Integer, F: Float, T: Timestamp, B: Binary

The UNSW-NB15 dataset provides two label features: ‘Attck\_cat’ represent the nine attack types (see table 3). It is used for a Multi-Class classification. The feature ‘label’ is used for binary classification (Normal and Attack). In this work, we choose to use a binary classification based on the second label feature.

Table 3: Types of attacks on UNSWNB-15

category	Sample
DOS	A DOS attack is a malicious activity in a computer network whose purpose is to prevent a network service from working for a given time.
Fuzzers	Is a method that randomly generates data to be injected into programs to detect possible vulnerabilities in a computer system.
Backdoors	A malicious program that provides access to a remote system for an unauthorized user.
Exploit	Consists of exploiting a software or hardware vulnerability in a computer network.
Analysis	This type of threat makes it possible to use vulnerabilities in a computer network to collect data useful for identifying future attacks.
Generic	This category opposes all block ciphers regardless of their structure. It concerns block cipher defined by given key sizes and blocks.
Recognition	This attack consists of discovering a computer system in order to detect vulnerabilities for future attacks.
Shellcode	A block of code to be injected into a computer program by a hacker. This makes it possible to exploit system deficiencies for malicious purposes.
Worms	A worm activates in a computer system by exploiting various vulnerabilities. It then spreads across the network by selecting other targets.

5. Proposed model

As we have previously shown, a signature-based-IDS is very efficient to detect the known attacks. However, it has a very low detection rate for recent attacks with unknown signatures. Figure 2 shows the Machine Learning steps followed in our model for the detection of attacks in a computer network. We used the SVM model to separate the two classes (Attack and No-Attack). Since in our case the data are non-linearly separable, we used the two functions with Polynomial and Gaussian kernel. These two functions are known by their good performances at the classification level in the SVM.

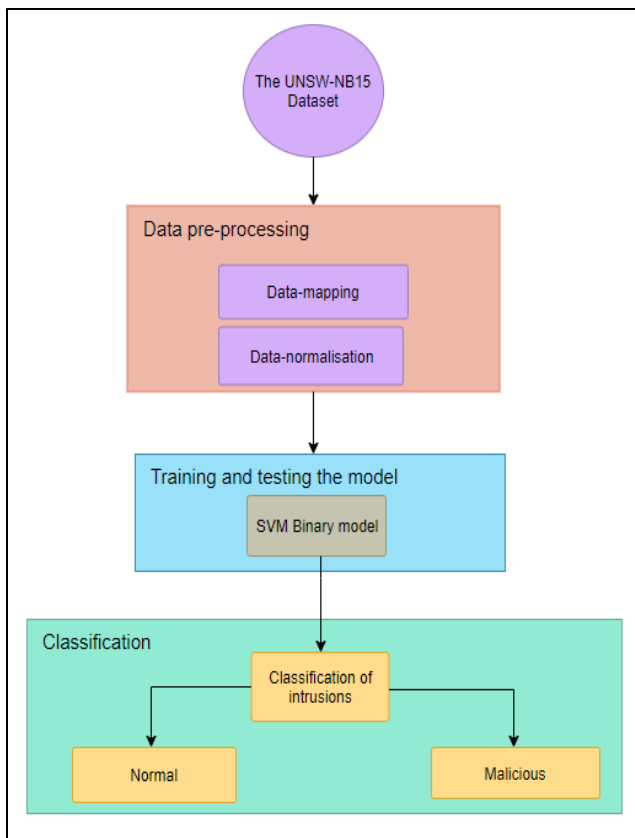


Figure 2: SVM binary classification proposed model

In our experience, the UNSWNB-15 dataset is composed of both CSV (Comma-separated values) files called UNSW\_NB15\_training\_set.csv and UNSW\_NB15\_testing\_set.csv. The first file contains 175 341 records for training the model (68%). However, the second contains 82 332 records for testing the model (32%). This dataset contains an integer, float, timestamp, and binary data type [23].

5.1. Data preprocessing

To train and test our model, a pre-processing step is necessary to select useful data ready for classification.

- Data-mapping: the first pre-processing step is data-mapping. It consists in transforming nominal data into numerical data (see figure 3). For example, the state field contains nominal data (INT, FIN, CON). In this phase, we converted these nominal values into numerical values (1,2,3).

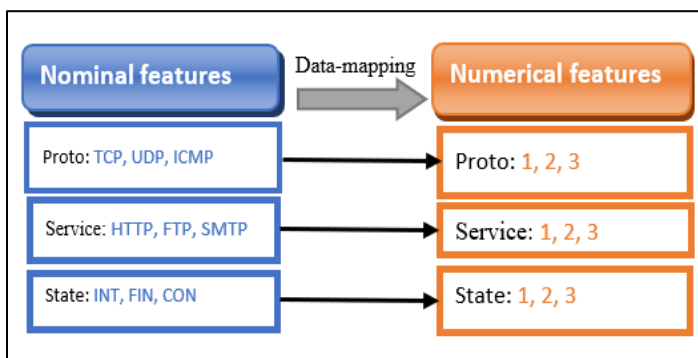


Figure 3: Data-mapping

- Data normalization: this step consists of increasing the speed of data processing. In this work, we applied the Min-Max method for data normalization. This consists in putting the numerical data in the margin [0-1] according to the following formula (1):

$$X_{normalised} = \frac{(x - \min(X))}{(\max(X) - \min(X))} \quad (1)$$

5.2. The SVM classifier

In this work, we have two observed classes (Attack and No-Attack). We used the SVM model for binary classification. This model is one of the algorithms of Machine Learning that have provided great performances either in classification or regression. It is a supervised learning model developed by Vapnik. In the case of classification, as in our case, the SVM model consists in finding a better boundary between the two observed classes (attack and No-Attack) (see figure 4). In this case, the data are non-linearly separable. For this purpose, we propose the kernel trick to have an optimal data separation as described in the work [24]. We assign two kernels one Gaussian and the other Polynomial because the data of the UNSWNB-15 dataset are also not linearly separable.

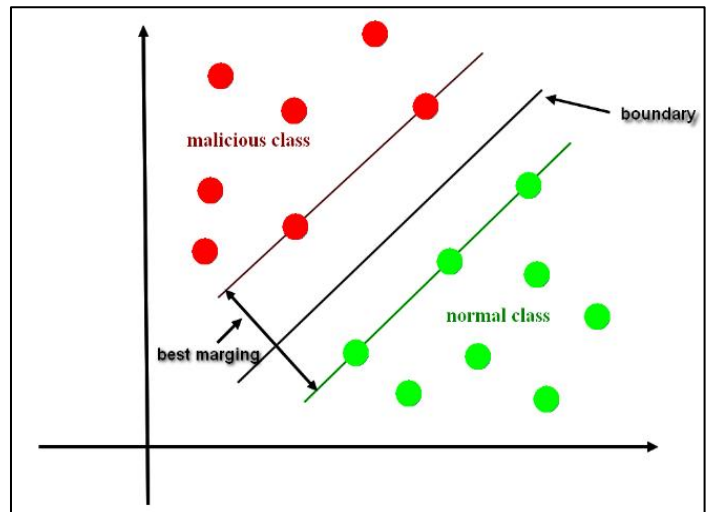


Figure 4: SVM Model hyperplane separation

6. Results and discussions

6.1. Criteria for Evaluating an IDS

The evaluation of an IDS is dependent on two main criteria:

- The Reliability of the IDS: The IDS must produce an alert for each incoming threat. An unreported intrusion is a system failure.
- The relevance of the IDS: Each alert must relate to a real threat. If we have a normal input event and it is classified as an intrusion also considered as well as a system failure. Table 4 presents the 4 possible cases.

Table 4: The confusion matrix

	Predicted Normal	Predicted Abnormal
Normal event	TN	FP
Abnormal event	FN	TP

TP (True Positive): Normal traffic is classified normal  
 FP (False positive): Normal traffic is classified abnormal.  
 FN (False Negative): abnormal traffic is classified normal  
 TN (True Negative): abnormal traffic is classified abnormal.

In our case, the confusion matrix gives the number of right and wrong predicted attacks, the same for the normal class (No-Attack).

6.2. Evaluation metrics of an IDS

To illustrate the effectiveness of our model, in this work we used six evaluation metrics.

We begin with the TPR (True Positive Rate) or Sensitivity. It describes the proportion of the positive class (Attack) correctly classified. The TPR calculates the quality of the IDS in terms of detection. To calculate the TPR, we use formula (2), note that the TPR must be close to 1 for a good IDS.

$$TPR = \frac{TP}{(TP+FN)} \tag{2}$$

On the opposite, the FPR (False Positive Rate) or Specificity. It describes the proportion of the negative class (No-Attack) correctly classified. To calculate the FPR, we use the formula (3), note that the FPR must be close to 0 for a good IDS.

$$FPR = \frac{FP}{(FP+TN)} \tag{3}$$

An effective IDS should provide a TPR close to 1 and a FPR close to 0.

The third evaluation metric is accuracy, which provides the true predictions for the entire test base. We calculate the accuracy according to the following formula (4):

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{4}$$

Accuracy calculation (PR) provides the proportion of true positive predictions in the forecast class (Attack). The accuracy is shown by the following formula:

$$Precision = \frac{TP}{TP+FP} \tag{5}$$

The fifth metric is RECALL (RE). It returns the proportion of correct predictions in the attack class. It is calculated according to the formula (6):

$$RE = \frac{TP}{TP+FN} \tag{6}$$

The last metric that we have used is F1-Score. This is a calculation of the harmonic average between precision (PR) and recall (RE). It is calculated according to formula (7):

$$F1 - score = \frac{2 \times (PR \times RE)}{(PR + RE)} \tag{7}$$

6.3. The obtained results

Our model is tested with the UNSWNB-15 dataset. The SVM classifier is used with two kernel functions Polynomial and

Gaussian for binary classification. With this model, we obtained good results in terms of DR for the classes normal and malicious. The figures 5 and 6 show, respectively the confusion matrix for the two kernels SVM-Gaussian and SVM-Polynomial.

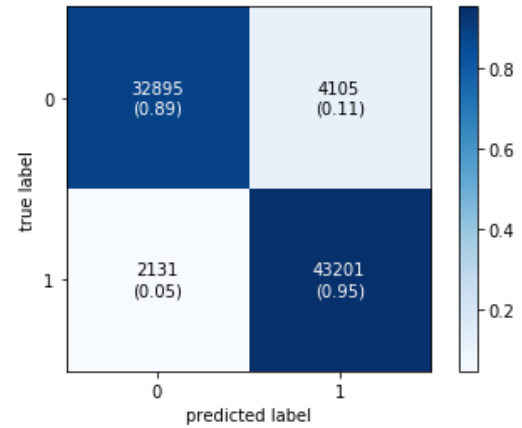


Figure 5: Confusion Matrix for the SVM- Gaussian model

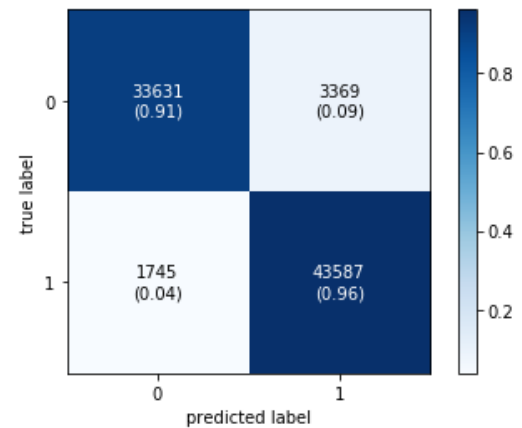


Figure 6: Confusion Matrix for the SVM-Polynomial model

To present our obtained results, we have plotted the ROC curves for the two kernels. These curves represent the TPR (Sensitivity) as a function of FPR (Specificity). They show that we have a very high rate of true positives (close to 1) against a low rate of false alarms (close to 0). This validates the performance of our model.

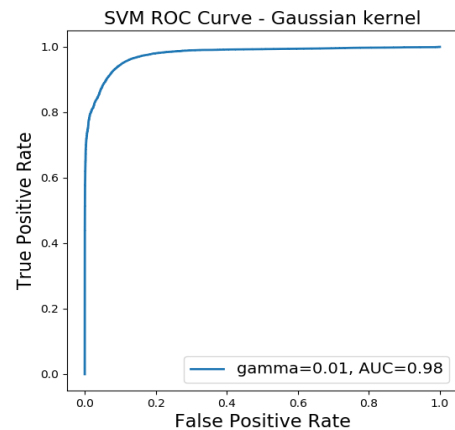


Figure 7: ROC curves correspond to Gaussian kernels

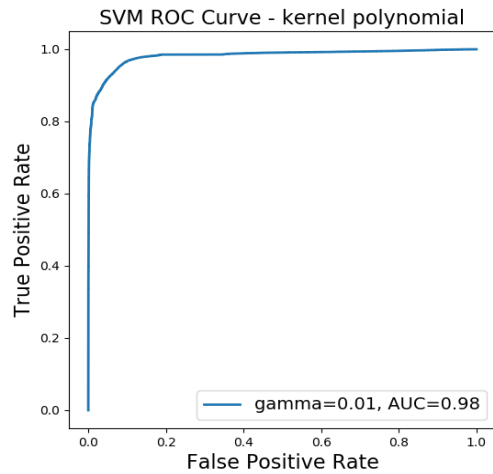


Figure 8: ROC curves correspond to Polynomial kernels

Tables 5 and 6 show the performance of the two kernels SVM-Gaussian and SVM-Polynomial:

- With the Gaussian kernel, we obtained a DR of 94% for the normal class and 91% for the malicious class.
- For the Polynomial kernel, we found slightly high results compared to the Gaussian function. We obtained a detection rate of 95% for the normal class and 93% for the abnormal class.

Table 5. SVM Gaussian kernel results

Class	Precision (5)	Recall (6)	F1-Score (7)
No-Attack	0.94	0.89	0.91
Attack	0.91	0.95	0.93
Accuracy (4)	0.93	0.92	0.92

Table 6. SVM Polynomial kernel results

Class	Precision (5)	Recall (6)	F1-Score (7)
No-Attack	0.95	0.91	0.93
Attack	0.93	0.96	0.94
Accuracy (4)	0.94	0.94	0.94

#### 6.4. Performance evaluation of the proposed model

To evaluate the effectiveness of our model, we compared our obtained results with those of other work using the same UNSWNB-15 dataset (Table 7). The comparison results of our SVM model with other classification models are good. We obtained for the Polynomial kernel an accuracy equal to 94% as shown in the following table:

Table 7. Comparison with of other models

Reference	Model	Accuracy	F1-score
[25]	RepTree	88.95	-
[26]	RandomForest	90.3	92.4
[27]	MLP	86.7	89.22
Proposed model	SVM Kernel-Gaussian	93	92
Proposed Model	SVM Kernel-Polynomial	94	94

## 7. Conclusions and perspective

In this work, we have developed a novel method for detecting attacks in computer networks. For this purpose, we presented a binary model for the classification of intrusions. We used the UNSWNB-15 data source. This source is widely used in the field of cybersecurity. We based on the SVM classifier with two kernel functions SVM-Polynomial and SVM-Gaussian. The evaluation results show high performance for both Gaussian (Accuracy = 93%) and Polynomial (Accuracy = 94%) kernels. In perspective, we will opt for a multi-class classification for the detection of each type of intrusion separately.

## References

- [1] A. BACHAR, N. E. MAKHFI and O. E. Bannay, "Towards a behavioral network intrusion detection system based on the SVM model", 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, pp. 1-7, 2020. doi: [10.1109/IRASET48871.2020.9092094](https://doi.org/10.1109/IRASET48871.2020.9092094).
- [2] L.A. Maglaras et al., "Cyber security of critical infrastructures. ICT Express, 4(1), 42-45, 2018. doi: [10.1016/j.icte.2018.02.001](https://doi.org/10.1016/j.icte.2018.02.001)
- [3] K. Choo, "The cyber threat landscape: Challenges and future research directions". Computers & Security 30:719-731, 2011. doi: [10.1016/j.cose.2011.08.004](https://doi.org/10.1016/j.cose.2011.08.004)
- [4] M. Aydin, M. Ali, A. Halim Zaim, and K. Gökhan Ceylan. "A hybrid intrusion detection system design for computer network security", Computers & Electrical Engineering, p 517-526, 2009. Doi: [10.1016/j.compeleceng.2008.12.005](https://doi.org/10.1016/j.compeleceng.2008.12.005)
- [5] D.K. Bhattacharyya, J.K. Kalita, "Network Anomaly Detection: A Machine Learning Perspective": CRC Press, 2013.
- [6] J. Cannady, "Artificial Neural Networks for Misuse Detection," Proc eedings, National Information Systems Security Conference (NISSC '98 ), October, Arlington ,VA, 443 -456. 1998.
- [7] P. Natesan, P. Balasubramanie, "Multi Stage Filter Using Enhanced Adaboost for Network Intrusion Detection. International Journal of Network Security & Its Applications", 4(3), 121-135, 2012. doi: [10.5121/ijnsa.2012.4308](https://doi.org/10.5121/ijnsa.2012.4308)
- [8] O. Depren et al., "An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks. Expert Systems with Applications", 29(4), 713-722, 2005. doi: [10.1016/j.eswa.2005.05.002](https://doi.org/10.1016/j.eswa.2005.05.002)
- [9] Divyatmika, M. Sreekesh, "A two-tier network based intrusion detection system architecture using machine learning approach. International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), 2016. doi: [10.1109/iceeot.2016.7755404](https://doi.org/10.1109/iceeot.2016.7755404)
- [10] G. Zhou, "An Effective Distance-Computing Method for Network Anomaly Detection. Security Technology Communications in Computer and Information Science", 177-182, 2011. doi: [10.1007/978-3-642-27189-2\\_19](https://doi.org/10.1007/978-3-642-27189-2_19)
- [11] G. Meena et al., "A review paper on IDS classification using KDD 99 and NSL KDD dataset in WEKA". International Conference on Computer, Communications and Electronics (Comptelix), 2017. doi: [10.1109/comptelix.2017.8004032](https://doi.org/10.1109/comptelix.2017.8004032)
- [12] L. Dhanabal and S. P. Shantharajah, "A study on NSL-KDD dataset for intrusion detection system based on classification algorithms,"International Journal of Advanced Research in Computer and Communication Engineering, 4(6), 446-452, 2015. doi: [10.17148/IJARCCCE.2015.4696](https://doi.org/10.17148/IJARCCCE.2015.4696)
- [13] N. Moustafa, J. Slay, "The Significant Features of the UNSW-NB15 and the KDD99 Data Sets for Network Intrusion Detection Systems". 4th International Workshop on Building Analysis Datasets and Gathering Experience Returns for Security (BADGERS), 2015. doi: [10.1109/badgers.2015.014](https://doi.org/10.1109/badgers.2015.014)
- [14] W. Hu, S. Maybank, "AdaBoost-Based Algorithm for Network Intrusion Detection. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 38(2), 577-583, 2008. doi: [10.1109/tsmcb.2007.914695](https://doi.org/10.1109/tsmcb.2007.914695)
- [15] A.L. Buczak, E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. IEEE Communications Surveys &amp; Tutorials, 18(2), 1153-1176, 2016. doi: [10.1109/comst.2015.2494502](https://doi.org/10.1109/comst.2015.2494502)
- [16] P. Garc'ia-Teodoro, J. D'iaz-Verdejo, G. Maci'a-Fern'andez, E. V'azquez, "Anomalybased network intrusion detection: Techniques, systems and challenges. Comput. Secur. 28(1-2), 18-28, Feb 2009. <https://doi.org/10.1016/j.cose.2008.08.003>
- [17] L. Lincoln, "DARPA datasets. In: DARPA 1998 & 1999. Datasets. <https://www.ll.mit.edu/r-d/datasets>. Accessed June 23, 2020



- [18] KDD Cup (1999) KDD Cup 1999 Data. In: KDD Cup 1999 Data. <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>. Accessed June 23, 2020
- [19] NSL-KDD (2009) NSL-KDD | Datasets | Research | Canadian Institute for Cybersecurity | UNB, <https://www.unb.ca/cic/datasets/nsl.html> Accessed June 23, 2020.
- [20] M. Tavallaei, E. Bagheri, W. Lu, A.A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set". 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, 2009. doi:[10.1109/cisda.2009.5356528](https://doi.org/10.1109/cisda.2009.5356528)
- [21] N. Moustafa, J. Slay, "The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set". Information Security Journal: A Global Perspective, **25**(1-3), 18-31, 2016. doi:[10.1080/19393555.2015.1125974](https://doi.org/10.1080/19393555.2015.1125974)
- [22] N. Moustafa, "UNSW\_NB15 dataset", IEEE Dataport, 2019. Online Available: <http://dx.doi.org/10.21227/8vf7-s525>. Accessed June 23, 2020.
- [23] N. Moustafa, J. Slay, UNSW-NB15: "A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)". 2015 Military Communications and Information Systems Conference (MilCIS), 2015. doi:[10.1109/milcis.2015.7348942](https://doi.org/10.1109/milcis.2015.7348942)
- [24] B.M. Aslahi-Shahri et al., "A hybrid method consisting of GA and SVM for intrusion detection system. Neural Computing and Applications", **27**(6), 1669-1676, 2015. doi:[10.1007/s00521-015-1964-2](https://doi.org/10.1007/s00521-015-1964-2)
- [25] M. Belouch, S. El, M. Idhammad, "A Two-Stage Classifier Approach using RepTree Algorithm for Network Intrusion Detection". International Journal of Advanced Computer Science and Applications, **8**(6), 2017. doi:[10.14569/ijacsa.2017.080651](https://doi.org/10.14569/ijacsa.2017.080651)
- [26] R. Vinayakumar et al., "Deep Learning Approach for Intelligent Intrusion Detection System". IEEE Access, **7**, 41525-41550, 2019 doi:[10.1109/access.2019.2895334](https://doi.org/10.1109/access.2019.2895334)
- [27] J. Yan, D. Jin, C. Lee, P. Liu, "A Comparative Study of Off-Line Deep Learning Based Network Intrusion Detection". 2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN), 2018. doi:[10.1109/icufn.2018.8436774](https://doi.org/10.1109/icufn.2018.8436774)

## COVID-19 Crises: Global Economic Shocks vs Pakistan Economic Shocks

Naveeda K. Katper<sup>1</sup>, Muhammad Nawaz Tunio<sup>2\*</sup>, Niaz Hussain<sup>3</sup>, Attaullah Junejo<sup>4</sup>, Faheem Gul Gilal<sup>5</sup>

<sup>1</sup>*Institute of Business Administration, University of Sindh, Jamshoro, 76080, Pakistan*

<sup>2</sup>*Doctoral Track Entrepreneurship, Innovation and Economic Development, Alpen Adria University Klagenfurt, Klagenfurt, 9020, Austria*

<sup>3</sup>*Department of Syriah and Management, University of Malaya, Kuala Lumpur, 50603, Malaysia*

<sup>4</sup>*Office of Accountant General Sindh, Karachi, 75400, Pakistan*

<sup>5</sup>*Department of Business Administration, Sukkur IBA University, Sukkur, 65200, Pakistan*

### ARTICLE INFO

*Article history:*

*Received: 01 June, 2020*

*Accepted: 23 July, 2020*

*Online: 25 August, 2020*

*Keywords:*

*Contagious virus*

*COVID-19*

*Economic Shocks*

*Socio-psychology*

*Behavioral economics*

### ABSTRACT

*This paper documents, information about the current outbreak of Corona virus and its multidimensional effects witnessed worldwide. The objective of the study is to disseminate information to policymakers, government and the general public, to facilitate the efforts for fighting against the contagious virus and aims to contribute to the economic and social welfare. Though the data changes rapidly by every passing minute, yet the current scenario may be analyzed to evaluate the present situation for a fruitful outcome. Hence, the discussion may be useful for all the stakeholders (government, policymakers and the public) for their reference. Various studies previously observed that widespread contagious diseases have a severe impact on the economy. It is observed that the faster and wider the outbreak of disease, the higher and severer the influence on the economies. The economies face shrinkage of income, revenue, earnings, while health expenditure increases for hospitals, medical kits, ventilators, uniforms, masks, gloves, medicines and other medical equipment. Therefore, it has a two-pronged adversely hitting effect on economies to manage the imbalance in the budget and fiscal matters. Another unfortunate situation is when the high populated country is poor too. The world has realized the effects of the different contagious viruses in the different times in history. Nevertheless, the virulence and the severity of the COVID-19 is far greater than the previous viral breakouts. We, in this research attempt to highlight the multidimensional effects of COVID-19 and create some possibilities how to deal and fight with this cruel and invisible enemy.*

## 1. Introduction

Where there are many advantages of globalization, there are disadvantages attributed to it, as well. What we are facing currently in the shape of contagious viral breakout worldwide is one such example. Therefore, one should not think that the disaster that happens in one country would not affect the other. Everything is affected because of the social interaction and shared economies around the globe. Education, skills, tourism, business and human resource mobilization are major factors which create a cause for connecting the people and making the world a global village.

\*Corresponding Author: Muhammad Nawaz Tunio. Mozartstrasse 61, Klagenfurt 9020, Austria. [m1tunio@edu.aau.at](mailto:m1tunio@edu.aau.at) Cell No. 004368860041579

What we are facing currently in the shape of contagious viral breakout worldwide is one such example. Therefore, one should not think that the disaster that happens in one country would not affect the other. Everything is affected because of the social interaction and shared economies around the globe. Education, skills, tourism, business and human resource mobilization are major factors which create a cause for connecting the people and making the world a global village.

In this connection whenever the good comes to one country, it reaches to other also by imports, social attitudes or demonstration effect. Accordingly, when evil occurs in one country, the same comes to the rest of the world. Today we the habitants of this global village are facing the same evil occurrence in the shape of

the novel coronavirus. It is only due to the interconnection of the countries of the whole world called global village. This strengthens the concept that everything has its pros and cons on its side.

Similarly, where the globalization has shared knowledge, expertise, technical skills and productive output of resources among countries, it shares bad events there too, which weakens the ability of workforce as being witnessed today as an aftermath of COVID-19 outbreak. Resultantly, cause deaths and morbidity in human resource globally. According to [1] “In a strongly connected and integrated world, the impacts of the disease beyond mortality (those who die) and morbidity (those who are unable to work for some time) have become apparent since the outbreak”. They further elaborate that due to the novel coronavirus, a panic is created between business and consumers. That has created market anomalies and slanted consumption patterns. International financial markets and global stock indices have plunged in the response of changing world conditions.

Moreover, the intensity of loss increases when the country is underdeveloped like Pakistan. The socio-economic effects of COVID-19 are dispersed with greater force across developing countries and their emerging markets. Such increasing effect has directly hit their income adversely. The loss of employment creates declining GDP, national income GNP and per capita income leading to the rise in poverty. As in 2019 more than 25% of the population lived below the poverty line and more than 29.5% of the population considered poor and depends on agriculture. Such a component of the population is expected to reach 40%, according to the World Bank report (see Figure 1). The graph shows the Poverty Headcount Ratio since 2001 to 2020f (forecast). “Poverty Headcount Ratio”:

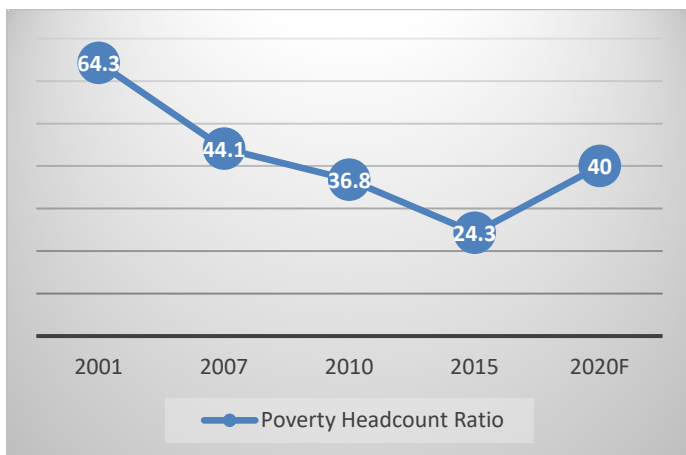


Figure.1 Poverty headcount ratio (Source: World Bank)

Pakistan’s per capita income is already lower. The average per capita income is estimated at 1250 USD. Such a lower income rate cannot even fulfil the basic needs of people. Higher unemployment rates around six per cent, according to “Trading Economics, global macro models and analysts expectations”<sup>1</sup>. Most of the labor force comprising of daily wage and temporary employment under the informal sector, worried about their bread and butter. The industry is not as strong as in developed countries.

<sup>1</sup><https://tradingeconomics.com/pakistan/unemployment-rate>  
[www.astesj.com](http://www.astesj.com)

Pakistan relies heavily on the imports for most technical products, automobiles, machinery and oil. Similarly, exports are more important for these developing countries whose foreign exchange is not at a satisfactory level, rather indebted to international institutions heavily and chronically. In this backdrop, the outbreak of COVID-19 is a big disaster for Pakistan threatening all spheres including its crippling economy.

Here the research question arises, due to COVID-19 what aspects of the economy of the world and Pakistan, in particular, are affected? To what extent the world is experiencing psychological and socioeconomic shocks? And how far their impact would be seen?

Therefore, our objective is to evaluate and analyze the effects of this epidemic COVID-19 on the world and Pakistan economy. To what extent, particularly psychological, social and economic hits the globe is facing currently. In this study, we try to focus on the emerging threats jeopardies and vulnerabilities in the world economy that may be considered. The economic costs of viral outbreaks were different in the past as compared with the current worldwide spread of viral disease. The economic cost is directly proportionate with the rise of the virus; rather the cost has a multiplying impact for the least developed economies. Deaths and illness or morbidity further create chaos in multi spheres. It also sprouts the disability and uncertainty at the workplace in future.

The paper is structured further as below; section 2 provides some previous studies on the pandemic. Section 3 briefs the method to articulate the study while section 4 discusses and analyses the different economic shocks worldwide and in Pakistan by highlighting its different aspects. Section 5 concludes the paper and section 6 suggests some recommendations on the issue.

## 2. Previous studies on pandemics

The economy cannot be separated from the health of the population as the human resource is the major factor of production. Therefore, the health of people and economies go together either grow/flourish or weaken/collapse. A lot of research is conducted on the issue that has found that the health of the population is closely concerned with economic welfare and economic growth [2-9]. Other studies on SARS, including [10] for Taiwan, [11] for China and [12] for Hong Kong also elaborate the worse situation of the affected regions. These studies directly or indirectly can be connected to this current issue, however the effects higher than the historical contagious diseases. We can also comprehend that the health of people and the health of the economy are directly connected.

In the 20<sup>th</sup> century, the historical breakouts of epidemics that world witnessed are: “Spanish Influenza of 1918, Asian flu of 1957 and the Hong Kong flu of 1968. In the 21<sup>st</sup> century, outbreaks are NIHI in 2009 (‘bird flu’), Severe Acute Respiratory Syndrome (SARS) in 2002, Middle East Respiratory Syndrome (MERS) in 2012, and Ebola which peaked in 2013-14” [13]. All these pandemics spread panic in the regions when they broke out and infected population at large scale by health and economic depressions. For example, in 2003 by the breakout of SARS, the economic problems increased. There was a huge reduction in the demand or consumption and the cost of operating a business which increased the uncertainties and risk premium. The spread

of disease crossed borders, according to the interaction of the host country's exposure and predisposition to the disease. Similarly, [14], describes that AIDS/HIV affects the economy. Government, business and households suffer through changing labor supply, labor efficiency, growing cost of business and confining investment in training of staff by firms, increased health care expenditure by supporting disabled people and orphan children due to AIDS in public sectors.

The researchers elaborate that despite a low level of health damage, mortality rate and the small number of deaths the world suffered a lot in terms of economic loss. The global cost witnessed that it is not only specific to the countries affected adversely [15]. Conversely, the impact of COVID-19 is far more than the formerly experienced contagions because this is not the story of a few countries, rather the disaster of the whole world consisting around 200 countries. This epidemic impact on the health of people and the economy is much more than World War I and World War II. These wars did not infect 200 countries all around. The unexpected and sudden blow-out of coronavirus has put the entire world into the shock. According to [15] "*The fear of 1918-19 Spanish influenza, the deadliest plague in history, with its extreme severity and gravity of clinical symptoms, is still present in the research and general community*". Likewise, at the start of breakout people think that COVID-19 has the same level of fear as the other former contagious diseases. On the contrary, the COVID-19 has infected at greater extent by multi spheres all over the world. Plenty of research is available about previously highly infectious observed diseases of epidemic for which no vaccine prepared yet. All that literature providing experience from these aforementioned epidemics can be informative to some extent how to think and deal regarding COVID-19 for implications [16].

However, this time the invisible enemy COVID-19 has played far beyond the previous epidemics did. According to CEPR Press, 2020, the number of people affected by the COVID-19 is much greater than the SARS cases. Hence, due to the fear of novel coronavirus, countries have started lockdown. Lockdown is creating a gap in the world level economy. The contagious virus is going to be affected so badly that countries went for lockdown and curfew.

To see the global situation, John Cochrane in the chapter of eBook indicates some bold queries about lockdown mechanism. That there will not be any disaster if a short pause is given to economize for a certain period for the sake of public health and socioeconomic prosperity. For this, there should be a systemic way and constructive approach in implementing it properly and completely. The researcher displays worry about huge financial issues even for firms' disbursements or payments of debts, mortgage, salaries, wages, rents, utilities. Subsequently, when the business shuts during lockdown under demand shock, how can household and firms manage their fixed costs? This fear is felt and John Cochrane writes that it is possible to occur economic shocks like bankruptcies, insolvencies as well as inability to pay the bills. Lockdown situation will prolong across economic fluctuations as happened in 2001 and 2008.

The great influence of the COVID-19 on the economy has been started due to the prolonged worldwide shut down. What repercussions the world is observing the matter of attention and focus so that some way out could be found to cope with this perilous condition.

### 3. Methodology

In this study, our objective is "to evaluate and analyze the effects of this epidemic COVID-19 on the world and Pakistan economic shocks and what particular psychological, social and economic hits the globe is facing currently". Therefore the methodology we used is mostly through collecting secondary data by desk research we referred published reports and statistics by different sources ranging from national and international surveys, reviews and publications [17]. For that purpose, we conducted desktop research, accordingly. We also reviewed and referred to the research papers on the subject. In this study, we have analyzed the current pandemic situation and discussed its multi sphere effects on economic factors during the occurrence of the pandemic situation. We evaluated the world economic shocks as well as Pakistan economic shocks being faced by the masses.

### 4. Analysis and discussion on the multi sphere, economic shocks

#### 4.1. Global economic shocks

As the first attack of COVID-19 in China, the second-largest economy behind US 1/3 of world manufacturer and world's largest exporter shut down after viral breakout. According to Statista expected report 2019, China's GDP contributes to 19.24% of the global total in 2019 and 19.72% in 2020; the same is expecting 21.39% in 2024. Despite that China's economic effect is widely noticed and spread globally.[18], in the ebook of CEPR Press, mention that the Chinese economy showed a sluggish pattern in the 1<sup>st</sup> quarter of 2020. While looking at the global economic situation and economic contribution, China stands high by contribution in global GDP by 17%, in world trade by 9%, in global tourism by 9.3% and in global demand by 40% which leaves not a positive impression and influence to the other countries. This does not affect china only, but also the supply and demand of the rest of the country.

However, it is observed that the World GDP is condensed up to 0.75%, by this the impact on world GDP growth is being around half a present point in 2020. The reason for this decline is due to the initial curtailment in the demand in China. Thus China's international trade has significantly declined by around 1.4% in the 1<sup>st</sup> half of and 0.9% in the whole year of 2020. China's cross border linkage is strong with other countries through commodity exporters of Japan, Korea, East and South-East Asian countries ranging from the supply chain, tourism and other travel-based services. All other countries of America, Europe, Middle East and Asia are growing seriously infected day by day. Therefore, the whole world is facing isolation and lockdown due to which all economic activities are stagnant. All together the countries affected over 70% of global GDP in terms of PPP. Thus, the big economies of the world are infected by this invisible enemy at their worst. The world's GDP trend of various countries since 2017 -2020 can be seen in the table 1.

According to [19]. "Taking just the US, China, Japan, Germany, Britain, France, and Italy, they account for 60% of world supply and demand (GDP), 65% of world manufacturing, and 41% of world manufacturing exports". Due to a strong supply chain in the regional dimension, China, Japan, Korea at hard-hitting and this shock of the supply chain will be extended to whole Asia. According to an estimate, the global GDP has



declined by 1.75% in the latter half of 2020, whereas it is to be declined around 1.5% in the full-year GDP growth impact in 2020. Global trade is, decreasing approximately by 3.75% in 2020, hammering exports in all economies.

The global panic of COVID-19 has created unprecedented uncertainty to a level which caused a massive reduction in oil prices, resulted by declining of investment in China and the rest of the world. IEA further foresees the global oil demand falling by “435,000 barrels per day when compared at the year-on-year basis. The current quarter compared with the corresponding period of last year reveals that the world has witnessed the highest shrinkage in the economy of the decade. It is expected further that the global oil demand will further touch to the low by “365,000 barrels per day”. It is said to be the worst behavior of demand since 2011. Consequently, “OPEC plus countries consider the reduction in oil production of 600,000 barrels a day as an emergency measure on the top of the 1.7 million barrels a day already pledged” [20]. This declining scenario is observed only from the direct effect of shrinkage of demand-supply side. However, the impact of the risk and future uncertainty is yet to cripple in slow the days to come.

Furthermore, the financial sector and the global stock markets have witnessed crashed down situations all over the world. Stock exchanges are experiencing a decline in market capitalization on an hourly basis. European stock indices, American stock markets witnessed worst drops since March 20, 2020. Financial crunch and stock market of developing countries would lead their economies into a hazard. Puts it: “*Developing economies in East Asia and the Pacific (EAP), recovering from a trade war and struggling with a viral disease, now face the prospect of a global financial shock and recession. Given the unprecedented nature of the economic shock to each country, and the fact that it is also affecting all other countries in the region and beyond, an exceptional policy response is needed*”. Hence, having more than 70% of the poor population in developing countries is more vulnerable to the situation. The production has declined and so the employment level due to the shutter down and bans imposed for halting the outbreak of this pandemic. The low-grade staff or daily wage labor is burdening the economy in two ways. From one side they are unproductive whereas from the other side the governments have to arrange their bread.

#### 4.2 Multidimensional hits of COVID-19

COVID-19 has created chaos and hitting in multi spheres globally. The effect of this contagious virus is discussed below along with the decoration of the graphical presentation.

#### 4.3 Demographic hits

The rampant and contagious coronavirus has created a global emergency. The demographic change has occurred dramatically. It has hit the health of people in millions across the world. The figures of infected people change quickly in minutes all over the world. The reasons are all other kinds of flu viruses do not transmit faster, the way it transmits from one person to another. Its mortality rate is 2% to 3%, which is greater than other viruses.

Thus, what people have faced already is terrible. Increasing deaths or mortality rates, illness and morbidity by damaging health have created a big demographic change in the world.

According to [21] who runs preparedness efforts for NYC Health and Hospitals, said: “This particular virus seems like it is highly transmissible, I think that it is certainly plausible that 40–70% of the world’s population could become infected with coronavirus disease, but a large number of cases are (expected to be) mild.”

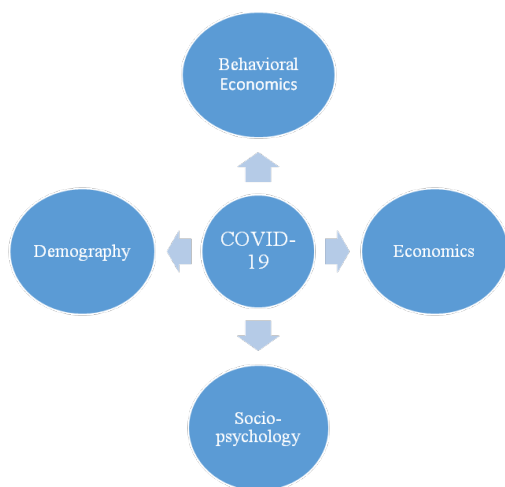
#### 4.4 Economic hit:

The novel coronavirus is witnessed as contagious economically as it is medically. The economy is suffering and gets a double hit in the following ways:

1. From one side “**revenue reduction**”: Coronavirus affects all type of enterprises specifically service sector and small & medium enterprises (SMEs). All forms of business organizations such as manufacturing, trade and service organizations are facing curtailment of operational activities. Millions of formal as well as informal employment and self-employment are affected. The business loss occurred as services have been curtailed globally. Transport being limited and even restricted among countries has further slowed down global economic activities. It has further minimized the consumption of oil and other fuels. Resultantly, Oil & petrol companies face revenue loss. The lockdowns resulted in low production, which has reduced the local and global trade. As the world economies also depend on international trade, which has dented the economies world over. Confined economic activities have resulted in revenue kerbing. For example, hotel booking cancellation, restaurants, marriage halls supermarkets, casinos, bars, pubs, saloons, and parlours shut down, limited supply chain, banned exports and imports, worldwide airline cancellation, and restricted tourism, all have affected the revenue generation and ultimately the economies of the world. Hence the loss of tax collection would also occur and the Governments would bear this burden on the exchequer. The fiscal and monetary policies, depending on government revenue collection, have witnessed major policy shifts. It ultimately shows a big economic depression for all economies.
2. From the other side “**expenditure extension**” For the national security purpose, huge expenditure is set for this battle to fight by purchasing expensive medical equipment in huge number. Every country is engaged in making budgets for vaccines, medicines, ventilators, isolation centres, hospitals and other medical apparatus. Addition to that spending on research and science in this particular direction to prepare apparatus, vaccines and medicines, as well as other health care expenditure, would be a severe problem for poor countries. Thus, revenue reduction and extending expenditure for poor economies is a big challenge. Because the loss of human lives and economies by COVID-19 is greater than the world witnessed before.

Table 1: Economic situation of the world

Real GDP (% change from previous years)	(% difference from January 2020 forecasts)						
	2017	2018	2019	2020f	2021f	2020f	2021f
<b>World</b>	3.3	3.0	2.4	-5.2	4.2	-7.7	1.6
<b>Advanced economies</b>	2.5	2.1	1.6	-7.0	3.9	-8.4	2.4
United States	2.4	2.9	2.3	-6.1	4.0	-7.9	2.3
Euro Area	2.5	1.9	1.2	-9.1	4.5	-10.1	3.2
Japan	2.2	0.3	0.7	-6.1	2.5	-6.8	1.9
<b>Emerging market and developing economies</b>	4.5	4.3	3.5	-2.5	4.6	-6.6	0.3
Commodity-exporting EMDEs	2.2	2.1	1.5	-4.8	3.1	-7.4	0.2
Other EMDEs	6.1	5.7	4.8	-1.1	5.5	-6.2	0.3
Other EMDEs excluding China	5.4	4.8	3.2	-3.8	3.5	-7.6	-0.8
<b>East Asia and Pacific</b>	6.5	6.3	5.9	0.5	6.6	-5.2	1.0
China	6.8	6.6	6.1	1.0	6.9	-4.9	1.1
Indonesia	5.1	5.2	5.0	0.0	4.8	-5.1	-0.4
Thailand	4.1	4.2	2.4	-5.0	4.1	-7.7	1.3
<b>Europe and Central Asia</b>	4.1	3.3	2.2	-4.7	3.6	-7.3	0.7
Russia	1.8	2.5	1.3	-6.0	2.7	-7.6	0.9
Turkey	7.5	2.8	0.9	-3.8	5.0	-6.8	1.0
Poland	4.9	5.3	4.1	-4.2	2.8	-7.8	-0.5
<b>Latin America and the Caribbean</b>	1.9	1.7	0.8	-7.2	2.8	-9.0	0.4
Brazil	1.3	1.3	1.1	-8.0	2.2	-10.0	-0.3
Mexico	2.1	2.2	-0.3	-7.5	3.0	-8.7	1.2
Argentina	2.7	-2.5	-2.2	-7.3	2.1	-6.0	0.7
<b>Middle East and North Africa</b>	1.1	0.9	-0.2	-4.2	2.3	-6.6	-0.4
Saudi Arabia	-0.7	2.4	0.3	-2.8	2.5	-5.7	0.3
Iran	3.8	-4.7	-8.2	-5.3	2.1	-5.3	1.1
Egypt	4.2	5.3	5.6	3.0	2.1	-2.8	-3.9
<b>South Asia</b>	6.5	6.5	4.7	-2.7	2.8	-8.2	-3.1
India	7.0	6.1	4.2	-3.2	3.1	-9.0	-3.0
Pakistan	5.2	5.5	1.9	-2.6	-0.2	-5.0	-3.2
Bangladesh	7.3	7.9	8.2	1.6	1.0	-5.6	-6.3
<b>Sub-Saharan Africa</b>	2.6	2.6	2.2	-2.8	3.1	-5.8	0.0
Nigeria	0.8	1.9	2.2	-3.2	1.7	-5.3	-0.4
South Africa	1.4	0.8	0.2	-7.1	2.9	-8.0	1.6
Angola	-0.1	-2.0	-0.9	-4.0	3.1	-5.5	0.7
<b>Memorandum items</b>							
<b>Real GDP</b>							
High-income countries	2.4	2.2	1.7	-6.8	3.8	-8.3	2.3
Developing countries	4.8	4.4	3.7	-2.4	4.7	-6.7	0.2
Low-income countries	5.4	5.8	5.0	1.0	4.6	-4.4	-0.9
BRICS	5.3	5.3	4.7	-1.7	5.3	-6.6	0.4
World (2010 PPP weights)*	3.9	3.6	2.9	-4.1	4.3	-7.3	1.0
<b>World trade volume*</b>	5.9	4.0	0.8	-13.4	5.3	-15.3	2.8
Commodity prices*							
Oil prices	23.3	29.4	-10.2	-47.9	18.8	-42.5	16.9
Non-energy commodity price index	5.5	1.8	-4.2	-5.9	3.0	-6.0	1.3



#### 4.5 Socio-psychological hit:

The socio-psychological effect of coronavirus is felt greater than ever in the history of viruses' breakouts at different times in the world before. Intensifying fears of a global recession also circulate in Governments, business communities and the general public of all countries. Panic has created a deep impression on the psychology of people. There is a challenge for doctors, nurses, police and army while dealing with patients, salute to them. The uncertainty of the future and risk of lives has created a great depression all over the world. People are suffering from depression in quarantine and isolation. Indeed, such pandemics spread a lot of fear and panic so that people cannot perform at their normal level. As the same effect was noticed in 1918 on the breakout of Spanish influenza and World Health Organization estimates a bigger number of people some more than 300 million people were suffering from stress and depression worldwide and becomes the leading cause of disability in human. Also, [22] describes that the invisible and lethal situation of virus creates not only the biological issue, but also a psychological pressure like a terrorism threat which leaves long-lasting consequences. However, due to COVID-19 in 2020, the extent of depression is far more than ever expected. [23] also opine the same by describing that a large proportion of the masses will observe the risk of pandemic whereas the risk of death will not be equal.

Due to physical and psychological distress in patients either they lose their appetite temporarily before and even after treatment because they face the traumatic situation and see the death closely or relish the meal when they get back to normality. Patients become reluctant, doubtful, angry, exhausted, regretted in stress. Therefore, they need support with greater zeal.

#### 4.6 Behavioral economics

The economy is also affected by the various behaviors/psychological reactions of different stakeholders through high/low demand and high/low supply shocks and other reasons. For example: The investment decision of investors in fluctuating certain/uncertain conditions, by the consumer behavior in varying situations such as high/low-risk taking, healthy/unhealthy conditions, fear of injury, infection, fear of death, fear of loss, due to precaution, government instructions, quarantine, lockdown, curfew and other various reasons. In the

result, demand curtails and suppliers also hold the production according to limitations in orders and apply wait and watch policy. Labor force is prone to be infected with such pandemic diseases. The contagious disease affects labor from many aspects, e.g. infections, disability and deaths. Therefore, from the human resource side the production process slows down and employment rate reduces, hence supply curtails. Every economic indicator becomes vulnerable.

During the attack of COVID-19 hand washing, hygiene, diet, clean environment, mask and gloves wearing is projected. Likewise, by changing behavior, goods and services change. Manufacturing takes place in that connection.

On the other hand, people in quarantine can save much living cost, no social interaction/ gathering, no tour or outing cost, no hotel, restaurant, travel, parlor and theatre/cinema cost, overall precautionary steps to reduce the viral effect too. As [24] expresses the quarantining period is social isolation and disintegrate from the period, which results in the social pressure and psychological down. Similarly, [25] state that undertaking hygiene measures are not a big deal because it minimizes the possibility of contagion, which is very favorable for healthy economic and social conditions.

From the experience of these cruel epidemics, if the world population adopts some hygiene based behaviors and make them their habit to keep themselves safe and secure would reduce the transmissions and decrease the economic cost. Likewise, [26] emphasize the habit formation for desirable behaviors during the pandemic outbreaks. Subsequently, these behaviors offset the cost of contagious outbreaks exerting positive health effects and decrease the risk of viruses like COVID-19. Contagious diseases and behavioral economics have progressed recently for developing the change in behavior and understanding the nature of outbreaks. Scientists and policy makers can coordinate between the two approaches which can lead to reducing the speed of transmission of contagious and overall loss [27].

#### 1. Pakistan economic shock

Pakistan being a developing country is already suffering from multiple socioeconomic problems. The outbreak of COVID-19 has dragged the poor country into more multifarious issues, to deal with.

Unfortunately, like global stock markets, freaking out Pakistan stock exchange faces curbing of the investment and production. The decrease in employment/ labor force is hitting the capital formation and production badly. Lockdown and restricted movement are disturbing channelization of raw material and finished goods. Plunging demand and sapping supply due to lock down and transportation restrictions has created a gap between suppliers and consumers. Limited movement of agricultural and industrial goods (raw material and finished goods) and nominal exploitation or uses of natural resources are increasing problem by narrowing economic activities. It also has slowed down the Investment process. This situation is responsible for the lower performance of the economy as a whole which leads to upsurge overall poverty This growing gap in the economy has created a big challenge for developing countries like Pakistan around the world.

Furthermore, Pakistan's economy is crumbling because of overall system shut down. For example, reduction in capital formation, production, revenue, trade imbalance and the health

care, safety due to COVID-19 related extra medical expenditure has further burdened limited resources of the crippling economy of the country. Pakistan can face the estimated economic loss of Rs1.3tr initially due to coronavirus according to the news report. However, in this current situation, Pakistan may take some advantage of the lower oil prices dropped worldwide.

However, she would face the adverse impact of lower exports and necessary imports /tariffs. Lower imports would lead to scarcity of goods. Because unlike China and America or any other developed country, Pakistan is unable to manufacture all products locally. America and other developed countries are trying to produce goods locally but no poor country is at this level of economic strength.

Consistently, the news Pakistan reports that “we cannot afford to practice such lockouts given that we can face a massive revenue shortfall as our resources are limited and the economy is collapsing. The FBR estimates show that the lockdown of Karachi is going to cause major revenue losses which, if persisted till June 2020, then the tax losses would go up to Rs380 billion”. Equally, according to the present federal secretary commerce “the exports might face loss in the range of \$2 to \$4 billion as export orders had got cancelled. Pakistan’s textile export sector relies on China for the bulk of its capital goods inputs, so there will be an impact if there is a protracted closedown of the Chinese economy. The imports would be reduced in the shape of declining POL prices as well as in quantity. Pakistan imported 80 billion barrels of POL products and keeping in view the low-ever prices in the international market in the last two decades, the import bill would shrink harming the FBR’s collection and petroleum levy might also be reduced if the consumption decreased because of the possible lockdown in different parts of the country”<sup>2</sup>.

Therefore, a complete lockdown situation may not be in favor of Pakistan. Pakistan being a poor country does not afford complete shutdown. The worrisome situation is 70% of the population depends on agriculture or earning by working on the sites. Moreover, the Pakistan economy is too weak to bear the burden arising in the result of stagnating and no production. Pakistan can lose from multi spheres what contributes to the GDP. For example, Pakistan’s GDP growth in 2017-18 was recorded as 5.5%, in 2018-19 was 1.9%, in 2019 to the first quarter of 2020 was -2.6% and in 2020 till the first quarter of 2021 is reported as -0.2%. We can notice a visible decline in GDP by these statistics. According to the estimation of the Planning Commission, the level of Pakistan GDP is Rs. 44trillion and 1/4<sup>th</sup> stood at the amount of Rs. 11 trillion. Thus the loss by the outbreak of COVID-19 is predicted to occur at least 10% means Rs1.1 trillion in the last Q of April-June. Some common macroeconomic indicators are given below with the comparison of the two years 2019 and 2020.

Furthermore, in Pakistan, the decreasing interest rate can play a role in lifting exports and other businesses. Still, it is criticized by economists that interest rates are not decreased at a satisfactory level. The interest rate has been decreased by 11%, according to the situation through the monetary policy of SBP. This is also a fact that the inflation rate in Pakistan has fluctuated widely over the years since 1984 and onwards (see appendix A). Nonetheless, this time as compared to 2019 which is 7.74% reached at 13.25%

in 2020 at that level of worry indicating the surging of prices more in the future along with a loss of labor especially daily wages and poor people working at lower levels. The increasing prices of routine products like sugar and wheat have created many problems for the lower and middle class to cope with the situation. Realizing the sensitivity of the time, producers in Pakistan must stop hoarding edible products in such critical time to let feed poor cheaply. Otherwise, all this leads to the short supply and rising inflation even more thus people may face starving at a greater level. To understand this reality a figure provided by SBP highlights the inflation rates of Pakistan in the global context (See the appendix B).

Table 2 Economic indicators of Pakistan

Economic indicators of Pakistan	2018-19	2019-20
GDP growth	1.9%	-2.6%
GDP per capita	\$1,482 (nominal)	\$1,244 (nominal)
Inflation rate	6.74%	11.12%
Interest rate	12.25% as of June-19	13.25% - 7.00%
Exchange rate USD	PKR 160.00	PKR 167.40
Unemployment rate	4.45%	6.2%
Poverty rate	29.5%	40.0% forecasted

How could poor and emerging economies be in the current scenario of the COVID-19 viral shock? Which has put even developed countries in greater trouble. Hence, the economy of Pakistan is not self-sufficient to deal with it without intensifying losses. If we compare Pakistan with the pandemic host country China and other regional countries, we can analyze much difference. China is the biggest economy and thus could do what they needed. Despite that, they face an economic loss substantially. Currently, China has started reviving her economy. Chinese factories are working and manufacturing the medical equipment for local and international market supply. Through producing and providing medical equipment and supplying widely to various countries in the world China is creating and reviving jobs too. Contrariwise Pakistan’s economy is more delicate to bear a huge loss. Pakistan is not self-sufficient economically therefore badly affected in terms of GDP, GNP, production, employment and per capita income. Consequently, the common man is living a miserable life and poverty increased badly. Besides, in Pakistan, most of the illiterate and unaware population does not understand the situation properly, they are afraid of everything even from quarantine and treatment. Though many of the infected people reacted wisely and recovered yet others lost their lives. The role of doctors and army/police is important many of front fighters in this regards lost their lives while providing treatment and dealing with infected crowds.

Similarly, COVID-19 outbreak drastically hit all over the Asian region. The manufacturing, service and trade activities are badly affected in the Asian region, including Pakistan. The disturbing trade led to exchange rate depreciation. Inflation along with demand and supply shocks in the whole region is alarming.

<sup>2</sup>Malik A. Rehman. (2020, March 26) Pak economy under dark shadow of coronavirus vs deep economic crisis. The News International.



The epidemic has distressed financial markets, especially capital markets (equity/bond markets) in the region. The drop in the stock exchanges in the whole region is exerting an adverse effect on the regional economic indicators (See appendix c).

## 5. Conclusion

In this paper, we discussed the outbreak of pandemic COVID-19 and its effects on the global economy as well as the Pakistan economy. We evaluate that along with the strong demographic losses of the contagious virus, the economic losses are also jeopardizing the world in its full capacity. We can judge the channel of globalization as the source of spreading any good or bad worldwide. The blessings of globalization in the shape of trade, business, employment, education, tourism and socialization we have been shared since long. However, we cannot avoid the curses of globalization in the shape of contagious, diseases and other disasters too. For example, the world witnessed the historical contagious diseases time to time in the past and this time the widespread of novel coronavirus is also the result of globalization. Thus, we can conclude that the greater the extent to which the world is becoming a global village by increasing mobilization means and ways the higher the blessing or curses are shared around the globe. We analyzed the multidimensional economic shocks worldwide. For example Demographic hits, socio-psychological hit and Economic hits from the revenue and expenditure side as well as demand and supply side. We also discuss how behavioral economies are important in the result of changing behaviors of people. Almost all the countries are observing losses in their business activities and every sector of industries tumbling in their revenue. Capital markets are shrinking their activities. The stock markets/ equity markets got hard hit so declining their performance around the globe. Apart from that, we elaborate that developing countries like Pakistan are suffering more from the repercussions of a cruel pandemic that poverty is getting even more strength. The process of economic development weakens further by the deteriorating performance of all socioeconomic indicators. We conclude that the Governments, policy makers and stakeholders along with realization must take soothing measures to uplift the dropped down economy and try to support markets with high/low supply and high/low demand creating imbalances from both sides of demand and supply in markets for different commodities.

### 5.1. Recommendations

In this critical time, the role of Government is very important to fight the battle against the microscopic enemy. [28], describes that the intensity of financial losses will be a responsibility of governments to control.

Government's expenditure on health care is inevitable to control the transmission and treatment of infected people. Government and the public must work hard and go together in this fight by moving forward and think of each other collectively and do not leave anyone behind.

The government may compromise for the production of luxury goods but the core sectors, which relates to basic needs, i.e. utility services, food, health and construction should not be compromised by the total lockdown. However, in the operating factories, applying SOPs is essential to provide health care.

Moreover, by deploying the teams of doctors to check employees frequently will ensure smooth sailing during these troubled times. Significant Business organizations should be given relaxation on their fixed costs, inject liquidity, where ever needed to sustain the pressure of abrupt breakout and sudden downfall.

Isolation centers and temporary hospitals should be created in sufficient number. According to the situation, retired doctors may be called in case of need. Increase the doctors and nurses by calling final- year-medical students and house jobholders. Meanwhile, foreign degree holders may be called by waving their PMDC exams. Special training may be given to new nurses and doctors. Sufficient stock of medical kits, uniforms of doctors and patients, ventilators, drugs and other clinical apparatus should be arranged.

Support the labour force and daily wages people who lost their jobs or income and health. They may be compensated through government by free treatment and distribution of food/groceries where necessary.

Extraordinary cleaning programs should be started. More importantly, behavioural change, social attitude, values, norms, and changing lifestyle is highly required in Pakistan. More and frequent awareness programs are needed and teams may be organized for public awareness.

Spend a handsome amount for research and development to improve scientific research and preparing vaccine and medicine and/or other necessary equipment against the contagious virus. Create multiple teams for this highly demanding cause and make plans to work hard to face the crises efficiently.

The government may take measures to protect the people residing in foreign countries by providing support to them and make sure of their safety by the country's embassies.

Many foreign returned people are unknowingly scattered in different regions of the country and they are mixed up with masses thus unnoticed. Government and public try their best together to detect and testify them and their connected circle to reduce transmissions.

The government may review fiscal and monetary policies frequently, according to the situation.

### Compliance with Ethical Standards:

It is to confirm that principles of ethical and professional conduct have been carefully followed and the following statements are to ensure the mandatory requirements of the journal.

**a. Funding:** It is to mention here that there is no sponsorship received from any institution nor any institution is requested for the funding.

**b. Ethical approval:** The Editors reserve the right to reject manuscripts that do not comply with the above-mentioned guidelines. The author will be held responsible for false statements or failure to fulfil the above-mentioned guidelines.

**c. Informed consent:** It is to ensure here that the research work does not involve human participation nor research on animals is involved.

**d. Conflict of the statement:** On behalf of all authors, the corresponding author states that there is no conflict of interest involved, neither financial nor non-financial.

### Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

Authors are thankful to the editor ASTES, its team to correspond and guide properly, many thanks to the anonymous reviewers for their productive feedback.

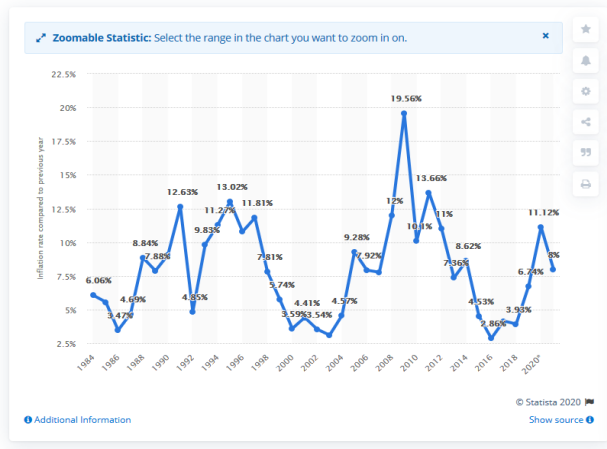
## References

- [1] W. J. Mckibbin, R. Fernando, "The global macroeconomic impacts of COVID-19: Seven scenarios", 2020. <http://dx.doi.org/10.2139/ssrn.3547729>
- [2] J. T. Cuddington, J. D. Hancock, "Assessing the Impact of AIDS on the Growth Path of the Malawian Economy" *Journal of Development Economics* **43**(2), 363-68, 1994. [https://doi.org/10.1016/0304-3878\(94\)90013-2](https://doi.org/10.1016/0304-3878(94)90013-2)
- [3] J. T. Cuddington, J. D. Hancock, "A Dynamic Aggregate Model of the AIDS Epidemic with Possible Policy Interventions" *Journal of Policy Modeling* **16**(5), 473-96, 1994. [https://doi.org/10.1016/0161-8938\(94\)90024-8](https://doi.org/10.1016/0161-8938(94)90024-8)
- [4] A. Deaton, "Health in an Age of Globalization" (No. w10669). National Bureau of Economic Research, 2004. 10.3386/w10669
- [5] O. Kodila-Tedika, S. A. Asongu, J. M. Kayembe, "Middle class in Africa: Determinants and consequences" *International Economic Journal*, **30**(4), 527-549, 2016. <https://doi.org/10.1080/10168737.2016.1204340>
- [6] T. Ensor, S. Cooper, L. Davidson, A. Fitzmaurice, W. J. Graham, "The impact of economic recession on maternal and infant mortality: lessons from history" *BMC Public Health*, **10**(1), 727, 2010. <https://doi.org/10.1186/1471-2458-10-727>
- [7] T. Ensor, S. Cooper, "Overcoming barriers to health service access: influencing the demand side" *Health policy and planning*, **19**(2), 69-79, 2004. <https://doi.org/10.1093/heapol/czh009>
- [8] D. A. Robalino, A. Voetberg, "The Macroeconomic Impacts of AIDS in Kenya Estimating Optimal Reduction Targets for the HIV/AIDS Incidence Rate" *Journal of Policy Modeling* **24**(2), 195-218, 2002. [https://doi.org/10.1016/S0161-8938\(02\)00097-2](https://doi.org/10.1016/S0161-8938(02)00097-2)
- [9] N. Ruggiano, T. E. Perry, "Conducting secondary analysis of qualitative data: Should we, can we, and how?" *Qualitative Social Work*, **18**(1), 81-97, 2019. <https://doi.org/10.1177%2F1473325017700701>
- [10] D. A. Robalino, A. Voetberg, "The Macroeconomic Impacts of AIDS in Kenya Estimating Optimal Reduction Targets for the HIV/AIDS Incidence Rate". *Journal of Policy Modeling* **24**(2), 195-218, 2002. [https://doi.org/10.1016/S0161-8938\(02\)00097-2](https://doi.org/10.1016/S0161-8938(02)00097-2)
- [11] D. tuckler, L. P. King, S. Basu, "International Monetary Fund programs and tuberculosis outcomes in post-communist countries" *PLoS Med*, **5**(7), e143, 2008. <https://doi.org/10.1371/journal.pmed.0050143>
- [12] J. Chou, N. F. Kuo, S. L. Peng, "Potential impacts of the SARS outbreak on Taiwan's economy". *Asian Economic Papers*, **3**(1), 84-99, 2004. <https://doi.org/10.1162/1535351041747969>
- [13] W. Hai, Z. Zhao, "The Short-Term Impact of SARS on the Chinese Economy" *Asian Economic Papers* **3**(1), 57-61, 2004. <https://doi.org/10.1162/1535351041747905>
- [14] A. Sui, Y. C. R., Wong, "Economic Impact of SARS: The Case of Hong-Kong", *Asian Economic Papers* **3**(1), 62-83, 2004. <https://doi.org/10.1162/1535351041747996>
- [15] C. Binder, "Coronavirus fears and macroeconomic expectations", *Review of Economics and Statistics*, 1-27, 2020. [https://doi.org/10.1162/rest\\_a\\_00931](https://doi.org/10.1162/rest_a_00931)
- [16] J. L. Kulzer, J. A. Penner, R. Marima, P. Oyaro, A. O. Oyanga, S. B. Shade, E. A. Bukusi, "Family model of HIV care and treatment: a retrospective study in Kenya" *Journal of the International AIDS Society*, **15**(1), 8-8. 2012. <https://doi.org/10.1186/1758-2652-15-8>
- [17] R. Keith, Z. Lin, "SARS in Chinese politics and law" *China Information*, **21**(3), 403-424, 2007. <https://doi.org/10.1177%2F0920203X07083321>
- [18] J. Barry, Supply chain risk in an uncertain global supply chain environment *international journal of physical distribution & logistics management* **34**(9), 695-697, 2004. DOI 10.1108/09600030410567469
- [19] W. J. Mckibbin, R. Fernando, "The global macroeconomic impacts of COVID-19: Seven scenarios", 2020. <http://dx.doi.org/10.2139/ssrn.3547729>
- [20] N. Ruggiano, T. E. Perry, "Conducting secondary analysis of qualitative data: Should we, can we, and how?" *Qualitative Social Work*, **18**(1), 81-97, 2019. <https://doi.org/10.1177%2F1473325017700701>
- [21] W. J. Mckibbin R. Fernando, The global macroeconomic impacts of COVID-19: Seven scenarios, 2020. <http://dx.doi.org/10.2139/ssrn.3547729>
- [22] L. Boone, J. Dewulf, G. Ruyschaert, T. D'Hose, H. Muylle, I. Roldán-Ruiz, "Assessing the consequences of policy measures on long-term agricultural productivity—Quantification for Flanders" *Journal of Cleaner Production*, **246**, 119000, 2020. <https://doi.org/10.1016/j.jclepro.2019.119000>
- [23] X. Zhang, J. Xie, R. Rao, Y. Liang, "Policy incentives for the adoption of electric vehicles across countries", *Sustainability*, **6**(11), 8056-8078, 2014. <https://doi.org/10.3390/su6118056>
- [24] S. S. SyraMadadMadad, J. Masci, S. N. V. Cagliuso, M. Allen, "Preparedness for zika virus disease—New York city, 2016" *Morbidity and Mortality Weekly Report*, **65**(42), 1161-1165, 2016. doi:10.2307/24859111
- [25] K. C. Hyams, F. M. Murphy, "Responding to Chemical, Biological, or Nuclear Terrorism: The Indirect and Long-Term Health Effects May Present the Greatest Challenge" *Journal of Health Politics, Policy and Law* **27**(2), 273-91, 2002. <https://doi.org/10.1215/03616878-27-2-273>
- [26] W. J. Mckibbin, R. Fernando, "The global macroeconomic impacts of COVID-19: Seven scenarios", 2020. <http://dx.doi.org/10.2139/ssrn.3547729>
- [27] W. J. McKibbin, R. Fernando, "The global macroeconomic impacts of COVID-19: Seven scenarios", 2020. <http://dx.doi.org/10.2139/ssrn.3547729>
- [28] R. Hussam, A., Rabbani, G. Reggiani, N. Rigol, "Habit formation and rational addiction: A field experiment in handwashing", *Harvard Business School BGIE Unit Working Paper*, 18-030, 2017. <http://dx.doi.org/10.2139/ssrn.3040729>

Appendices

Appendix A Pakistan's inflation 1984 -2021

Pakistan: Inflation rate from 1984 to 2021 (compared to the previous year)

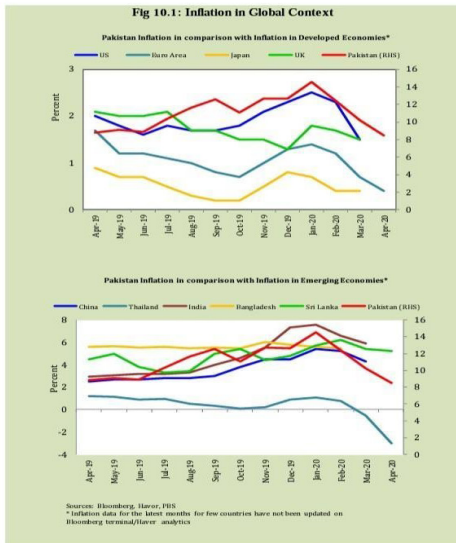


Appendix B: Pakistan's Inflation to the global context

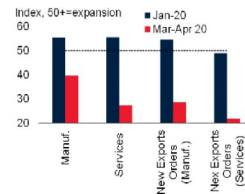
State Bank of Pakistan

10. Inflation in Global Context

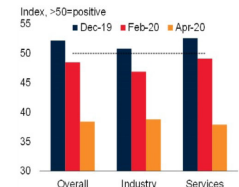
In global context, Pakistan witnessed highest inflation not only in comparison with developed economies but also with emerging economies (See figure 10.1).



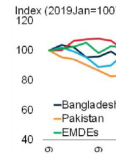
Appendix C:



A. India: Purchasing managers' indexes



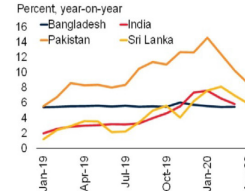
B. Pakistan: Business confidence survey



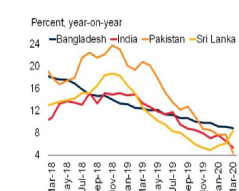
C. Equity indices SAR



D. Exchange rates



E. Inflation



F. Private sector credit growth

## Social-Interactive Learning Concept Used for Electronic Educational Resource “Post-Graduate Foreign Language” and the Obtained Learning Curve

Natalya Chernova<sup>\*1</sup>, Victor Chernov<sup>1</sup>, Margarita Emelianova<sup>2</sup>, Raisa Akhunzianova<sup>3</sup>, Danil Sukhopluev<sup>1</sup>

<sup>1</sup>Kazan (Volga Region) Federal University, Naberezhnye Chelny Institute, Naberezhnye Chelny, 423812, Russian Federation

<sup>2</sup>I.N. Ulianov Chuvash State University, Faculty of Foreign Languages, Cheboksary, 428015, Russian Federation

<sup>3</sup>Naberezhnye Chelny State Pedagogical University, Chair of Foreign Languages, Naberezhnye Chelny, 423806, Russian Federation

### ARTICLE INFO

Article history:

Received: 27 June, 2020

Accepted: 04 August, 2020

Online: 25 August, 2020

Keywords:

Educational process

Second language acquisition

E-learning

Participation

Zone of proximal development

### ABSTRACT

The article touches upon an urgent question of social-interactive learning. The resource under study deals with one of the possible ways to solve this problem, in the context of global socio-cultural transformations and a new paradigm for the development of society. It is the formation of a subject-subject dialogue between a teacher and a student. The second language research study showed how pre- and intermediate speakers' performance breaks down in the face of a difficult narrative task and self-regulation and control over the mediational means are lost. More advanced speakers are able to guide themselves through the task. The main idea of such education is that good learning leads to development. It seems to us important to refer the concept of the zone of proximal development to the development of the individual, which affects the boundaries of the zone of proximal development. Electronic educational resource with the help of mediating means or sign operations makes external social interactions become "internalized", namely, internally reconstructed psychological processes - ways of thinking and learning. A student identifies active personality development prospect through the actual experience. The students' learning curve let us prove the idea that IT technologies intensify the process of studying, but should provide not only language accomplishments but active communication with a teacher as well.

## 1. Introduction

This paper is an extension of work originally presented in 2017 10th International Conference on Developments in E-Systems Engineering (DeSE). The desire to find the right balance between the usage of electronic educational resources, direct participation on the part of a student and a learning curve determined the problem of our research. In theoretical terms, this is the problem of substantiating social-interactive learning concept in the electronic-digital form. In practical terms, it is the problem of determining the relationship between the success of training and the formation of a student's personal qualities as a subject of educational activity [1]. We would like to develop the idea that constant transitions of external actions into internal ones and vice versa take place, not only constantly, but also contribute to

penetration into the zones of proximal development, due to the expansion of the amplitude of oscillations, i.e. development and self-development of the subject itself (Popov L.M.) [2].

In Vygotsky's understanding, this means that the personality is realized and manifested in all spheres of mental development in ontogenesis [3]. The actualization of personal characteristics, in particular, social orientation and personal abilities and qualities contributes to the subjective position formation (Leontiev A.N.) [4]. The development of the semiotic function is concretized in the ability to highlight in the text (in the broad sense of the word) the signified and the signifier; in the signified - the object and meaning, in the signifier - the form and content (meaning), which, in its turn, is the development of new levels of means mastery. Thus, it is necessary not only for the formation of scientific knowledge, but also for any spheres [5]. The “Post-graduate Foreign Language” electronic educational resource was developed according to the idea implying that a student should act, first of all,

\*Corresponding Author: Natalya Chernova, Naberezhnye Chelny 423812, Russian Federation, +79033190461, Email: natalchem@gmail.com



as a researcher of thinking, activity. Having mastered which, a student can successfully bring his/her random impressions about different phenomena of the surrounding world into a single system of verified judgments, grounded in understanding the essence of a particular sphere of reality (Luria A.R.) [6]. The main goal of a person is self-development, first of all, of personal qualities, without which the fulfillment of functional goals by a person is either impossible or of poor quality. This logic supposes the subjective activity of the student, who him/herself creates the teaching and makes him/herself, while the line between the processes of training and education is erased.

As it turns out, only instructional conversations relate human mental functioning to positioning themselves as individuals in carrying out a task in a foreign language (R. Donato). Computer assisted learning helps to move a focus from form to meaning. So, societal context participation (Zinchenko V.V.) [7] is implemented as social-interactive learning concept. The authors' aim is to provide evidence of the relationship between the success of training and the usage of "Post-graduate Foreign Language" electronic educational resource, which provides additional opportunity for a student to be a subject of educational activity when learning a foreign language. "Participation" metaphor is the main tool used in the process of the program creation according to Social-interactive Learning Concept.

The "Post-graduate Foreign Language" electronic educational resource is an automated learning system, providing a socio-cultural approach (Vygotsky L.S.) to language learning [3]. In the first place, it is the opportunity of extemporaneous communication (i.e. Figure 4). Further, a student does not only gain some new knowledge in the subject but receives skills of various kinds. As a result, a student has an opportunity of realizing in practice a fundamental way of learning [8].

The study shows that "activity theory" (Leontiev A.N.) finds its application in e-teaching as well as classroom language learning [4]. The paradigm of "zone of proximal development" is followed through: 1) communicative language studying as a part of cognition; 2) semiotic mediation; 3) private speech. In analyzing the recorded interaction between the students and the teacher, it shows that instructional conversations are developmentally sensitive to the students' growing ability to use the language required to carry out a specific task. The evidence of this is the shift from other-regulated to self-regulated error correction.

This computer training program can be presented as a systematic presentation of the above-mentioned approach to the process of studying, which compiles the fundamental provisions of the subject-activity approach and the socio-cultural theory with modern technologies in e-learning, such as Moodle software platform [1]. The choice of the studied subject is due to the necessity of fluent English for any scientist or a post-graduate.

The "Post-graduate Foreign Language" resource shows that Social-interactive Learning Concept can be reflected through metacognitive strategies such as predicting, planning, and monitoring. So, mediating learning with the help of IT technologies shows that the Learning Curve levels up in comparison with traditional learning strategies providing instructions alone. The succession and number of exercises help

students and teachers as well help in the process of communicative and cognitive activities in an ESL class [8].

## 2. Background to the Problem

The program under study provides for independent work on the part of a student and work under the guidance of a teacher as well [2]. It helps in getting the necessary knowledge, gives an opportunity of providing effective dealing with the educational problem and allows active learning in natural environment. This computer-training program is a means of studying "Academic English" within the Post-graduate curriculum and meets all educational and methodological demands [7].

Referring to the private speech in his study, J. Smith (1996) operationalizes it as a 'verbal attempt of self-regulation during problem-solving tasks', which is, the result of stress that accompanies construction of the meaning in L1 or L2, the private speech is distinguishable from the interpersonal communication. Smith examines a grammar class of high intermediate ESL students and claims that private speech can arise in a discussion [9]. The thing is that classroom discourse is usually analyzed for its social, communicative value, neglecting the cognitive function of instructional talk.

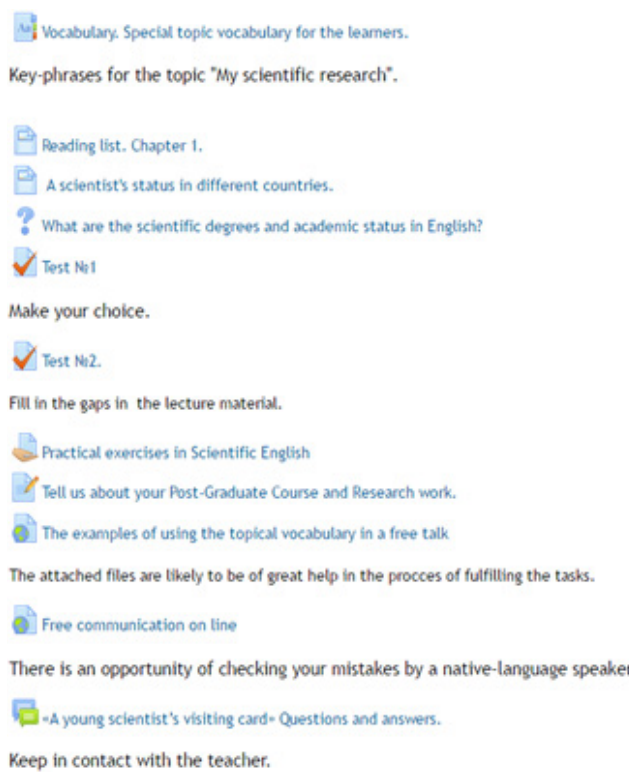


Figure 1: "My scientific research" item program

Classroom conversations depend on spontaneity, unpredictability and focus on new information. If the discussion is shaped toward a curricular goal, and teachers in their turns build or activate background knowledge in students, these conversations can also be instructional. They are relevant to language classrooms because they provide opportunities for experiencing language outside. The concept of instructional conversations was explored within the frames of a PhD study, compiled by N.A. Chernova. The motivation for the study was the following observation. It turned out that instructional conversations usually arose with fairly

advanced ESL learners. The question remained, was it possible to have an instructional conversation in a beginning foreign language class? Limited linguistic resources are not available for a topically coherent, extended, and meaningful conversation. The following data were received. If a teacher followed a predictable pattern a) setting up the vocabulary practice exercise; b) completing textbook exercises using the new vocabulary; and c) spontaneous questions and comments by students and teacher embedded within the practice exercise, conversational communication was out of the question (Chernova & Mustafina, 2016) [7]. The possibility of such conversation depended on management talk and extension activities, which show features of it most consistently and impressively [10].

Here is the screen shot to illustrate Chapter 1 with the topic “My scientific research”. The tasks “Tell us about your Post-Graduate Course and Research work”, “Free communication online”, and “«A young scientist’s visiting card» Questions and answers” provide for Instructional conversations. The last one let the students keep in contact with the teacher.

The tests are checked automatically, but two-way communication is possible when fulfilling training tests. So, you keep in contact with the teacher in English. The fact that every stage is under control is the most prominent achievement with this resource.

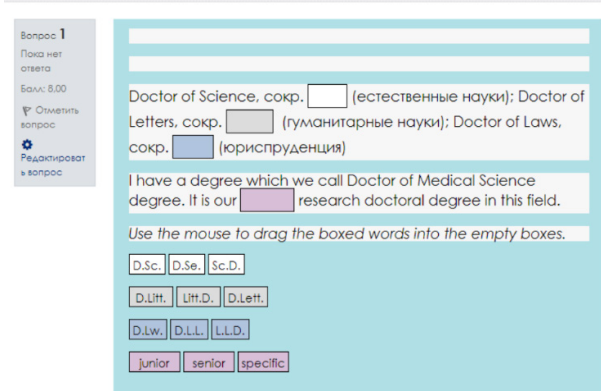


Figure 2: Chapter 1. Test №2. Fill in the gaps in the lecture material



Figure 3: The available marks for the fulfilled test

There is some material in Russian in the test due to the need to check understanding. The items in the upper left-hand corner are given in Russian because of the application interface.

The “Post-graduate Foreign Language” resource makes it possible to estimate the test results at once. Thus, the process of training is intensified.

### 3. Methods of Research

The methodological basis for creating a program is the theory of D. Wood, J.S. Bruner, J. Ross and R. Donato about spontaneous communication in the classroom [11]. At the initial stage, students simply reproduce monosyllabic statements, repeating them after the teacher, and the invariable model of teaching in the classroom is "a teacher – students – a teacher", or otherwise "subject-object-subject". Over time, students have spontaneous responses to the teacher's statements, as well as mutual assistance and joint construction of statements, the teacher's help decreases, and students play an increasingly active role in the lesson, using the same techniques as the teacher [12].

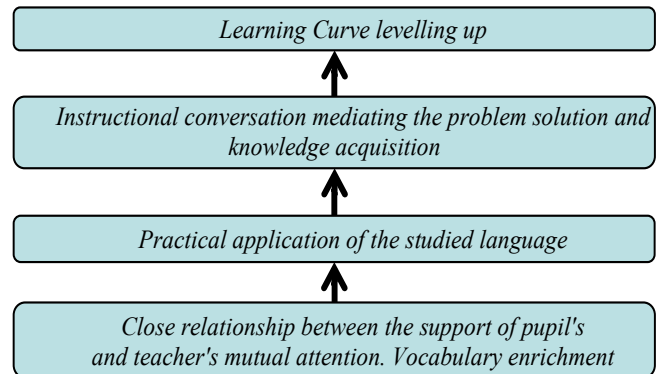


Figure 4: Educational technique used with “Post-graduate Foreign Language” electronic educational resource

The Instructional conversation is a dialogue in which the speakers are involved in solving the problem in the process of acquiring knowledge [4]. We would like to emphasize that it mediates both the solution of the problem and the acquisition of knowledge. Growth in the quality of knowledge is the main result achieved. This is the tracked indicator presented in the study.

In the framework of the experimental program, the study was carried out. The experimental group, within the framework of this strategy, discussed their actions in a dialogue and corrected each other's mistakes, explaining them. The results were compared with those of the control group the students of which studied the subject “Academic English” within the Post-graduate curriculum but didn't use the “Post-graduate Foreign Language” electronic educational resource. The topics and the volume of material studied was the same. The class periods in both groups are fully alike. The quantitative values of the quality of knowledge obtained after the experimental study are presented in Figures 5, 6.

The study took place in Kazan Federal University. The participants were the first-year post-graduate students. The level of language proficiency was rather high in the second term by the time of the study. By this time a student can already use the specific vocabulary on their specialty in spontaneous dialogues. The study was aimed at testing the influence of “Post-graduate Foreign Language” electronic educational resource on the correct use of special language forms in oral speech.

#### 4. Results of the Study and Discussion

The obtained changes in the quantitative characteristics of the quality of knowledge are statistically significant. The system of evaluation begins with 1 (equals to “Bad” mark) and finishes with 5 (equals to “Excellent” mark).

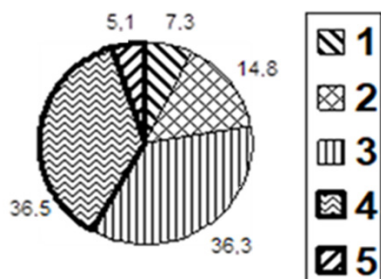


Figure 5: Quality of knowledge. Control Groups. Control Survey

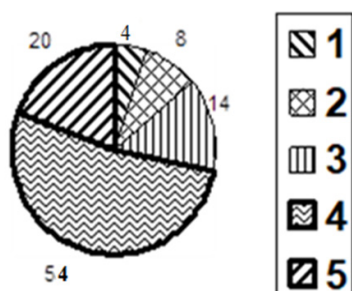


Figure 6: Quality of knowledge. Experimental Groups. Control Survey

Taking into account that levels 4 and 5 are sufficient, it can be concluded that 74% of students in the Experimental Groups achieved this indicator, and 41.6% in the Control groups. Thus, the quality of knowledge in the Experimental Groups is almost 1.8 times higher than in the Control Groups.

This distribution allows us to speak about an increase in knowledge among students of Experimental Groups, which is expressed in the ability to notice and correct errors with minimal or no hint, take full responsibility for correcting errors, and be consistent in the use of keywords in all contexts.

The study shifts the governing concept of learning away from the acquisition metaphor toward the participation one with an individual’s activity. In other words, students are taking part in a collaborative dialogue. Communicative collaboration is based upon the active and purposeful agent on the part of a student. The authors offer the following model of ESL classes: 1) self-directed activity comprising self-dependence, activity, social direction, self-government, reflection → 2) reflective, problem-solving orientation with the help of the “Post-graduate Foreign Language” electronic educational resource → 3) internalization of social interactive process in the zone of proximal development with the English language as a cognitive tool for the individual under the guidance of a teacher → 4) the second language mediated process of an agent formation in learning and professional activities (direct solution of professional tasks).

This model allows to be sensitive to students’ needs and abilities and support the overpowering and transformative agency embodied in the learner. When working with e-learning, it is important to arrange the teaching process based on the strengths of

the students developing their confidence and ability to work and study independently and in a team. The sociocultural context of the teaching process contributes to the development of the students’ identity within the society [10]. The diversity of sociocultural contexts used during the classes has to teach them to think critically and creatively, solve problems and recognize the advantages. These approaches applied during the teaching process provide students not only with the knowledge and skills, but also with emotional comfort and sufficient self-esteem.

By the end of the studying course, “free talk” technique is being practiced. The students in this study were tested individually. Firstly, they were asked a series of individual questions in the form of interviews, and secondly, they were asked three discussion questions, responding to which students had to express their opinion, tell the story and come up with a statement. The questions were formulated in such a way that, when Answering the questions provided for compiling complex sentences with specific vocabulary. The quality evaluation was carried out in three stages: ascertaining, control and post experimental ones. The last stage was held in four weeks. The ascertaining stage showed that the level of knowledge was alike at the end of the first term. The control stage determined significant improvements in the use of specific language forms and complex sentences at the end of the second term in the Experimental groups as a result of the “Post-graduate Foreign Language” electronic educational resource usage (Table 1). The students from Experimental groups showed a good progress in all four control tasks in comparison with the ascertained stage. The progress in Control groups was not so considerable. Moreover, the level of responses with the Experimental groups could be considered as a prolonged one after taking a post-experimental test within the period of four weeks.

Table 1: Levels of Specific Vocabulary and Complex Sentences Usage with “Free Talk” Technique

Groups		Specific language vocabulary and Complex sentences Usage			
		1	2	3	4
		Inter view	Opinion expression	Story telling	Discourse statement
Control Groups	Ascertaining stage	12.3%	17.7%	25.4%	18.2%
	Control stage	38.8%	38.3%	50.4%	38.9%
<b>Measurement odds</b>		<b>+26.5%</b>	<b>+20.6%</b>	<b>+25.0%</b>	<b>+20.7%</b>
Post-experimental stage		26.1%	27.4%	37.6%	27.5%
Experimental Groups	Ascertaining stage	13.4%	18.1%	24.6%	19.3%
	Control stage	64.2%	57.3%	73.8%	58.6%
<b>Measurement odds</b>		<b>+50.8%</b>	<b>+39.2%</b>	<b>+49.2%</b>	<b>+39.3%</b>
Post-experimental stage		60.1%	54.3%	70.6%	54.9%

Observations clearly show that students’ learning curve proves the idea that IT technologies intensify the process of studying, but should provide not only language accomplishments but active communication with a teacher as well. Firstly, the help is provided only by the teacher with the help of Instructional conversations,

then gradually it is internalized by the students, and at last it is used as an instrument of mediation, which provides good progress in knowledge [8].

The concept of participation is used in this methodology [11]. The “Post-graduate Foreign Language” electronic educational resource was made by using these methods of teaching. The data received prove the idea, that fulfilling different tests and practical exercises only doesn’t let students intensify the process of studying. They need active communication with a teacher as well.

## 5. Conclusion

To access the course, you need to follow the link provided <http://edu.kpfu.ru/enrol/index.php?id=1494> on the site of Kazan Federal University. Teaching the discipline solves the following tasks: improvement and further development of speech and language skills and abilities in all types of a foreign language speech activity (reading, speaking, listening, writing). That includes the highly specialized field of foreign languages as well. Development of skills and abilities you need for independent work help post-graduate students (applicants) to improve the level of proficiency in a foreign language in order to use it for scientific and professional activities. They are able to work with the world information resources in a foreign language in the profile of the specialty in order to prepare written (abstracts, annotations, abstracts, articles, motivational presentation) and oral (reports) scientific texts. Technology-Enhanced Learning should include such IT technologies which promote triadic and dynamic interaction rather than dyadic interaction at the same time. Indication language process is made complicated in a dyadic format. Electronic educational resource has a good future ahead of him. And, it is necessary on the part of teachers to benefit from this perspective of educational process upgrading.

## Conflict of Interest

The authors declare no conflict of interest.

## References

- [1] A.I. Bashmachnikov, I.A. Bashmachnikov, IT textbooks and educational systems development, Moscow: Philin, 2003.
- [2] N. F. Talyzina, *Upravlenie protsessom usvoeniya znaniy* [Managing the acquisition of knowledge], Moscow State University Press, 1975.
- [3] L. S. Vygotsky, *Thought and Language*, Cambridge MA: MIT Press, 1986.
- [4] J.P. Lantolf, *Sociocultural Theory and Second Language Learning*, Oxford University Press, 2000.
- [5] J. Lave, E. Wenger, *Situated Learning: Legitimate Peripheral Participation*, Cambridge University Press, 1991.
- [6] A.R. Luria, *Language and Cognition*, New York: Wile, 1981.
- [7] N. A. Chernova, J.N. Mustafina, “Instructional conversation as a socially-constructed cognitive tool serves second language learning by mediating its own construction, and the construction of knowledge about itself” *TURK ONLINE J DES ART COMMUN*, **6**, 3114-3117, 2016. <https://doi.org/10.7456/1060NVSE/130>
- [8] W. Horton, K. Horton, *Electronic education: instruments and technologies*, Moscow: Kudits Publishing House, 2005.
- [9] Smith, J., “A seven-minute slice of chaos or I am puzzling through now,” Unpublished research report, University of Pittsburgh, 1996.
- [10] J. V. Wertsch, *Sociocultural Studies of Mind*, Cambridge University Press, 1995.
- [11] D. Wood, J.S. Bruner, G. Gross, “The role of tutoring in problem-solving” *J CHILD PSYCHOL PSYC*, **17(2)**, 89-100, 1976. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- [12] P. Zukow-Goldring, K. R. Ferko, “An ecological approach to the emergence of the lexicon: Socializing attention”. In V. John-Steiner, C.3. Panofsky, & L. W. Smith (Eds.), *Sociocultural approaches to language and literacy: An*

interactionist perspective, Cambridge University Press, 1994. <https://doi.org/10.1017/CBO9780511897047.008>



## Correlation-Based Incremental Learning Network for Gas Sensors Drift Compensation Classification

Panida Lorwongtrakool<sup>\*1</sup>, Phayung Meesad<sup>2</sup>

<sup>1</sup>Department of Information Technology, Faculty of Information Technology and Digital Innovation, King Mongkut's University of Technology, North Bangkok, 10800, Thailand

<sup>2</sup>Department of Information Technology Management, Faculty of Information Technology and Digital Innovation, King Mongkut's University of Technology, North Bangkok, 10800, Thailand

### ARTICLE INFO

#### Article history:

Received: 19 June, 2020

Accepted: 04 August, 2020

Online: 25 August, 2020

#### Keywords:

Classification

Correlation Distance

Gas Sensor

Incremental Learning

Radial Basis Function

### ABSTRACT

A gas sensor array is used for gas analysis to aid in an inspection. The signals from the sensor array are fed into machine learning models for learning and classification. These signals are characterized by time series fluctuating according to the environment or drift. When an unseen pattern is entered, the classification may be incorrect, resulting in decreased model performance. Creating a new model results in the problem of forgetting the old knowledge called Catastrophic Forgetting. Accordingly, this research proposes Correlation-Based Incremental Learning Network (CILN) using the Correlation Distance method to measure similarity and the Gaussian membership function to determine membership of each node. The gas sensor array data is used to verify the proposed algorithm by choosing 16 steady-state features (DR) from 13,910 records which are divided into 6 classes: 1) Ethanol, 2) Ethylene, 3) Ammonia, 4) Acetaldehyde, 5) Acetone, and 6) Toluene. The data are normalized and divided as the training sets into 10%, 20%, 30%, 40%, and 50%, respectively. The proposed algorithm was compared with well-known classifiers. CILN experiment results yield the highest accuracy of 98.96% using 50% of the training data set. It shows that CILN has the incremental learning ability and can be used with data that fluctuate according to the situation.

## 1. Introduction

The odor inspection, such as food quality check, environmental perfume check, chemical leak check or even weapons or drugs inspection require experts. However, there are limitations of smelling by human nose, and the effectiveness of smelling depends on individual health. Humans may feel tired, and importantly, the human nose is not suitable for smelling various toxins [1].

Many researchers have tried to devise a variety of inspection methods, such as chemical properties or flavor tests, and these methods require direct contact. Another interesting method is to test odors or gases using an electronic nose or array sensor since this method does not require immersion into the sample but inspect the gas response. Currently, the array sensor is extensively used for inspections, such as potential contagious or chemical contamination inspection during the production process, checking the freshness of pork [2], fungus inspection in strawberries [3], evaluating the quality of black tea [4], identify the type of wine and

chinese liquor [5], distinguishing wine making techniques, [6] and medical diagnoses [7-8].

Machine learning is a useful tool for data analysis and data learning. Machine learning is divided into 2 approaches: Batch Learning and Incremental Learning. 1) Batch learning refers to a learning method which learn data at once. 2) Incremental learning (also called online learning) refers to learning methods which is applied to streaming collected over time. In this method, learning functions can be updated when new data is entered into the system [9].

In the real world, input data can be dynamic or streaming depending on the situation or environment, as data change or drift. Therefore, if the model is unable to learn new patterns, the performance of the model will be reduce [4]. Finally, all earlier data sets will no longer be available, and a new model must be created when there are new data. This process also leads to a phenomenon known as catastrophic forgetting [10].

\*Corresponding Author: Panida Lorwongtrakool, panida.l@email.kmutnb.ac.th

Therefore, to improve the algorithm for learning latest information while keeping the old knowledge of earlier data sets, this research proposes Correlation-Based Incremental Learning (CILN). The proposed algorithm can learn new patterns and adapt itself automatically while keeping existing knowledge.

## 2. Literature Reviews

### 2.1. Gas Sensor array

The gas sensor array is important to the electronic nose and is used to detect gas molecules from electrical signals. Each sensor has a different sensitivity to gas. When the gas sensor contact with gas molecules, it responds to the gas, forming a spectrum of different gases called odor fingerprint. The response is recorded and sent to the signal processing system for analysis using proper analysis methods to determine the type of gas [11].

### 2.2. Incremental Learning

In general, creating a model for classification requires a training set for patterns recognition. When the model has an acceptable performance, it will be used for prediction. However, the general limitation of the model is that it will work well for a period only and, inevitably, a new model needs to be created because the algorithm is not created for incremental learning. Therefore, it is extremely sensitive to continuous data in the form of streaming. A new model needs to be created to solve this problem, and it is unable to use the existing knowledge.

In addition, data in the new situation are still a problem, and there is a need to incremental learning capability [12] by gradually learning knowledge without abandoning or forgetting the existing one without retraining. To solve this problem, incremental learning algorithms must be able to combine new knowledge with previously bought knowledge in a way like human learning methods which are based on earlier learning [13].

Therefore, algorithms that can learn from new data without having to access the earlier set of data and support prior knowledge would be a good method of classification to support both static data and data stream, especially when new data samples are added continuously.

#### 1) Characteristics of Incremental learning

- Incremental learning algorithms handle with continuous data and non-stationary distributions.
- It adapts to new data without forgetting the existing knowledge; it does not need to retrain.
- It is compatible with data streams or big data to create machine learning models faster.
- It uses instance windows or instance weighting mechanism without making modifications to the algorithms. New models are calculated based on time periods of using windows or weights by considering new data received.
- All or some parts of the data are used to create an initial model, check for changes in data (using the detection function), and rebuild models as needed based on new data.
- It can automatically change the learning mechanism to increasingly learn new data. For example, the weights of an

artificial neural network are adjusted every time there is a new pattern coming into the system.

#### 2) Types of Incremental Learning

- Instance-incremental learning refers to the system that receives data at Step  $t$  - the input point  $\mathbf{x} \in X^n$ , where  $X$  represents the input domain in  $n$  dimensional space and predicts the output  $y \in Y$ . The output,  $Y$ , can be either continuous in regression or in classification in one dimensional space.
- Batch incremental learning receive batches of data  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K)$  and must specify the label  $(y_1, y_2, \dots, y_K)$  for each input point, where  $K$  is the number of data points.

Traditional Artificial Neural Networks (ANNs) work well in jobs with static data, where there is no incremental learning needed. Adding new capabilities to ANNs often results in catastrophic forgetting [10]. Therefore, researchers have attempted to solve this problem. In [14] developed a hybrid system including supervised learning and unsupervised learning online by using Fuzzy and Neural Networks together with Euclidean distance.

In addition, learning algorithms have been developed based on Support Vector Machine (SVM) together with Mahalanobis distance which is an elliptical kernel method for multidimensional data. The ability to classify data was compared with the traditional method using Euclidean distance, and no difference was found [15]. However, scattering of data should be considered when adopting the similarity measurement method.

Using threshold is another way to help learning, adjust the model, and support the integrity of earlier knowledge, while only adding a small number of parameters [16]. The important for a creating an incremental learning model is that it must be done quickly by using a small amount of data and gradually adjusting the model according to the new data while keeping old knowledge without access to the initial training data set [17].

However, it was found that the attempt to improve the recognition system for greater accuracy often result in more complex problems. Hence, Incremental Similarity (IS) has been presented, as it yields high accuracy and low complexity. Incremental Similarity is used for incremental online learning. The system can learn from sample data received, and only some parameters need to be updated. It was found that the efficiency of Incremental Similarity was higher than that of the traditional model [18].

At present, deep learning has received great attention, and it has been used in recognition, such as face recognition. However, the face recognition model without incremental learning after training results in problems with new data during the operation. In [19] introduced the Incremental SVM method that allowed the system to update the classification model in real-time, resulting in the increased accuracy of the system and reduced training time.

However, deep learning processing consists of many connected parameters, and it takes quite a lot of time in the training process, as well as retraining if the structure is not enough. Adding incremental learning capabilities to the system with deep learning can be quite difficult. Therefore, researchers have attempted to invent and adjust the Broad Learning System (BLS) [20] which

was created in the form of a flat network where the input is transferred and mapped into feature node. Then, the structure will be expanded broadly to add nodes. The incremental learning algorithm has been developed for rapid change without retraining.

The shallow learning network is more suitable for incremental learning than the deep learning network. It was also found that Radial Basis Function (RBF) is used in incremental learning and is combined with other techniques, such as Neuro-Fuzzy [14]. In addition to RBF, SVM is combined with Mahalanobis distance [15], and Self Organizing Map (SOM) is combined with Euclidean distance [21]. The popular methods used to similarity measure are Euclidean distance and Mahalanobis distance. The Euclidean distance is simple but limited because it is sensitive to scales of variables and suitable for data scattered in circles. The Mahalanobis distance is suitable for elliptically scattered data by considering the covariance matrix to solve the problem of Euclidean distance. Therefore, the similarity measure should be selected based on data scattering. Another interesting similarity measurement method is Correlation distance, as it considers the relationship of variables and is suitable for continuous data. In this research, the researchers proposed incremental learning with Correlation distance, which is a way of measuring similarity to support data that change according to the environment.

### 2.3. Radial Basis Function

Radial Basis Function is a feed forward neural network with only 1 hidden layer. RBF does not have a complicated structure and is more flexible and faster than Multilayer Perceptron (MLP) architecture, each kernel function is connected to hidden nodes which means the connection to one cluster. Incremental learning by increasing hidden nodes and updating weight and relevant parameters [4]. The norm value between two data points can be calculated by using the general calculation formulas, such as Euclidean distance, Mahalanobis distance, Correlation distance or others as shown in Figure 1.

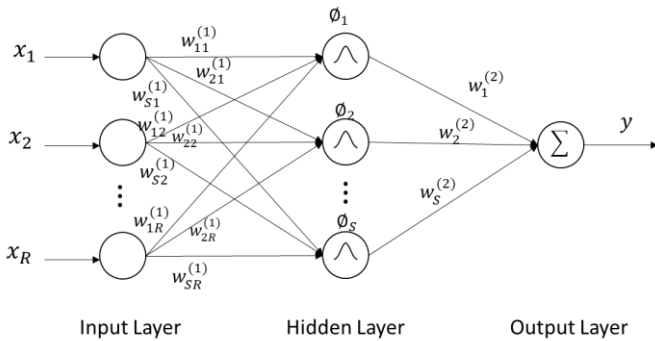


Figure 1: The Architecture of RBF

### 2.4. Correlation Distance

Correlation distance is a statistical measure used to measure the independence of two values or any two vectors. Correlation distance values are between 0 and 1, which can be measured by the variance or standard deviation. It can be calculated according to (1).

$$d_c(\mathbf{p}, \mathbf{q}) = 1 - \frac{cov(\mathbf{p}, \mathbf{q})}{(std_p)(std_q)} \quad (1)$$

$$cov(\mathbf{p}, \mathbf{q}) = \sum_{j=1}^n (\mathbf{p}_j - \bar{\mathbf{p}}) (\mathbf{q}_j - \bar{\mathbf{q}}) \quad (2)$$

$$std(\mathbf{p}) = \sqrt{\frac{1}{n} \sum_{j=1}^n (\mathbf{p}_j - \bar{\mathbf{p}})^2} \quad (3)$$

$$\bar{\mathbf{p}} = \frac{1}{n} \sum_{j=1}^n \mathbf{p}_j \quad (4)$$

where  $d_c(\mathbf{p}, \mathbf{q})$  is correlation distance from  $\mathbf{p}$  to  $\mathbf{q}$ ;  $\mathbf{p}$  and  $\mathbf{q}$  are any data points and  $n$  is number of dimension.

### 3. Method

The objective of this research is to propose the Correlation-Based Incremental Learning Network by using the Correlation Distance and membership function. The operational structure of CILN, as shown in Figure 2.

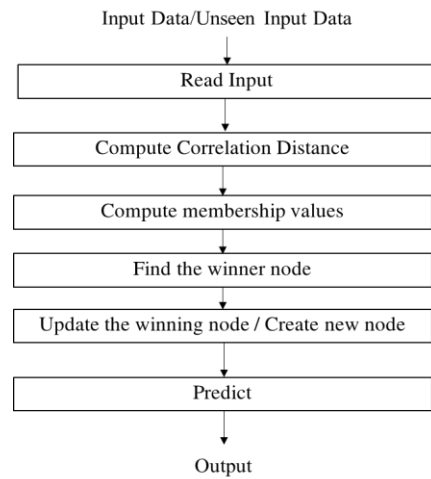


Figure 2: Operational Structure of CILN

Step 1) Set the membership threshold parameter ( $m_{th}$ ) for determining new neuron, where  $m_{th}$  is a value between 0 and 1.

Step 2) Read input data

- Read in a pair of an input pattern  $\mathbf{p}$  and target  $t$ .
- If new knowledge is found set a new neuron  $\mathbf{W}_P$  using  $\mathbf{p}$  and  $\mathbf{W}_T$  according target  $t$  of the input pattern in  $\mathbf{W}_P$ .

Step 3) Read in the next input pattern and its corresponding target, if any.

Step 4) Measure the Correlation distance between the input  $\mathbf{p}$  and the prototype  $\mathbf{W}_P$ , which are centroids using (5).

$$d_i(\mathbf{p}, \mathbf{W}_{Pi}) = 1 - \frac{cov(\mathbf{p}-\mathbf{W}_{Pi})}{(std_p)(std_{\mathbf{W}_{Pi}})} \quad (5)$$

Step 5) Compute membership values of each node using the Gaussian-type radial basis function.

$$m_i = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(d(\mathbf{p}-\mathbf{W}_{Pi}))^2}{2\sigma^2}}; i = 1, 2, \dots, C \quad (6)$$

where  $m_i$  is membership value, and  $\sigma$  is Standard Deviation in cluster  $i$ .  $\sigma$  will be used to show the scatter of the data in the cluster.  $\sigma$  value between 0.001 and 0.05 and after the patterns near the prototype are included in the same prototype, the standard deviation is updated accordingly.

Step 6) Find the winner node that has highest membership value [14].

$$\text{winner} = \arg \max_i (m_i); i = 1, 2, \dots, C \quad (7)$$

$$J = \text{winner index} = \arg \max_i (m_i) \quad (8)$$

Step 7) Update the Winning node [14]

If  $\text{winner} > m_{th}$ , the instance is similarity to the winning node, then update weight of the winning node.

$$\mathbf{W}_{P_{J,\text{new}}} = \frac{\mathbf{W}_{P_{J,\text{old}}}(C_{J,\text{new}} - 1) + \mathbf{p}}{C_{J,\text{new}}} \quad (9)$$

$$\sigma_{J,\text{new}} = \begin{cases} \sqrt{\left(1 - \frac{1}{C_{J,\text{new}}}\right) \sigma_{J,\text{old}}^2 + \left(\frac{(\sigma_{J,\text{old}} - \mathbf{p})^2}{C_{J,\text{new}}}\right)} & \text{if } C_{J,\text{new}} > 1, \\ \sigma_0 & \text{otherwise;} \end{cases} \quad (10)$$

where  $\mathbf{W}_{P_{J,\text{new}}}$  is new weight,  $\mathbf{W}_{P_{J,\text{old}}}$  is original weight,  $\mathbf{p}$  is latest input data, and  $C_j$  is number of members in the cluster.

If  $\text{winner} < m_{th}$ , the instance is considered node a member of the winning node; then a new node  $\mathbf{W}_T$  is created.

Step8) If in prediction mode, i.e. no target, assign the predicted class to the unseen pattern.

$$y = \mathbf{W}_{T_j} \quad (11)$$

where  $y$  is predicted class,  $\mathbf{W}_{T_j}$  is target of the Winner.

If in training mode, i.e. there is target of the input  $\mathbf{p}$ , compute error.

$$e = t - y \quad (12)$$

where  $e$  is error,  $t$  is target, and  $y$  is output.

Step 9) Continue process to step 3, until stop condition is met.

## 4. Experimental Results

### 4.1. Experimental Data

The data set used in the experiment is the gas sensor data set obtained from the UCI Machine Learning Repository [22]. This

data set includes measurements using 16 chemical sensors to measure the gas response at different concentrations. The data set consists of 16 attributes with a total of 13,910 records. This data set has been collected for 36 months from January 2008 to February 2011. The data set was collected from 6 types of pure gas: 1) Ethanol, 2) Ethylene, 3) Ammonia, 4) Acetaldehyde, 5) Acetone, and 6) Toluene. Each gas has added characteristics extracted with different values, making each sensor have 8 features. Therefore, the data set consists of 128 features and is divided into 10 batches in time sequence. The details are shown in Table 1.

The response of the said sensor is read in the form of resistance. Each measurement creates a 16-channel time series data set that responds to the chemicals being measured. In the experiment, the steady-state feature (DR), was selected which means the maximum resistance change compared to the base line.

The signals from 16 sensors showed the characteristics of data consisting of Multivariate Time-series. Since the data in each batch were collected several times, and each batch has a different number of classes (Imbalance class). Therefore, the data obtained are at different concentrations as shown in Figure 3.

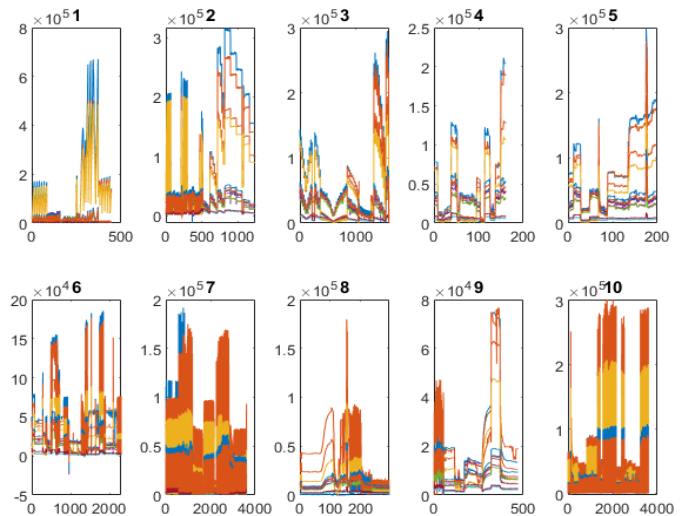


Figure 3: Gas Sensor Array Drift at Different Concentrations Data Set

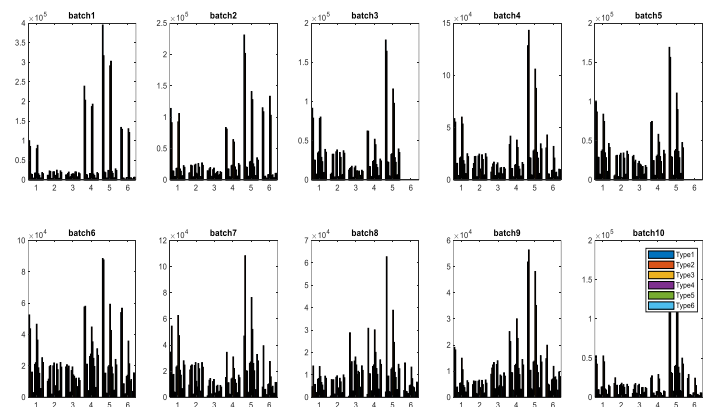


Figure 4: The Odor's Fingerprint of Gas Sensor Drift Compensation Data Set



Table 1: Experimental data

Batch ID	Month	Acetone	Acetaldehyde	Ethanol	Ethylene	Ammonia	Toluene	Total
Batch 1	1, 2	90	98	83	30	70	74	445
Batch 2	3-10	164	334	100	109	532	5	1,244
Batch 3	11, 12, 13	365	490	216	240	275	0	1,586
Batch 4	14, 15	64	43	12	30	12	0	161
Batch 5	16	28	40	20	46	63	0	1997
Batch 6	17, 18, 19, 20	514	574	110	29	606	467	2,300
Batch 7	21	649	662	360	744	630	568	3,613
Batch 8	22, 23	30	30	40	33	143	18	294
Batch 9	24, 30	61	55	100	75	78	101	470
Batch 10	36	600	600	600	600	600	600	3,600

The characteristic response spectrum is called odor’s fingerprint. Therefore, according to the characteristics of response spectrum, different odors can be distinguished showing the fingerprint of each gas type at various times of each batch, as shown in Figure 4.

#### 4.2. Experimental Setting

To evaluate the effectiveness of the proposed CILN, we selected only 16 steady-state (DR) features that have not been extracted from a total of 128 features. Therefore, the data set has 13,910 records, divided into 6 classes: 1) Ethanol, 2) Ethylene, 3) Ammonia, 4) Acetaldehyde, 5) Acetone, and 6) Toluene. The data were normalized and split into training sets and test sets. The training sets were randomly selected for 10%, 20%, 30%, 40%, and 50%; the rest of the data were used for test sets. The proposed CILN algorithm, which uses Correlation distance measurement, was compared with Euclidean distance and Mahalanobis distance measurement, as well as other well-known classifiers, including NaiveBayes, BayesNet, RBF, SVM, MLP, and Simple Logistics. Comparison of experimental results by Accuracy, Precision, Recall, F-Measure. The experiment steps are shown in Figure 5.

#### 4.3. Experimental Results

The experimental results under certain settings show the comparison of the efficiency of the proposed CILN algorithm which uses Correlation distance measurement method and the incremental learning algorithm which uses Euclidean distance and Mahalanobis distance measurement methods. According to Table 2, CILN yielded top 5 accuracy scores: 98.96%, 98.74%, 98.51%, 97.87% and 96.08% by using 50%, 40%, 30%, 20%, and 10% of the training sets, respectively. Considering all the 3 measurement methods, Correlation distance method gave the highest mean of 98.03%, followed by Euclidean distance method of 93.16% and Mahalanobis distance method of 93.10%. Euclidean distance and Mahalanobis distance methods supply similar accuracy, while Correlation distance method gives a high accuracy. Therefore, the similarity measurement method affects accuracy, and the proposed method is suitable for the gas sensor array drift at different concentrations data set which contains continuous data.

Table 2: Comparison the performance of incremental learning using different similarity measurement methods.

Similarity Measure	Training (%)	Accuracy (%)	Precision	Recall	F-Measure
Correlation distance	10	96.08	0.85	0.96	0.90
	20	97.87	0.90	0.99	0.94
	30	98.51	0.94	0.98	0.96
	40	98.74	0.94	0.99	0.97
	50	98.96	0.95	0.99	0.97
	Average	98.03	0.92	0.98	0.95
Euclidean distance	10	90.13	0.69	0.84	0.76
	20	92.44	0.74	0.91	0.82
	30	93.75	0.77	0.93	0.85
	40	93.82	0.79	0.91	0.84
	50	95.64	0.84	0.94	0.89
	Average	93.16	0.77	0.91	0.83
Mahalanobis distance	10	87.86	0.63	0.85	0.72
	20	93.22	0.75	0.95	0.84

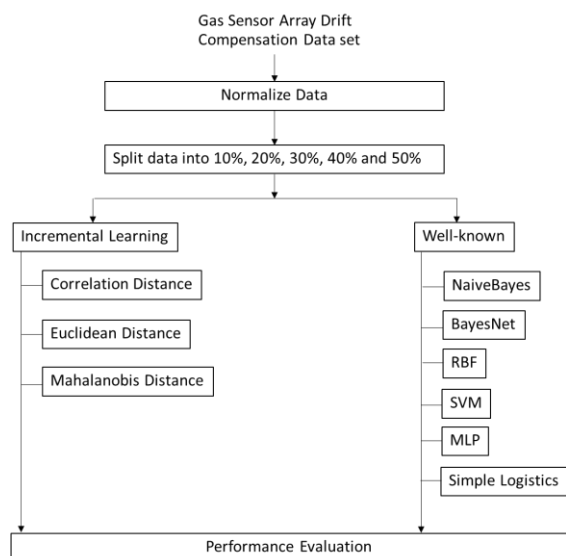


Figure 5: Experimental Steps

Similarity Measure	Training (%)	Accuracy (%)	Precision	Recall	F-Measure
	30	94.38	0.80	0.94	0.86
	40	94.36	0.79	0.94	0.86
	50	95.64	0.85	0.94	0.89
	Average	93.10	0.76	0.92	0.83

Table 3: Comparison of Performance with well-known classifiers

Algorithm	Training (%)	Accuracy (%)	Precision	Recall	F-Measure
CILN	10	96.08	0.85	0.96	0.90
	20	97.87	0.90	0.99	0.94
	30	98.51	0.94	0.98	0.96
	40	98.74	0.94	0.99	0.97
	50	98.96	0.95	0.99	0.97
Simple Logistics	10	93.52	0.94	0.94	0.94
	20	93.66	0.94	0.94	0.94
	30	95.48	0.96	0.96	0.96
	40	95.70	0.96	0.96	0.96
	50	96.05	0.96	0.96	0.96
MLP	10	91.37	0.92	0.91	0.91
	20	93.37	0.94	0.93	0.94
	30	94.64	0.95	0.95	0.95
	40	94.42	0.95	0.94	0.95
	50	95.64	0.96	0.96	0.96
SVM	10	70.71	0.80	0.71	0.71
	20	76.23	0.85	0.76	0.78
	30	79.48	0.87	0.80	0.81
	40	82.19	0.88	0.82	0.84
	50	83.89	0.89	0.84	0.85
BayesNet	10	59.41	0.65	0.59	0.60
	20	65.05	0.69	0.65	0.66
	30	65.27	0.69	0.65	0.66
	40	66.30	0.70	0.66	0.67
	50	66.48	0.70	0.67	0.67
RBF	10	64.29	0.67	0.64	0.64
	20	61.94	0.65	0.62	0.60
	30	63.09	0.67	0.63	0.61
	40	62.73	0.66	0.63	0.61
	50	62.78	0.66	0.63	0.61
NaiveBayes	10	51.56	0.60	0.52	0.52
	20	51.24	0.60	0.51	0.52
	30	51.43	0.61	0.51	0.52
	40	51.72	0.61	0.52	0.53
	50	51.45	0.60	0.52	0.52

Table 3 shows the comparison of the performance of the proposed CILN algorithm with the well-known classifiers, such as NaiveBayes, BayesNet, RBF, SVM, MLP, and Simple

Logistics. Overall, CILN still yielded the highest accuracy of 98.96%, Simple Logistics of 96.05%, MLP of 95.64%, SVM of 83.89%, BayesNet of 66.48%, RBF of 64.29%, and NaiveBayes of 51.72% by using 50%, 50%, 50%, 50%, 50%, 10%, and 40% of the training sets, respectively. It was found that the number of training data sets can affect the accuracy. However, new data entered the classifiers may change according to the environment. Without incremental learning, it will also affect accuracy. Therefore, the algorithm should have incremental learning ability for effective classification.

### 5. Conclusions and Future Work

Machine learning is a useful tool for analysis and learning. However, data imported into the system are both static data and dynamic data which fluctuate and change according to the environment. Conventional algorithms still lacks incremental learning new data. When new data are added, the algorithm will not be able to learn and adapt, resulting in reduced efficiency. Moreover, all the trained data sets are no longer available, and a new model needs to be created for new data. This process leads to a phenomenon known as catastrophic forgetting, resulting in poor classification performance. Therefore, to solve this problem, this research has proposed Correlation-Based Incremental Learning algorithm that allows the model to learn and improve automatically while maintaining old knowledge using the Correlation distance to measure similarities and membership functions using Gaussian-type Radial Basis Function to determine membership of each node.

Gas sensor data from the UCI machine learning repositories used to evaluate the performance of the proposed algorithm. This data set holds 13,910 measurements from 16 chemical sensors exposed to 6 gases at different concentration levels. This data set has been collected for 36 months from January 2008 to February 2011. The data were collected from 6 types of pure gas: 1) Ethanol, 2) Ethylene, 3) Ammonia, 4) Acetaldehyde, 5) Acetone, and 6) Toluene. In the experiment, only 16 steady-state features (DR) were chosen. The data were normalized and split into training sets and test sets. The training sets were randomly selected for 10%, 20%, 30%, 40%, and 50%; the rest of the data were used for test sets.

The results show that CILN allows the system to learn new patterns while maintaining the old knowledge. The proposed CILN algorithm supplies an initial accuracy of 96.08% by using only the 10% of the training data set which is higher than all classifiers. In addition, CILN yields the highest accuracy of 98.96% when 50% of the training data set was used. It shows that CILN can learn from a small sample size and can adapt and learn new data automatically while keeping the existing knowledge. Therefore, CILN can increase the accuracy of classification and support the time series data which are dynamic data and can be used for environmental or other inspections. Moreover, it was found that using only 16 steady-state features (DR) was sufficient for gas classification without additional feature extraction. In the future work, we will consider reducing the dimensions of data by selecting features, removing noise, and selecting the proper signal range.

### Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

We would like to gratefully acknowledge the King Mongkut's University of Technology North Bangkok and Rajamangala University of Technology Isan.

## References

- [1] A. Loutfi, S. Coradeschi, G. K. Mani, P. Shankar, J.B.B. Rayappan, "Electronic noses for food quality: A review" *Journal of Food Engineering*, **144**, 103-111, 2015. <http://dx.doi.org/10.1016/j.jfoodeng.2014.07.019>
- [2] L. Huang, J. Zhao, Q. Chen, Y. Zhang, "Nondestructive measurement of total volatile basic nitrogen (TVB-N) in pork meat by integrating near infrared spectroscopy" computer vision and electronic nose techniques *Food Chemistry*, **145**, 228-236, 2014. <http://dx.doi.org/10.1016/j.foodchem.2013.06.073>
- [3] L. Pan, W. Zhang, N. Zhu, S. Mao, K. Tu, "Early detection and classification of pathogenic fungal disease in post-harvest strawberry fruit by electronic nose and gas chromatography-mass spectrometry" *Food Research International*, **62**, 162-168, 2014. <http://dx.doi.org/10.1016/j.foodres.2014.02.020>
- [4] B. Tudu, A. Jana, A. Metla, D. Glosch, N. Bhattacharyya, "Electronic nose for black tea quality evaluation by an incremental RBF network" *Sensors and Actuators B: Chemical*, **138**(1), 90-95, 2009. <http://dx.doi.org/10.1016/j.snb.2009.02.025>
- [5] Y. Yang, H. Liu, Y. Gu, "A Model Transfer Learning Framework With Back-Propagation Neural Network for Wine and Chinese Liquor Detection by Electronic Nose" *IEEE Access*, **8**, 105278-105285, 2020. <https://doi.org/10.1109/ACCESS.2020.2999591>
- [6] M. Aleixandre, J. P. Santos, I. Sayago, M. C. Horrillo, J. M. Cabellos, T. Arroyo, "Use of an electronic nose as a tool to differentiate winemaking techniques" in 2015 10th Spanish Conference on Electron Devices (CDE), Madrid, Spain, 2015. <https://doi.org/10.1109/CDE.2015.7087490>
- [7] B. Schmekel, F. Winquist, A. Vikström, "Analysis of breath samples for lung cancer survival" *Analytica Chimica Acta*, **840**, 82-86, 2014. <https://doi.org/10.1016/j.aca.2014.05.034>
- [8] E. Westenbrink, R.P. Arasaradnam, N. O'Connell, C. Bailey, C. Nwokolo, K.D. Bardhan, J.A. Covington, "Development and application of a new electronic nose instrument for the detection of colorectal cancer" *Biosensors and Bioelectronics*, **67**, 733-738, 2015. <http://dx.doi.org/10.1016/j.bios.2014.10.044>
- [9] Q. Wu, D. Snell, "An Introduction to Incremental Learning" *Predictive Analytics and Futurism*, **13**, 18-19, 2016.
- [10] R. M. French, "Catastrophic forgetting in connectionist networks" *Trends in Cognitive Sciences*, **3**(4), 128-135, 1999. [http://dx.doi.org/10.1016/S1364-6613\(99\)01294-2](http://dx.doi.org/10.1016/S1364-6613(99)01294-2)
- [11] Z. Zheng, X. Lin, "Study on Application of Medical Diagnosis by Electronic Nose" *World Science and Technology*, **14**(6), 2115-2119, 2012. [http://dx.doi.org/10.1016/S1876-3553\(13\)60016-2](http://dx.doi.org/10.1016/S1876-3553(13)60016-2)
- [12] V. Losing, B. Hammer, H. Wersing, "Incremental on-line learning: A review and comparison of state of the art algorithms" *Neurocomputing*, **275**, 1261-1274, 2018. <https://doi.org/10.1016/j.neucom.2017.06.084>
- [13] S. Wan, and L.E. Banta, "Parameter Incremental Learning Algorithm for Neural Networks" *IEEE Transactions on Neural Networks*, **17**(6), 1424-1438, 2006. <https://doi.org/10.1016/j.neucom.2017.06.084>
- [14] G. G. Yen, P. Meesad, "An effective neuro-fuzzy paradigm for machinery condition health monitoring" *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **31**(4), 523-536, 2001. <https://doi.org/10.1109/3477.938258>
- [15] H. O. Myint, P. Meesad, "Incremental Learning Algorithm based on Support Vector Machine with Mahalanobis distance (ISVMM) for intrusion prevention" in 2009 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology. Pattaya, Chonburi, Thailand, 2009. <https://doi.org/10.1109/ECTICON.2009.5137129>
- [16] Y. Pang, J. Deng, Y. Yuan, "Incremental threshold learning for classifier selection" *Neurocomputing*, **89**, 89-95, 2012. <https://doi.org/10.1016/j.neucom.2012.01.012>
- [17] K. Diaz-Chito, K. Georgouli, A. Koidis, J. M. del Rincon, "Incremental model learning for spectroscopy-based food analysis" *Chemometrics and Intelligent Laboratory Systems*, **167**, 123-131, 2017. <https://doi.org/10.1016/j.chemolab.2017.06.002>
- [18] M. Režnáková, L. Tencer, M. Cheriet, "Incremental Similarity for real-time on-line incremental learning systems" *Pattern Recognition Letters*, **74**, 61-67, 2016. <https://doi.org/10.1016/j.patrec.2016.01.010>
- [19] L. Li, Z. Jun, J. Fei, S. Li, "An incremental face recognition system based on deep learning" in 2017 15th IAPR International Conference on Machine Vision Applications (MVA), Nagoya, Japan, 2017. <https://doi.org/10.23919/MVA.2017.7986845>
- [20] C. L. P. Chen and Z. Liu, "Broad Learning System: An Effective and Efficient Incremental Learning System Without the Need for Deep Architecture" *IEEE Transactions on Neural Networks and Learning Systems*, **29**(1), 10-24, 2018. <https://doi.org/10.1109/TNNLS.2017.2716952>
- [21] L. Tian and W. Liu, "Incremental intrusion detecting method based on SOM/RBF" in 2010 International Conference on Machine Learning and Cybernetics. Qingdao, China, 2010. <https://doi.org/10.1109/ICMLC.2010.5580770>
- [22] A. Vergaraa, S. Vembu, T. Ayhan, M.A. Ryan, M.L. Homer, R. Huerta, "Chemical gas sensor drift compensation using classifier ensembles" *Sensors and Actuators B: Chemical*, **166-167**, 320-329, 2012. <https://doi.org/10.1016/j.snb.2012.01.074>

## Design of a Flapping Wings Butterfly Robot based on Aerodynamics Force

Kanjanapan Sukvichai\*, Kan Yajai

Faculty of Engineering, Kasetsart University, Bangkok, 10900, Thailand

---

### ARTICLE INFO

Article history:

Received: 06 July, 2020

Accepted: 11 August, 2020

Online: 25 August, 2020

---

Keywords:

Flapping Wings

Butterfly Robot

Aerodynamics

---

### ABSTRACT

*Insect robots are always amazed by humans due to their ability to fly using a wing flapping mechanism. The butterfly robot was designed in this research based on aerodynamics and aeroelastic especially for designing a flapping mechanism due to its complexity. A butterfly wing structure was designed by considering aerodynamics forces based on assumptions. Aerodynamic equations were derived in order to obtain lift and thrust forces that acted on a small wing section. The wing was assumed to be in the Quasi-steady state when it was analyzed based on the thin airfoil theorem. Airflow was simulated in order to obtain air pressure and vertexes acting on the wing surface when it swings in the still air. By integrating the wing section's lift force for a flapping cycle motion trajectory, the average lift force was obtained. The robot wing structure was designed based on the average lift. The real butterfly wing was used as the guideline for designing the robot wing. Each wing was fabricated from a laminar plastic sheet. Carbon fiber robs were used as an internal reinforced support structures for wing frames. The reinforced wing achieved the wing's rigidity and was considered as a thin airfoil. The flapping mechanism was designed by using two separated servo motors because of its flexibility and performance. This mechanism enables the robot's rotation without an extra actuator. The butterfly robot body was manufactured from the 3D printer using PLA material. The experiments were conducted to identify the robot performance. The designed butterfly robot can take off from the support platform and fly up to a certain height.*

---

## 1. Introduction

Flying robots has the ability to fly up into the air and amazes human because of their behavior. There are two major types of flyable animals which are insects and birds. They use different flying mechanisms. The flyable robot uses several aerodynamic properties benefits for creating aerodynamics forces. The robot that mimics the insect is called micro air vehicles. This manuscript is an extension of work originally presented in the International Conference on Electronics, Information, and Communication 2020 (ICEIC2020) [1]. The flapping winged type robot has an advantage over a regular drone robot that uses rotating propellers. In general, flapping motion produces a larger lift force compared to a regular fixed wing at a specific angle of attack (AoA) which is important for designing a tiny flying robot with a high efficiency. Furthermore, a flapping wing robot model has the advantages over other flying robot models which are a vertical taking off from and landing to a surface, hovering with a specific height, and

maneuvering in a complex environment. The autonomous flying robots were interested because it is complex to analyze, model and design robot mechanisms and hard to determine control algorithms because of the nonlinearity term. Delfly [2] Nanohumming bird [3], Purdue Hummingbird robot [4] and RoboBee [5] were popular developed flying robots that can taking off and hovering based on flapping wing mechanisms. Delfly and Nanohumming bird robot's flapping mechanism were designed based on a crank mechanism that required one actuator to drive both wings while RoboBee and Purdue Hummingbird robot's flapping mechanism used two separated actuators for wings. The crank flapping mechanism is easy to develop but both wings share the same angle and angular velocity trajectories yield the less degree of freedom than the two separated actuators mechanism. A crank mechanism flying robot requires an extra actuator to rotate a robot. Although these robots were successful to taking-off and hovering but the robots are small and cannot carry any large payload and the wing flapping speed is very high that consumes a lot of energy.

---

\*Corresponding Author: Kanjanapan Sukvichai, Faculty of Engineering, Kasetsart University, Bangkok, Thailand, Tel: +662-797-0999, Email: fengkpsc@ku.ac.th



The aerodynamics analysis of the butterfly wing section was explained and derived at the beginning of this manuscript in order to reveal the aerodynamics of the wing. The wing aerodynamics equations are derived based on assumptions that wing was in the Quasi-steady state with thin airfoil properties. Then, the derived equations were used to design the butterfly robot's wing. Navier-Stokes equations are used to obtain the numerical solution of the airflow against the wing. The flapping mechanism was designed based on two separate actuators mechanism which enables a rotation. The designed robot was based on a butterfly animal that has bigger wings and flaps its wings at a slower speed than other insects yield less energy consumption and more payload. Then, the designed robot can be used in the applications in Thailand such as the survey drone in the agriculture field and in the anti-terrorist application. Lastly, the butterfly wing structure, materials, flapping mechanism and the butterfly robot are designed. The robot parts were manufactured and constructed. The designed robot was experimented and the results were explained at the end of this manuscript. The overall robot design process is displayed in Figure 1.

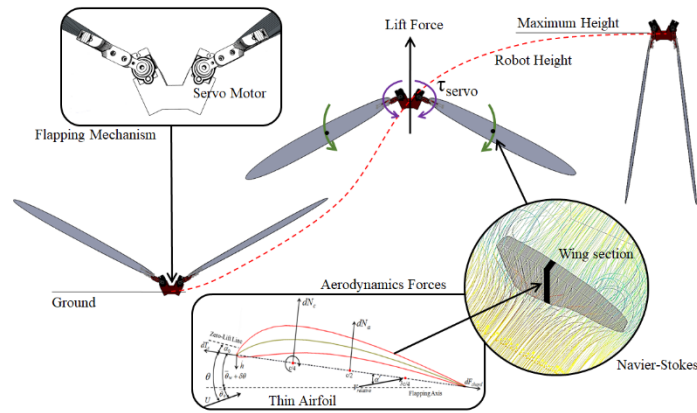


Figure 1: Overall robot design process.

## 2. Wing Aerodynamic Force Analysis

Aerodynamics of the butterfly wing can be analyzed by considering the small wing section along the wing span. A real butterfly wing consists of very thin membranes and a body wall made from thousands of scales and hairs, then, butterfly wing was considered as the thin airfoil. The butterfly robot model was modeled and aerodynamic forces that applied to the wing section are shown in Figure 2.

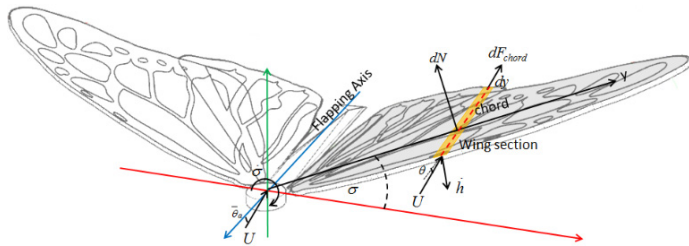


Figure 2: The butterfly robot model.

Due to the assumption that the wing's aspect ratio is large, the air flow over each section can be considered as a chordwise flow. The wing is in the Quasi-steady state when parameters and time-varying components are considered as time invariant constants. The wing section is shown in Figure 3. The aerodynamics and

aeroelastic forces are orthogonal to the wing section's chord line and equal to (1). Aeroelastic happens in this calculation because the wing motion induces an added mass effect into the wing section.

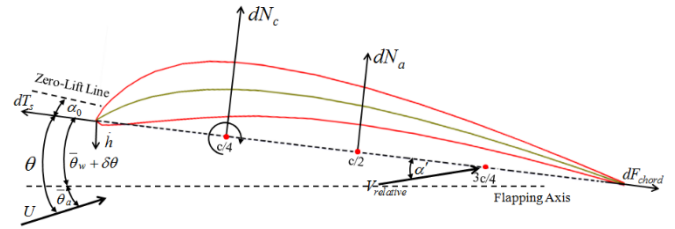


Figure 3: Wing Section

$$dN = dN_c + dN_a \quad (1)$$

where,

$dN_c$  is the circulation normal force.

$dN_a$  is the add mass normal force.

A plunging velocity at the leading edge  $\dot{h}$ , the local pitch angle  $\theta$  are created by a swing motion specific for each section. For simplicity, the wing is considered as the fixed structure and the non-span-wised bending wing. A plunging motion is a pure sinusoid function in this research and it equals to (2)

$$h(y, t) = \Gamma \cdot y \cdot \cos(\omega t) \quad (2)$$

where,

$\Gamma$  is the maximum magnitude of a flapping motion.

$\omega$  is the flapping frequency.

The section aerodynamic force,  $dN_c$ , acts at the one-quarter of the chord length. A real butterfly wing that consists of a very thin membrane with the support structure was used as the designed reference for the butterfly wing. Then, the butterfly wing is modeled as the thin airfoil because a ratio of the wing thickness to the chord length is small. The normal-force was derived as (3).

$$dN_c = \rho U V_{0.25} \pi \left( \alpha' + \left( \alpha_0 + \bar{\theta}_a + \bar{\theta}_w \right) \right) c dy \quad (3)$$

where,

$U$  is the free stream velocity.

$V_{0.25}$  is the flow's relative velocity at a 0.5 chord.

$\rho$  is the air density.

$c$  is the wing chord length.

$\alpha'$  is the flow's relative angle of attack at the chord.

$\alpha_0$  is the angle of the section's zero-lift line.

$\bar{\theta}_a$  is the angle of flapping axis respected to the free stream.

$\bar{\theta}_w$  is a mean pitch of chord respected to the flapping axis.

Due to the properties of the aeroelasticity, the added mass of the wing section generates an apparent normal force when a wing is flapping. This normal force acts at the middle of the chord length of the wing section and equals to (4).

$$dN_a = \frac{\rho\pi c^2}{4}(U\dot{\alpha} - 0.25c\ddot{\theta})dy \quad (4)$$

where,

$V_{0.5}$  is the flow's relative velocity at half-chord.

$\alpha$  is the relative AoA at location of chord.

$$\theta = \bar{\theta}_a + \bar{\theta}_w$$

The relative angle of attack (AoA) was approximated by a diagram of velocities displayed in Figure 3 and equals to (4)

$$\alpha = \frac{\dot{h} \cos(\theta - \bar{\theta}_a) + \frac{3}{4}c\dot{\theta} + U\Delta\theta}{U} \quad (4)$$

where,

$\Delta\theta$  is the pitch angle change from a wing's motion.

Clearly, the pitch angle was changed according to time, flapping frequency and the twisted angle constant per unit distance along the wing span. By assumption, the wing is an untwisted and not bended wing, thus, the flow's relative AoA at the 3/4 chord position was derived as (5)

$$\alpha' = C(k)_{Jones} \alpha - \frac{w_0}{U} \quad (5)$$

where,

$w_0$  is the downwash velocity at the chord.

$C(k)_{Jones}$  is the Jones's coefficient.

The Jones's coefficient [6] for the finite aspect ratio ( $AR$ ) solution was derived based on the strip theory in case of the wing's finite span unsteady vortex wake [7].  $AR$  equals to the wingspan<sup>2</sup> divided by the wing area. The dimensionless reduced frequency,  $k$ , was introduced for an unsteady aerodynamics and aeroelasticity condition. Reduced frequency was related to the flapping frequency as (6).

$$k = \frac{c\omega}{2U} \quad (6)$$

Jones's coefficient consists of the real part and imaginary parts, thus, coefficient is a complex function. The alternative approximated formulation was proposed [8] as (7)

$$C(k)_{Jones} = \left(\frac{AR}{2+AR}\right)(C'(k)) = \left(\frac{AR}{2+AR}\right)(F'(k) + iG'(k)) \quad (7)$$

The dimensionless constants  $F'(k)$  and  $G'(k)$  are depending on  $AR$  and  $k$ . Scherer proposed the approximated equations of  $F'(k)$  and  $G'(k)$  as (8) and (9) respectively.

$$F'(k) = 1 - \frac{C_1 k^2}{(k^2 + C_2)} \quad (8)$$

$$G'(k) = -\frac{C_1 C_2 k}{(k^2 + C_2)} \quad (9)$$

where,

$$C_1 = \frac{0.5AR}{2.32 + AR} \quad \text{and} \quad C_2 = 1.81 + \frac{0.772}{AR}$$

The flapping motion was designed as a sinusoidal equation (10)

$$\alpha = \Gamma e^{i\omega t} \quad (10)$$

By putting (6) to (10) into (5), the flow's relative AoA at the 3/4 chord location can be approximated by (11)

$$\alpha' = -\frac{w_0}{U} + \left(\frac{AR}{2+AR}\right) \left(F'(k)\alpha + \frac{c}{2Uk}G'(k)\dot{\alpha}\right) \quad (11)$$

The untwisted elliptical-planform wing downwash and a free stream velocity have a relationship as shown in (12)

$$\frac{w_0}{U} = \frac{2(\alpha_0 + \bar{\theta})}{2+AR} \quad (12)$$

Therefore, the flow's relative velocity at half-chord was calculated by (13)

$$V_{2.5} = \sqrt{\left(-\dot{h} \sin(\theta - \bar{\theta}_a) + U \cos(\theta)\right)^2 + \left(-\frac{1}{2}c\dot{\theta} + U(\alpha' + \bar{\theta})\right)^2} \quad (13)$$

The chordwise forces acts on the wing section. The total chordwise forces,  $dF_{chord}$ , was the combination of 3 forces which are the leading-edge suction, camber force, and chordwise viscosity friction drag was explained as (14).

$$dF_{chord} = dT_s - dF_{camber} - dF_{friction} \quad (14)$$

The leading-edge suction,  $dT_s$ , was given in [6] as (15)

$$dT_s = \eta_s 2\pi \left(\alpha' + \bar{\theta} - \frac{1}{4}c \frac{\dot{\theta}}{U}\right)^2 \frac{\rho U V_{2.5}}{2} c dy \quad (15)$$

where,

$\eta_s$  is the leading edge suction efficiency factor

Two other chordwise forces were calculated by (16) and (17)

$$dF_{camber} = -2\pi\alpha_0 (\alpha' + \bar{\theta}) \frac{\rho U V_{2.5}}{2} c dy \quad (16)$$

$$dF_{friction} = \frac{1}{2} \rho V_{chord}^2 C_{df} c dy \quad (17)$$

where,

$V_{chord}$  is the section normal flow speed.

$C_{df}$  is the skin friction drag coefficient.

$$V_{chord} = U \cos(\theta) + \dot{h} \sin(\bar{\theta}_a - \theta) \quad (18)$$

The skin friction drag coefficient  $C_{df}$  for a laminar flat plate skin was approximated as a function of the Reynolds number [9] which is calculated by (19) where  $\mu$  is the air dynamic viscosity.

$$C_{df} = \frac{0.664}{\sqrt{\frac{\rho U c}{\mu}}} \quad (19)$$

The horizontal force  $dF_{chord}$  and the vertical force  $dN$  at each strip  $dy$  were altered into the normal lift and parallel thrust forces equations of the free stream velocity as (20) and (21) respectively.

$$dL = dF_{chord} \sin(\theta) + dN \cos(\theta) \quad (20)$$

$$dT_s = dF_{chord} \cos(\theta) - dN \sin(\theta) \quad (21)$$

Wing's lift and thrust at any particular instant time can be calculated by integrating (20) and (21) along the wing span. The instant section's dihedral angle  $\sigma(t)$  in the swing motion is displayed in Figure 4 and it equals to  $\Gamma \cos(\omega t)$ . The average wing lift and thrust were calculated by integrating every instant wing's lift and thrust function over one flapping period ( $1/f$ ), equals to (22) and (23).

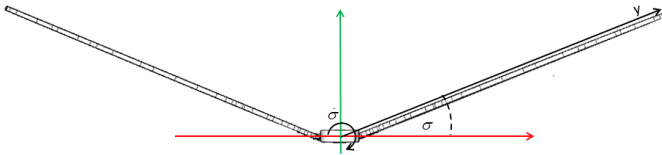


Figure 4: The section's dihedral angle at that instant in the flapping cycle.

$$\bar{L} = \frac{1}{1/f} \int_0^{1/f} \left( \int_{-b/2}^{b/2} (\cos(\sigma(t)) dL) \right) dt \quad (22)$$

$$\bar{T}_s = \frac{1}{1/f} \int_0^{1/f} \left( \int_{-b/2}^{b/2} (dT_s) \right) dt \quad (23)$$

where,

$f = \omega/2\pi$  is the flapping frequency.

The average lift and trust shown in (22) and (23) are the function of the flapping frequency. Therefore, wing swings with the high flapping frequency, it produces large lift force for lifting robot body up to the air. Unfortunately, when wing flaps with symmetry magnitude for the upward and downward swing, wing generated down force and lift force in the motion. Unlike the bird's wing, the insect wing cannot be folded because the wing structure is fixed. To overcome this issue in this research, the nonsymmetric motion for the upward and downward swing is considered. The wing starts at position (a) and moves down to position (b) as shown in Figure 5. The wing has a high speed for the downward direction and a slow speed for the upward direction to generate the maximum possible lift force.

Butterfly robot will fly in the controlled wind-free indoor environment, then, the velocity of the air can be negligible. Thus, the wing motion can be considered as the laminar flat plate moves

perpendicular to the air. In order to analyze the designed wing structure, the airflow against a wing is simulated. This flow is the non-reactive flow condition because the wing motion in this research was not generated from a fuel or heat which can alters the physical property of the air flow causing more shear layers in the flow. The governing Navier-Stokes (NS) equations were used as the basic flow equations. The conservation of air can be calculated from (24)

$$\frac{D(\rho)}{Dt} + \rho \nabla \cdot v = 0 \quad (24)$$

where,

$\frac{D}{Dt} \square \frac{\partial}{\partial t} + v \cdot \nabla$  is material derivative.

$v$  is the velocity of air in x-, y- and z-axis.

$\nabla$  is the divergence in x-, y- and z-axis.

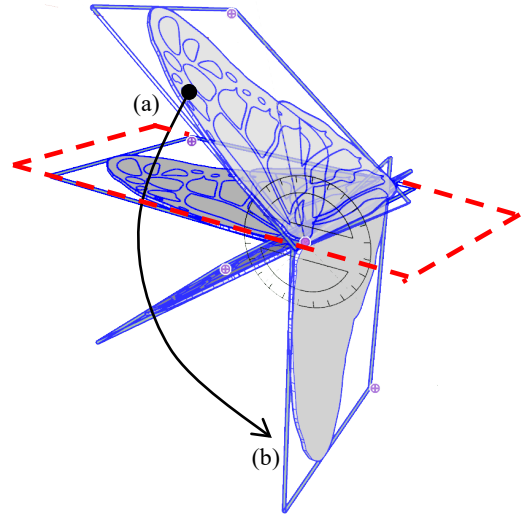


Figure 5: Butterfly wing motion

The Cauchy momentum equation of the air flow can be described as (25)

$$\rho \frac{D(v)}{Dt} = -\nabla P + \nabla \cdot \tau + (\rho - \rho_\infty) g \quad (25)$$

where,

$P$  is a pressure.

$\tau$  is a viscous stress tensor.

$g$  is an external force vector due to the buoyancy.

For simplicity, in this research the fluid was considered as Newtonian fluid and the fluid was incompressible. Therefore, the viscous stress tensor can be estimated as (26)

$$\tau = \mu(\nabla v + (\nabla v)^T) - \frac{2}{3} \mu(\nabla v) I \quad (26)$$

Navier-Stokes equations cannot be analytically solved due to the complexity [10]. The Semi-Implicit Method for Pressure Linked Equations Revised (SIMPLER) [11] method was selected



in order to obtain the numerical solution of a flow. The first step in this method is to calculate the divergence-free condition pseudo-time stepping by considering (26) with the pressure-free condition and equals to (27)

$$\rho \frac{D(\hat{v})}{Dt} = \left( \frac{\partial(\rho\hat{v})}{\partial t} + \hat{v} \cdot \nabla(\rho\hat{v}) \right) = \nabla \cdot \hat{\tau} + (\rho - \rho_\infty)g \quad (27)$$

where,

$\hat{v}$  and  $\hat{\tau}$  are the estimated values of  $v$  and  $\tau$

Equation (27) is needed to be discretized for numerical method, yield (28)

$$\left( \frac{\rho(\hat{v} - v_0)}{\Delta t} + \hat{v} \cdot \nabla(\rho\hat{v}) \right) = \nabla \cdot \hat{\tau} + (\rho - \rho_\infty)g \quad (28)$$

The next step, the estimated pressure is estimated from pressure Poisson equation as (29)

$$\rho \nabla \left( \frac{\hat{v}}{\Delta t} \right) = \nabla \cdot \nabla \hat{P} = \nabla^2 \hat{P} \quad (29)$$

where,

$\hat{P}$  is the estimated values of  $P$

Then, value of  $\hat{P}$  is used in the discrete version of equation (25) to obtain the updated velocity value  $v^*$  and used this value to update  $\hat{P}$  value to  $P^*$  value via (29). Finally, the next step velocity is obtained by (30)

$$v = v^* - \frac{\Delta t}{\rho} \nabla P^* \quad (30)$$

There are several techniques to solve the numerical flow equations such as Euler, Lax-Wendroff, and 3<sup>rd</sup>–4<sup>th</sup> orders Runge-Kutta method. In this research, the Runge-Kutta method was selected as the solver with proper boundary conditions such as flow inlet and outlet. For simplicity, the simulation is based on the Flowsquare+ flow simulator [12]. A designed wing STL mesh file was imported into the simulator, placed at the center of the airflow chamber and set perpendicularly to the x-axis. The number of grid points were 128, 64 and 64 for the x-axis, y-axis, and z-axis respectively. Air flowed in +x-direction with 5 m/s to make the simulation reaches the Quasi-steady state quickly. Dynamic viscosity was set at 1.8E-5 (kg/m/s) to simulate the atmospheric air at room temperature. The simulation results show in Figure 6 and Figure 7. Figure 6 shows that there are several vertexes happen at the back of the butterfly wing tip and the airflow simulation proves the assumption that when wing flap perpendicular to the air in the still air, it acts like a laminar flat plate. The pressure density is shown in Figure 7. From this result, the pressure condenses around the wing's center of mass. This information will be used for the analysis of dynamics based on masses system via Euler-Lagrange equation of motion. It is clear that if one wing flaps, the reaction force is perpendicular to the wing surface and has components in both x- and y-axis. This phenomenon generates pros and cons for controlling the butterfly robot. If two wings have a symmetrical synchronized motion, then, wings generate only the left force to the robot body but if wings motion is not symmetric or

asynchronized then the robot will be rotated. The diagram of wings motion is displayed in Figure 8. The averaged reaction force acting point is estimated from the pressure simulation result.

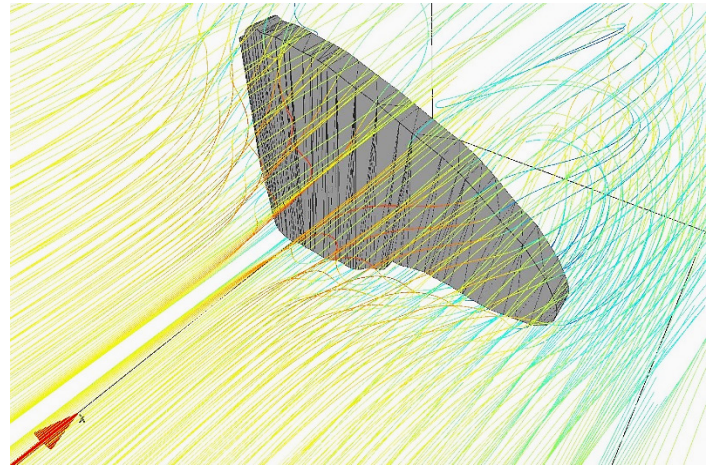


Figure 6: Air flow to a wing.

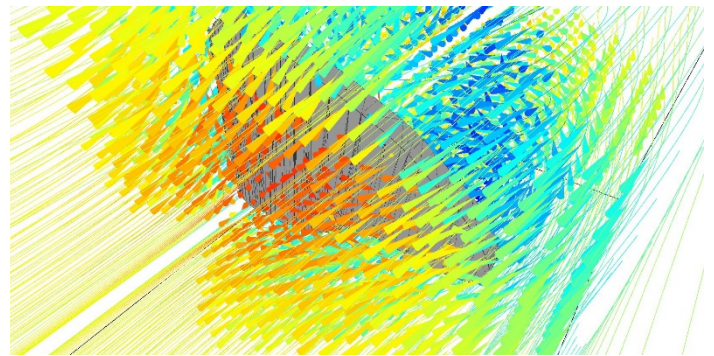


Figure 7: Pressure at a wing

### 3. Butterfly Robot Construction

In this section, the butterfly robot construction is explained. The first butterfly robot wing has been designed in such the way that it is 10 times larger than the real butterfly but keeps the wing physical properties such as a shape and the width to height proportion as shown in Figure 8. The first robot wing has a wingspan and area about 51 cm and 539 cm<sup>2</sup> respectively. Due to the assumption from the second topic, the wing was made from a reinforced carbon fiber rods structure with a transparent laminar plastic sheet. This reinforced structure creates the wing rigidity and ensures thin airfoil properties of a flat plate skin shape. Due to the thin wing, the chord line is approximately overlap to the wing camber line. Butterfly robot has one left-wing and one right-wing, unlike a real butterfly that has two left-wings and two right-wings which generate the 3 degrees of freedoms. The designed butterfly robot creates a rotation about the flapping axis while the real butterfly creates a rotation around a single rotation point. There are two major flapping mechanisms which are one actuator and two actuators driven mechanisms. A crank-and-piston is the one actuator flapping mechanism that creates the same motion for both left and right-wings while two separated actuators mechanism can create different motions for the left-wing and the right-wing. In this research, in order to achieve the advantages of unsymmetrical aerodynamics from two wings, two servos driven mechanism was selected in order to create mismatch motions and forces between



two wings. Servo motors were placed in the servos holder and attached to each wing by using a wing locker as shown in Figure 10. The complete butterfly robot is shown in Figure 11. The designed mechanism can enable robot rotation without an extra actuator such as a tail.

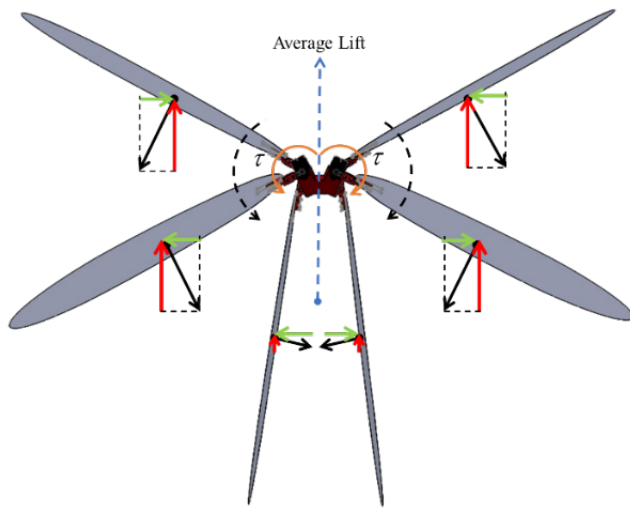


Figure 8: Two wings flapping downward motion.

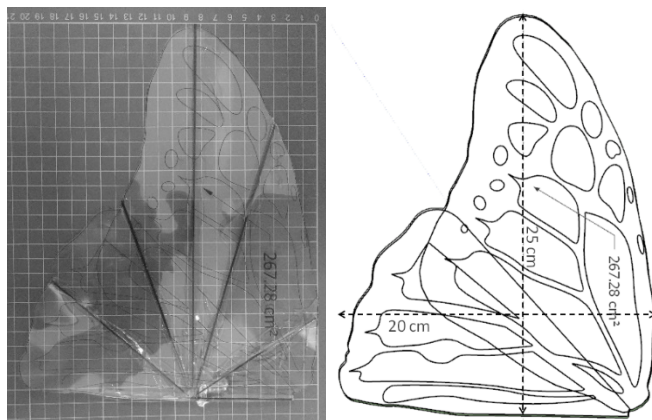


Figure 9: Prototype Butterfly Wings

The KST MS325 micro high-speed servo motors were used for the robot. The motor can generate 5.2 Kg.cm torque and achieve 60° per 0.07 second rotating speed when operated at 8.4 volts. Trust and lift forces were obtained by equations and simulations from the previous section. The completed butterfly robot dynamics was not analyzed in this manuscript because it requires masses system dynamics analysis via Euler-Lagrange or Newton-Euler methods. The completed butterfly robot mathematic model, then, will be explained in the future research. STM32 ARM Cortex-M3 is selected as the robot's main controller operated at 72 MHz frequency via a PLL module. Servo motor and ARM controller are shown in Figure 12. A servo motor shaft position can be controlled by using the RC control signal. This control signal has a specific pulse train that has a frequency of 50 Hz, and a duty cycle varies between 5 to 10 percent of the period depends on the desired shaft location. The selected servo motor consists of a local control circuit based on a PID control algorithm, a h-bridge motor driver and the shaft position sensor. Thus, servo motor's shaft can rotate to any desired position by using the RC control signal as long as a load torque doesn't greater than a maximum servo torque. If the

maximum torque is much greater than a load torque, the servo shaft can rotate with a constant angular velocity. Therefore, in this research, the control algorithm in the ARM controller is used for the trajectory and RC control signal generators. The servo shaft will start at the start position and then rotates to the end position with the desired speed. In this research, the swing down speed is greater than the swing up speed which will be explained in the experiments.

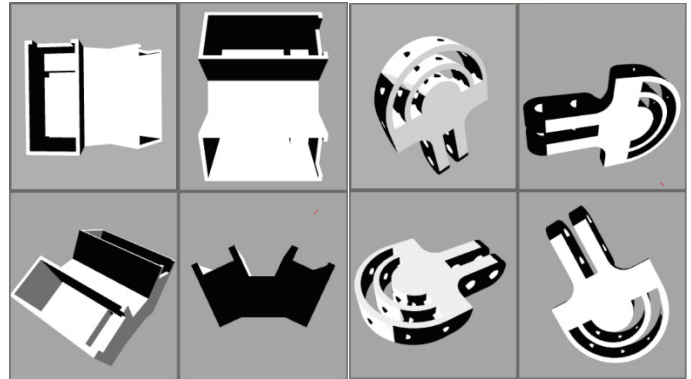


Figure 10: Servo holder (left) and Wing locker (right)

#### 4. Experiments

The experiments on the designed butterfly robot were conducted. No power system was installed into the robot body in this research in order to reduce robot payload then the robot system was wired to the external power source. The butterfly robot was installed into a sliding rod with a support platform in order to make the robot move freely when the wings are flapped.



Figure 11: Butterfly robot model (top) and manufactured butterfly robot (bottom)



Figure 12: KST MS325 high speed servo (left) and ARM controller (right)

In the first experiment, the robot flapping mechanism was tested. The robot mechanism could flap the wings simultaneously as the desired motion but two problems occurred which were bended wing and robot rotation situation. The reinforced wings were bended due to servo torques and air pressure as shown in Figure 12. To understand this behavior each wing has to be model as a flexible robot arm which was not the aim of the wing design proposed in this research. Although, the wing was based on the real non-symmetry wing butterfly wing but the number of wings on each side was reduced from 4 to 2 wings, yield, the extra aerodynamics force and wake vortex rotate the robot body as shown in the experiment. Therefore, wings have to be further reinforced by another layer of a carbon fiber rods structure with duct tape to make a wing rigid as much as possible. The wing was redesigned to be a symmetry wing to make the air pressure condense around the center of the wing's mass and reduce the rotation as shown in Figure 13. The modified wing was simulated in order to confirm the location of the air pressure. Then, the butterfly robot was tested again. The experiment result shows that two problems are solved as shown in Figure 14.



Figure 12: Bended wings situation

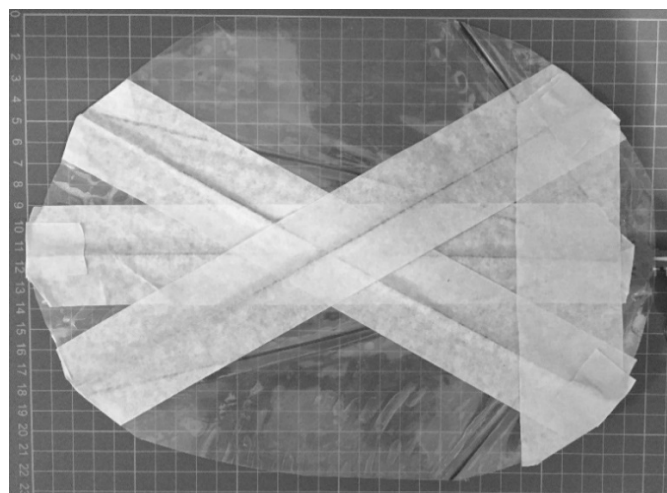


Figure 13: Redesigned and reinforced wing

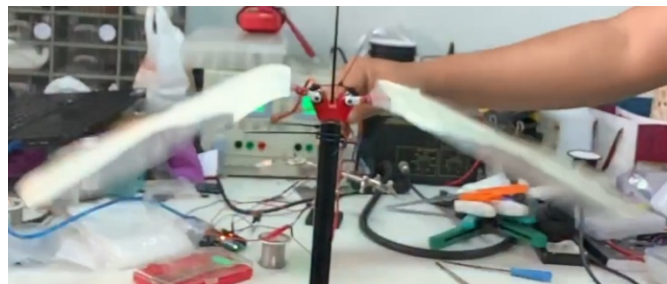


Figure 14: Unbended wings

Next, the butterfly robot was experimented. Servo motors were driven by the 50 Hz PWM signal with specified duty cycles depended on the wing locations. The driver circuit was installed to the robot power system since the ARM controller uses 3.3 volts logic while servo motors require 8.4 volts. In this experiment, the friction between the robot body and the sliding rod were neglected. The servo motor and wing lockers were printed by the 3D printer and the PLA plastic was selected as the printing material. Nuts and bolts were used to attach a wing locker to a servo horn. The total butterfly robot weight is 110 grams and the robot parameters are shown in table 1. The experiment was recorded by the Huawei P20 pro smartphone at 960 fps. In this research, the completed dynamics equation of the robot including the mass system was not considered. The coupling velocity effect and robot inertia also were neglected.

Table 1: Robot Parameters

	Part	weight
1	Wing with wing locker	25 g
2	KST servo motor with a plastic horn	23 g
3	Servos holder	9 g
4	Nuts and bolts	5 g

From the experiment, the butterfly robot can take-off from the support platform, fly up when wings swing down and fall back to the platform when wings swing up. It is clear that the robot wings can generate the average lift force is greater than 1.0791 N which is enough lift force to lift robot up from the support platform. The lift force can be simply estimated from the combination of the maximum torque generated from the servo motors acting at the center of the air pressure estimated from the simulation. The robot starts moving slowly from the platform then the speed is increasing and reaching the maximum acceleration when both wings are parallel to the ground and robot. Then, the robot starts moving slower until it stops at the maximum height as shown in Figure 15. The estimated lift force, if the coupling velocity effect and inertia is neglected, is about 5.4 N. At this point, the robot can reach the maximum body acceleration of 0.93 m/s. The robot can fly up to about 10 centimeters above the support platform in 1.02 seconds as shown in Figure 16. Although, in the experiment, the robot shows the success of take-off and fly up to certain height, but when the robot swings up wings, the robot drops down quickly because wing's downforce combined with the robot weight as shown in Figure 17. This situation makes the robot drops to the platform faster when wings swing up than the robot reaches the maximum height when wings swing down. This problem has a severe impact



if the butterfly carries payload such as battery and other electronic circuits. In order to overcome this problem, the wings motion is needed to be modified.

### 5. Conclusion and Future Work

A butterfly robot wing section's aerodynamics and aeroelastic are analyzed, estimated and derived in order to obtain a wing behavior. Lift and thrust are important for this research because they are critical of designing a robot wing. After wing section's lift and thrust forces were obtained, the average lift and thrust can be calculated by integrating all small wing section's forces in particular time over one flapping motion period. Navier-Stokes equations are used to simulated the air pressure acts on the wing when swings in the still air.

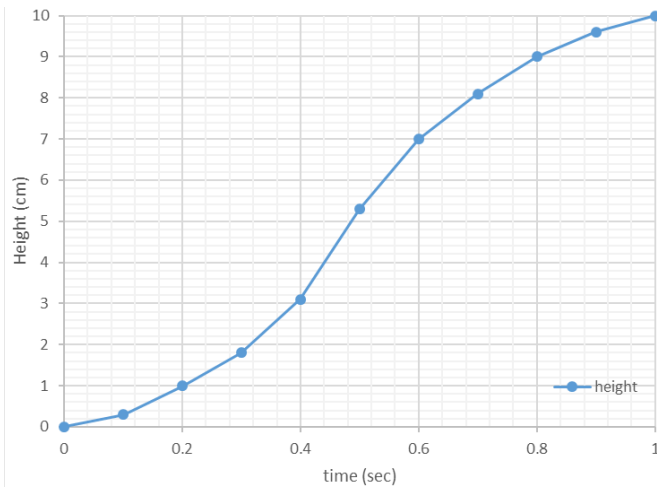


Figure 15: Height of the butterfly robot in the experiment.

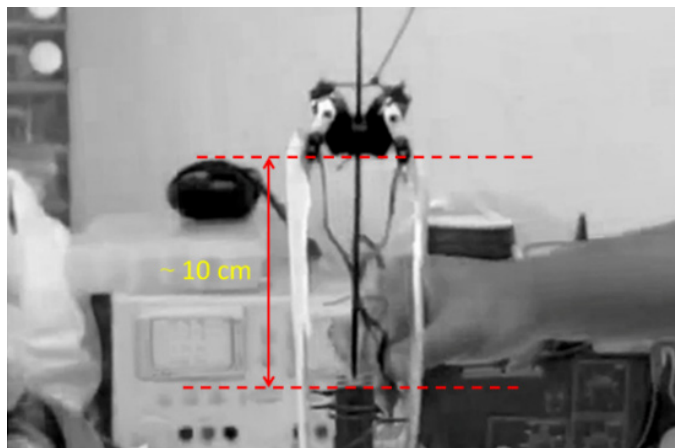


Figure 16: Maximum height.

The SIMPLER algorithm is used to obtain the numerical solution of Navier-Stokes equations. The real butterfly wing properties and dimension was used the reference for designing the desired robot wing. A laminar plastic sheet was used to mimic butterfly wing membranes properties while the reinforced fiber carbon rods was used to create wing wall structure like the scales and hairs in the butterfly wing. Thus, the designed wing was considered as a thin airfoil with laminar skin properties. A flapping mechanism was constructed from two servo motors to create flexible motions between two wings and enhance overall robot performance. The butterfly robot parts were designed and

manufactured. In the first experiment, the designed wing was bended occurred from the servo torques and air pressure due to the wing material. Then, the wing was redesigned to be a symmetry wing and reinforced to reduce a bend. The experiment shows that the butterfly robot can take off from the platform and fly up to a certain height before it goes back down to the platform.

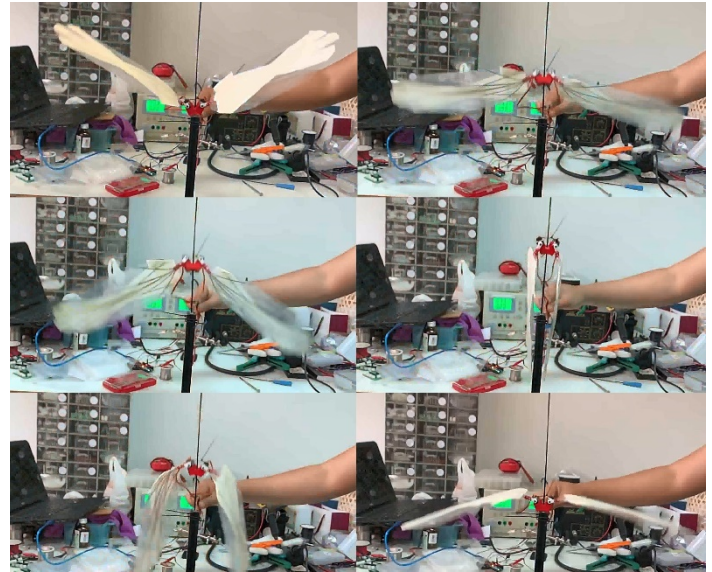


Figure 17: Experimental Result.

In this research, masses, inertias and the coupling velocity effect on the butterfly robot system were not considered. The standalone butterfly robot with payload will be designed and the completed butterfly robot dynamics will be derived using the Euler-Lagrange equation of motion. More experiments will be conducted in future work.

### Conflict of Interest

The authors declare no conflict of interest.

### References

- [1] K. Sukvichai, K. Yajai, "Aerodynamics Force Analysis for Designing a Flapping Butterfly Robot Wing" in International Conference on Electronics, Information, and Communication 2020 (ICEIC2020), Barcelona, Spain, 2020. <https://doi.org/10.1109/ICEIC49074.2020.9051389>
- [2] G. De Croon, K. De Clercq, R. Ruijsink, B. Remes, C. de Wagter, "Design, aerodynamics, and vision-based control of the delfly" International Journal of Micro Air Vehicles, **1** (2), 71–97, 2009. <https://doi.org/10.1260/175682909789498288>
- [3] M. Keennon, K. Klingebiel, H. Won, "Development of the nano hummingbird: A tailless flapping wing micro air vehicle," in 50th AIAA aerospace sciences meeting including the new horizons forum and aerospace exposition, Nashville, TN, USA, 2012. <https://doi.org/10.2514/6.2012-588>
- [4] J. Zhang, F. Fei, Z. Tu, X. Deng, "Design optimization and system integration of robotic hummingbird," in 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, Singapore, 2017. <https://doi.org/10.1109/ICRA.2017.7989639>
- [5] K. Y. Ma, P. Chirarattananon, S. B. Fuller, R. J. Wood, "Controlled flight of a biologically inspired, insect-scale robot" Science, **340**(6132), 603–607, 2013. <https://doi.org/10.1126/science.1231806>
- [6] R. T. Jones, The Unsteady Lift of a Wing of Finite Aspect Ratio, NACA Report No. 681, 1940
- [7] T. Theodorsen, General theory of aerodynamic instability and the mechanism of flutter, NACA Rept. 496, 1935.
- [8] J. O. Scherer, Experimental and Theoretical Investigation of Large Amplitude Oscillating Foil Propulsion Systems, Hydronautics Incorporated,

Research in Hydrodynamics, Technical Report 662-1, Laurel, Maryland, USA, 1968.

- [9] S. F. Hoerner, Fluid-Dynamic Drag: Theoretical, experimental and statistical information, Hoerner Fluid Dynamics; Second Edition Reprint edition , 1992.
- [10] L. C. Fefferman, “Existence and Soothness of the Navier–Stokes equation”, Millennium Prize Problems—Navier–Stokes Equation, Clay Mathematics Institute, March 27, 2017, retrieved 2017-04-02, [Online] Available: <http://www.claymath.org/sites/default/files/navierstokes.pdf> [Accessed: 7- Jul- 2020]
- [11] S. Hosseinzadeh1, R. Ostadhossein, H.R. Mirshahvalad, J. Seraj, “Using Simpler Algorithm for Cavity Flow Problem” Mechatronics and Applications: An International Journal (MECHATROJ), 1(1), 55–63, 2017. <https://doi.org/10.2139/ssrn.3438116>
- [12] Flowsquare+, Available: <https://fsp.norasci.com/> [Accessed: 7- Jul- 2020]



## **Clustering of Mindset towards Self-Regulated Learning of Undergraduate Students at the University of Phayao**

Pratya Nuankaew\*

*School of Information and Communication Technology, University of Phayao, Phayao, 56000, Thailand*

---

### **ARTICLE INFO**

*Article history:*

*Received: 22 June, 2020*

*Accepted: 11 August, 2020*

*Online: 25 August, 2020*

---

*Keywords:*

*Educational Data Mining*

*Learning Analytics Model*

*Learning Styles*

*Self-Regulated Learning*

---

### **ABSTRACT**

*The effects of Covid-19 severely affected the Thai higher education model. Therefore, there are three significant objectives in this research: (1) to cluster the mindsets and attitudes toward self-regulated learning styles of undergraduate students at the University of Phayao. (2) to construct a predictive model for recommending an appropriate student learning clusters. (3) to evaluate the predictive model that has been constructed. Samples collected a compilation of 472 student satisfaction with questionnaires from three schools, with seven disciplines at the University of Phayao, Thailand. Research tools consisted of statistical and machine learning techniques as follows: frequency, percentage, average, standard deviation, k-means clustering, decision tree techniques, cross-validation methods, confusion matrix performance, accuracy, precision, and recall measurement. Researcher found that the k-means model with the highest accuracy is the decision tree model that was classified into three clusters by dividing the model testing into the leave-one-out cross-validation method with a depth of seven levels of the decision tree model and an accuracy of 98.73%. From the results and studies, it can be concluded that the developed model is effective and reasonable to be further developed as an application for further organizational development.*

---

## **1. Introduction**

Nowadays, the learning behavior of youth and the new generation has changed dramatically which have made educational system unable to cope with acts that wants to inquire more about the issue than just following what is instructed. The results have created learning that is more aligned with the new generation who demand to know more and expand their interests [1]–[5]. The learning styles of the new generation of children are highly identified as having a limited attention, known as ADHD: Attention Deficit Hyperactivity Disorder [6], [7]. In addition, the device addiction and mobile addiction symptoms are more intensifying and widespread among youth and the younger generation [8].

These type of behaviors often lead into internet addiction [9]. In 2009, a medical study found that the average age of internet addicts was 17.6 years (range: 12-27 years), as the use of internet was nine hours a day and increasing in proportion [9]. In 2016,

there was a study of personality and positive orientation in Internet and Facebook addiction [10]. It has been found that age has a significant effect on factors distinguishing both Internet and Facebook addiction. Moreover, young people more often have problem with excessive use with the Internet and Facebook than adults [10]. Therefore, it can be summarized primarily that young people and the new generation are driven by the changes in technology, which has completely changed the learning behavior of students.

At the same time, the social patterns of the new generation have cause changes in the way of communication, making contact, and having a dialogue. Due to the addiction from the mobile phones and the internet, their habit has made them choose to be more associated with online communities rather than interacting with people in a normal society [11]. From the patterns and behaviors of the aforementioned young generation, the results of formal education, which is a basic education for everyone, are not consistent with the proper standards for student behavior of learning in a classroom environment [11], [12]. It is possible that the traditional teaching has been outdated for a long period of time thus presenting itself to be a problem in developing a healthy and

---

\*Pratya Nuankaew, School of Information and Communication Technology, University of Phayao, Phayao, 56000, Thailand, +66 89 961 4832 & [pratya.nu@up.ac.th](mailto:pratya.nu@up.ac.th)

sound method for students to make any real progress in learning, and acquiring the necessary employable skills and abilities after graduation. The best solution is to change the style of knowledge management and teaching strategies to suit students.

The learning theory that is consistent and suitable for solving the above problem is the theory of Self-Regulated Learning (SRL), which is a widely accepted theory [13]–[16]. Self-regulated learning strategies can be applied to learning in the new era. It can also be applied to the promotion of Technology Enhanced Learning (TEL), which provides opportunities to increase the learning skills necessary for students [11], [17]. In addition, an important principle of self-regulated strategies is the development of learning that aims learning towards achieving the goals set by the learners themselves [15].

However, learning can happen anywhere and anytime, with each person learning differently, because each day in life presents itself with different situations for exploring something new. Moreover, each event may have similarities or differences with an experience that results in the behavior of the learners who are learning. In addition, when humans learn and achieve academic achievement, the result is a change in learning behavior that is the result from past events or situations. But changing human behavior may not always be learning, due to changes in a certain period of time whereas the person may have to find the self-motivation to get themselves to take part in the learning process.

From the benefits of self-regulated learning strategies, the researchers can use this theory to solve problems and design the learning processes that are appropriate for learners in the new normal education system. This is of vital interest and persuades the researcher to conduct research. The background of the researcher's past work is the study of the behavior of learners at the tertiary [5], [18], [19] and secondary [20], [21] levels. In addition, the researcher is also interested in developing educational models in order to create a learning model that is truly suitable for the learners [22], [23]. The success and achievement that researchers have found is to support learners to achieve learning success and to combine academic achievement. These are the forces that support and continually motivate researchers to pursue our research.

### 1.1. Research Objectives

There are three significant objectives: The first objective is to cluster the mindsets and attitudes toward self-regulated learning styles of undergraduate students at the University of Phayao. The second objective is to construct a predictive model for recommending an appropriate student learning clusters. Lastly, the third objective is to evaluate the predictive model that has been constructed.

The expectation of this research is to know the impact of the Covid-19 pandemic, which has had a severe impact on the education model by studying the perception and attitudes towards online and online learning styles of learners at the University of Phayao. In addition, the expected result is knowing the group of learners according to the attitude and self-learning style which will be used for developing the quality and potential of students in the future.

### 1.2. Research Approach

The research approach has been conducted according to the process of the CRISP-DM methodology [24], [25]. It consists of six steps as follows: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The details of the research process are presented on the topic of research methodology. Data collection was carried out in a compilation of 472 student satisfaction with questionnaires from three schools, with seven disciplines at the University of Phayao which is stored on the website: <https://bit.ly/2BobB8l>.

The research consisted of statistical tools and machine learning tools as follows: percentage, mean, average, standard deviation, k-means clustering, decision tree techniques, cross-validation methods, confusion matrix performance, accuracy, precision, and recall measurement.

### 1.3. Research Ethic

This research has requested permission from the School of Information and Communication Technology, the University of Phayao, and related agencies, by implementing the regulations of the university.

## 2. Literature Reviews

### 2.1. Self-Regulated Learning

Self-regulated learning refers to the process of setting goals, controlling, and managing the sources of knowledge based on the motivation of the learners to set learning goals and expect success in intellectual learning [15], [17], [26]–[29].

The composition of the self-regulated learning consists of three important phases [13], [27], [29]. The first phase is the forethought phase, which consists of two important processes as follows: (1) task analysis, and (2) self-motivation beliefs. The second phase is the performance phase, which consists of two important processes as follows: (1) self-control, and (2) self-observation. Finally, the third phase is the self-reflection phase, which consists of two important processes as follows: (1) self-judgment, and (2) self-reaction. Details of the components of the Self-regulated learning are shown in Figure 1.

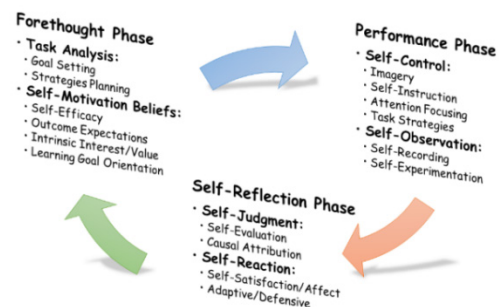


Figure 1: Phases and sub-process of self-regulated learning

From the theory of self-regulated learning, there is a lot of interest in research [13], [14], [17], [28]. It can therefore be concluded that this self-regulated learning theory is appropriate for this research.

## 2.2. Student Academic Performance

Student academic performance is a method of studying the effectiveness arising from graduation or learners receiving academic achievement. Many researchers have studied learning styles that encourage learners to receive high efficiency [5], [18], [30]–[32]. For example, there has been research works done to study the relationships of multiple variables that affect academic achievement using a method named “Structural Equation Modeling (SEM)” [30]. SEM is a type of statistical model that searches for and describes the relationships between multiple variables. The second example is research which studies the compatibility of different characteristics [32]. In the second example, they attempted to present the concept of the compatibility of the mentor and the receiver by comparing it to a jigsaw. The last example is research that studies the impact of unsuccessful studies or dropouts, in which these researches are discussed in many dimensions and perceptions of the researchers [19], [33].

There are also researchers who study the tools for applying the concepts for measuring and evaluating student performance [33]–[35]. Tools used in their studies include the use of statistical tools and data mining tools: basic statistical tools, decision tree techniques, k-means and k-medoids algorithms, confusion matrix performance, and cross-validation methods.

From the above example, it shows that various researchers clearly value the study of student academic performance. Therefore, it corresponds to the purpose of this research which aims to discover the pattern of relationships affecting graduation or non-graduation as specified by the curriculum.

## 2.3. Improving Academic Achievement

Generally, researchers in various fields have studied research and development in the science of educational quality development [5], [14], [36]–[38], in which the objectives of each research group are different in methods and perspectives. Some research groups want to study the factors that support and change learners’ behavior and instructional methods [3], [13], [31], [33]. Some research groups want to study and develop tools that support the efficiency of teaching and learning [35], [37], [39]. Some research groups want to apply modern technology to be used in teaching and learning activities [40], [41]. Finally, all research groups have a similar objective, which is to encourage learners to have a positive impact with their studies toward graduation.

However, the concept of improving academic achievement is aimed on raising the level of student achievement performance values [5], [15], [38], [42]. The main goal of this research should focus on the three main components: The first target is to cluster the characteristics of mindset towards self-regulated learning styles of undergraduate students at the University of Phayao. The second target is to construct the predictive model for suggesting the appropriate student cluster. The third target is to evaluate the predictive model that has been developed.

## 2.4. Educational Data Mining

Educational data mining (EDM) is the science of combining the use of data science tools and educational technology for

educational data analysis. It consists of machine learning applications, data mining tools and advanced statistics to carry out the process of educational system success [43]–[45]. In addition, the educational data mining refers to techniques, tools, and research designed to automate the definition of large data sources that are related to learning activities in the educational system [5], [19], [35], [45].

Examples of research in this field include research in application development to be recommended as suitable educational institutions for students [4], [35], [39]. They study and research on learners’ factors, educational institutions’ factors, educational data mining models that encourage learners to study in a suitable educational institution, and develop them into applications. In addition, there are other research studies such as Ahmad’s research which analyzes the educational model to see where the best fit are for a particular program [36]. They study the behavior of learners occurring in online activities through a learning management system known as MOOC. Firdausiah Mansur’s research [37] proposed a personalized learning model to find suitable learning methods based on a deep learning algorithm. Both of their results are impressive because they have applied machine learning applications, data mining tools, and advanced statistics which enabled them to discover some facts from the data.

Finally, it can be concluded that in the analysis of data mining for education it is necessary to find a new perspective that allows teachers to truly understand the learners.

## 3. Research Methodology

The research methodology was conducted according to the CRISP-DM method [24], [25]. It has six operations in accordance with the principles of CRISP-DM techniques, which include (1) business understanding, (2) data understanding, (3) data preparation, (4) modeling, (5) evaluation, and (6) deployment, as shown in Figure 2 below.

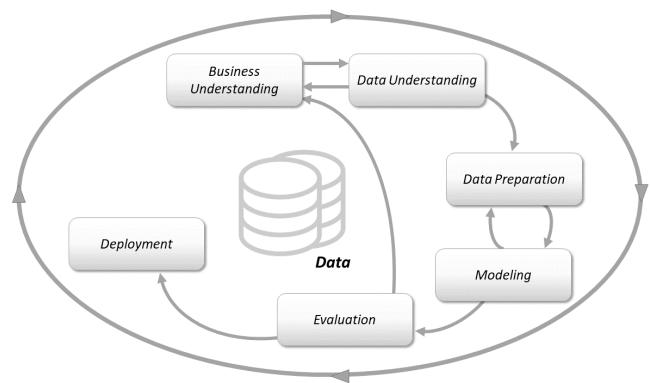


Figure 2: Research Methodology

### 3.1. Business Understanding

The business understanding phase focuses on understanding the audit objectives and needs from a case perspective and converting this knowledge into a definition of evidence mining problems and a preliminary plan designed to achieve the objectives [24], [25], [46]. For this reason, the researcher aimed

at the three objectives of the research as follows: The first objective is to cluster the characteristics of mindset towards self-regulated learning styles of undergraduate students at the University of Phayao. The second objective is to construct the predictive model for suggesting the appropriate student cluster. Finally, the third objective is to evaluate the predictive model that has been developed.

### 3.2. Data Understanding

The purpose of data understanding phase is to identify data quality problems, construct questions for finding patterns of data, and to examine interesting subset that create assumptions for hidden data [24], [25], [46]. The studied data was on the students of the University of Phayao, Phayao, Thailand. The data collected consisted of students' data from the School of Information and Communication Technology (ICT), the School of Management and Information Sciences (MIS), and the School of Law at the University of Phayao. In addition, the data collected were from seven disciplines of students majoring in the courses of accounting, business computer, law and accounting, management, marketing, and tourism. The purpose of the various data collection is to get information that covers the attitude of data providers.

### 3.3. Data Preparation

The data preparation phase aims to design activities that cover all activities to enable data collection for analysis and development of models. It details the sub-steps to prepare the data that will be fed into the research tool to create a model by initial management at raw data. It has five sub-steps to implement: selecting data, cleaning data, constructing data, integrating data, and formalizing data [24], [25], [46].

The data collected were on 472 students from the University of Phayao, Thailand. The data collection is divided into two categories according to the survey type. It consists of 319 students who responded from regular surveys and 153 people who answered from online surveys; details are shown in Table 1 to Table 6. In addition, the data collected in this research is defined in a digital format, which is stored on the website: <https://bit.ly/2BobB8l>.

Table 1: Data collected is classified by survey types

Survey Types	Absolute Count	Fraction (Percentage)
Regular Survey	319	67.58%
Online Survey	153	32.42%
<i>Total:</i>	<i>472</i>	<i>100.00%</i>

Table 1 shows that the regular survey has the largest number of respondents, with 319 students representing 67.58 percent of all respondents.

Table 2: data collected is classified by gender.

Gender	Absolute Count	Fraction (Percentage)
Male	112	23.73%
Female	360	76.27%
<i>Total:</i>	<i>472</i>	<i>100.00%</i>

Table 2 shows that Female have the largest number of respondents, with 360 students representing 76.27 percent of all respondents.

Table 3: Data collected is classified by disciplines.

Disciplines	Absolute Count	Fraction (Percentage)
Accounting	208	44.07%
Business Computer	73	15.47%
Law and Accounting	28	5.93%
Management	3	0.64%
Marketing	43	9.11%
Tourism	117	24.79%
<i>Total:</i>	<i>472</i>	<i>100.00%</i>

Table 3 shows that Accounting have the largest number of respondents, with 208 students representing 44.07 percent of all respondents. The second group that answered the most questionnaire was Tourism, with 117 students representing 24.79 percent.

Table 4: Data collected is classified by affiliation.

Affiliation	Absolute Count	Fraction (Percentage)
School of Information and Communication Technology (ICT)	73	15.47%
School of Law	28	5.93%
School of Management and Information Sciences (MIS)	371	78.60%
<i>Total:</i>	<i>472</i>	<i>100.00%</i>

Table 4 shows that the School of Management and Information Sciences (MIS) have the largest number of respondents, with 371 students representing 78.60 percent of all respondents.

Table 5: Data collected is classified by learning styles.

Learning Styles	Absolute Count	Fraction (Percentage)
Online Learning Style	81	17.16%
Offline Learning Style	391	82.84%
<i>Total:</i>	<i>472</i>	<i>100.00%</i>

Table 5 shows that the collected data is categorized according to the learning styles that students are interested in. Table 5 shows that students are interested in the traditional learning or in front of the classroom, with 391 students representing 82.84 percent of all respondents.

Table 6: Data collected is classified by acceptance of SRL style.

Acceptance of SRL style	Absolute Count	Fraction (Percentage)
Low Level (0-30%)	18	3.81%
Medium Level (31-70%)	418	88.56%
High Level (71-100%)	36	7.63%
<i>Total:</i>	<i>472</i>	<i>100.00%</i>



Table 6 shows that the level of acceptance to self-regulated learning style is at a medium level (31-70%), with 418 students representing 88.56 percent of all respondents.

### 3.4. Modeling

Various modeling techniques are selected and implemented, with their parameters being compared to the best values. In general, there are many techniques for the same type of data mining problem. Some techniques require a specific data format data [24], [25], [46]. In addition, modeling is the process of creating a suitable prototype. It consists of four important parts: selecting the modeling techniques, generating test design, building the model, and assessing the model [31], [35].

As mentioned above, the machine learning tools selected are k-means clustering and decision tree techniques. The benefit of the k-means is that it can be recommended for clustering with similar data patterns [31], while the benefit of the decision tree is that it is a structural decision consisting of nodes (features) and leaves (decisions) [35].

In this research, the analysis for clustering and charting of decision tree was based on data from questionnaires filled out by students who were assigned to take part in the research activities. The final result of the modeling is a set of variables (characteristics) that are important for predicting a reasonable cluster for the learners, which will be used to suggest activities for the learners in the next academic year.

### 3.5. Evaluation

The goal of the evaluation is to evaluate the performance of the results, which aims to construct the significant relationship models [19], [33]. The tools are used in the research, including the cross-validation techniques as shown in Figure 3, and the calculation of confusion matrix as mentioned in Figure 4.

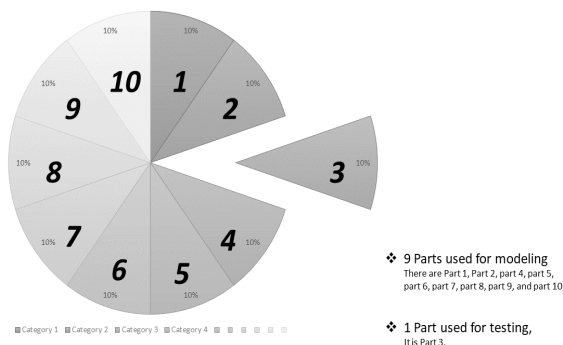


Figure 3: Cross-Validation Techniques

Figure 3 displays the separation of data for evaluating the model. The cross-validation method divides the data into two parts. The first part is used for modeling and the remainder is for testing the model. In addition, model evaluation requires a tool called a confusion matrix to test the model's performance, with the principles shown in Figure 4.

		Actual Class		Precision
		Positive	Negative	
Predicted Class	Positive	True Positive : TP	False Positive: FP (Type 1 Error)	Positive Predictive Value : $\frac{TP}{TP+FP}$
	Negative	False Negative : FN (Type 2 Error)	True Negative: TN	Negative Predictive Value : $\frac{TN}{TN+FN}$
Recall		Sensitivity : $\frac{TP}{TP+FN}$	Specificity : $\frac{TN}{TN+FP}$	

Figure 4: Confusion Matrix Calculation

Figure 4 presents the composition of the confusion matrix performance, which is composed of the actual class and the predicted class. An important benefit of the performance of the confusion matrix is the ability to determine the model's ability to predict results, such as the predictive ability or accuracy, model precision, model sensitivity, and model specificity (recall measurement). These values are used to determine the actual performance model. Moreover, Figure 4 also demonstrates the formulas and methods for calculating the various performance parameters in detail.

### 3.6. Deployment

The deployment is intended to bring results and discoveries in order to establish relationships and analyze the relationships that are discovered.

#### 3.6.1. Testing Model Results

As mentioned early, the tools used to evaluate the model consist of two parts: cross-validation method and confusion matrix performance. This section presents the use of evaluation tools in research. The testing process divides the data into two parts according to the cross-validation method principles.

There are three methods of cross-validation in this research. The first method is 10-Fold cross-validation, which used 9-Fold for modeling and 1-Fold for testing. The second method is 50-Fold cross-validation, which used 49-Fold for modeling and 1-Fold for testing. The last method is leave-one-out cross-validation, which used 99 percent of data for modeling and 1 percent of data for testing. However, each time the cross-validation test is reported, the model results are also tested by using the confusion matrix performance.

#### 3.6.2. Applying Model Results

The purpose of applying model results is to manage the developed models. It has four sections as follows: (1) decision tree model, (2) decision tree model applying results, (3) cluster model, and (4) number of members in each cluster.

## 4. Research Results

In the research results, the researchers classified the research report into four parts which are satisfaction level towards questionnaire, modeling results, model testing results, and model applying results.

### 4.1. Satisfaction level towards questionnaire

This section summarizes the satisfaction levels of the 472 students, which contain data from three schools and seven

disciplines at the University of Phayao. The details of the summary results are shown in Table 8.

In interpreting the data according to the characterization criteria, the interpretation is based on a five-level interpretation method by comparing it with the criteria that divides the level estimation into five equal levels, as followed in Equation (1). The result of the calculation is shown in Equation (2).

$$\text{Width of the level} = \frac{\text{Maximum} - \text{Minimum}}{\text{Number of levels}} \quad (1)$$

$$\text{Width of the level} = \frac{5 - 1}{5} = 0.8 \quad (2)$$

From the calculation results in Equation (2), the interpretation results can be specified as shown in Table 7.

Table 7: Interpretation and Meaning

Width of the level	Interpretation	Meaning
1.00 – 1.80	Highly Unacceptable	The lowest level of satisfaction
1.81 – 2.60	Unacceptable	Low level of satisfaction
2.61 – 3.40	Acceptable	Satisfaction
3.41 – 4.20	Highly Acceptable	High level of satisfaction
4.21 – 5.00	Maximum Acceptable	The highest level of satisfaction

Table 8: Satisfaction level towards questionnaire

Question	Average	S.D.	Interpretation
<i>Question 1:</i> What level of self-regulated learning can you apply to your studies?	3.38	0.64	Highly Acceptable
<i>Question 2:</i> What level do you agree that self-regulated learning is important for online learning?	3.58	0.78	Highly Acceptable
<i>Question 3:</i> What level do you agree that finding your own knowledge can be more useful than studying in the classroom?	3.45	0.81	Highly Acceptable
<i>Question 4:</i> What level do you agree that inspiration and motivation have more influence on learning than the advice of teachers?	3.73	0.56	Highly Acceptable
<i>Total Average:</i>	3.54	0.79	Highly Acceptable

Table 8 shows the level of satisfaction with the four questionnaires. It can be seen that the overall level of satisfaction is high (3.54). It can therefore be concluded that the students of the University of Phayao are satisfied with the high level of self-regulated learning styles.

#### 4.2. Modeling Results

Modeling results are the various models on different criteria, such as defining the unequal depth of the decision tree, and determining the different types of cross-validation method tests, which have the results shown in Table 9 and Table 10.

Table 9: Modeling results from k-Means clustering

Cluster Number	Depth of Decision Tree	Cross-validation method		
		10-Fold	50-Fold	Leave-one-out
3 Clusters	Level 2	69.06%	69.09%	69.07%
	Level 3	89.38%	89.47%	89.41%
	Level 4	90.65%	91.60%	91.53%
	Level 5	92.56%	92.40%	92.37%
	Level 6	95.34%	95.73%	95.55%
	Level 7*	98.30%	98.69%	98.73%*
	Level 8	98.30%	98.69%	98.73%
4 Clusters	Level 2	59.12%	59.22%	59.11%
	Level 3	78.40%	78.98%	78.81%
	Level 4	87.08%	84.22%	82.63%
	Level 5	91.52%	92.73%	91.10%
	Level 6	93.01%	92.07%	89.41%
	Level 7*	95.55%*	95.00%	92.58%
	Level 8	95.55%	95.00%	92.58%
5 Clusters	Level 2	57.42%	57.51%	57.42%
	Level 3	80.93%	81.00%	80.93%
	Level 4	95.35%	95.62%	95.55%
	Level 5*	96.83%	97.07%*	97.03%
	Level 6	96.83%	97.07%	97.03%
	Level 7	96.83%	97.07%	97.03%
	Level 8	96.83%	97.07%	97.03%
6 Clusters	Level 2	53.61%	53.62%	53.60%
	Level 3	56.34%	58.16%	58.05%
	Level 4	77.12%	77.84%	77.97%
	Level 5	86.24%	85.51%	84.75%
	Level 6	93.21%	93.71%	93.64%
	Level 7*	97.24%	97.71%	97.67%*
	Level 8	97.24%	97.71%	97.67%
7 Clusters*	Level 2	45.77%	46.24%	46.19%
	Level 3	56.16%	56.18%	56.14%
	Level 4	71.42%	73.53%	77.12%
	Level 5	82.23%	85.69%	86.65%
	Level 6	88.58%	90.36%	89.83%
	Level 7	94.29%	94.13%	94.28%
	Level 8	96.20%	95.24%	95.13%
Level 9*	97.90%*	97.00%	97.03%	

From Table 9, it shows that the k-means model with the highest accuracy is the decision tree model that is classified into 3 clusters by dividing the model testing into the leave-one-out cross-validation method with a depth of 7 levels of the decision tree model and has an accuracy of 98.73%. In Addition, the testing results classified by other clusters yield a lower accuracy. For example, the 4 clusters with the highest results are 95.55%, the 5 clusters with the highest results are 97.07%, the 6 clusters with the highest results are 97.67%, and the 7 clusters with the highest results are 97.90%.

#### 4.3. Model Testing Results

From the results of the prototype model development, it can be concluded that the model with the highest accuracy is the development of the model from k-means clustering, with the appropriate number of 3 clusters and the leave-one-out cross-validation result with an accuracy of 98.73%. Details of the developed model are shown in Table 10.

Table 10: Model testing results

Precision Class	Actual Class							Recall Class
	True Cluster_0	True Cluster_1	True Cluster_2	True Cluster_3	True Cluster_4	True Cluster_5	True Cluster_6	
97.37%	111	0	0	0	1	1	1	36
97.50%	0	39	1	0	0	0	0	1
100.00%	0	0	85	0	0	0	0	22
98.37%	0	0	1	121	1	0	0	1
97.96%	0	1	0	0	48	0	0	0
91.67%	0	2	0	0	0	0	0	0
97.30%	0	0	0	0	0	0	36	1
	100.00%	92.86%	97.70%	100.00%	96.00%	91.67%	97.30%	
	Accuracy: 97.90% (+/- 2.41%)							
	Predicted Class							

Table 10 shows that the model performance testing is at the highest level, which can confirm the suitability of the model as well.

#### 4.4. Model Applying Results

From the model that has been selected and demonstrated the performance, this section presents the decision tree model as shown in Table 11, the decision tree rules model for self-testing as shown in the test results in Table 12, the cluster model as shown in Table 13, and number of members in each cluster as shown in Table 14.

Table 11: Decision tree model

Decision Tree Model
Question 2 > 3.500
Question 4 > 2.500
Question 3 > 3.500
Question 4 > 3.500
Question 1 > 4.500
Question 4 > 4.500: cluster_1 {cluster_1=11}
Question 4 ≤ 4.500: cluster_5 {cluster_1=1, cluster_5=2}
Question 1 ≤ 4.500
Question 2 > 4.500
Question 1 > 3.500: cluster_1 {cluster_1=20}
Question 1 ≤ 3.500
Question 3 > 4.500: cluster_1 {cluster_1=3}
Question 3 ≤ 4.500: cluster_0 {cluster_0=7}
Question 2 ≤ 4.500
Question 3 > 4.500
Question 1 > 3.500: cluster_1 {cluster_1=6}
Question 1 ≤ 3.500: cluster_0 {cluster_0=6}
Question 3 ≤ 4.500: cluster_0 {cluster_0=71}
Question 4 ≤ 3.500
Question 3 > 4.500: cluster_4 {cluster_1=1, cluster_4=3}
Question 3 ≤ 4.500
Question 1 > 3.500: cluster_2 {cluster_2=10}
Question 1 ≤ 3.500: cluster_4 {cluster_2=1, cluster_4=3}
Question 3 ≤ 3.500
Question 4 > 4.500: cluster_0 {cluster_0=17, cluster_6=1, cluster_5=1}
Question 4 ≤ 4.500: cluster_2 {cluster_2=75}
Question 4 ≤ 2.500
Question 3 > 3.500: cluster_2 {cluster_2=1, cluster_4=1}
Question 3 ≤ 3.500: cluster_3 {cluster_3=4}
Question 2 ≤ 3.500
Question 3 > 3.500
Question 1 > 3.500
Question 4 > 3.500: cluster_5 {cluster_5=20}
Question 4 ≤ 3.500: cluster_4 {cluster_4=7}
Question 1 ≤ 3.500
Question 4 > 4.500: cluster_0 {cluster_0=10, cluster_4=1}
Question 4 ≤ 4.500: cluster_4 {cluster_4=35}
Question 3 ≤ 3.500
Question 4 > 3.500: cluster_6 {cluster_6=36, cluster_5=1}
Question 4 ≤ 3.500: cluster_3 {cluster_3=117}

Table 11 shows the decision tree model. It can be developed into a decision tree rules model, as shown in Table 12. Decision tree rules model, it is used to test models using the data collected in the developed model tests.

Table 12: Decision tree rules model applying results

Rule	Condition (If)	Prediction (Then)
1	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 > 3.5 and Question 1 > 4.5 and Question 4 > 4.5	Then, suitable for cluster_1 = 100%
2	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 > 3.5 and Question 1 > 4.5 and Question 4 ≤ 4.5	then cluster_5 (0 / 1 / 0 / 0 / 0 / 2 / 0)
3	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 > 3.5 and Question 1 ≤ 4.5 and Question 2 > 4.5 and Question 1 > 3.5	Then, suitable for cluster_1 = 100%
4	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 > 3.5 and Question 1	Then, suitable for cluster_1 = 100%

Rule	Condition (If)	Prediction (Then)
	$\leq 4.5$ and Question 2 $> 4.5$ and Question 1 $\leq 3.5$ and Question 3 $> 4.5$	
5	If Question 2 $> 3.5$ and Question 4 $> 2.5$ and Question 3 $> 3.5$ and Question 4 $> 3.5$ and Question 1 $\leq 4.5$ and Question 2 $> 4.5$ and Question 1 $\leq 3.5$ and Question 3 $\leq 4.5$	Then, suitable for cluster_0 = 100%
6	If Question 2 $> 3.5$ and Question 4 $> 2.5$ and Question 3 $> 3.5$ and Question 4 $> 3.5$ and Question 1 $\leq 4.5$ and Question 2 $\leq 4.5$ and Question 3 $> 4.5$ and Question 1 $> 3.5$	Then, suitable for cluster_1 = 100%
7	If Question 2 $> 3.5$ and Question 4 $> 2.5$ and Question 3 $> 3.5$ and Question 4 $> 3.5$ and Question 1 $\leq 4.5$ and Question 2 $\leq 4.5$ and Question 3 $> 4.5$ and Question 1 $\leq 3.5$	Then, suitable for cluster_0 = 100%
8	If Question 2 $> 3.5$ and Question 4 $> 2.5$ and Question 3 $> 3.5$ and Question 4 $> 3.5$ and Question 1 $\leq 4.5$ and Question 2 $\leq 4.5$ and Question 3 $\leq 4.5$	Then, suitable for cluster_0 = 100%
9	If Question 2 $> 3.5$ and Question 4 $> 2.5$ and Question 3 $> 3.5$ and Question 4 $\leq 3.5$ and Question 3 $> 4.5$	Then, suitable for cluster_1 = 25%, and suitable for cluster_4 = 75%
10	If Question 2 $> 3.5$ and Question 4 $> 2.5$ and Question 3 $> 3.5$ and Question 4 $\leq 3.5$ and Question 3 $\leq 4.5$ and Question 1 $> 3.5$	Then, suitable for cluster_2 = 100%
11	If Question 2 $> 3.5$ and Question 4 $> 2.5$ and Question 3 $> 3.5$ and Question 4 $\leq 3.5$ and Question 3 $\leq 4.5$ and Question 1 $\leq 3.5$	Then, suitable for cluster_2 = 25%, and suitable for cluster_4 = 75%
12	If Question 2 $> 3.5$ and Question 4 $> 2.5$ and Question 3 $\leq 3.5$ and Question 4 $> 4.5$	Then, suitable for cluster_0 = 89.47%, suitable for cluster_5 = 5.26%, and suitable for cluster_6 = 5.26%
13	If Question 2 $> 3.5$ and Question 4 $> 2.5$ and Question 3 $\leq 3.5$ and Question 4 $\leq 4.5$	Then, suitable for cluster_2 = 100%
14	If Question 2 $> 3.5$ and Question 4 $\leq 2.5$ and Question 3 $> 3.5$	Then, suitable for cluster_2 = 50%, and suitable for cluster_4 = 50%
15	If Question 2 $> 3.5$ and Question 4 $\leq 2.5$ and Question 3 $\leq 3.5$	Then, suitable for cluster_3 = 100%
16	if Question 2 $\leq 3.5$ and Question 3 $> 3.5$ and Question 1 $> 3.5$ and Question 4 $> 3.5$	Then, suitable for cluster_5 = 100%
17	if Question 2 $\leq 3.5$ and Question 3 $> 3.5$ and Question 1 $> 3.5$ and Question 4 $\leq 3.5$	Then, suitable for cluster_4 = 100%
18	if Question 2 $\leq 3.5$ and Question 3 $> 3.5$ and Question 1 $\leq 3.5$ and Question 4 $> 4.5$	Then, suitable for cluster_0 = 90.90%, and suitable for cluster_4 = 9.10%
19	if Question 2 $\leq 3.5$ and Question 3 $> 3.5$ and Question 1 $\leq 3.5$ and Question 4 $\leq 4.5$	Then, suitable for cluster_4 = 100%
20	if Question 2 $\leq 3.5$ and Question 3 $\leq 3.5$ and Question 4 $> 3.5$	Then, suitable for cluster_6 = 97.30% Then, suitable for cluster_5 = 3.70%
21	if Question 2 $\leq 3.5$ and Question 3 $\leq 3.5$ and Question 4 $\leq 3.5$ then	Then, suitable for cluster_3 = 100%

Correct: 464 out of 472 training examples (98.31%)

Table 12 shows the decision tree rules model. It consists of 21 rules. The test results show that the rules can be predicted at a high level, which can accurately predict 464 data sets from a total of 472 data (98.31%). It can therefore be concluded that the developed models are appropriate by showing details of each cluster in Table 13.

Table 13: Cluster model

Cluster / Attribute	Question 1	Question 2	Question 3	Question 4
Cluster_0	3.42	4.00	3.91	4.47
Cluster_1	4.24	4.76	4.67	4.71
Cluster_2	3.39	4.18	3.02	3.46
Cluster_3	3.00	2.90	2.72	2.81
Cluster_4	3.14	3.10	4.12	3.36
Cluster_5	4.25	3.08	4.04	4.33
Cluster_6	3.22	2.78	2.84	4.19

Table 13 shows details of each cluster. In addition, Table 14 shows the number of members in each cluster.

Table 14: Number of members in each cluster

Cluster	Absolute Count (Member)	Fraction (Percentage)
Cluster_0	111	23.52%
Cluster_1	42	8.90%
Cluster_2	87	18.43%
Cluster_3	121	25.64%
Cluster_4	50	10.59%
Cluster_5	24	5.08%
Total:	472	100.00%

From Table 9 to Table 10, it details of testing and selection of suitable models. Moreover, Tables 11 to Table 14 are showing the members and the centroid of each cluster. Finally, based on the data and the results of this research it can be concluded that the developed model is very suitable for this study.

## 5. Research Discussion

In this research, the researcher has divided the discussion process into two sections as follows: the first section is the discussion report of data collection, and the second section is the discussion of the model, testing results and model effectiveness.

### 5.1. Data Collection Discussion

Based on the summary of the research data collection report from the University of Phayao there were 472 students divided into two groups as follows: the first group is to collect data directly (regular surveys collection) with 319 students. The second group is 153 students from online data collection.

From the data gathered, it can be concluded that this data is small. The researchers should expand the results of the study and gather more data for further analysis in order to comply with data mining principles that require large amounts of data.



## 5.2. Identify the Headings

Table 9 to Table 14 show the model development process by presenting the analysis results for selecting the model in Table 9, which was done by showing the model performance test in Table 10 and testing the model by data collection in Table 12. It can be concluded that the model is effective and accepted by this research. In addition, Table 13 and Table 14 show the clustering and membership model for each cluster. It can be concluded that each cluster has members distributed appropriately.

## 6. Conclusion

This research achieved three objectives as follows: The first objective is to cluster the mindsets and attitudes toward self-regulated learning styles of undergraduate students at the University of Phayao. The second objective is to construct a predictive model for recommending an appropriate student learning clusters. Lastly, the third objective is to evaluate the predictive model that has been constructed. Data collection is a compilation of 472 student satisfaction with questionnaires from three schools, and majoring in seven disciplines at the University of Phayao, Thailand. The data consisted of students from the School of Information and Communication Technology (ICT), the School of Management and Information Sciences (MIS), and the School of Law.

The research consisted of statistical tools and machine learning tools as follows: percentage, mean, average, standard deviation, k-means clustering, decision tree techniques, cross-validation methods, confusion matrix performance, accuracy measurement, precision measurement, and recall measurement. The results of the research found that the k-means model with the highest accuracy is the decision tree model that is classified into three clusters by dividing the model testing into the leave-one-out cross-validation method with a depth of seven levels of the decision tree model and has an accuracy of 98.73%. From the results and studies, it can be concluded that the developed model is effective and reasonable to be further developed as an application for further organizational development.

For future studies, the researchers found that the results of this research could be clearly further enhanced. For example, the researcher could use results to cluster learners' instruction based on their perceptions and attitudes towards managing online instruction based on Self-Regulated Learning theory. In addition, the researcher can use the results of the research into a computer program to facilitate the learners and teachers to use the results of the research in a useful way.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

This research is part of a study on "The Study of Trends and Impacts of Disruptive Technology on the Education System at the School of Information and Communication Technology, the University of Phayao". This research is supported by the University of Phayao and Rajabhat Mahasarakham University. The authors would like to thank Mrs. Wongpanya Nuankaew, the

advisor, lecturers, students, technicians, and all respondents for their entire support.

## References

- [1] P. Olivos, A. Santos, S. Martín, M. Cañas, E. Gómez-Lázaro, Y. Maya, "The relationship between learning styles, motivation to transfer of learning in a vocational training programme" *Suma Psicológica*, **23**(1), 25–32, 2016. <http://doi.org/10.1016/j.sumpsi.2016.02.001>
- [2] A. Viloría, J. R. López, K. Payares, C. Vargas-Mercado, S. E. Duran, H. Hernández-Palma, M. A. David, "Determinating Student Interactions in a Virtual Learning Environment Using Data Mining" *Procedia Comput. Sci.*, **155**, 587–592, 2019. <http://doi.org/10.1016/j.procs.2019.08.082>
- [3] A. I. Adekitan, O. Salau, "The impact of engineering students' performance in the first three years on their graduation result using educational data mining" *Heliyon*, **5**(2), 2019. <http://doi.org/10.1016/j.heliyon.2019.e01250>
- [4] P. Nuankaew, W. Nuankaew, T. Thamma, "The Recommended System for the Relationship between Educational Programs, Students' Interests" in *International Conference on Digital Arts, Media, Technology*, 2-3, 2016
- [5] P. Nuankaew, W. Nuankaew, K. Phanniphong, S. Imwut, S. Bussaman, "Students Model in Different Learning Styles of Academic Achievement at the University of Phayao, Thailand" *Int. J. Emerg. Technol. Learn. IJET*, **14**(12), 133–157, 2019.
- [6] D. P. Cantwell, L. Baker, "Association between attention deficit-hyperactivity disorder, learning disorders" *J. Learn. Disabil.*, **24**(2), 88–95, 1991.
- [7] P. N. Pastor, C. A. "Reuben, Diagnosed Attention Deficit Hyperactivity Disorder and Learning Disability: United States, 2004-2006. Data from the National Health Interview Survey" *Vital and Health Statistics* **10**(237), 1-14, 2008
- [8] M. Salehan, A. Negahban, "Social networking on smartphones: When mobile phones become addictive" *Comput. Hum. Behav.*, **29**(6), 2632–2639, 2013. <http://doi.org/10.1016/j.chb.2013.07.003>
- [9] R. Tao, X. Huang, J. Wang, H. Zhang, Y. Zhang, M. Li, "Proposed diagnostic criteria for internet addiction" *Addiction*, **105**(3), 556–564, 2010. <http://doi.org/10.1111/j.1360-0443.2009.02828.x>
- [10] A. Błachnio, A. Przepiorka, "Personality, positive orientation in Internet, Facebook addiction. An empirical report from Poland" *Comput. Hum. Behav.*, **59**, 230–236, 2016. <http://doi.org/10.1016/j.chb.2016.02.018>
- [11] S. K. Sharma, A. Joshi, H. Sharma, "A multi-analytical approach to predict the Facebook usage in higher education" *Comput. Hum. Behav.*, **55**, 340–353, 2016. <http://doi.org/10.1016/j.chb.2015.09.020>
- [12] M. Ozkan, B. Solmaz, "Mobile Addiction of Generation Z, its Effects on their Social Lives: (An Application among University Students in the 18-23 Age Group)" *Procedia - Soc. Behav. Sci.*, **205**, 92–98, Oct. 2015. <http://doi.org/10.1016/j.sbspro.2015.09.027>
- [13] W. S. Nuankaew, P. Nuankaew, D. Teeraputon, K. Phanniphong, S. Bussaman, "Perception, Attitude Toward Self-Regulated Learning of Thailand's Students in Educational Data Mining Perspective" *Int. J. Emerg. Technol. Learn. IJET*, **14**(09), 2019.
- [14] P. J. Gaskill, A. Woolfolk Hoy, "Self-efficacy and self-regulated learning: The dynamic duo in school performance. In J. Aronson (Ed.) *Improving academic achievement: Impact of psychological factors on education*, 185–208, 2002. <https://doi.org/10.1016/B978-012064455-1/50012-9>
- [15] B. J. Zimmerman, "Models of self-regulated learning, academic achievement" *Self-regulated learning, academic achievement*, Springer, 1–25, 1989.
- [16] P. R. Pintrich, "Understanding self-regulated learning" *New Dir. Teach. Learn.*, **63**, 3–12, 1995.
- [17] H. C. Schmitz, A. Mikroyannidis, T. Connolly, L. C. E. Law, H. Vieritz, A. Nussbaumer, M. Berthold, C. Ullrich, A. Dhir, "Self-regulated learning in formal education: perceptions, challenges, opportunities" *Int. J. Technol. Enhanc. Learn.*, **6**(2), 145–163, 2014.
- [18] P. Nuankaew, W. Nuankaew, S. Bussaman, P. Jedeejit, "Education mining in the relationship between general knowledge, deep knowledge for lifelong learning" in *2017 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications, Information Technology (ECTI-CON)*, 694–697, 2017. <http://doi.org/10.1109/ECTICon.2017.8096333>
- [19] P. Nuankaew, "Dropout Situation of Business Computer Students, University of Phayao" *Int. J. Emerg. Technol. Learn. IJET*, **14**(19), 2019.
- [20] K. Pupara, W. Nuankaew, P. Nuankaew, "An institution recommender system based on student context, educational institution in a mobile environment" in *2016 International Computer Science, Engineering*

- Conference (ICSEC), 1–6, 2016. <http://doi.org/10.1109/ICSEC.2016.7859877>
- [21] S. Bussaman, W. Nuankaew, P. Nuankaew, N. Rachata, K. Phanniphong, P. Jeedejit, “Prediction models of learning strategies, learning achievement for lifelong learning” in 2017 IEEE 6th International Conference on Teaching, Assessment, Learning for Engineering (TALE), 192–197, 2017. <http://doi.org/10.1109/TALE.2017.8252332>
- [22] P. Nuankaew, W. Nuankaew, P. Temdee, “Institution recommendation using relationship optimisation between program, student context” *Int. J. High. Educ. Sustain.*, **2**(4), 279–302, 2019.
- [23] P. Nuankaew, W. Nuankaew, T. Thamma, “The Recommended System for the Relationship between Educational Programs, Students’ Interests” in International Conference on Digital Arts, Media, Technology, 2-3, 2016.
- [24] R. Wirth, J. Hipp, “CRISP-DM: Towards a standard process model for data mining” in Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery, Data Mining, 29–39, 2000.
- [25] J. Venter, A. de Waal, C. Willers, “Specializing CRISP-DM for Evidence Mining” in *Advances in Digital Forensics III*, New York, USA, 303–315, 2007.
- [26] B. L. McCombs, “Self-regulated learning, academic achievement: A phenomenological view” *Self-regulated learning, academic achievement*, Springer, 51–82, 1989.
- [27] B. J. Zimmerman, “Self-regulated learning, academic achievement: An overview” *Educ. Psychol.*, **25**(1), 3–17, 1990.
- [28] B. J. Zimmerman, D. H. Schunk, “Self-regulated learning, academic achievement: Theoretical perspectives” Routledge, 2001.
- [29] B. J. Zimmerman, “Becoming a self-regulated learner: An overview” *Theory Pract.*, **41**(2), 64–70, 2002.
- [30] E. Shahzadi, Z. Ahmad, “A study on academic performance of university students” *Recent Adv. Stat.*, 255, 2011.
- [31] P. Nuankaew, P. Temdee, “Matching of compatible different attributes for compatibility of members, groups” *Int. J. Mob. Learn. Organ.*, **13**(1), 4–29, 2019.
- [32] P. Nuankaew, P. Temdee, “Determining of compatible different attributes for online mentoring model” in 2014 4th International Conference on Wireless Communications, Vehicular Technology, Information Theory, Aerospace & Electronic Systems (VITAE), 1–5, 2014.
- [33] P. Nuankaew, W. Nuankaew, K. Phanniphong, R. Fooprateepsiri, S. Bussaman, “Analysis Dropout Situation of Business Computer Students at University of Phayao” in *The Impact of the 4th Industrial Revolution on Engineering Education*, Cham, 2020, 419–432. [http://10.1007/978-3-030-40274-7\\_42](http://10.1007/978-3-030-40274-7_42)
- [34] A. M. Shahiri, W. Husain, N. A. Rashid, “A Review on Predicting Student’s Performance Using Data Mining Techniques” *Procedia Comput. Sci.*, **72**, 414–422, 2015. <http://doi.org/10.1016/j.procs.2015.12.157>
- [35] W. Nuankaew, P. Nuankaew, “The Study of the Factors, Development of Educational Model: The Relationship between the Learner Context, the Curriculum Context in Higher Education” *Int. J. Emerg. Technol. Learn. IJET*, **14**(21), 2019.
- [36] N. B. Ahmad, U. F. Alias, N. Mohamad, N. Yusof, “Principal Component Analysis, Self-Organizing Map Clustering for Student Browsing Behaviour Analysis” *Procedia Comput. Sci.*, **163**, 550–559, 2019. <http://doi.org/10.1016/j.procs.2019.12.137>
- [37] A. B. Firdausiah Mansur, N. Yusof, A. H. Basori, “Personalized Learning Model based on Deep Learning Algorithm for Student Behaviour Analytic” *Procedia Comput. Sci.*, **163**, 125–133, 2019. <http://doi.org/10.1016/j.procs.2019.12.094>
- [38] E. Aronson, “Chapter 10 - Building Empathy, Compassion, Achievement in the Jigsaw Classroom” *Improving Academic Achievement*, J. Aronson, Ed. San Diego: Academic Press, 209–225, 2002.
- [39] K. Pupara, W. Nuankaew, P. Nuankaew, “An institution recommender system based on student context, educational institution in a mobile environment” in 2016 International Computer Science, Engineering Conference (ICSEC), 1–6, 2016.
- [40] J. A. Meirink, P. C. Meijer, N. Verloop, T. C. Bergen, “Understanding teacher learning in secondary education: The relations of teacher activities to changed beliefs about teaching, learning” *Teach. Teach. Educ.*, **25**(1), 89–100, 2009.
- [41] G. Hampden-Thompson, J. Bennett, “Science teaching, learning activities, students’ engagement in science” *Int. J. Sci. Educ.*, **35**(8), 1325–1343, 2013.
- [42] W. Nuankaew, P. Nuankaew, S. Bussaman, P. Tanasirathum, “Hidden academic relationship between academic achievement, higher education institutions” in 2017 International Conference on Digital Arts, Media, Technology (ICDAMT), 308–313, 2017. <http://doi.org/10.1109/ICDAMT.2017.7904982>
- [43] C. Romero, S. Ventura, “Educational data mining: a review of the state of the art” *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, **40**(6), 601–618, 2010.
- [44] S. K. Mohamad, Z. Tasir, “Educational data mining: A review” *Procedia-Soc. Behav. Sci.*, **97**(2013), 320–324, 2013.
- [45] S. Slater, S. Joksimović, V. Kovanovic, R. S. Baker, D. Gasevic, “Tools for educational data mining: A review” *J. Educ. Behav. Stat.*, **42**(1), 85–106, 2017.
- [46] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, Reinartz, T., C. Shearer, R. Wirth, “CRISP-DM 1.0 Step-by-step data mining guide” SPSS inc, **9**, 13, 2000.

# Fuzzy Recognition by Logic-Predicate Network

Tatiana Kosovskaya\*

Department of Mathematics and Mechanics, St. Petersburg State University, St. Petersburg, 199034, Russia

---

## ARTICLE INFO

*Article history:**Received: 28 May, 2020**Accepted: 04 August, 2020**Online: 28 August, 2020*

---

*Keywords:**Predicate calculus**logic-predicate approach to AI**predicate network**fuzzy predicate network*

---

---

## ABSTRACT

The paper presents a description and justification of the correctness of fuzzy recognition by a logic-predicate network. Such a network is designed to recognize complex structured objects that can be described by predicate formulas. The NP-hardness of such an object recognition requires to separate the learning process, leaving it exponentially hard, and the recognition process itself. The learning process consists in extraction of groups of features (properties of elements of an object and the relations between these elements) that are common for objects of the same class. The main result of a paper is a reconstruction of a logic-predicate recognition cell. Such a reconstruction allows to recognize objects with descriptions not isomorphic to that from a training set and to calculate a degree of coincidence between the recognized object features and the features inherent to objects from the extracted group.

---

## 1 Introduction

This paper is an extension of work originally presented in conference CSIT-2019 [1].

The term “logical approach to solving Artificial Intelligence (AI) problems” is usually understood as data notation in the form of a binary string that defines the values of some object properties under study. In this case, if the object is represented as a set of its elements that have some properties and are connected by given relations, then when describing it as a binary string, firstly, the structure of the object itself is lost, and secondly, you have to store a large number of “unnecessary” information that the element  $a_j$  does not have the properties  $p_{i_1}, \dots, p_{i_k}$ , that the elements  $a_{j_1}, a_{j_2}$  are not in the relations  $q_{i_1}, \dots, q_{i_l}$ , etc.

Consider objects composed of smaller elements with preassigned properties, and some smaller object having preassigned relations. In such a case predicate formulas are an adequate description language.

The use of predicate calculus and automatic proof of a theorem for AI problems solving was offered by many authors [2]-[4] in the 70-th years of the XX century. At the same time the notion of NP-completeness was introduced and began to be actively used by the scientific community [5].

In 2003 the use of predicate calculus and automatic proof of a theorem for AI problems solving (without complexity bounds) was described in [6].

In particular, in that book it is shown that if a binary string simulates a description of an economic problem in which interaction

agents have given properties and are in given relations then the notation length of such a string is exponential of the length of the description of the same problem input data by setting properties of these agents and relations between them, i.e. in fact, using the language of predicate calculus.

In 2007 the author of this paper has proved NP-completeness of recognition problems described in the terms of predicate formulas and upper bounds of number of steps for two solving algorithms [7]. When such problems are solved by an exhaustive search algorithm, their computational complexity coincides with the length of their input data encoding using a binary string [6].

A hierarchical level description of classes was suggested in [8] to decrease the computational complexity of these problems. Formulas which are isomorphic to “frequently appeared” sub-formulas of “small complexity” are extracted from class descriptions in order to construct such a level description. This allows to decompose the main problem into a series of similar problems with input data with the smaller length. An instrument for this extraction is partial deduction of a predicate formula [9],[10].

Later, the author noticed that the level description is actually a recognition network, which, after training (having exponential computational complexity), can quickly solve the recognition problem.

The concept of “network” is increasingly used both in theoretical and in practical research, particularly, when solving pattern recognition problems. Research and practical applications of neural networks [11],[12], Bayesian networks [13]-[15], technical networks [16] are widespread. The inputs of such networks are usually signals characterizing the properties of the studied objects or pro-

---

\*Corresponding Author: Tatiana Kosovskaya, St. Petersburg State University, St. Petersburg, 199034, Russia, +7 921 323 23 07 & kosovtm@gmail.com

cesses, and expressed, as a rule, by numerical characteristics, which can be encoded by binary strings. The processing of binary strings in the most cases is carried out in linear or polynomial (usually quadratic) number of steps under the length of these strings. A convenient model for processing such strings is propositional logic (Boolean functions).

As noted above, the use of binary strings in the recognition of complex structured objects has disadvantages. To recognize such objects, the author offered the notion of a logic-predicate network using level description of classes [17]. Such a network contains two blocks: a training block and a recognition block. The training block run is based on the extraction from a class of objects description such fragments, which are inherent in many objects of the class. Recognition block run is reduced to the sequential solution of problems of the same type with a smaller length of the input data.

The disadvantage of such a recognition network is that it is able to recognize only objects that differ from those included in the training set (TS) by renaming the elements of the object (objects with isomorphic descriptions). But such a network can be retrained by adding unrecognized objects to the TS. This was the reason for modification of the logic-predicate network permitting to recognize objects, that not coincide but only similar to the objects from a training set. Such a modification was offered in [1, 18]. The degree of coincidence for description of an object part and the formula, satisfiability of which is checking in the cell, is calculated. Such a fuzzy network allows not only quickly enough to recognize objects isomorphic to those represented in the TS, but also to calculate the degree of coincidence that a “new” object belongs to one or another class. If necessary, it can be retrained using such a “new” object.

The structure of the paper is as follows.

Section 2 “Logic-predicate approach to AI problems” contains 2 subsections: “Problem Setting” and “Important Definitions”. This section describes the results previously obtained by the author and defines the terminology previously introduced by the author, which are necessary for understanding the further presentation.

Section 3 “Level description” presents the basic idea of constructing a level description of classes. It contains two sub-sections “Construction of a level description of classes” and “The use of level description of a class”. They give algorithms for constructing a level description of classes and the use of such a description for object recognition. These algorithms are further used in constructing the logic-predicate network. It is shown that the computational complexity of an object recognition decreases while using a level description of a class.

Section 4 “Logic-predicate network” describes the structure of such a network.

Section 5 “Example of logic-predicate network” describes an example illustrating the formation of a network, recognizing of an object with the use of a network, and its retraining. It consists of 4 sub-sections “Training Block Run: Extraction of Sub-Formulas”, “Training Block Run: Forming a Level Description”, “Recognition Block Run” and “Retraining the Network”.

Section 6 “Fuzzy recognition by a logic-predicate network” describes the changes that need to be made to the logic-predicate network in order to recognize objects that are not isomorphic to

those presented in the TS. A description of the contents of a fuzzy network cell is given. A cell of this type replaces each cell in the logic-predicate network in which the logical sequence of the target formula from the description of an object is checked. In this case, the degree of coincidence is calculated that the fragment being tested partially (and to what extent) satisfies the target formula.

Section 7 “Model example of fuzzy recognition” presents an example of fuzzy recognition of a “new” object.

In comparison with [1], the setting of problems that can be solved in the framework of logic-predicate approach are described in the presented paper in more details. A justification for the introduction of a level description is given. Algorithms for a level description construction and a level description use are described. Scheme 1 of a common up to the names of arguments sub-formulas extraction is added. A section 5 has been added with an example of constructing and modifying a logic-predicate network. It describes in details how a network is formed by the training set, a new object is recognized, and the network is rebuilt. The scheme of a fuzzy logic-predicate network cell is presented. An example of fuzzy recognition is described in more details.

## 2 Logic-Predicate Approach to AI Problems

A detailed presentation of the logic-predicate approach to solving AI problems is available in [19]. In this paper only a general setting of problems and main methods for their solving are formulated.

### 2.1 Problem Setting

As it was mentioned in the Introduction, if an object to be recognized is a complex structured one, then predicate formulas are a convenient language for its description and recognition. Let an investigated object  $\omega$  be represented as a set of its elements  $\omega = \{\omega_1, \dots, \omega_t\}$  and be characterized by predicates  $p_1, \dots, p_n$  which define some properties of its elements or relations between them. The description  $S(\omega_1, \dots, \omega_t)$  of the object  $\omega$  is a set of all constant literals (atomic formulas or their negations) with predicates  $p_1, \dots, p_n$  which are valid on  $\omega$ .

Let the set of all investigated objects  $\Omega$  be divided into classes  $\Omega_1, \dots, \Omega_K$  such that  $\Omega = \bigcup_{k=1}^K \Omega_k$ . Logical description of the class  $\Omega_k$  is such a formula  $A_k(\bar{x})$ <sup>1</sup> that if the formula  $A_k(\bar{\omega})$  is true then  $\omega \in \Omega_k$ . The class description may be represented as a disjunction of elementary conjunctions of atomic formulas.

Many AI problems may be formulated as follows with the use of such descriptions.

**Identification problem.** To pick out all parts of the object  $\omega$  which belong to the class  $\Omega_k$

$$S(\omega) \Rightarrow \exists \bar{x}_{k \neq} A_k(\bar{x}_k)^2. \quad (1)$$

**Classification problem.** To find all such class numbers  $k$  that  $\omega \in \Omega_k$

$$S(\omega) \Rightarrow \bigvee_{k=1}^K A_k(\bar{x}_k). \quad (2)$$

<sup>1</sup>Here and below the notation  $\bar{x}$  is used for an ordered list of the set  $x$ .

<sup>2</sup>To denote that there exist distinct values for variables from the list  $\bar{x}$  the notation  $\exists \bar{x}_{k \neq} A_k(\bar{x})$  is used.



**Analysis problem.** To find and classify all parts  $\tau$  of the object  $\omega$  which may be classified

$$S(\omega) \Rightarrow \bigvee_{k=1}^K \exists \bar{x}_{k \neq} A_k(\bar{x}_k). \quad (3)$$

In fact, instead of the quantifier  $\exists$  (*whether exists*) a symbol ? (*for what*) must be written. But if one uses a constructive method of its proof (an exhaustive search algorithm or logical methods such as derivation in a sequential predicate calculus or resolution method for predicate calculus) then not only existence of such arguments would be proved but the values of them would be found.

It was proved that the problems (1) and (3) are NP-complete [7] and the problem (2) is GI-complete [20], i.e. it is polynomially equivalent to the “open” problem of Graph Isomorphism, for which it is not proved its NP-completeness and a polynomial in time algorithm is not found.

As the formulas  $A_k(\bar{x}_k)$  may be represented as a disjunction of elementary conjunctions of atomic formulas, the checking of each of the problems (1), (2) and (3) may be reduced to sequential checking of the formula

$$S(\omega) \Rightarrow \exists \bar{x}_{\neq} A(\bar{x}), \quad (4)$$

where  $A(\bar{x})$  is an elementary conjunction. Note that in the case of the problem (2) the number of variables in  $\bar{x}$  equals to the number of constants in  $\omega$ .

The formula (4) may be checked, for example, by an exhaustive algorithm and an algorithm based on the derivation in sequential calculus or on the use of resolution method.

The upper bound of number of steps for an exhaustive algorithm is  $O(t^m)^3$ , where  $t$  is the number of elements in  $\omega$ ,  $m$  is the number of variables in  $A(\bar{x})$  [7]. This upper bound coincides with the upper bound of a binary string simulating input data in the form of  $S(\omega)$  and  $A(\bar{x})$ .

The upper bound of the number of steps for a logical algorithm is  $O(s^a)$  where  $s$  is the maximal number of atomic formulas in  $S(\omega)$  with the same predicate symbol having occurrences in  $A(\bar{x})$ ,  $a$  is the number of atomic formulas in the elementary conjunction  $A(\bar{x})$  [21].

Obviously, these estimates depend exponentially on the parameters of the formula  $A(\bar{x})$ . That’s why it will be useful to break the solution of the problem into a series of sub-problems of the type (4) with a shorter right-hand side.

## 2.2 Important Definitions

The objects and notions satisfying the following definitions will be used later in the text.

**Definition 1:** Elementary conjunctions  $P(a_1, \dots, a_m)$  and  $Q(b_1, \dots, b_m)$  are called **isomorphic**

$$P(a_1, \dots, a_m) \sim Q(b_1, \dots, b_m),$$

if there are an elementary conjunction  $R(x_1, \dots, x_m)$  and substitutions of arguments  $a_{i_1}, \dots, a_{i_m}$  and  $b_{j_1}, \dots, b_{j_m}$  instead of the variables  $x_1, \dots, x_m$  such that the results of these substitutions  $R(a_{i_1}, \dots, a_{i_m})$  and  $R(b_{j_1}, \dots, b_{j_m})$  coincide with formulas  $P(a_1, \dots, a_m)$  and  $Q(b_1, \dots, b_m)$ , respectively, up to the order of literals.

The substitutions  $(a_{i_1} \rightarrow x_1, \dots, a_{i_m} \rightarrow x_m)$  and  $(b_{j_1} \rightarrow x_1, \dots, b_{j_m} \rightarrow x_m)$  are called **unifiers** of  $R(x_1, \dots, x_m)$  with  $P(a_1, \dots, a_m)$  and  $Q(b_1, \dots, b_m)$  and are denoted as  $\lambda_{R,P}$  and  $\lambda_{R,Q}$ , respectively. [1], [20]

Note that concept of isomorphism of elementary conjunctions of atomic predicate formulas differs from the concept of equivalence of these formulas, because they can have significantly different arguments. In fact, for isomorphic formulas there are permutations of their arguments such that they define the same relationship between their arguments.

**Definition 2:** Elementary conjunction  $C(x_1, \dots, x_n)$  is called a **common up to the names of arguments sub-formula** of two elementary conjunctions  $A(a_1, \dots, a_m)$  and  $B(b_1, \dots, b_k)$  if it is isomorphic to some sub-formulas  $A'(a'_1, \dots, a'_n)$  and  $B'(b'_1, \dots, b'_n)$  of  $A(a_1, \dots, a_m)$  and  $B(b_1, \dots, b_k)$ , respectively.

The unifiers of  $C(x_1, \dots, x_n)$  with  $A'(a'_1, \dots, a'_n)$  and  $B'(b'_1, \dots, b'_n)$  will be denoted as  $\lambda_{C,A}$  and  $\lambda_{C,B}$ , respectively [19].

For example, let

$$A(x, y, z) = p_1(x) \& p_1(y) \& p_1(z) \& p_2(x, y) \& p_3(x, z),$$

$$B(x, y, z) = p_1(x) \& p_1(y) \& p_1(z) \& p_2(x, z) \& p_3(x, z).$$

The formula

$$P(u, v) = p_1(u) \& p_1(v) \& p_2(u, v)$$

is their common up to the names of variables sub-formula with the unifiers  $\lambda_{P,A} = (x \rightarrow u, y \rightarrow v)$  and  $\lambda_{P,B} = (x \rightarrow u, z \rightarrow v)$  because

$$P(x, y) = p_1(x) \& p_1(y) \& p_2(x, y)$$

is a sub-formula of  $A(x, y, z)$  and

$$P(x, z) = p_1(x) \& p_1(z) \& p_2(x, z)$$

is a sub-formula of  $B(x, y, z)$ .

**Definition 3:** Elementary conjunction  $C(x_1, \dots, x_n)$  is called a **maximal common up to the names of arguments sub-formula** of two elementary conjunctions  $A(a_1, \dots, a_m)$  and  $B(b_1, \dots, b_k)$  if it is their common up to the names of arguments sub-formula and after adding any literal to it, it ceases to be one [19].

For further presentation, the concept of partial sequence [9] is important.

The problem of checking if the formula  $A(\bar{x})$  or some its sub-formula  $A'(\bar{y})$  is a consequence of the set of formulas  $S(\omega)$  is under consideration in [9]. Here the list of arguments  $\bar{y}'$  is a sub-list of the list of arguments  $\bar{y}$ .

**Definition 4:** Let  $A(\bar{x})$  and  $B(\bar{y})$  be elementary conjunctions. If  $A(\bar{x}) \Rightarrow \exists \bar{y}_{\neq} B(\bar{y})$  is not valid but for some sub-formula  $B'(\bar{y}')$  of  $B(\bar{y})$  the sequence  $A(\bar{x}) \Rightarrow \exists \bar{y}'_{\neq} B'(\bar{y}')$  is true, we will say that  $B(\bar{y})$  is a **partial sequence** from  $A(\bar{x})$  and denote this by  $A(\bar{x}) \Rightarrow_P \exists \bar{y}_{\neq} B(\bar{y})$  [1].

A constructive algorithm for a proof of partial sequence is in [22]. While using a constructive algorithm to prove  $A(\bar{x}) \Rightarrow_P \exists \bar{y}_{\neq} B(\bar{y})$  we can find the maximal sub-formula  $B'(\bar{y}')$  such that  $A(\bar{x}) \Rightarrow \exists \bar{y}'_{\neq} B'(\bar{y}')$  and such values  $\bar{x}'$  ( $\bar{x}'$  is a permutation of a sub-string of  $\bar{x}$ ) for  $\bar{y}'$  that  $B'(\bar{x}')$  is a sub-formula of  $A(\bar{x})$ . It means that we find a maximal common up to the names of arguments

<sup>3</sup>  $f(x) = O(g(x))$  means that there is such a constant  $C$  that for every  $x$  the inequality  $f(x) \leq C \cdot g(x)$  is valid.

sub-formula  $B'(\bar{y}')$  of two elementary conjunctions  $A(\bar{x})$  and  $B(\bar{y})$  and its unifier with  $B'(\bar{x}')$  (a sub-formula of  $A(\bar{x})$ ).

Further, for fuzzy recognition of an object we need an extension of the partial sequence concept.

Every sub-formula  $A'(\bar{x}')$  of the formula  $A(\bar{x})$  is called its **fragment** [1].

Let  $a$  and  $a'$  be the numbers of atomic formulas in  $A(\bar{x})$  and  $A'(\bar{x}')$ , respectively,  $m$  and  $m'$  be the numbers of objective variables in  $\bar{x}$  and  $\bar{x}'$ , respectively.

Numbers  $q$  and  $r$  are calculated by the formulas  $q = \frac{a'}{a}$ ,  $r = \frac{m'}{m}$  and characterize the degree of coincidence between  $A(\bar{x})$  and  $A'(\bar{x}')$ . For every elementary conjunction  $A(\bar{x})$  and its fragment  $A'(\bar{x}')$  it is true that  $0 < q \leq 1$ ,  $0 < r \leq 1$ . Besides,  $q = r = 1$  if and only if  $A'(\bar{x}')$  coincides with  $A(\bar{x})$ .

Under these notations, the formula  $A'(\bar{x}')$  will be called a  $(q, r)$ -**fragment of the formula**  $A(\bar{x})$  [1].

If  $S(\omega) \Rightarrow \exists \bar{x}_{\neq} A(\bar{x})$  is not valid but for some  $(q, r)$ -fragment  $A'(\bar{x}')$  ( $q \neq 1$ ) of  $A(\bar{x})$  the sequence  $S(\omega) \Rightarrow \exists \bar{x}'_{\neq} A'(\bar{x}')$  is true, we will say that  $S(\omega) \Rightarrow_p \exists \bar{x}_{\neq} A(\bar{x})$  is a **partial**  $(q, r)$ -**sequence** for description  $S(\omega)$  [1].

A  $(q, r)$ -fragment  $A'(\bar{x}')$  of the formula  $A(\bar{x})$  with maximal value of  $q$  satisfying  $S(\omega) \Rightarrow \exists \bar{x}'_{\neq} A'(\bar{x}')$  will be called a **maximal fragment of the formula**  $A(\bar{x})$  for description  $S(\omega)$  [1].

As for a maximal fragment of the formula  $A(\bar{x})$  the checking whether  $S(\omega) \Rightarrow \exists \bar{x}'_{\neq} A'(\bar{x}')$  may be done by some constructive method then such values  $\bar{\tau}$  ( $\tau \subseteq \omega$ ) for the list of variables  $\bar{x}'$  that  $S(\omega) \Rightarrow A'(\bar{\tau})$  will be found.

**Definition 5:** Conjunction of literals from  $A(\bar{x})$  which are not in  $A'(\bar{x}')$  is called a **complement** of  $A'(\bar{x}')$  up to  $A(\bar{x})$  [1].

A complement of  $A'(\bar{x}')$  up to  $A(\bar{x})$  will be denoted by  $C_{A(\bar{x})}A'(\bar{x}')$ .

**Definition 6:** A  $(q, r)$ -fragment  $A'(\bar{x}')$  of the formula  $A(\bar{x})$  is called **contradictory** to the description  $S(\omega)$  on the list of constants  $\bar{\tau}$  if  $S(\omega)$  and  $C_{[A(\bar{x})]_{\bar{\tau}}}A'(\bar{\tau})$  lead to the contradiction, i.e.,  $S(\omega) \Rightarrow \neg C_{[A(\bar{x})]_{\bar{\tau}}}A'(\bar{\tau})$  [1].

Here the denotation  $[A(\bar{x})]_{\bar{\tau}}$  is used for the result of substitution of the constants from the list  $\bar{\tau}$  instead of the corresponding variables from the list  $\bar{x}'$ .

### 3 Level Description

As noted above, estimates of the number of steps of the proof of (4) exponentially depends on the number of variables or on the number of literals of the formula  $A(\bar{x})$ . Moreover, the proof of formulas (1), (2) and (3) in which  $A_k(\bar{x}_k) = A_{k,1}(\bar{x}_{k,1}) \vee \dots \vee A_{k,m_k}(\bar{x}_{k,m_k})$ , with the available algorithms for solving (4), can be reduced to a sequential proof of (4) with  $A_{k,1}(\bar{x}_{k,1}), \dots, A_{k,m_k}(\bar{x}_{k,m_k})$  on the right side.

Due to the unprovable, but repeatedly confirmed in practice, statement of Einstein "God is subtle, but he is not malicious" elementary conjunctions included in the description of one class must have common up to the names of arguments sub-formulas. Therefore, it is natural to break the proof of this formula into a series of formulas of the same kind, but with lower values of the essential parameters. This may be done by means of the following procedure.

– Extract "frequently occurred" common up to the names of arguments sub-formulas  $P_i^1(\bar{y}_i^1)$  ( $i = 1, \dots, n_1$ ) of goal formulas

$A_{k,1}(\bar{x}_{k,1}), \dots, A_{k,m_k}(\bar{x}_{k,m_k})$  with "small complexity". Simultaneously we find unifiers of  $P_i^1(\bar{y}_i^1)$  and sub-formulas of  $A_k(x_1, \dots, x_m)$ .

– For every sub-formula  $P_i^1(\bar{y}_i^1)$  a new 1st level predicate  $p_i^1$  with one 1st level variable  $y_i^1$ , defined by the equivalence  $p_i^1(y_i^1) \Leftrightarrow P_i^1(\bar{y}_i^1)$ , is introduced. This 1st-level variable  $y_i^1$  is a variable for a string of initial variables.

– Replace each occurrence of a formula, isomorphic to  $P_i^1(\bar{y}_i^1)$  ( $i = 1, \dots, n_1$ ), into  $A_{k,j}(\bar{x}_{k,j})$  ( $j = 1, \dots, m_k$ ) with a literal  $p_i^1(y_i^1)$ . The earlier found unifier gives us the particular string of initial variables which are in this occurrence of  $y_i^1$ . Denote the received from  $A_{k,j}(\bar{x}_{k,j})$  formula by means of  $A_{k,j}^1(\bar{x}_{k,j}^1)$ .

For  $l = 1, \dots, L - 1$  repeat this procedure for  $A_{k,1}^l(\bar{x}_{k,1}^l), \dots, A_{k,m_k}^l(\bar{x}_{k,m_k}^l)$  and receive formulas  $A_{k,1}^{l+1}(\bar{x}_{k,1}^{l+1}), \dots, A_{k,m_k}^{l+1}(\bar{x}_{k,m_k}^{l+1})$ . The process will end because the number of literals in  $A_{k,j}^l(\bar{x}_{k,j}^l)$  decreases with increasing  $l$ .

Level description of formulas  $A_{k,1}(\bar{x}_{k,1}), \dots, A_{k,m_k}(\bar{x}_{k,m_k})$  has the following form [8].

$$\left\{ \begin{array}{l} A_{k,j}^L(\bar{x}_{k,j}^L) \quad (j = 1, \dots, m_k) \\ p_1^1(y_1^1) \Leftrightarrow P_1^1(\bar{y}_1^1) \\ \vdots \\ p_{n_1}^1(y_{n_1}^1) \Leftrightarrow P_{n_1}^1(\bar{y}_{n_1}^1) \\ \vdots \\ p_i^l(y_i^l) \Leftrightarrow P_i^l(\bar{y}_i^l) \\ \vdots \\ p_{n_l}^L(y_{n_l}^L) \Leftrightarrow P_{n_l}^L(\bar{y}_{n_l}^L) \end{array} \right. \quad (5)$$

The described procedure deals with such intuitive notions as "frequently occurred" and having "small complexity" sub-formula. In the following algorithms it is supposed that "frequently occurred" means that a sub-formula appears at least twice and "small complexity" means that the number of a sub-formula literals is less than such a number of at least one of the initial formulas.

#### 3.1 Construction of a Level Description of Classes

An algorithm for construction of a level description of classes is presented in [19], [23]. This algorithm contains two parts: 1) extraction of common up to the names of arguments sub-formula from the class description, 2) forming a level description. Here it is presented with more accuracy.

##### 1) Extraction of common up to the names of arguments sub-formulas from the class description

Let the description of a class be  $A(\bar{x}) = A_1(\bar{x}_1) \vee \dots \vee A_K(\bar{x}_K)$ , where  $A_1(\bar{x}_1), \dots, A_K(\bar{x}_K)$  are elementary conjunctions of literals.

- For every  $i$  and  $j$  ( $i < j$ ) check whether  $A_i(\bar{x}_i) \Rightarrow_p \exists \bar{x}_{j \neq} A_j(\bar{x}_j)$ .

Using the notion of partial sequence we can obtain a maximal common up to the names of arguments sub-formula  $Q_{i,j}^1(\bar{z}_{i,j}^1)$  of two elementary conjunctions  $A_i(\bar{x}_i)$  and  $A_j(\bar{x}_j)$  and unifiers  $\lambda_{Q_{i,j}^1, A_i}$  and  $\lambda_{Q_{i,j}^1, A_j}$ . Let  $s_i$  be the lists of indices of  $Q_{s_i}^l(\bar{z}_{s_i}^l)$ .

- For  $l = 1, \dots, L - 1$  do.
- For every  $s_i$  and  $s_j$  ( $s_i \neq s_j$ ) do.

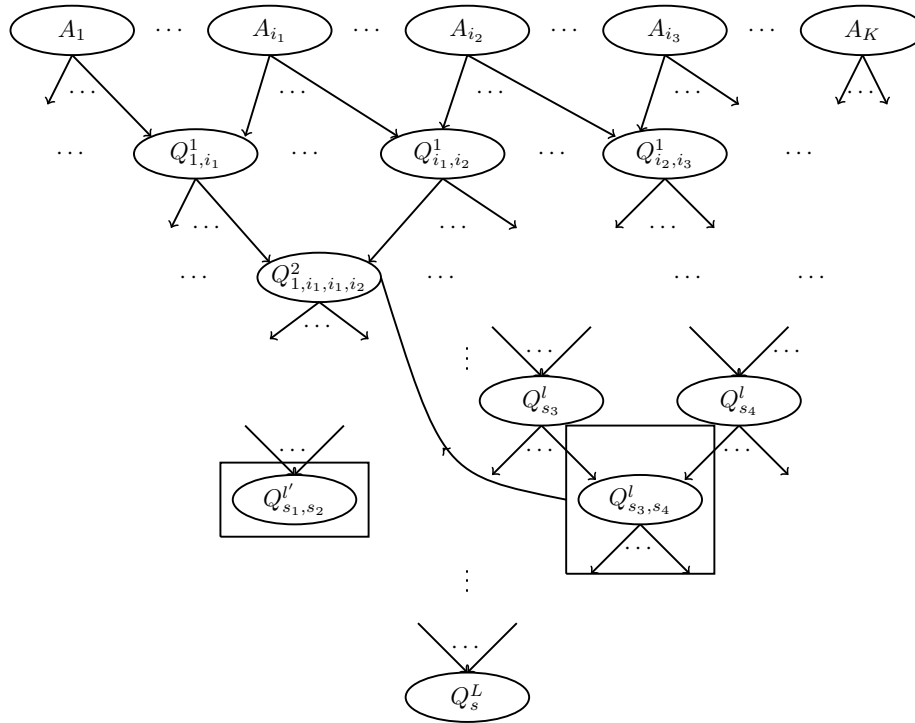


Figure 1: Scheme of common up to the names of arguments sub-formulas extraction.

– In the similar way find a maximal common up to the names of arguments sub-formula  $Q_{s_i,s_j}^{l+1}(\bar{z}_{s_i,s_j}^{l+1})$  of the formulas  $Q_{s_i}^l(\bar{z}_{s_i}^l)$  and  $Q_{s_j}^l(\bar{z}_{s_j}^l)$  and unifiers  $\lambda_{Q_{s_i,s_j}^{l+1}, Q_{s_i}^l}$  and  $\lambda_{Q_{s_i,s_j}^{l+1}, Q_{s_j}^l}$ .

– If a formula  $Q_{s_i,s_j}^{l+1}(\bar{z}_{s_i,s_j}^{l+1})$  is isomorphic to some previously obtained  $Q_s^l(\bar{z}_s^l)$  for some  $s$  and  $l' \leq l + 1$  then it is deleted from the set of the  $(l + 1)$ th level formulas and we write down unifiers  $\lambda_{Q_s^l, Q_{s_i}^l}$  and  $\lambda_{Q_s^l, Q_{s_j}^l}$ .

Note that the length of  $Q_{s_i,s_j}^{l+1}(\bar{z}_{s_i,s_j}^{l+1})$  decreases with increasing  $l$ . That is why the process would stop.

A scheme of the extraction is presented in Figure 1. In order not to overload the scheme, the arguments of the formulas are not written on it and instead of  $Q_{s_i}^l(\bar{z}_{s_i}^l)$  it is written  $Q_{s_i}^l$ .

In this scheme it is supposed that highlighted together with its connections formula  $Q_{s_3,s_4}^l$  is isomorphic to some previously obtained formula, for example, to  $Q_{1,i_1,i_2}^2$  and that's why it is deleted from the set of the  $l$ th level formulas and all its connections are rewritten to  $Q_{1,i_1,i_2}^2$ .

A formula  $Q_{s_1,s_2}^l$  that does not have a common sub-formula with any other formulas, is highlighted on it.

## 2) Forming a level description.

Further, we assume that  $Q_i^0(\bar{z}_i^0)$  is  $A_i(\bar{x}_{k,i})$ .

• Every formula of the type  $Q_i^l(\bar{z}_i^l)$  which has no sub-formula of the type  $Q_{s,s}^{l+1}(\bar{z}_{s,s}^{l+1})$  for any  $s_j$  is declared as  $P_i^l(\bar{y}_i^l)$  ( $i = 1, \dots, n_1$ ).<sup>4</sup>

A new 1st level predicate defined as  $p_i^1(y_i^1) \Leftrightarrow P_i^1(\bar{y}_i^1)$  is introduced and  $p_i^1(y_i^1)$  is substituted (using the correspondent unifies)

instead of  $P_i^l(\bar{y}_i^l)$  into all formulas of the type  $Q_s^l(\bar{z}_s^l)$  ( $l \geq 0$ ) such that there is an oriented path from  $Q_s^l(\bar{z}_s^l)$  to  $P_i^l(\bar{y}_i^l)$ .

• For  $l = 1, \dots, L - 1$  do.

For every  $i = 1, \dots, n_l$  if  $P_i^l(\bar{y}_i^l)$  is a maximal common up to the names of arguments sub-formula of formulas  $Q_{s_{i'}}^l(\bar{z}_{s_{i'}}^l)$  and  $Q_{s_{j'}}^l(\bar{z}_{s_{j'}}^l)$  then these formulas are declared as  $P_{i''}^{l+1}(\bar{y}_{i''}^{l+1})$  and  $P_{i''+1}^{l+1}(\bar{y}_{i''+1}^{l+1})$  ( $i'' = 1, \dots, n_{l+1} - 1$ ).

New  $(l + 1)$ th level predicates  $p_{i''}^{l+1}(y_{i''}^{l+1})$  and  $p_{i''+1}^{l+1}(y_{i''+1}^{l+1})$  defined as  $p_{i''}^{l+1}(y_{i''}^{l+1}) \Leftrightarrow P_{i''}^{l+1}(\bar{y}_{i''}^{l+1})$  and  $p_{i''+1}^{l+1}(y_{i''+1}^{l+1}) \Leftrightarrow P_{i''+1}^{l+1}(\bar{y}_{i''+1}^{l+1})$  are introduced and every of them is substituted (using the correspondent unifies) instead of  $P_{i''}^{l+1}(\bar{y}_{i''}^{l+1})$  and  $P_{i''+1}^{l+1}(\bar{y}_{i''+1}^{l+1})$ , respectively, into all formulas of the type  $Q_s^l$  such that there is an oriented path from  $Q_s^l$  to at least one of them.

If  $P_i^l(\bar{y}_i^l)$  corresponds to the cell marked as  $Q_{1,i_1,i_2}^2$  in Figure 1 then  $p_{i''}^{l+1}(y_{i''}^{l+1})$  is substituted instead of  $P_{i''}^{l+1}(\bar{y}_{i''}^{l+1})$  into the formulas corresponding the cells marked as  $Q_{1,i_1}^1$ ,  $A_1$  and  $A_{i_1}$ . The literal  $p_{i''+1}^{l+1}(y_{i''+1}^{l+1})$  is substituted instead of  $P_{i''+1}^{l+1}(\bar{y}_{i''+1}^{l+1})$  into the formulas corresponding the cells marked as  $Q_{1,i_2}^1$ ,  $A_{i_1}$  and  $A_{i_2}$ .

The words “using the correspondent unifies” may be explained in such a way. If  $P_i^l(\bar{y}_i^l)$  corresponds to the cell marked as  $Q_{1,i_1,i_2}^2$  in Figure 1 then  $p_{i''}^{l+1}(y_{i''}^{l+1})$  is substituted instead of  $P_{i''}^{l+1}(\bar{y}_{i''}^{l+1})$  into the formulas corresponding the cells marked as  $Q_{1,i_1}^1$ ,  $A_1$  and  $A_{i_1}$ . The literal  $p_{i''+1}^{l+1}(y_{i''+1}^{l+1})$  is substituted instead of  $P_{i''+1}^{l+1}(\bar{y}_{i''+1}^{l+1})$  into the formulas corresponding the cells marked as  $Q_{1,i_2}^1$ ,  $A_{i_1}$  and  $A_{i_2}$ .

It may be easily proved that the problem of level description

<sup>4</sup>These are formulas  $Q_s^l$  and  $Q_{s_1,s_2}^l$  in Figure 1.

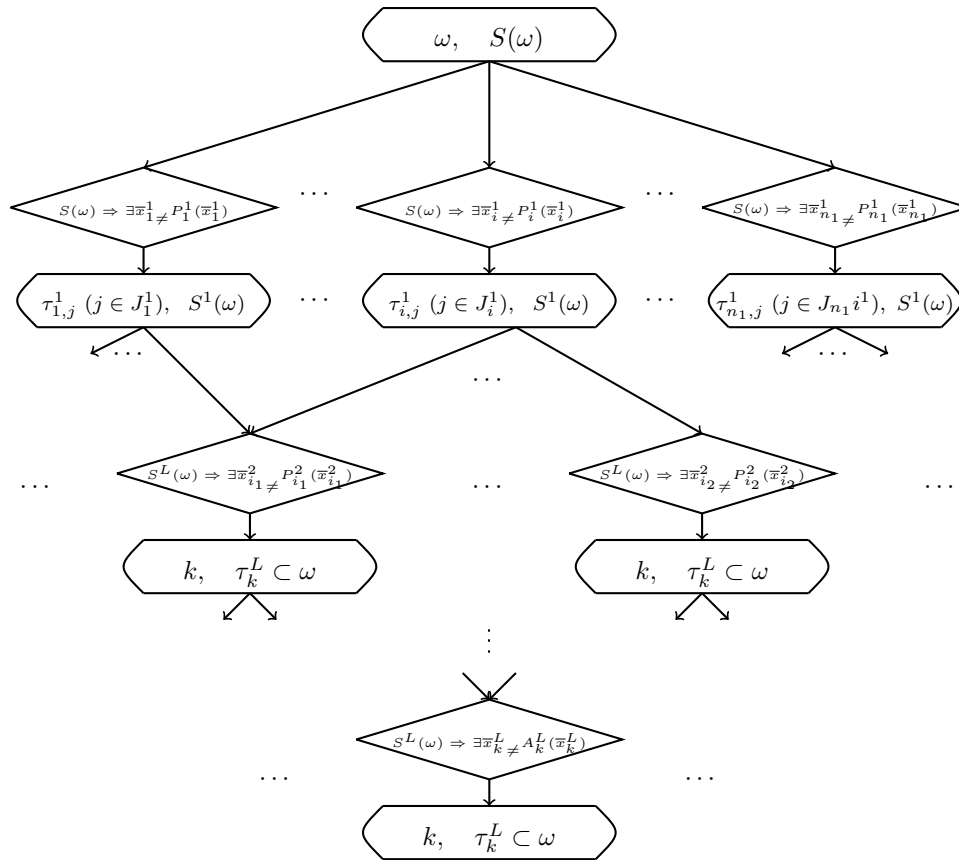


Figure 2: Scheme of level recognition.

construction is NP-hard. In particular, the problem of checking if there exists a level description is NP-complete and the upper bounds of number of steps for traditional algorithms exponentially depends on parameters of  $A_1(\bar{x}_1), \dots, A_K(\bar{x}_K)$ . At the same time such a description must be constructed only once and then may be used as many times as you like.

### 3.2 The Use of Level Description of a Class

Checking for what  $k$  the sequence  $S^l(\omega) \Rightarrow \exists \bar{x}_k \neq A_k(\bar{x}_k)$  is valid and if “yes” then finding for what value of  $\bar{x}_k$  it is valid with the use of level description of a class, may be done according to the following algorithm.

- For  $i = 1, \dots, n_1$  check  $S(\omega) \Rightarrow \exists \bar{y}_i \neq P_i^1(\bar{y}_i^1)$  and find all lists of values  $\bar{\tau}_{i,j}^1$  for the variables  $\bar{y}_i^1$  such that  $S(\omega) \Rightarrow P_i^1(\bar{\tau}_{i,j}^1)$ .<sup>5</sup>

Let  $\tau_{i,j}^1$  be a notation for the list  $\bar{\tau}_{i,j}^1$ . All atomic formulas of the form  $p_i^1(\tau_{i,j}^1)$  with  $\tau_{i,j}^1$  a notation for the list  $\bar{\tau}_{i,j}^1$  are added to  $S(\omega)$  to obtain  $S^1(\omega)$  – a 1st level description of  $\omega$ .

- For  $l = 1, \dots, L - 1$  do.
  - For  $i = 1, \dots, n_l$  check  $S^l(\omega) \Rightarrow \exists \bar{y}_i \neq P_i^l(\bar{y}_i^l)$  and find all lists of values  $\bar{\tau}_{i,j}^l$  for the variables  $\bar{y}_i^l$  such that  $S^l(\omega) \Rightarrow P_i^l(\bar{\tau}_{i,j}^l)$ .

- All atomic formulas of the form  $p_i^l(\tau_{i,j}^l)$  with  $\tau_{i,j}^l$  a notation for the list  $\bar{\tau}_{i,j}^l$  are added to  $S^l(\omega)$  to obtain  $S^{l+1}(\omega)$  – a  $l$ th level description of  $\omega$ .

- – For  $k = 1, \dots, K$  check  $S^L(\omega) \Rightarrow \exists \bar{y}_k \neq A_k^L(\bar{y}_k^L)$  and find all lists of values  $\bar{\tau}_{k,j}^L$  for the variables  $\bar{y}_k^L$  such that  $S^L(\omega) \Rightarrow A_k^L(\bar{\tau}_{k,j}^L)$ .

The object  $\omega$  satisfies such formulas  $A_k(\bar{x}_k)$  for which the corresponding sequence  $S^L(\omega) \Rightarrow \exists \bar{y}_k \neq A_k^L(\bar{y}_k^L)$  is valid.

The scheme of level recognition is presented in Figure 2.

It is proved in [8] that for a two-level description ( $L = 1$ ) the number of steps with the use of 2-level description decreases. For an exhaustive algorithm the number of steps decreases from  $O(m^k)$  to  $O(n_1 \cdot r^r + r^{\delta_k + n_1})$ , where  $r$  is a maximal number of arguments in the 1st level predicates,  $n_1$  is the number of such predicates,  $\delta_k$  is the number of initial variables which are presented in  $A_k(\bar{x}_k)$  and are absent in  $A_k^1(\bar{x}_k^1)$ . Both parameters  $r$  and  $\delta_k + n_1$  are less than  $m_k$ .

Number of steps for an algorithm based on the derivation in predicate calculus decreases from  $O(s_k^{a_k})$  to  $O(s^{a_k^1} + \sum_{j=1}^{n_1} s^{p_j^1})$ , where  $a_k$  and  $a_k^1$  are maximal numbers of literals in  $A_k(\bar{x}_k)$  and  $A_k^1(\bar{x}_k^1)$ , respectively,  $s$  and  $s^1$  are the numbers of literals in  $S(\omega)$  and  $S^1(\omega)$ , respectively,  $p_j^1$  is the numbers of literals in  $P_i^1(\bar{y}_i^1)$ . Both parameters  $a_k^1$  and  $p_j^1$  are less than  $a_k$ .

<sup>5</sup>Note that for every predicate  $P_i^1$  there may be many lists of values from  $\omega$  satisfying it.



So, the recognition of an object with the use of a level description requires less time than that with the use of initial description.

### 4 Logic-Predicate Network

The inputs of a traditional neuron network are binary or many-valued characteristics of an object. The neuron network configuration is fixed that does not correspond to biological networks.

Below a logic-predicate network is described. Such a network inputs are atomic formulas setting properties of the elements composing an investigated object and relations between them [17], [19]. Further it will be shown that its configuration may be changed during its retraining.

A level description of a class of objects permits quickly enough to recognize an object which description is isomorphic to the description of an element of a training set.

A logic-predicate network consists of two blocks: a training block and a recognition block. Training block forms a level description of a class of objects according to a training set (or some formalized description of the class). Its content is presented in Figure 1.

Recognition block implements the algorithm of the use of level description of a class. It must be marked that the level description may be represented as an oriented graph presented in Figure 2.

Scheme of logic-predicate network is presented in Figure 3. Here inputs and outputs are situated in ovals and checking the conditions are situated in rhombus.

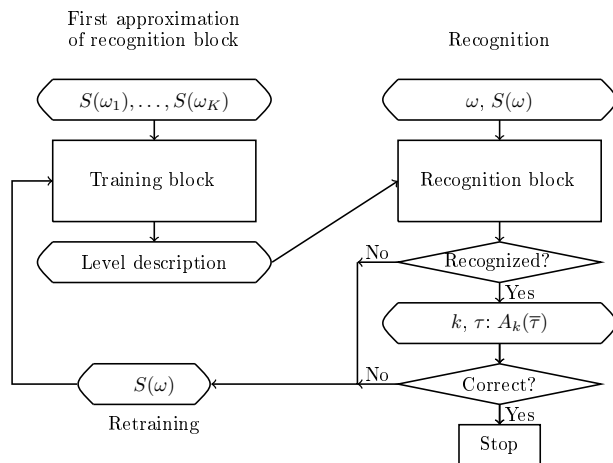


Figure 3: Scheme of logic-predicate network.

Let a training set of objects  $\omega^1, \dots, \omega^K$  be given. An initial variant of the class description may be formed by replacement of every constant  $\omega_j^k$  in  $S(\omega^k)$  by a variable  $x_j^k$  ( $k = 1, \dots, K, j = 1, \dots, t^k$ ) and substitution of the sign  $\&$  between the atomic formulas. Construct a level description for these goal formulas. The first approximation to the recognition block is formed.

If after the recognition block run an object is not recognized or has wrong classification, then it is possible to retrain the network. The description of the “wrong” object must be added to the input set of the training block. The training block extracts common up to

the names sub-formulas of this description and previously received formulas forming a new recognition block. Some sub-formulas, the number of layers and the number of formulas in every layer may be changed in the level description. Then the recognition block is reconstructed.

### 5 Example of Logic-Predicate Network Construction, its Running and Retraining

Let a set of contour images of “boxes” be to recognize. Every image may be described in the terms of two predicates  $V$  and  $L$  defined as it is presented in Figure 4.

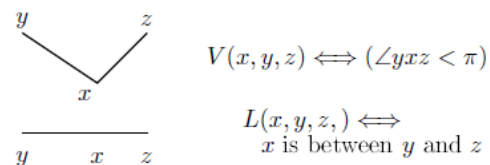


Figure 4: Initial predicates.

A training set  $\{\omega_1, \omega_2, \omega_3\}$  of such images is presented in Figure 5. Here  $\omega_1 = \{a_1, \dots, a_{10}\}$ ,  $\omega_2 = \{b_1, \dots, b_8\}$ ,  $\omega_3 = \{c_1, \dots, c_{10}\}$ .

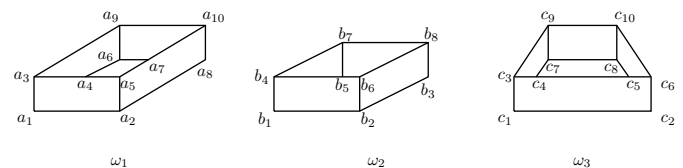


Figure 5: Training set.

In order not to overload the text, we write down only the formula  $A_2(y_1, \dots, y_8)$ , which is obtained from the description  $S(\omega_2)$  by replacing the names  $b_i$  with the variables  $y_i$  and placing the sign  $\&$  between the literals.

$$A_2(y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8) = V(y_1, y_4, y_2) \& V(y_2, y_1, y_6) \& V(y_2, y_1, y_3) \& V(y_2, y_6, y_3) \& V(y_3, y_2, y_8) \& V(y_4, y_7, y_6) \& V(y_4, y_7, y_5) \& V(y_4, y_7, y_1) \& V(y_4, y_5, y_1) \& V(y_4, y_6, y_1) \& V(y_5, y_4, y_7) \& V(y_5, y_4, y_6) \& V(y_5, y_7, y_6) \& L(y_5, y_4, y_6) \& V(y_6, y_4, y_8) \& V(y_6, y_5, y_8) \& V(y_6, y_8, y_2) \& V(y_6, y_2, y_5) \& V(y_6, y_2, y_4) \& V(y_7, y_4, y_5) \& V(y_7, y_5, y_8) \& V(y_7, y_4, y_8) \& V(y_8, y_3, y_6) \& V(y_8, y_6, y_7) \& V(y_8, y_3, y_7)$$

Further, for clarity, we will write down not formulas, but drawings.

### 5.1 Training Block Run: Extraction of Sub-Formulas

For constructing a level description, find pairwise maximal common up to the names of arguments sub-formulas  $Q_i^1(\bar{y}_i^1)$  of formulas  $A_1(\bar{x}), A_2(\bar{y}), A_3(\bar{z})$  and their unifiers. See Figure 6.

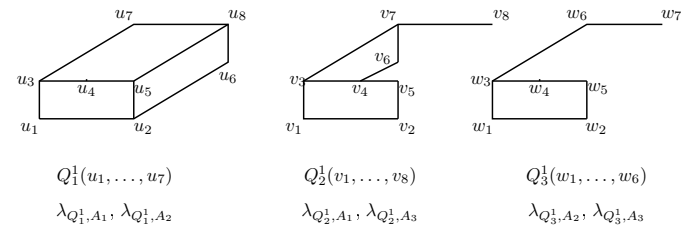


Figure 6: First extraction of pairwise maximal common up to the names of arguments sub-formulas.

Here the following unifier are formed.

Let  $u^1 = (u_1, \dots, u_8), v^1 = (v_1, \dots, v_8), w^1 = (w_1, \dots, w_7)$ .

$$\lambda_{Q_1^1, A_1} = (u_1 \rightarrow x_1, u_2 \rightarrow x_2, u_3 \rightarrow x_3, u_4 \rightarrow x_4, u_5 \rightarrow x_5, u_6 \rightarrow x_6, u_7 \rightarrow x_7, u_8 \rightarrow x_8) = (u^1 \rightarrow x_1^1),$$

$$\lambda_{Q_1^1, A_2} = (u_1 \rightarrow y_1, u_2 \rightarrow y_2, u_3 \rightarrow y_3, u_4 \rightarrow y_4, u_5 \rightarrow y_5, u_6 \rightarrow y_6, u_7 \rightarrow y_7, u_8 \rightarrow y_8) = (u^1 \rightarrow y_1^1),$$

$$\lambda_{Q_2^1, A_1} = (v_1 \rightarrow x_1, v_2 \rightarrow x_2, v_3 \rightarrow x_3, v_4 \rightarrow x_4, v_5 \rightarrow x_5, v_6 \rightarrow x_6, v_7 \rightarrow x_7, v_8 \rightarrow x_8) = (v^1 \rightarrow x_2^1),$$

$$\lambda_{Q_2^1, A_3} = (v_1 \rightarrow z_1, v_2 \rightarrow z_2, v_3 \rightarrow z_3, v_4 \rightarrow z_4, v_5 \rightarrow z_5, v_6 \rightarrow z_6, v_7 \rightarrow z_7, v_8 \rightarrow z_8) = (v^1 \rightarrow z_1^1),$$

$$\lambda_{Q_3^1, A_2} = (w_1 \rightarrow y_1, w_2 \rightarrow y_2, w_3 \rightarrow y_3, w_4 \rightarrow y_4, w_5 \rightarrow y_5, w_6 \rightarrow y_6, w_7 \rightarrow y_7) = (w^1 \rightarrow y_2^1),$$

$$\lambda_{Q_3^1, A_3} = (w_1 \rightarrow z_1, w_2 \rightarrow z_2, w_3 \rightarrow z_3, w_4 \rightarrow z_4, w_5 \rightarrow z_5, w_6 \rightarrow z_6, w_7 \rightarrow z_7) = (w^1 \rightarrow z_2^1).$$

Pairwise extractions of maximal common up to the names of arguments sub-formulas of  $Q_1^1(u_1, \dots, u_8), Q_2^1(v_1, \dots, v_8)$  and  $Q_3^1(w_1, \dots, w_6)$  give only just obtained formula  $Q_3^1(w_1, \dots, w_6)$ . That's why we have a situation presented in Figure 7.

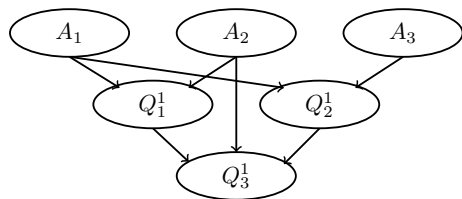


Figure 7: Connections between the extracted formulas.

### 5.2 Training Block Run: Forming a Level Description

The formula  $Q_3^1(w_1, \dots, w_7)$  has no common sub-formulas with the other ones. Therefore, it is the 1st level predicate  $P_1^1(w_1, \dots, w_7)$  of the network. A new 1st level predicate  $p_1^1(w^1)$  defining as  $p_1^1(w^1) \Leftrightarrow P_1^1(w_1, \dots, w_7)$  is introduced. Remind that  $w^1$  is a new 1st level variable for the string of initial variables  $(w_1, w_2, w_3, w_4, w_5, w_6, w_7)$ .

Formula  $P_1^1(w_1, \dots, w_7)$  (which corresponds to  $Q_3^1(w_1, \dots, w_7)$ ) is on the end of edges from  $A_2(y_1, \dots, y_8), Q_1^1(u_1, \dots, u_8)$  and  $Q_2^1(v_1, \dots, v_8)$  in Figure 7. The unifier of  $P_1^1(w_1, \dots, w_7)$  with  $A_2(y_1, \dots, y_8)$  coincides with  $\lambda_{Q_3^1, A_2}$  obtained above.

The unifiers of  $P_1^1(w_1, \dots, w_7)$  with  $Q_1^1(u_1, \dots, u_8)$  and  $Q_2^1(v_1, \dots, v_8)$  are the following.

$$\lambda_{P_1^1, Q_1^1} = (w^1 \rightarrow u^2), \text{ where } u^2 = (u_1, u_2, u_3, u_4, u_5, u_7, u_8).$$

$$\lambda_{P_1^1, Q_2^1} = (w^1 \rightarrow v^2), \text{ where } v^2 = (v_1, v_2, v_3, v_4, v_5, v_7, v_8).$$

The formulas  $Q_1^1(u_1, \dots, u_8)$  and  $Q_2^1(v_1, \dots, v_8)$  with the substitutions of  $p_1^1(u^1)$  and  $p_1^1(v^1)$ , respectively, instead of sub-formulas isomorphic to  $P_1^1(w_1, \dots, w_7)$  will form the 2nd level of the network and are denoted by  $P_2^1(u_1, \dots, u_8; u^1)$  and  $P_2^2(v_1, \dots, v_8; v^1)$ .

It must be marked that if we know the value of the 1st level variable  $u^1$  then we know the values of  $(u_1, u_2, u_3, u_4, u_5, u_7, u_8)$ . That's why the essential variables of  $P_2^1(u_1, \dots, u_8; u^1)$  are  $(u_6; u^1)$ . Similarly the essential variables of  $P_2^2(v_1, \dots, v_8; v^1)$  are  $(v_6; v^1)$ .

New 2nd level predicates  $p_2^1(u^2)$  and  $p_2^2(v^2)$  defining as  $p_2^1(u^2) \Leftrightarrow P_2^1(u_1, \dots, u_8; u^1)$  and  $p_2^2(v^2) \Leftrightarrow P_2^2(v_1, \dots, v_8; v^1)$  are introduced.

Taking into account only essential variables and earlier received unifiers we have the last layer formulas.

$$A_{1,1}^2(x_7; x^1, x_1^2), \text{ where}$$

$$x^1 = (x_1, x_2, x_3, x_4, x_5, x_8, x_9, x_{10}), x_1^2 = (x_6; x^1).$$

$$A_{1,2}^2(x_7; x^1, x_2^2), \text{ where}$$

$$x^1 = (x_1, x_2, x_3, x_4, x_5, x_8, x_9, x_{10}), x_2^2 = (x_6; x^1).$$

$$A_{2,1}^2(y_1, y_1^2), \text{ where}$$

$$y^1 = (y_1, y_2, y_4, y_5, y_6, y_7, y_8), y_1^2 = (y_3, y_8; y^1).$$

$$A_{2,2}^2(y_3, y_7; y^1), \text{ where}$$

$$y^1 = (y_1, y_2, y_4, y_5, y_6, y_7, y_8).$$

$$A_3^2(z_5, z_8; z^1, z^2), \text{ where}$$

$$z^1 = (z_1, z_2, z_3, z_4, z_6, z_9, z_{10}), z^2 = (z_4, z_7; z^1).$$

The contents of the last layer cells are the following:

$$A_{1,1}^2(\bar{x}; w^1, u^2) \vee A_{1,2}^2(\bar{x}; w^1, v^2),$$

$$A_{2,1}^2(\bar{y}; w^1, u^2) \vee A_{1,2}^2(\bar{y}; w^1, v^2),$$

$$A_3^2(\bar{z}; w^1, v^2).$$

### 5.3 Recognition Block Run

The recognition block is formed according to the level description constructed in the previous subsection. The scheme of recognition block is presented in Figure 8. In this scheme only names of a checking formula and results of checking are written down.

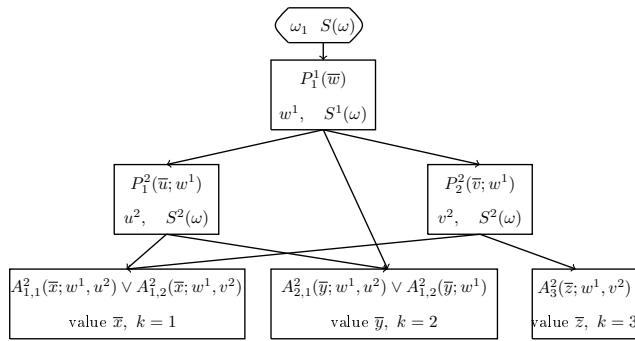


Figure 8: Predicate network.

It is obvious that if an input object  $\omega$  is isomorphic to one from the training set, then it will be recognized. It means that the number  $k$  and such a permutation  $\bar{\omega}$  of  $\omega$  that  $S(\omega) \Rightarrow A_k(\bar{\omega})$  will be found.

Let a control object with the description not isomorphic to any one from the training set is to be recognized. Such a control object  $\{d_1, \dots, d_8\}$  is presented in Figure 9.

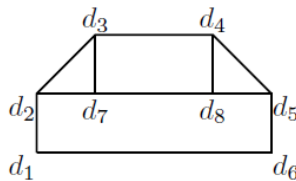


Figure 9: Control object.

Checking  $S(\omega) \Rightarrow \exists \bar{\omega} \neq P_1^1(\bar{\omega})$  in the first layer (see Figure 8) gives us  $\bar{\omega} = (d_1, d_6, d_2, d_7, d_5, d_3, d_4)$ . A new 1st level variable  $w^1$  takes this list of constants as a value. Add new atomic formula to  $S(\omega)$  and obtain  $S^1(\omega) = S(\omega) \cup \{p_1^1(w^1)\}$ .

Neither  $P_1^2(\bar{u}, w^1)$  or  $P_2^2(\bar{v}, w^1)$  are a consequence of  $S^1(\omega)$  because, in fact, 7 variables in  $\bar{u}$  and  $\bar{v}$  are changed by constants from  $w^1$ .

This control object is not recognized.

### 5.4 Retraining the Network

Let's retrain the network with the object presented in Figure 9.

Replace every constant  $d_i$  ( $i = 1, \dots, 8$ ) in its description by a variable  $x_i$  and substitute the sign  $\&$  between the atomic formulas. The formula  $A_4(\bar{x})$  is obtained.

Pairwise checking of partial sequence between  $A_4(\bar{x})$  and  $A_j(\bar{x}_j)$  ( $j = 1, 2, 3$ ) permits to obtain maximal common up to the names of arguments formulas  $Q_4^1$ ,  $Q_5^1$  and  $Q_6^1$ . The images of these fragments are shown in Figure 10.

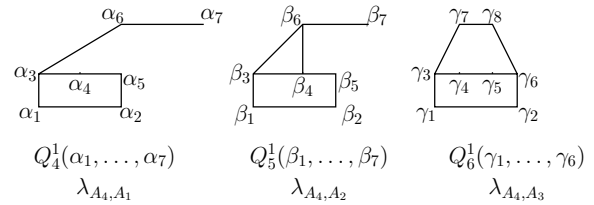


Figure 10: Pairwise extraction of maximal common up to the names of arguments sub-formulas between  $A_4(\bar{x}_4)$  and  $A_j(\bar{x}_j)$  ( $j = 1, 2, 3$ ).

One of the obtained formulas (formula  $Q_4^1(\alpha_1, \dots, \alpha_7)$ ) is isomorphic to the earlier obtained formula  $Q_3^1(w_1, \dots, w_7)$ . That's why it would not be added to the 2nd layer of the retrained network. But all its connections with the cells of another layers would be saved.

Pairwise checking of partial sequence between  $Q_5^1$ ,  $Q_6^1$  and  $Q_j^1(\bar{x}_j)$  ( $j = 1, 2, 3$ ) does not give any new formulas.

So, the 1st layer would have only the formula  $Q_3^1(w_1, \dots, w_7)$  renamed as  $P_1^1(w_1, \dots, w_7)$ .

On the 2nd layer the cell with formula  $P_4^2(\bar{\beta}; w^1)$  and output  $t^2$  and the cell with formula  $P_5^2(\bar{\gamma}; w^1)$  and output  $r^2$  would be added on the end of the edge from the cell with  $P_1^1(\bar{w})$ . Edges from these cells to the cells with  $A_j^2(\bar{x}_j; x_j^1, x_j^2)$  ( $j = 1, 2, 3$ ) would be added as it is shown in Figure 11.

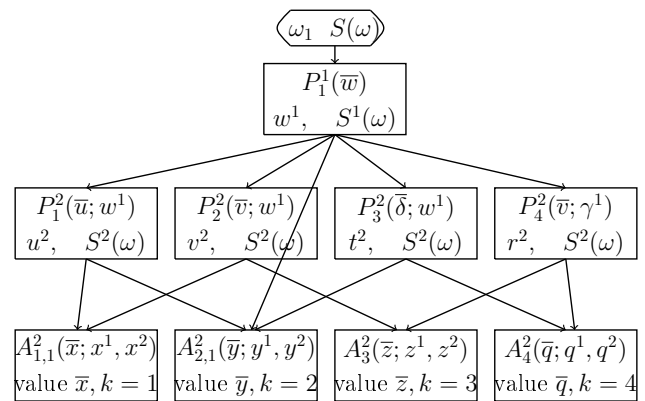


Figure 11: Pairwise extraction of maximal common up to the names of arguments sub-formulas between  $A_4(\bar{x}_4)$  and  $A_j(\bar{x}_j)$  ( $j = 1, 2, 3$ ).

The network is retrained.

It should be noted that the examples of images discussed in the article are taken from [3]. The original predicates  $W$ ,  $Y$  and  $T$  in that book for describing the images differ from the predicates  $V$  and  $L$  used here. This is due to the fact that  $W$ ,  $Y$  and  $T$  are expressed in terms of the predicates  $V$  and  $L$  and do not significantly reduce the computational complexity of recognition when constructing a level description.

Both  $\{W, Y, T\}$  and  $\{V, L\}$  are not enough to describe real images of "boxes". So, for example, the description of the image in Figure 12 is isomorphic to the description of  $A_2(y_1, \dots, y_8)$ , but you need to have a rich imagination to guess the box in it. The choice

of initial predicates in each specific problem should be chosen very carefully. also be changed.

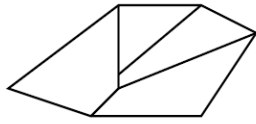


Figure 12: Is it a box?

Here we do not consider an example when, while retraining the network, not only additional cells in a layer appear, but the number of layers also changes. This could be done by retraining the network using the example of a mirror (relative to the vertical axis) image of an object from the training set. In this case, rectangle recognition will appear in the 1st layer. The contents of the remaining cells will

## 6 Fuzzy Recognition by a Logic-Predicate Network

The disadvantage is that the proposed logic-predicate network does not recognize objects with descriptions that are not isomorphic to the descriptions that participate in its formation. This can be eliminated by slight modifications of the network.

Further the  $i$ th cell of the  $l$ th level will be denoted as  $cell_i^l$ .

Let's change the contents of the network cells presented in Figure 2 by replacing the checking of  $S^{l-1}(\omega) \Rightarrow \exists \bar{x}_{i \neq}^l P_i^l(\bar{x}_i^l)$  in the  $cell_i^l$  with the **partial** sequence checking  $S^{l-1}(\omega) \Rightarrow_P \exists \bar{x}_{i \neq}^l P_i^l(\bar{x}_i^l)$ . The degree of coincidence of a recognized object with the description of a class and the degree of recognition correctness certainty is calculated during partial sequence checking.

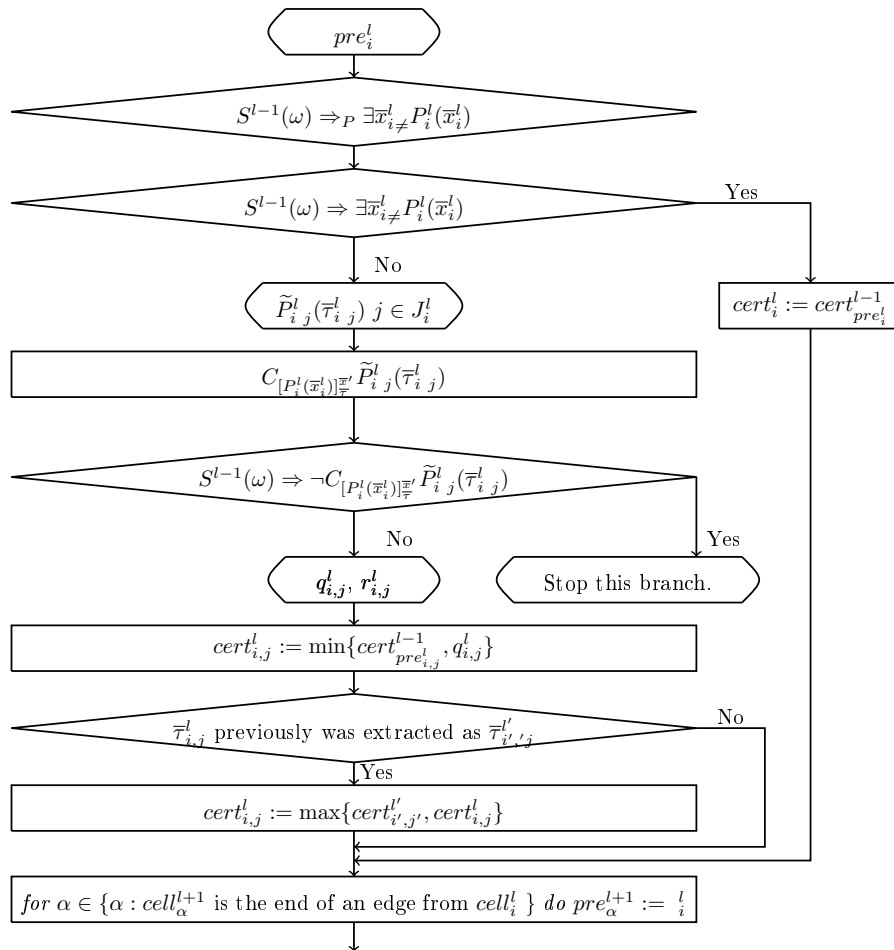


Figure 13: The fuzzy network  $i$ th cell of the  $l$ th level contents.



For a fully identified object, both of these degrees equal to 1. For an object from a given class for which these degrees are essentially small (this is determined by an expert), you can still retrain the network.

A fuzzy network has two additional parameters:

$cert_i^l$  ( $l = 0, \dots, L, i = 1, \dots, n_l$ ) – a **degree of certainty in the correctness of recognition** by the  $cell_i^l$ ;

$pre_i^l$  ( $l = 1, \dots, L, i = 1, \dots, n_l$ ) – the **cell number of a cell preceding** the  $cell_i^l$  in the current traversal of a network. The cell number has the form  $l_j$ , where  $l$  is the number of a layer and  $j$  is the number of a cell in this layer

Initially,  $cert_1^0 := 1, pre_1^1 := 1$  ( $i = 1, \dots, n_1$ ).

The scheme of  $cell_i^l$  contents of a fuzzy network is presented in Figure 13.

First of all, a partial sequence  $S^{l-1}(\omega) \Rightarrow_P \exists \bar{x}_{i \neq}^l P_i^l(\bar{x}_i^l)$  is checked in this cell. If this sequence is not partial but total, then the parameter  $cert_i^l$  is not changed and the calculation goes to the end of the cell.

Otherwise, all lists of constants  $\bar{\tau}_{i j}^l$  and maximal not contradictory on  $\bar{\tau}_{i j}^l$  with  $S^{l-1}(\omega)$  sub-formulas  $\bar{P}_{i j}^l(\bar{x}_{i j}^l)$  of the formula  $P_i^l(\bar{x}_i^l)$  are found while partial sequence checking.

If the sub-formulas  $\bar{P}_{i j}^l(\bar{x}_{i j}^l)$  is contradictory to the description  $S(\omega)$  on the list of the found constants, then this recognition branch stops.

Remind that  $S(\omega)$  and  $C_{[P_i^l(\bar{x}_i^l)]_{\bar{\tau}_{i j}^l}} \bar{P}_{i j}^l(\bar{\tau}_{i j}^l)$  are in the contradiction means that  $S(\omega) \Rightarrow \neg C_{[P_i^l(\bar{x}_i^l)]_{\bar{\tau}_{i j}^l}} \bar{P}_{i j}^l(\bar{\tau}_{i j}^l)$ , where  $C_{[P_i^l(\bar{x}_i^l)]_{\bar{\tau}_{i j}^l}} \bar{P}_{i j}^l(\bar{\tau}_{i j}^l)$  is a complement (conjunction of literals from  $P_i^l(\bar{x}_i^l)$  which are not in  $\bar{P}_{i j}^l(\bar{x}_{i j}^l)$ ) of  $\bar{P}_{i j}^l(\bar{x}_{i j}^l)$  up to  $P_i^l(\bar{x}_i^l)$ .<sup>6</sup>

Parameters  $q_{i j}^l$  and  $r_{i j}^l$ <sup>7</sup> are calculated with the use of the full form of formulas  $P_i^l(\bar{x}_i^l)$  and  $\bar{P}_{i j}^l(\bar{y}_{i j}^l)$ , i.e., with the replacement in them of each atomic formula of the levels  $l'$  ( $l' < l$ ) by the defining elementary conjunction.

Except this, a degree of certainty  $cert_i^l$  that the recognition would be valid is calculated in every cell. While first visit to the  $cell_i^l$  the value of  $cert_i^l$  will be  $cert_i^l := \min\{cert_{pre_i^l}^{l-1}, q_i^l\}$ . This corresponds to conjunction of degrees of certainty while successive passage along one branch of network traversal.

While next visit of the  $cell_i^l$  the value of  $cert_i^l$  will be  $cert_i^l := \max\{cert_i^l, \min\{cert_{pre_i^l}^{l-1}, q_i^l\}\}$ . This corresponds to disjunction of degrees of certainty while parallel passage along different branches of network traversal.

Why can we visit a cell not only once? For example, in Figure 11 the 3rd layer cell with  $A_1^2$  may be visited by a path including the cells with  $P_1^1, P_2^1$  and  $A_1^2$  and by a path including the cells with  $P_1^1, P_2^2$  and  $A_1^2$ .

In the last layer of the network, objects  $\tau_k^L$  would be received and for each object the degree of certainty  $q_k^L$ , that it is the  $r_k^L$ th part of an object satisfying the description  $A_k(\bar{x}_k)$ , would be calculated.

A flowchart of fuzzy recognition network can not be presented in the paper because it is very large. It consists of the “Scheme of logic-predicate network” shown in Figure 3, in which instead of the “Training block” the “Scheme of common up to the names of

arguments sub-formulas extraction” (Figure 1) is inserted, and the “Scheme of level recognition” (Figure 2) is inserted instead of the “Recognition block”. At the same time, in Figure 2, each element enclosed in a rhombus is replaced by “The fuzzy network  $i$ th cell of the  $l$ th level contents” (Figure 13).

## 7 Model Example of Fuzzy Recognition

Let’s look how a fuzzy network (corresponding to that constructed in the section 6 and presented in Figure 8) recognizes the control object presented in Figure 9.

To recognize it, a partial sequence  $S(d_1, d_2, \dots, d_8) \Rightarrow_P \exists w_1 \dots w_6 P_1^1(w_1, \dots, w_6)$  is checked in the **1st layer**. There is a total sequence and the 1st level variable  $w^1 = (w_1, w_2, w_3, w_4, w_5, w_6, w_7)$  takes the value  $(d_1, d_6, d_2, d_7, d_5, d_3, d_4)$ .

$cert_1^1 := 1, pre_2^2 := 1, pre_2^2 := 1, pre_3^3 := 1$ .

The result of the 1st layer fuzzy cell run is presented in Figure 14.<sup>8</sup>

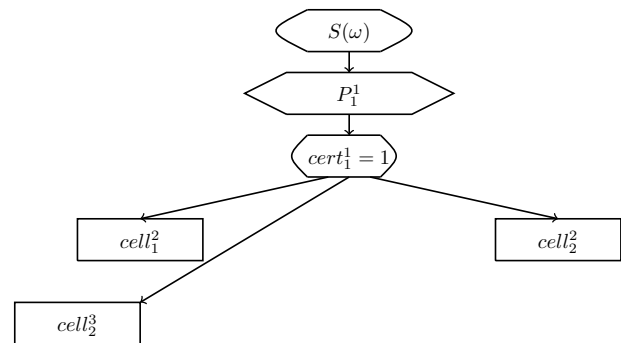


Figure 14: The result of the 1st layer fuzzy cell run.

In the **2nd layer**, partial sequence  $S^1(d_1, d_2, \dots, d_8, w^1) \Rightarrow_P \exists u_1 \dots u_7 \neq P_1^2(u_1 \dots u_8; w^1)$  is checked. It must be noted that in fact  $S^1(d_1, d_2, \dots, d_8, w^1) \Rightarrow_P \exists u_6 P_1^2(d_1, d_6, d_2, d_7, d_5, u_6, d_3, d_4; w^1)$  is checked. That is, there is only one variable at the right-hand side of the formula.

Since we need to check formulas for consistency, we note that  $V(x, y, z) \Rightarrow \neg V(x, z, y)$ .

There is only a partial sequence. A formula  $\bar{P}_1^2(d_1, d_6, d_2, d_7, d_5, d_3, d_4; w^1)$  is the maximal sub-formula such that  $S^1(d_1, d_2, \dots, d_8, w^1) \Rightarrow \bar{P}_1^2(d_1, d_6, d_2, d_7, d_5, d_3, d_4; w^1)$ . Its complement up to  $P_1^2(d_1, d_6, d_2, d_7, d_5, d_3, d_4, u_6; w^1)$  contains three literals with predicate  $V$ , one of them is  $V(u_6, d_4, d_6)$ . This contradicts  $S^1(d_1, d_2, \dots, d_8, w^1)$  containing  $V(u_6, d_6, d_4)$ . This branch of the network stops.

There is another cell in the 2nd layer on the end of an edge from  $cell_1^1$ . This is  $cell_2^2$ .

Check partial sequence  $S^1(d_1, d_2, \dots, d_8, w^1) \Rightarrow_P \exists v_1 \dots v_8 P_2^2(v_1 \dots v_8; w^1)$ . In fact,  $S^1(d_1, d_2, \dots, d_8, w^1) \Rightarrow_P \exists v_6 P_2^2(d_1, d_6, d_2, d_7, d_5, v_6, d_3, d_4; w^1)$  is checked.

<sup>6</sup>See the definitions of a complement and contradictory formulas in the section 2.

<sup>7</sup>Here  $q_{i j}^l$  and  $r_{i j}^l$  are ratios of the number of literals and variables, respectively, in  $\bar{P}_{i j}^l(\bar{x}_{i j}^l)$  and  $P_i^l(\bar{x}_i^l)$ . See definition of  $(q, r)$  fragment in the section 2.

<sup>8</sup>Not to overload the following schemes, only the name of the formula is written in the cell instead of the formula of corresponding partial sequence.

There is only partial sequence. A formula  $\tilde{P}_2^2(d_1, d_6, d_2, d_7, d_5, v_6, d_3, d_4; w^1)$  is the maximal sub-formula such that  $S^1(d_1, d_2, \dots, d_8, w^1) \Rightarrow \tilde{P}_2^2(d_1, d_6, d_2, d_7, d_5, v_6, d_3, d_4; w^1)$ . Its complement up to  $P_2^2(d_1, d_6, d_2, d_7, d_5, v_6, d_3, d_4; w^1)$  contains 5 literals with predicate  $V$

$V(v_6, d_7, d_3), V(d_3, v_6, d_2), V(d_7, v_6, d_2), V(d_7, d_8, v_6), V(d_7, d_5, v_6)$  and 3 literals with predicate  $L$

$L(d_7, d_2, d_8), L(d_8, d_2, d_5), L(d_8, d_7, d_5)$ , but no one of them is in contradiction with  $S^1(d_1, d_2, \dots, d_8, w^1)$ . The value of  $v^2$  is  $(v_6; w^1)$ . It means that one 1st level variable  $v_6$  does not take value.

The number of literals in  $P_2^2(v_1 \dots v_8; w^1)$  (taking into account its full form in which the 1st level predicate is changed by the defining formula only with initial variables) is  $a_2^2 = 16$ . The number of literals in  $\tilde{P}_2^2(d_1, d_6, d_2, d_7, d_5, v_6, d_3, d_4; w^1)$  is  $\tilde{a}_2^2 = 8$ .

That's why  $q_2^2 = \frac{8}{16} = \frac{1}{2}$  and  $cert_2^2 := \min\{1, \frac{1}{2}\} = \frac{1}{2}$ .

There are two edges from the cell with  $P_2^2$  to the 3rd layer cells with  $A_{1,2}^2$  and with  $A_3^2$ . That's why  $pre_1^3 := \frac{2}{2}, pre_3^3 := \frac{2}{2}$ .

The result of the 2nd layer fuzzy cells run is presented in Figure 15.

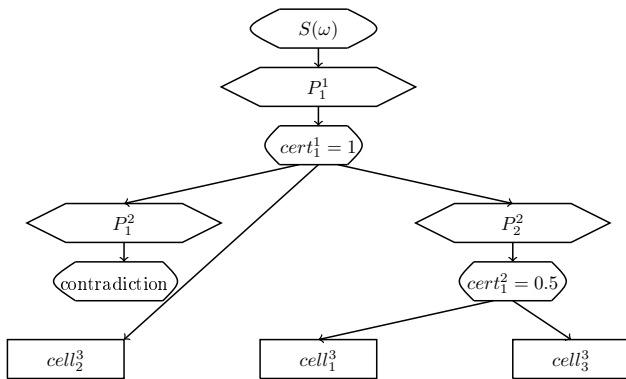


Figure 15: The result of the 2nd layer fuzzy cell run.

In the **3rd layer** while going from the cell with  $P_1^1$  to the  $cell_2^3$  with  $A_{2,1}^2$  partial sequence  $S^1(d_1, \dots, d_8; w^1) \Rightarrow \exists y_1 \dots y_8 \neq A_{2,1}^2(y_1 \dots y_8; v^1)$  is checked. In fact, a partial sequence  $S^1(d_1, \dots, d_8; w^1) \Rightarrow \exists y_3 A_{2,1}^2(d_1, d_6, y_3, d_2, d_7, d_5, d_3, d_4; v^1)$  with the only one initial variable  $y_3$  is checked.

There is only a partial sequence. The same string of values for variables in  $v^2$ , as it was in the second layer, partially satisfies  $A_{1,2}^2(d_1, d_6, y_3, d_2, d_7, d_5, d_3, d_4; v^1)$ . But the description of this fragment has two additional literals  $V(d_7, d_2, d_3)$  and  $V(d_7, d_3, d_5)$ .

The number of literals in  $A_2(y_1 \dots y_8; v^1)$  is  $a_2 = 22$ . The number of literals in  $\tilde{A}_{1,2}^2(y_1 \dots y_8; v^1, v^2)$  is  $\tilde{a}_{1,2}^2 = 10$ . That's why  $q_2^3 = \frac{10}{22} \approx 0.455$  and  $cert_2^3 := \min\{1, \frac{10}{22}\} \approx 0.455$ .

While going from the cell with  $P_2^2$  to the  $cell_1^3$  with  $A_{1,2}^2$  partial sequence  $S^2(d_1, d_2, \dots, d_8; v^1, v^2) \Rightarrow \exists v_1 \dots v_{10} \neq A_{1,2}^2(v_1 \dots v_{10}; v^1, v^2)$  is checked. In fact  $S^2(d_1, d_2, \dots, d_8; v^1, v^2) \Rightarrow \exists v_6, v_7, v_8 \neq A_{1,2}^2(d_1, d_6, d_2, d_7, d_5, v_6, v_7, v_8, d_3, d_4; v^1, v^2)$  with 3 initial variables  $v_6, v_7, v_8$  is checked.

There is only a partial sequence. The same string of values for variables in  $v^2$ , as it was in the second layer, partially satisfies

$A_{1,2}^2(v_1 \dots v_{10}; v^1, v^2)$ .

The number of literals in  $A_1(x_1, \dots, x_{10})$  is  $a_1 = 34$ . The number of literals in  $\tilde{A}_{1,2}^2(v_1 \dots v_{10}; v^1, v^2)$  is  $\tilde{a}_{1,2}^2 = 8$ . That's why  $q_1^3 = \frac{8}{34} \approx 0.235$  and  $cert_1^3 := \min\{1, \frac{8}{34}\} \approx 0.235$ .

While going from the cell with  $P_2^2$  to the  $cell_2^3$  with  $A_3^2$  partial sequence  $S^2(d_1, d_2, \dots, d_8; v^1, v^2) \Rightarrow \exists z_1 \dots z_{10} \neq A_3^2(z_1 \dots z_{10}; v^1, v^2)$  is checked. In fact  $S^2(d_1, d_2, \dots, d_8; v^1, v^2) \Rightarrow \exists z_7, z_8 \neq A_3^2(d_1, d_6, d_2, d_7, d_8, d_5, z_7, z_8, d_3, d_4; v^1, v^2)$  with 2 initial variables  $z_7, z_8$  is checked.

There is only a partial sequence. The variables  $z_7, z_8$  have not taken values. All literals with these variables are absent in  $\tilde{A}_3^2(d_1, d_6, d_2, d_7, d_8, d_5, z_7, z_8, d_3, d_4; v^1, v^2)$ . The number of such literals is 14.

The number of literals in  $A_3^2(z_1 \dots z_{10})$  is  $a_3 = 32$ . The number of literals in  $\tilde{A}_3^2(z_1 \dots z_{10}; v^1, v^2)$  is  $\tilde{a}_3^2 = 32 - 14 = 18$ . That's why  $q_2^3 = \frac{18}{32} \approx 0.563$  and  $cert_2^3 := \min\{1, \frac{18}{32}\} \approx 0.563$ .

The full degree of certainty that it is a representative of the class of "boxes" that is given as a control image is  $cert = \max\{\frac{8}{34}, \frac{10}{22}, \frac{18}{32}\} = \frac{18}{32} \approx 0.563$ .

Moreover,  $\frac{7}{8}$  elements of this image coincide with  $\frac{8}{10}$  elements of the image  $\omega_3$  from the training set.

The result of the fuzzy network run is presented in Figure 16.

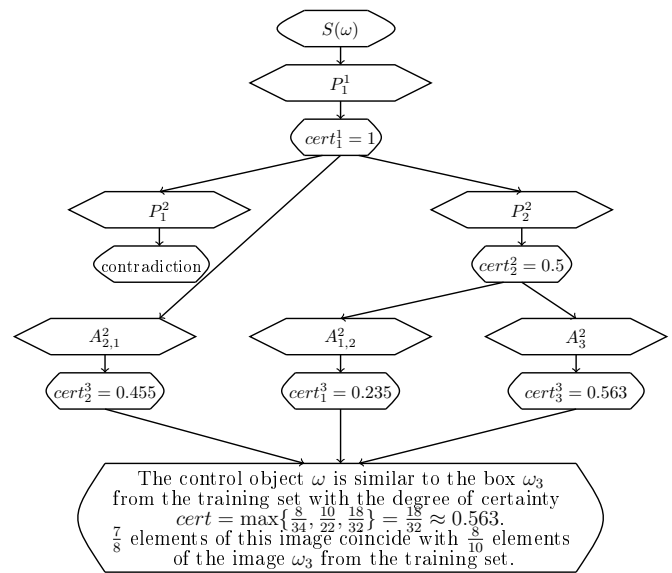


Figure 16: The result of the fuzzy network run.

## 8 Results and Discussion

The main result of this article is the proposal of fuzzy recognition of complex structured objects by logic-predicate network. For a detailed presentation of this result, it was necessary to describe in detail a logic-predicate approach to the recognition of complex structured objects.

A significant disadvantage of this approach is that the problems arising are NP-hard, which to a large extent "scares off" some researchers. On the other hand, the use of predicate logic does

not always seem legitimate, since the propositional logic is much simpler for understanding by non-mathematicians.

The use of level descriptions is not something essentially original to decrease the computational complexity of the main problem. It is a well-known technique in which sufficiently complex problems are divided into a sequence of problems of the same type of lower dimension. The main problem was HOW to build such a description. To this end, the mathematical apparatus for extracting the general sub-formulas of two elementary conjunctions of predicate formulas was developed and briefly presented in this paper.

Based on the level description of classes, it was possible to formulate the concept of a logic-predicate network. Unlike existing networks, for example, neural or Bayesian, when retraining such a network, a logic-predicate network can (and, as a rule, this happens) change its configuration. This corresponds to the training of a person in whose brain connections between one neural cells break and connections between others arise.

This is one of the essential features of the proposed approach, which distinguishes it from the use of various other networks. Its other essential feature is to calculate the degree of coincidence of the object being recognized with those ones for which the network was trained.

A trained logic-predicate network quickly recognizes objects with descriptions that are isomorphic to the ones on which it was trained. But its training and retraining processes remain to be NP-hard. This corresponds to the process of teaching a person. But in some cases, it is not necessary to accurately recognize an object or process, but to find out what it looks like and to what extent.

To answer such a question, the concept of a fuzzy logic-predicate network was proposed. The structure of its cell, shown in Figure 13, is quite complex and contains a partial sequence check. Such verification is NP-hard, but after repeated retraining of the network, the formulas  $P_i^l(\bar{x}_i^l)$ , setting intermediate goal formulas, are quite short and have a small number of essential arguments, as it was shown by the example in section 7. This means that an exponent of the computational complexity of the partial sequence check is a fairly small value.

Special importance has the choice of “good” initial features. As already mentioned in Subsection 5.4, the predicates  $W$ ,  $Y$  and  $T$ , which were used to describe contour images in [3], do not give a good decreasing of computational complexity for the logic-predicate network, constructed from them. Apparently, this is due to the fact that they reflect the perception of a person and are essentially composite.

## 9 Conclusion

In development of the proposed approach, many-valued predicates may be considered. Especially this relates to the properties of an object elements. Such multi-valued predicates  $p(x) = a$  with  $a \in D$  can be interpreted as binary ones with a numerical argument:  $q(x; a) \Leftrightarrow p(x) = a$ . The inclusion of such predicates in the formulation of the problem may serve as the subject of further research.

It is also planned to consider the possibility of introducing cell weights in a logic-predicate network. Such weights can be deter-

mined depending on the frequency with which parts of an object description satisfy the formula checked in the corresponding cell. Other weighting strategies are possible, similar to strategies for determining the cell weights of a classical artificial neural network.

In addition, if in the problem setting global features of the object that characterize the entire object are considered, then such an object is usually described by a binary or multi-valued string. Class descriptions can be presented as a propositional formula in disjunctive normal form. Extraction the “frequently encountered” sub-formulas of such propositional formulas [24] allows to form a level description of classes. Using this description in a manner similar to that described in this article, it is possible to build a logic-algebraic network, which changes its configuration in the process of learning. This corresponds to a change of connections in biological neural networks.

**Acknowledgments** The author is grateful to Vakhtang Kvaratskhelia from Muskhelishvili Institute of Computational Mathematics of the Georgian Technical University, whose question at the conference in 2017 allowed this paper to be born. The author thanks the anonymous referees for their useful suggestions.

## References

- [1] T. Kosovskaya . “Implementation of Formula Partial Sequence for Rough Solution of AI Problems in the Framework of the Logic-Predicate Approach” in 2019 Computer Science and Information Technologies (CSIT), Yerevan, Armenia, 2019. <https://doi.org/10.1109/CSITechnol.2019.8895153>
- [2] N. J. Nilson, Problem-solving methods in Artificial Intelligence, McGraw-Hill Book Company, 1971.
- [3] R. Duda, P. Hart, Pattern Classification and Scene Analysis, John Wiley & Sons, 1973.
- [4] E.B. Hunt, Artificial Intelligence, Academic Press, 1975.
- [5] M.R. Garey, D.S Jonson, Computers and Intractability: A Gide to the Theory of NP-Completeness, W.H. Freeman and Company, 1979.
- [6] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, Third edition, Prentice Hall Press Upper Saddle River, 2009.
- [7] T.M. Kosovskaya. “Estimating the number of steps that it takes to solve some problems of pattern recognition which admit logical description” in VESTNIK ST PETERSBURG UNIVERSITY-MATHEMATICS, **40**(4), 82–90, 2007. (In Russian) <https://doi.org/10.3103/S1063454107040061>
- [8] T.M. Kosovskaya, “Multilevel descriptions of classes for decreasing of step number of solving of a pattern recognition problem described by predicate calculus formulas” VESTNIK SANKT-PETERBURGSKOGO UNIVERSITETA SERIYA 10 PRIKLADNAYA MATEMATIKA INFORMATIKA PROTSESSY, **1**, 64–72 (2008). (In Russian)
- [9] T.M. Kosovskaya, “Partial deduction of predicate formula as an instrument for recognition of an object with incomplete description” VESTNIK SANKT-PETERBURGSKOGO UNIVERSITETA SERIYA 10 PRIKLADNAYA MATEMATIKA INFORMATIKA PROTSESSY, **1**, 74–84, 2009. (In Russian) <https://doi.org/10.24412/FgeTsEutG6w>
- [10] T.M. Kosovskaya, “Partial Deduction in Predicate Calculus as a Tool for Artificial Intelligence Problem Complexity Decreasing” in 2015 IEEE 7th International Conference on Intelligent Computing and Information Systems, ICICIS 2015, 73–76, 2016. <https://doi.org/10.1109/IntelCIS.2015.7397199>
- [11] J. Schmidhuber, “Deep learning in neural networks: An overview” Neural Networks, **61**, 85–117, 2015. <https://doi.org/10.1016/j.neunet.2014.09.003>
- [12] T. Chunwei, X. Yong, Z. Wangmeng, “Image denoising using deep CNN with batch renormalization” Neural Networks, **121**, 461–473, 2020. <https://doi.org/10.1016/j.neunet.2019.08.022>
- [13] E.H. Dawn, C. J. Lakhmi (Eds), Innovations in Bayesian Networks. Theory and Applications, Springer-Verlag, 2008. <https://doi.org/10.1007/978-3-540-85066-3>

- [14] R. Agrahari, A. Foroushani, T.R. Docking et al. "Applications of Bayesian network models in predicting types of hematological malignancies", *Sci Rep* 8, 6951 (2018). <https://doi.org/10.1038/s41598-018-24758-5>
- [15] A. Tulupiev, C. Nikolenko, A. Sirotkin, *Bayesian Networks: logic-probabilistic approach*, Nauka, 2006. (in Russian)
- [16] A. Shahzad, R. Asif, A. Zeeshan, N. Hamad, A. Tauqeer, "Underwater Resurrection Routing Synergy using Astucious Energy Pods" *Journal of Robotics and Control (JRC)*, 1(5), 173–184, 2020. <https://doi.org/10.18196/jrc.1535>
- [17] T. Kosovskaya, "Self-modificated predicate networks" *International Journal on Information Theory and Applications*, 22(3), 245–257, 2015.
- [18] T. Kosovskaya, "Fuzzy logic-predicate network" in *Proceedings of the 11th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2019)*, Part of series ASUM (Atlanties Studies in Uncertainty Modeling), 1, 9–13, 2019. <https://doi.org/10.2991/eusflat-19.2019.2>
- [19] T. Kosovskaya, "Predicate Calculus as a Tool for AI Problems Solution: Algorithms and Their Complexity", 1–20. 2018. Chapter 3 in "Intelligent System", open access peer-reviewed Edited volume. Edited by Chatchawal Wongchoosuk Kasetsart University <https://doi.org/10.5772/intechopen.72765>
- [20] T. M. Kosovskaya, N. N. Kosovskii, "Polynomial Equivalence of the Problems Predicate Formulas Isomorphism and Graph Isomorphism" *VESTNIK ST PETERSBURG UNIVERSITY-MATHEMATICS*, 52(3), 286–292, 2019. <https://doi.org/10.1134/S1063454119030105>
- [21] T.M. Kosovskaya, "Some artificial intelligence problems permitting formalization by means of predicate calculus language and upper bounds of their solution steps" *SPIIRAS Proceedings*, 14, 58–75, 2010. (In Russian) <https://doi.org/10.15622/sp.14.4>
- [22] T. M. Kosovskaya, D. A. Petrov, "Extraction of a maximal common sub-formula of predicate formulas for the solving of some Artificial Intelligence problems" *VESTNIK SANKT-PETERBURGSKOGO UNIVERSITETA SERIYA 10 PRIKLADNAYA MATEMATIKA INFORMATIKA PROTSESSY*, 3, 250–263, 2017. (In Russian) <https://doi.org/10.21638/11701/spbu10.2017.303>
- [23] T. Kosovskaya, "Construction of Class Level Description for Efficient Recognition of a Complex Object" *International Journal "Information Content and Processing"*, 1(1), 92–99, 2014.
- [24] D.A. Koba, "Algorithm for constructing a multi-level description in pattern recognition problems and estimating the number of its run steps" in *Proceedings of the XVII International School-Seminar "Synthesis and Complexity of Control Systems" named after Acad. O.B. Lupanov (Novosibirsk, October 27 - November 1, 2008)*, Publishing House of the Institute of Mathematics, Novosibirsk, 55–59, 2008. (In Russian)



# Deep Learning Approach for Automatic Topic Classification in an Online Submission System

Tran Thanh Dien, Nguyen Thanh-Hai, Nguyen Thai-Nghe\*

College of Information and Communication Technology, Can Tho University, Can Tho city, 900000, Vietnam

---

## ARTICLE INFO

Article history:

Received: 8 July, 2020

Accepted: 11 August, 2020

Online: 28 August, 2020

---

Keywords:

Article classification

Deep learning

Knowledge search

Online submission system

---

---

## ABSTRACT

Topic classification is a crucial task where knowledge categories exist within hierarchical information systems designed to facilitate knowledge search and discovery. An application of topic classification is article (e.g., journal/conference paper) classification which is very useful for online submission systems. In fact, numerous online journals/magazine submission systems usually receive thousands of article submissions or even more for each month. This leads to a huge amount of time-consumption of editors to process and categorize the submissions aiming to look for and assign appropriate reviewers to the submitted articles. In this study, we propose an approach based on natural language processing techniques and machine learning algorithms (both classic machine learning and deep learning) to automatically classify the topics of articles in an online submission system. We show by promising performance collected from prediction tasks to present that the proposed method is a potential approach for applying to the real system.

---

## 1 Introduction

With the rapid development of data sources and computational algorithms, the classification tasks, especially in text classification, have revealed an important role in numerous fields [1]. Text classification is a supervised learning technique that has been deployed popularly in practical cases [2]. More specific, text classification tasks are classical text processing problems attempting to categorize an unseen text into a group of known texts based on its similarities to the considered group [3]. The authors in [4] stated that text classification tasks are the assignment or categorization of labels on a new text based on the similarities of the considered text to the labeled texts in the training set. The framework using text classification automatically allows information to be processed and searched easier. Furthermore, sorting each of them takes a lot of time and effort with a large number of texts, not to mention the possibility of inaccurate categorization due to the subjectivity of the people. From the proposed studies, there are numerous real applications of text classification tasks including news classification by topics in online newspapers, knowledge management, spam email filtering, and supportable tools for search engines on the Internet, etc. [1].

The text classification tasks are attracting numerous scientists with a vast of the studies proposed with different algorithms including machine learning as well as mathematical-statistical model. In recent years, machine learning has been widely implemented and in-

vestigated with numerous advancements. Many learning algorithms include k-nearest neighbors (kNN), Naïve Bayes, support vector machines (SVM), decision tree, and artificial neural network, etc. which are leveraged to solve text classification problem with the published studies in [5–12].

An important application of text classification is article/journal classification where not only the authors but also the editorial boards of the magazines/journals would like to know how to classify a document into a relevant topic of those magazine/journal to search or/and submit. More specifically, the submission system enables us to disseminate texts and extract relevant information automatically when a manuscript is submitted to the system. From the predicted category, editors can find appropriate reviewers for the submission faster and help to speed up the review process.

This paper is an extension of work originally presented in the International Conference on Advanced Computing and Applications (ACOMP) 2019 in Nha Trang, Vietnam [13]. We present an approach using machine learning algorithms to classify automatically the articles which were submitted via an online submission system. More specifically, when we submit an article (the extension can be doc(x) or pdf, etc.) to the online submission system, the system not only extract automatically the information on the author, title, and abstract but stratify and assign the topic to the submission. For this procedure, we can use the natural language processing techniques to pre-process the data before fetching them into a machine learning

---

\*Corresponding Author: Nguyen Thai-Nghe, Can Tho city, Vietnam, Contact: +84918028402 & email: ntnghe@cit.ctu.edu.vn

algorithm to do the classification tasks. Comparing to the work in [13], this study provides contributions as follows:

- After pre-processing the data, we proposed to use Deep Learning (MLP) for the classification tasks. Experimental results show that by using this approach, the results even get better than using the SVM in our previous work.
- For experiments, we have collected 5 data sets (instead of 2 Vietnamese data sets as in [13]). These data sets represent for 3 languages (English, Turkey, and Vietnamese) and have multi-classes (2, 4, 6, 9, and 10 classes as presented in Table 1)
- In this work, since the topic classes are imbalanced, we have used AUC (Area Under the ROC Curve) as a metric for comparison. Previous works show that when the data sets are imbalanced, the AUC metric is a better measurement for evaluation [14, 15]. We will analyze this point in the result section.
- For faster experiments, the pre-processing data steps are combined to a single procedure as presented in Algorithm 1.
- We have reviewed more and up-to-date related works.

In the next sections of the paper, we present a literature review on robust machine learning algorithms used in text classification tasks in Section 2. In Section 3, we introduce the proposed framework including classification models and steps for pre-processing. Section 4 describes the empirical results and Section 5 provides insightful remarks of the study.

## 2 Related works

In the context of the enormous development of digital information in recent years, text mining techniques hold a crucial role in information and knowledge management and mining, attracting the attention of scientists [1]. Text classification through modern techniques is the division of a dataset including documents into two or more topics. The text classification purposes to assign a predefined label or category to a document. For example, a new article published on an online newswire system can be assigned to one of the given topics while each submission sent to an online journal system can be automatically stratified into its topics, etc.

### 2.1 Related works on text classification

Numerous papers have proposed methods and learning architectures on text classification to enhance performance and apply in practical cases. For example, authors in [16] employed the Chi-square feature selection (referred to, hereafter, as ImpCHI) to enhance the classification performance for Arabic documents classification. They also compared this improved chi-square with three traditional features selection metrics namely mutual information, information gain and Chi-square. Another study in [17] introduced a new firefly algorithm based feature selection method which achieved a precision value equals to 0.994 on an Open Source Arabic Corpora (OSAC) dataset.

Some techniques for data pre-processing phase have been introduced in numerous studies. Authors in [18] presented the maximum matching segmentation (MMSEG) algorithm to segment words. When the segmentation completed, the text was fetched into a vector form, with the vectorized TF\*IDF. Next, data were classified using decision trees and Support Vector Machines with the Weka package. The experimental results of [18] were evaluated on the dataset with 7,842 texts categorized in 10 different topics. About 500 texts of each topic were selected randomly for training phase and the remaining were used to verify independence. As reported in the paper, the performance with Support Vector Machines algorithm was greater than the algorithm of decision tree. Besides, the authors deployed singular value decomposition to analyze and reduce the characteristic space dimension and hence, it can improve classification performance of Support Vector Machines algorithm.

Another work [19] studied the semantic relation extraction and classification in scientific paper abstracts. The authors presented the steps of setup procedures and experimental results of semantic relation extraction and classification on scientific papers datasets. The task included three sub-tasks: the first classification was performed on the clean data while the second one was on noisy data, and the final task combined extraction and classification scenario. Some datasets which were used in the challenges such as a subset of abstracts of published papers in the ACL Anthology Reference Corpus, annotated for domain specific entities and semantic relations were introduced in this study. Root Cause Analysis of Incidents using Text Clustering was proposed in [20]. The authors studied the use of two machine learning (ML) algorithms, namely random forest (RF), and support vector machine (SVM) and found that SVM performed best in classifying the accident narratives. Multinomial Naive Bayes (MNB), Logistic Regression (LR), Support Vector Machines were also deployed in [21] to prediction task on Twitter Data.

Approaches based on deep learning techniques are also carried out in numerous studies. The authors in [22] presented A Comprehensive Review on deep learning techniques for text classification. The authors in [23] proposed three fundamental architectures of deep learning models for the tasks of text classification: Deep Belief Neural (DBN), "Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). The work also introduced basic guidance about the deep learning models that which models is best for the task of text classification. The authors [24] presented a novel angle to further improve this representation learning, i.e., feature projection, and projected existing features into the orthogonal space of the common features. Another paper is [25] which illustrated a model of statistics like TF-IDF, to exploit pre-trained SOTA DL models (such as the Universal Sentence Encoder) without any need for traditional transfer learning or any other expensive training procedure on Text Classification tasks of UNGA Resolutions. The authors in [26] introduced a model with Gated recurrent unit (GRU) and support vector machine. The method implemented a linear support vector machine (SVM) as the replacement of Softmax in the final output layer of a GRU model to obtain comparative, remarkable results. The work in [27] introduced a new improved algorithm of the original Sine Cosine Algorithm for feature selection with the software written in Matlab2013 to achieve high performance on the datasets of Reuters-21578 collection<sup>1</sup>. Some studies [28, 29] have

<sup>1</sup>D.D., 2004. Reuters-21578. Retrieved November 6, 2013, <http://www.daviddlewis.com/0Aresources/testcollections/reuters21578/>

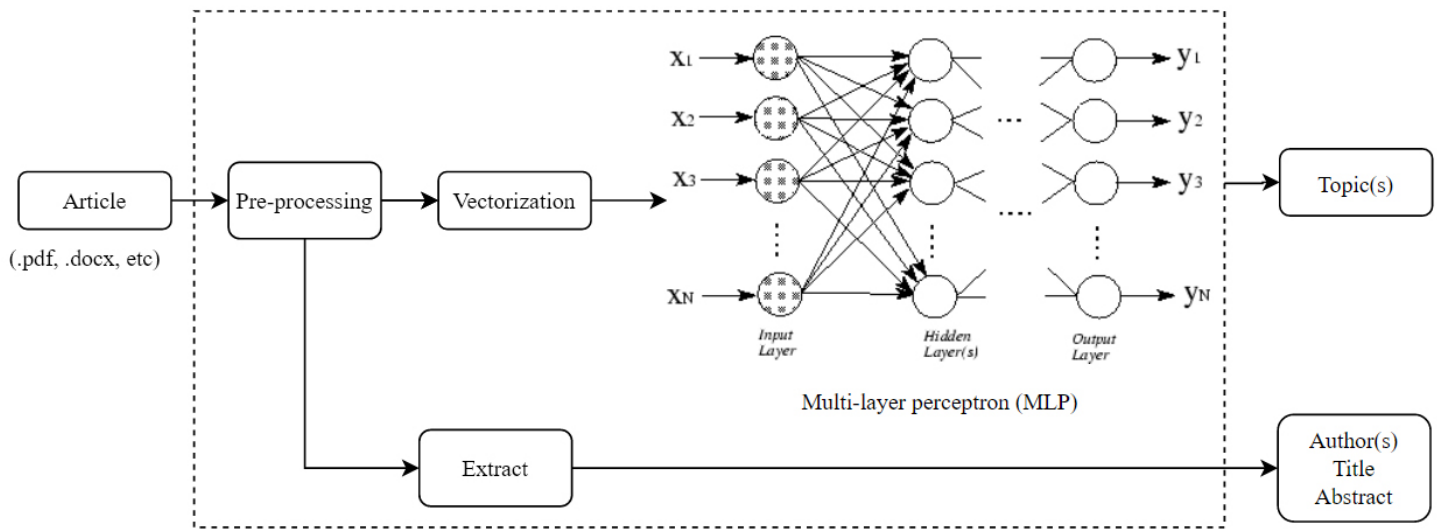


Figure 1: The proposed architecture for pre-processing and classification of articles

attempted to use Reinforcement Learning for text classification.

### 2.2 Text vectorization

In real word, some text representation models have been proposed including vector space model, bag of words model, and graph-based model. The vector space model [30] is deployed in this work because it can represent unformatted documents as simple and formulaic notation. Due to its advantages, many researches based on vector space model are implemented [31]. According to the details of this method, each document or article is represented as a vector while each component of such vector is a separated term that is assigned to a value namely “weight” of that term.

TF measures the frequency of a word in a document. It depends on the document length and the generality of word. For measuring the weight of a word, the number of occurrences of the word is divided by the length of the document (the number of words) as equation 1:

$$TF(t,d) = \frac{\text{number of occurrences of term } t \text{ in } d}{\text{Total number of terms in } d} \quad (1)$$

There is difference between TF and IDF. TF counts frequency of a term t in document d, where as DF counts occurrences of term t in the document set N. For calculating the TF, all the terms reveal the same importance. However, it is found that not all terms in a dataset are important, for instance, connecting terms, determiners; and prepositions. It is necessary to reduce the importance of such terms with computing IDF by the formula as Equation 2:

$$IDF(t,D) = \log \frac{\text{Total document in } D}{\text{Number of document including } t} \quad (2)$$

We compute TF\*IDF which integrate between TF and IDF. The method calculates the TF\*IDF value of a term via its importance in a document belonging to a document set. Using such method, we are able to filter out common words and retain high value words as

equation 3.

$$TF * IDF(t,d,D) = TF(t,d) \times IDF(t,D) \quad (3)$$

In this study, we propose an approach of automated classification of articles submitted via an online submission system. When the submission (with extension such as \*.doc(x), \*.pdf, etc.) is uploaded, the system, then, extracts the author’s information and abstract to stratify the submission.

## 3 Proposed method

The proposed overall system including extracting information, pre-processing data methods and categorizing articles is exhibited in Fig. 1.

In proposed model, when a new article under formatting of .docx, .pdf, etc. is submitted to the system, its information consisting of abstract, author(s), title can be automatically extracted, and especially categorized into an appropriate topic based on the previous data trained by machine learning models. Due to article’s pre-formatted template, article’s information extraction is easy to get. Therefore, this work only focuses on how to classify article’s topics when it is submitted to the system.

In the following sections, we will present how to pre-process the data and setting up the classification models.

### 3.1 Data pre-processing

Data pre-processing described as Algorithm 1. This algorithm receives a document as input, then the document is converted and normalized with techniques such as changing to lower cases, removing blanks, etc. The algorithm also separates documents into words and eliminate noises then transform to the vector. The details are in Algorithm 1.

**File format conversion and word standardization:** If the documents of datasets have the extension of .doc(x), they will be converted to plain text (.txt) for easy use in most of classification mod-

els. After converting file format, word standardization performs to convert all text characters into lowercase while spaces are also eliminated.

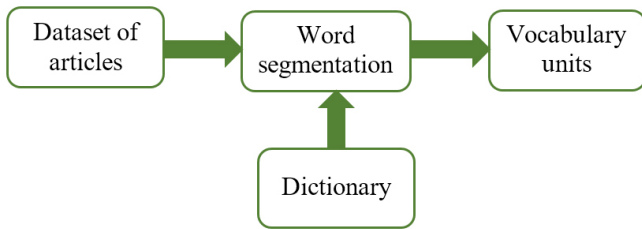


Figure 2: Process of word segmentation

---

#### Algorithm 1: Pre-processing for documents

---

**Data:** InputDocument  $d$

**Result:** Vectorized-Documents  $D$

- 1 *Conversion*( $d$ ): convert the input document (word/pdf) to text
  - 2 *WordNormalization*( $d$ ): changed to lower cases, removing blanks
  - 3 *WordSegmentation*( $d$ ): separate document to words
  - 4 *RemovingStopWords*( $d$ ): remove noise words
  - 5 *Vectorization*( $d$ ): convert documents to respectively vectors
  - 6 **Return** sets of Vectorized-Documents  $D$
- 

**Word segmentation:** In some languages including Vietnamese, spaces do not segment words but separate syllables so the segmentation is an important phase in NLP. There are a lot of tools that have been successfully developed to segment words with relatively high accuracy. For instance, the *VnTokenizer* segmentation tool<sup>2</sup> was used for Vietnamese word segmentation. It was based on the integrated methods of maximum matching, weighted finite-state transducer and regular expression parsing, running on the dataset of Vietnamese syllabary and Vietnamese vocabulary dictionary. This tool segments Vietnamese documents into vocabulary units (words, names, numbers, dates and other regular expressions) with over 95% accuracy described in Fig. 2

**Removing stop words:** As mentioned in [32] that stop words are the words that widely appear in all documents of the considered dataset, or the words that appear only in one and several documents. Therefore, they do not contain useful information or make sense. For tasks of text classification, the appearance of such types of words not only do not help examine the classification but also cause noises leading to a decrease in performance of the classification process.

**Text vectorization** In this study, we use TF-IDF (term frequency-inverse document frequency) as a statistical measure that evaluates how relevant a term is to a document in a collection of documents.

**Text classification algorithms** From a set of documents  $\{d_1 \dots d_n\}$  called a training set, in which document  $d_i$  is labeled under  $c_j$  belonging to set of categories  $C = \{c_1 \dots c_m\}$ , classification model is determined for classifying any document  $d_k$  into an appropriate category of set  $C$ .

In our experiments, some text classification algorithms which are implemented for comparison include SVM algorithms, tree

decision and deep learning techniques.

### 3.2 Deep learning model

The proposed model is presented in Fig. 1 where the input attributes are selected from Table 1 and the output (prediction) of the model including classes depending the selected dataset. The proposed MultiLayer Perceptron (MLP) architecture includes one hidden layer with 16 neurons (see an illustration in Figure 3) which is conducted from fine-tune hyperparameters experiments (see Section 4.2) running  $n$  hidden layer(s) with various ( $m$ ) numbers of neurons.

In order to investigate the difference in the performance, starting at one hidden layer, we increase the number of neurons from 2 to 128. Each multiplies two times the preceding one ( $2^n$  with  $n = 1..7$ ). When we obtain the best number of neurons, let say  $k$ , for example, we begin to increase the number of hidden layers from 2 to 5 with  $k$  neurons for each hidden layer to observe the changes in the prediction results. The experiments for hyperparameters search are done on Scientific\_Articles dataset. As exhibited from Figure 3, the network receives 3431 attributes of Scientific\_Articles dataset as the input following by a hidden layer including 16 neurons and produces 9 outputs corresponding to the predicted probabilities of 9 topics for the classification. After selecting hyperparameters from the experiments, we keep the number of neurons and one hidden layer for prediction tasks on five datasets while the number of nodes of the input layer and output layer can be vary depending on the considered dataset.

The MLP models which perform the binary classification are implemented Sigmoid function to do prediction tasks. The Sigmoid function [33] usually appears in the output layers of Deep learning architectures. It transforms the input values which lie in the domain  $\mathbb{R}$  to outputs have the domain in  $[0,1]$ . The Sigmoid function is also called "squashing" because this function squashes any input in the range of  $(-\infty, \infty)$  to the range of  $[0,1]$ . When we shifted to gradient-based learning, the Sigmoid was considered as a natural selection due to its smooth and differentiable approximation to a thresholding unit. The Sigmoid function is given by the formula:

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

where,  $x$  denotes data after being computed by the preceded neural layer.

For multi-classification problems, we use softmax function (Equation 5) with  $k$  classes. The Softmax function normalizes an input value into a vector of values that follows a probability distribution whose total sums up to 1.

$$\text{Softmax}(x_i) = \frac{e^{x_i}}{\sum_j^k e^{x_j}} \quad (5)$$

The activation function namely, ReLU [34], is also implemented in our architecture. ReLU follows the formula:

$$f(x) = \max(0, x) \quad (6)$$

where,  $x$  denotes data after being processed by the preceded neural layer.

<sup>2</sup>N. T. M. Huyen, V. X. Luong and L. H. Phuong, "VnTokenizer", 2010. <http://vntokenizer.sourceforge.net/>



Table 1: Five considered datasets descriptions

Data set	#Instances	#Attributes	#Classes	Language
Reuters_Corn	2,158	1,503	2	English
School_Text_Books	1,786	2,566	4	English
Turkish_News_Articles	3,600	5,693	6	Turkish
Scientific_Articles	650	3,431	9	Vietnamese
VnExpress_Newsletters	10,000	3,266	10	Vietnamese

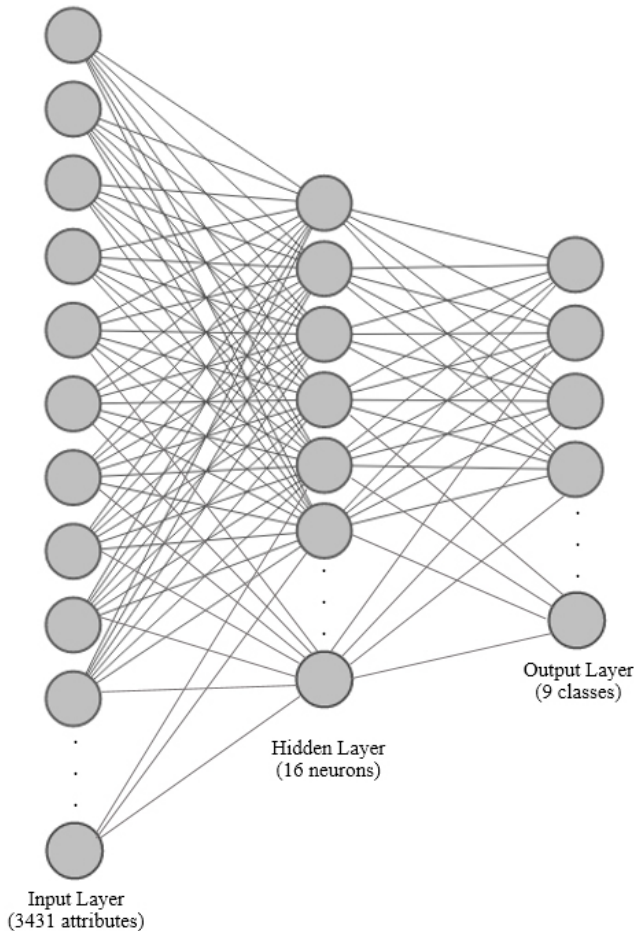


Figure 3: The proposed MLP architecture conducted from tune parameters experiments on Scientific\_Articles dataset

ReLU is the most widely used activation function for deep learning architectures with state-of-the-art results to date. ReLU helps models to produce better performance and generalization in deep learning compared to the Sigmoid and Tanh activation functions. It represents a nearly linear function, so this activation function preserves the properties of linear models that made them easy to optimize, with the gradient-descent method [35,36].

ReLU holds a role as an activation function to transform the output of the preceded hidden neural layer. ReLU helps to improve neural networks by speeding up training. Gradients of logistic and hyperbolic tangent models are lower than the positive portion of the

ReLU. That means the positive portion is updated more rapidly as training progresses.

In order to reduce overfitting issues, we consider implementing Early Stopping technique with a patience epoch of 5. This means that the loss cannot be improved after 5 consecutive epochs, the learning will be stopped. Otherwise, the learning will be continued to run to 10 epochs. The network is implemented with Adam optimizer function, use a batch size of 100 and a default learning rate of 0.001.

## 4 Experimental results

### 4.1 Data Description

This study used five experimental datasets in three various languages (including English, Turkish, and Vietnamese) as described in Table 1. The Reuters corn is available at UCI repository<sup>3</sup> for binary classification tasks. The School text books of 11<sup>th</sup> and 12<sup>th</sup> grade which is available at Kaggle Website<sup>4</sup> with four topics. A collection of Turkish news and articles dataset can be downloadable at UCI repository<sup>5</sup> including 3600 samples on 6 categories. The Scientific articles of a university and VnExpress Newsletters in Vietnam were used in our previous work at ACOMP 2019 [13] include 650 samples, 3431 features and 10000 samples, 3266 features, respectively. The considered numbers of classes also vary from binary classification to 10-class classification. The largest dataset is VnExpress\_Newsletters with 10000 samples on 10 different topics.

We also face imbalanced datasets issues where the number of samples of some classes is much more than other classes. For example, a class of Reuters\_corn dataset occupies to 97% while only 3% is for the other.

The performances of classifiers are examined by average AUC on 3-fold cross-validation. The folds are the same for all classifiers, i.e. training and test sets were identical for each classifier. AUC is a reliable metric for evaluating classifiers where data are not balanced. AUC is widely used in numerous studies to examine the performance of prediction tasks and it is reliable metric to measure the performance of prediction.

### 4.2 Hyper-parameter turning

In order to select appropriate parameters for the MLP models, we run the experiments with various configurations of MLP architec-

<sup>3</sup> <https://storm.cis.fordham.edu/~gweiss/data-mining/datasets.html>

<sup>4</sup> <https://www.kaggle.com/deepak711/4-subject-data-text-classification>

<sup>5</sup> <https://archive.ics.uci.edu/ml/datasets/TTC-3600:%20Benchmark%20dataset%20for%20Turkish%20text%20categorization>

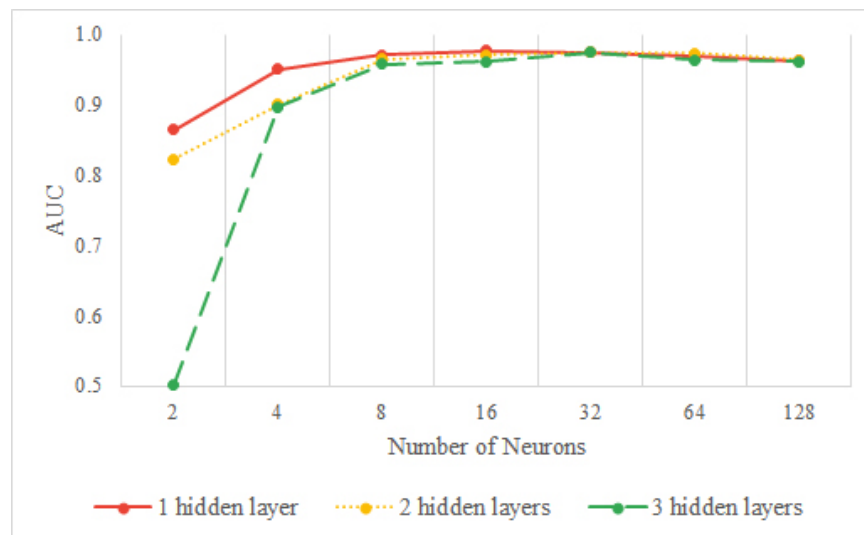


Figure 4: Hyper-parameter search for the Number of Hidden Layers

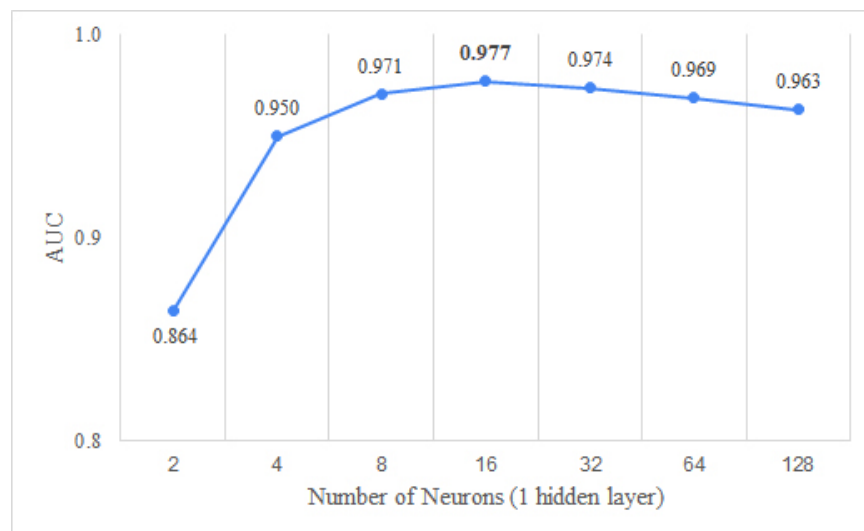


Figure 5: Hyper-parameter search for the Number of Neurons

tures.

In Figure 4 and Figure 5, we reveal the performance of different configurations of MLP on Scientific\_Articles dataset. The results show the performance is rising according to the width of the MLP. However, the performances reach peaks which are vary depending on the number of hidden layers used. After reaching the peak, the performance tends to go down when we keep increasing the number of neurons. As seen from Figure 4, with one hidden layer, we obtain the best with 16 neurons while using two or three hidden layers achieve the highest performance with 32 neurons per hidden layer. We noted that the performance decreases when we add more layers due to overfitting, but these differences are trivial when the number of neurons per hidden layer is high.

In most of the cases, the architectures of MLP with one neural layer obtain greater performance than the architectures with more hidden layers (see details performance of the number of neurons with one hidden layer in Figure 5). However, with a large number of neurons, we can see the performance of various numbers of neural

layers are nearly the same. Our results exhibit that the shallow architecture of MLP including one hidden layer with 16 neurons reaches the best performance. The number of hidden layers and the number of neurons conducted from the experiments in hyperparameter tuning, we implement an MLP architecture with one hidden layer including 16 neurons to run the learning and validation on the other 4 datasets.

The experiments of five considered datasets are presented and compared with various machine learning methods in Section 4.3.

#### 4.3 Topic classification Results of various machine learning algorithms on five considered datasets

Previous work [13] we showed that SVM works very well for automatic topic classification in an online submission system, however, in this work we have continued to improve and showed that using Deep Learning approach the results even better. Since the data sets are imbalanced, we report the AUC instead of the accuracy metric as

Table 2: Performance Comparison in AUC (Area under the ROC Curve) of various machine learning approaches on five considered datasets

Data set	Classifier	AUC
Reuters_Corn	MLP	<b>0.991</b>
	SVM	0.811
	Decision Tree	0.813
School_Text_Books	MLP	<b>0.999</b>
	SVM	0.991
	Decision Tree	0.928
Turkish_News_Articles	MLP	<b>0.962</b>
	SVM	0.949
	Decision Tree	0.871
Scientific_Articles	MLP	<b>0.977</b>
	SVM	0.965
	Decision Tree	0.819
VnExpress_Newsletters	MLP	<b>0.990</b>
	SVM	0.985
	Decision Tree	0.876

reported in Table 2. We have also used Decision Tree as a baseline for comparison.

Table 2 reveals the topic classification performances in AUC of three different machine learning algorithms on five considered datasets. It is easy to see that MLP outperforms other algorithms. In most cases, Decision Tree algorithm gives the worst result while SVM holds the second place. The classification performances of MLP are promising results which all achieve over 0.960 in AUC. Three datasets of them reach over 0.990 while the article classification in Turkish reveals the lowest performance but this result is still high with an AUC of 0.962. An example of visualization for the AUC and precision-recall are presented in Figures 6 and 7, other datasets are quite similar.

We also present the results in confusion matrices of School\_Text\_Books dataset to observe how different in the performances between MLP and SVM are in Table 3.

As shown from the results above, MLP outperforms SVM and we might like to know how different between them. We can see the difference is that SVM performs worse than MLP on the class where owns the minimum number of samples among the considered classes (these numbers are formatted with blue and bold, revealed in Table 3). With 98 samples for the class of "geography", this number is compared to other classes to see that we face imbalanced issues in data. However, MLP achieves a promising classification result comparing to SVM on the class with much fewer samples. Similar results are also revealed in other datasets. This is expected to bring to a reliable result in practical cases where we usually meet imbalanced dataset issues.

Binary classification tasks on Reuters\_Corn dataset reveal the same results where Deep learning approach also reaches a better performance on the class with fewer samples (see the results in Table 4). The Class indicates whether the entry is related to corn ( $b = 1$ ) or not ( $a = 0$ ). As observed from the results, we collected fewer

samples which related to "corn" so we also face imbalanced issue in the dataset. In this case, the class of 1 only occupies 3% compared to 97% samples belonging to the class of 0. The same result with multi-classes classification tasks, we obtain better performance for the binary classification tasks with MLP on the class with fewer samples (these numbers are formatted with blue and bold, revealed in Table 4) when we compare to SVM algorithm.

Experimental results in this work showed that the MLP is more suitable than the SVM in case of imbalanced data where the minority class is more interesting to predict. This is also the reason why we have selected the AUC as a measure instead of the Accuracy [14, 15].

For the training times, the MLP was completed the training stage in a couple of minutes for the datasets using in this study, so it does not a matter for an online system where we can automatically set-up a training schedule after a time interval (e.g., the model can be automatically updated after one day or other intervals depending on the real number of submissions).

## 5 Conclusion

Leveraging techniques of natural language processing and machine learning algorithms, we presented a solution to the automated classification of articles to support authors/editors saving their efforts and time for processing articles on the system. Data pre-processing techniques with steps introduced in this work are significantly improved to make the dataset in a standardized format for learning with the three considered algorithms of Multilayer Perceptron, Support Vector Machines and Decision Tree. The experiments are done on five various datasets. The data used vary in the number of features, attributes as well as the number of classes. All datasets reach the performance of over 0.960 in AUC with deep learning models.

As shown from experimental results, deep learning algorithm

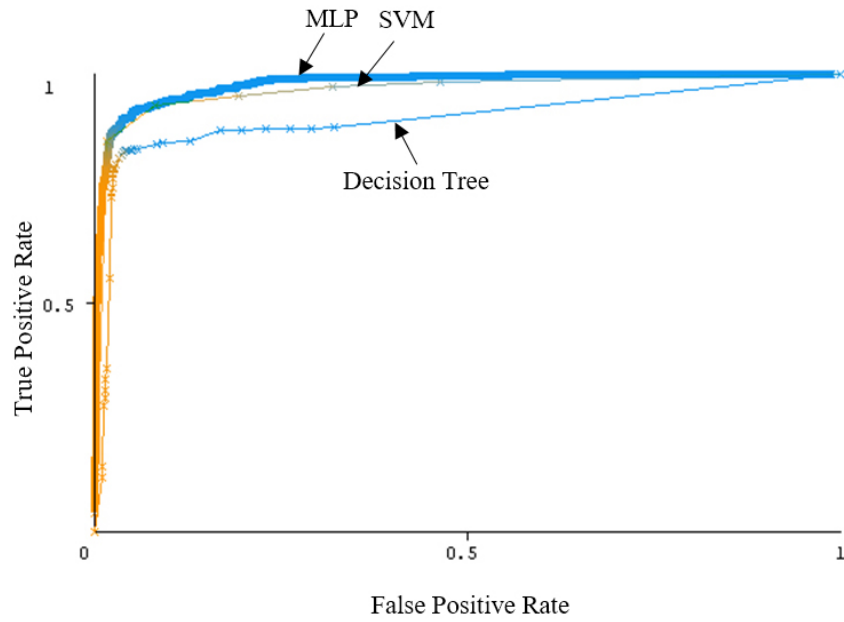


Figure 6: The Area under the ROC (AUC) for Turkish\_News\_Articles dataset

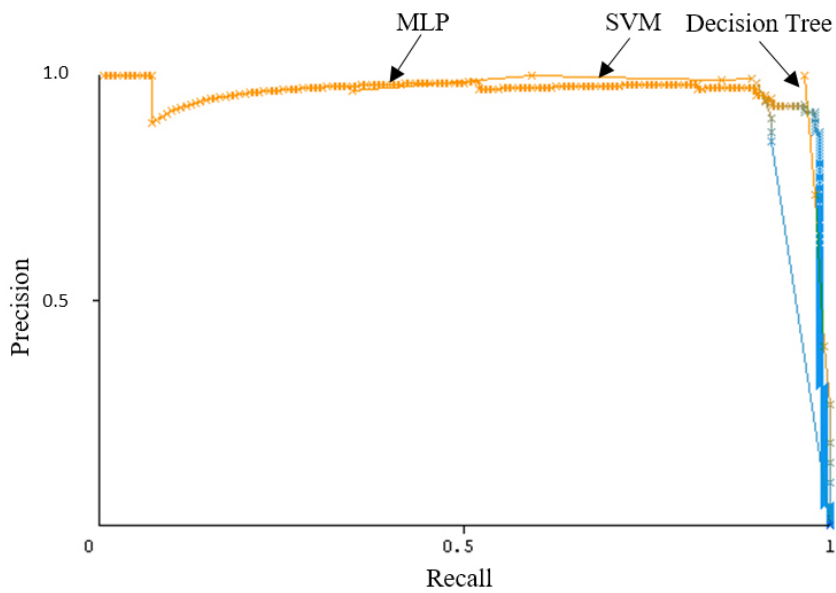


Figure 7: The Precision-Recall for Turkish\_News\_Articles dataset

Table 3: Confusion matrix comparison of predictions between Multilayer Perceptron and Support Vector Machines on School\_Text\_Books dataset

		Deep learning					SVM			
		Predicted classes					Predicted classes			
		a	b	c	d		a	b	c	d
Actual classes	a = accounts	283	0	0	1	a = accounts	281	1	0	2
	b = biology	0	626	7	2	b = biology	0	632	3	0
	c = geography	0	8	87	3	c = geography	0	16	78	4
	d = physics	0	4	3	762	d = physics	0	4	1	764

with Multilayer perceptron exhibits better classification performance than the classic machine learning such as Support Vector Machines.



Table 4: Confusion matrix comparison of predictions between Multilayer Perceptron and Support Vector Machines on Reuters\_Corn dataset with binary classification

	Deep learning		SVM		
	Predicted classes		Predicted classes		
	a	b	a	b	
Actual classes	2064	25	a = 0	2088	1
	13	<b>56</b>	b = 1	26	<b>43</b>

Some parameters are also evaluated to reach promising results in classification tasks. The results show that the proposed model is feasible to extract information and stratify articles automatically whenever a document is submitted to the system. We continue to find a solution for larger datasets in further research.

The proposed architecture of Multilayer perceptron is rather shallow with one neural layer. A vast of hyper-parameters of MLP are evaluated and we see that the performance tends to be saturated when we increase both the number of hidden layers as well as the number of neurons. Further studies should take into sophisticated architectures to improve performance in document categorization tasks.

**Conflict of Interest** The authors declare no conflict of interest.

## References

- [1] K. Thaoroijam, "A Study on Document Classification using Machine Learning Techniques," *International Journal of Computer Science Issues*, **11**(1), 217–222, 2014.
- [2] Y. Li, L. Zhang, Y. Xu, Y. Yao, R. Y. K. Lau, Y. Wu, "Enhancing Binary Classification by Modeling Uncertain Boundary in Three-Way Decisions," *IEEE Transactions on Knowledge and Data Engineering*, **29**(7), 1438–1451, 2017, doi:https://doi.org/10.1109/TKDE.2017.2681671.
- [3] F. Sebastiani, "Machine Learning in Automated Text Categorization," **34**(1), 1–47, 2002, doi:https://doi.org/10.1145/505282.505283.
- [4] Y. Yang, J. O. Pedersen, "A Comparative Study on Feature Selection in Text Categorization," in *Proceedings of the Fourteenth International Conference on Machine Learning, ICML '97*, 412–420, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1997, doi:https://dl.acm.org/doi/10.5555/645526.657137.
- [5] C. C. Aggarwal, C. Zhai, *A Survey of Text Classification Algorithms*, 163–222, Springer US, Boston, MA, 2012, doi:10.1007/978-1-4614-3223-4\_6.
- [6] M. A. Bijaksana, Y. Li, A. Algarni, "A Pattern Based Two-Stage Text Classifier," in P. Perner, editor, *Machine Learning and Data Mining in Pattern Recognition*, 169–182, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, doi:https://doi.org/10.1007/978-3-642-39712-7\_13.
- [7] J. Chen, H. Huang, S. Tian, Y. Qu, "Feature selection for text classification with Naïve Bayes," *Expert Systems with Applications*, **36**(3, Part 1), 5432–5435, 2009, doi:https://doi.org/10.1016/j.eswa.2008.06.054.
- [8] M. Haddoud, A. Mokhtari, T. Lecroq, S. Abdeddaïm, "Combining supervised term-weighting metrics for SVM text classification with extended term representation," *Knowledge and Information Systems*, **49**, 909–931, 2016, doi:https://doi.org/10.1007/s10115-016-0924-1.
- [9] A. Chouchoulas, Q. Shen, "A Rough Set-Based Approach to Text Classification," in N. Zhong, A. Skowron, S. Ohsuga, editors, *New Directions in Rough Sets, Data Mining, and Granular-Soft Computing*, 118–127, Springer Berlin Heidelberg, Berlin, Heidelberg, 1999, doi:https://doi.org/10.1007/978-3-540-48061-7\_16.
- [10] A. P. Gopi, R. N. S. Jyothi, V. L. Narayana, K. S. Sandeep, "Classification of tweets data based on polarity using improved RBF kernel of SVM," *International Journal of Information Technology*, 2020, doi:https://doi.org/10.1007/s41870-019-00409-4.
- [11] N. Thai-Nghe, Q. D. Truong, "An Approach for Building a Semi-automatic Online Consultancy System," in *2015 International Conference on Advanced Computing and Applications (ACOMP)*, 51–58, 2015, doi:https://doi.org/10.1109/ACOMP.2015.11.
- [12] B. E. Boser, I. M. Guyon, V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers," in *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, 144–152, ACM Press, 1992, doi:https://doi.org/10.1145/130385.130401.
- [13] T. T. Dien, B. H. Loc, N. Thai-Nghe, "Article Classification using Natural Language Processing and Machine Learning," in *2019 International Conference on Advanced Computing and Applications (ACOMP)*, 78–84, 2019, doi:https://doi.org/10.1109/ACOMP.2019.00019.
- [14] G. Haixiang, L. Yijing, J. Shang, G. Mingyun, H. Yuanyue, G. Bing, "Learning from class-imbalanced data: Review of methods and applications," *Expert Systems with Applications*, **73**, 220–239, 2017, doi:https://doi.org/10.1016/j.eswa.2016.12.035.
- [15] N. Thai-Nghe, Z. Gantner, L. Schmidt-Thieme, "A new evaluation measure for learning from imbalanced data," in *The 2011 International Joint Conference on Neural Networks*, 537–542, 2011, doi:https://doi.org/10.1109/IJCNN.2011.6033267.
- [16] S. Bahassine, A. Madani, M. Al-Sarem, M. Kissi, "Feature selection using an improved Chi-square for Arabic text classification," *Journal of King Saud University - Computer and Information Sciences*, **32**(2), 225–231, 2020, doi:https://doi.org/10.1016/j.jksuci.2018.05.010.
- [17] S. Larabi Marie-Sainte, N. Alalyani, "Firefly Algorithm based Feature Selection for Arabic Text Classification," *Journal of King Saud University - Computer and Information Sciences*, **32**(3), 320–328, 2020, doi:https://doi.org/10.1016/j.jksuci.2018.06.004.
- [18] C.-H. Tsai, "MMSEG: A word identification system for Mandarin Chinese text based on two variants of the maximum matching algorithm," Available on internet at <http://www.geocities.com/hao510/mmseg>, 2000.
- [19] K. Gábor, D. Buscaldi, A.-K. Schumann, B. QasemiZadeh, H. Zargayouna, T. Charnois, "SemEval-2018 Task 7: Semantic Relation Extraction and Classification in Scientific Papers," in *Proceedings of The 12th International Workshop on Semantic Evaluation*, 679–688, Association for Computational Linguistics, New Orleans, Louisiana, 2018, doi:http://dx.doi.org/10.18653/v1/S18-1111.
- [20] S. Sarkar, N. Ejaz, M. Kumar, J. Maiti, "Root Cause Analysis of Incidents Using Text Clustering and Classification Algorithms," in P. K. Singh, B. K. Panigrahi, N. K. Suryadevara, S. K. Sharma, A. P. Singh, editors, *Proceedings of ICETIT 2019*, 707–718, Springer International Publishing, Cham, 2020, doi:https://doi.org/10.1007/978-3-030-30577-263.
- [21] P. Harjule, A. Gurjar, H. Seth, P. Thakur, "Text Classification on Twitter Data," in *2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE)*, 160–164, 2020, doi:https://doi.org/10.1109/ICETCE48199.2020.9091774.
- [22] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. A. Chenaghlu, J. Gao, "Deep Learning Based Text Classification: A Comprehensive Review," *ArXiv, abs/2004.03705*, 2020.
- [23] M. Zulqarnain, R. Ghazali, Y. M. M. Hassim, M. Rehan, "A comparative review on deep learning models for text classification," *Indonesian Journal of Electrical Engineering and Computer Science*, **19**(1), 325–335, 2020, doi:https://doi.org/10.11591/ijeecs.v19.i1.pp325-335.
- [24] Q. Qin, W. Hu, B. Liu, "Feature Projection for Improved Text Classification," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 8161–8171, Association for Computational Linguistics, Online, 2020, doi:http://dx.doi.org/10.18653/v1/2020.acl-main.726.

- [25] F. Sovrano, M. Palmirani, F. Vitali, "Deep Learning Based Multi-Label Text Classification of UNGA Resolutions," ArXiv, **abs/2004.03455**, 2020.
- [26] M. Zulqarnain, R. Ghazali, Y. M. M. Hassim, M. Rehan, "Text classification based on gated recurrent unit combines with support vector machine," *International Journal of Electrical and Computer Engineering*, **10**(4), 3734–3742, 2020, doi:<http://doi.org/10.11591/ijece.v10i4.pp3734-3742>.
- [27] M. Belazzoug, M. Touahria, F. Nouioua, M. Brahimi, "An improved sine cosine algorithm to select features for text categorization," *Journal of King Saud University - Computer and Information Sciences*, **32**(4), 454 – 464, 2020, doi:<https://doi.org/10.1016/j.jksuci.2019.07.003>.
- [28] E. Lin, Q. Chen, X. Qi, "Deep reinforcement learning for imbalanced classification," *Applied Intelligence*, **50**(8), 2488–2502, 2020, doi:[10.1007/s10489-020-01637-z](https://doi.org/10.1007/s10489-020-01637-z).
- [29] D. Chai, W. Wu, Q. Han, F. Wu, J. Li, "Description Based Text Classification with Reinforcement Learning," ArXiv, **abs/2002.03067**, 2020.
- [30] C. S. Perone, "Machine learning: Cosine similarity for vector space models (Part III)," URL:<http://blog.christianperone.com>, 2019.
- [31] J.-Y. Chang, I.-M. Kim, "Analysis and Evaluation of Current Graph-Based Text Mining Researches," *Advanced Science and Technology Letter*, **42**, 100–103, 2013, doi:<http://dx.doi.org/10.14257/astl.2013.42.2>.
- [32] H. Saif, M. Fernandez, Y. He, H. Alani, "On Stopwords, Filtering and Data Sparsity for Sentiment Analysis of Twitter," in *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 810–817, European Language Resources Association (ELRA), Reykjavik, Iceland, 2014.
- [33] C. Nwankpa, W. Ijomah, A. Gachagan, S. Marshall, "Activation Functions: Comparison of trends in Practice and Research for Deep Learning," ArXiv, **abs/1811.03378**, 2018.
- [34] V. Nair, G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," in J. Fürnkranz, T. Joachims, editors, *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, June 21-24, 2010, Haifa, Israel, 807–814, Omnipress, 2010.
- [35] M. D. Zeiler, M. Ranzato, R. Monga, M. Mao, K. Yang, Q. V. Le, P. Nguyen, A. Senior, V. Vanhoucke, J. Dean, G. E. Hinton, "On rectified linear units for speech processing," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 3517–3521, 2013, doi:<https://doi.org/10.1109/ICASSP.2013.6638312>.
- [36] G. E. Dahl, T. N. Sainath, G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 8609–8613, 2013, doi:<https://doi.org/10.1109/ICASSP.2013.6639346>.

# Bathtub-Shape Failure Rate Life Time Model in a New Version with the use of Bayesian Prediction Bounds for the Presence of Outliers

Ayed Rheal A. Alanzi\*

Mathematics Department, Faculty of Science and Human Studies of Hotat Sudair, Majmaah University, Majmaah 11952, Saudi Arabia.

## ARTICLE INFO

Article history:

Received: 06 May, 2020

Accepted: 24 July, 2020

Online: 28 August, 2020

Keywords:

Outliers

Censored Samples

Bayesian Prediction

Bivariate Prior

Bathtub-Shaped Model

Density Function

## ABSTRACT

The present article explains obtaining of the Bayesian prediction bounds at the maximum and minimum rate taking into account the results of future observation from a new version of a bathtub-shape failure rate distribution of life time type in the presence of outliers. The Type-II censored sample serves as a basis for the intervals of the prediction with the numerical examples as illustrations of the studied procedure.

## 1 Introduction

Researchers study numerous statistical problems with the requirement to apply the same distribution on the basis of the previously obtained data to predict the future data. Previous analysis of various practical applications required to satisfy the mentioned needs have been conducted in the studies of [1]-[3] with the most prominent suggestions and analysis done in the study of [4]. The latter is devoted to analysis of the bathtub-shaped distribution of life-time type, taking into account the function of increasing failure rate or two parameters. Provided that the distribution is shown with  $\lambda$  and  $\beta$  as the available two parameters by Chen, it is possible to present the functions of probability density and cumulative distribution in the equations as follows:

$$f(x; \lambda, \beta) = \lambda\beta x^{\beta-1} e^{x^\beta - \lambda(e^{x^\beta} - 1)}, \quad x > 0, \lambda > 0, \beta > 0, \quad (1)$$

$$F(x; \lambda, \beta) = 1 - e^{-\lambda(e^{x^\beta} - 1)}. \quad (2)$$

The calculations of [5] presented Bayesian estimations on the basis of bathtub-shaped life time distribution of two parameters with the use of record values. The study of [6] is devoted to obtaining bounds of Bayesian prediction for a life time distribution with the function of bathtub-shaped failure rate on the basis of two parameters. Based on censored data, [7] suggested a two-parameter

Lifetime Distribution with Bathtub Shaped Hazard. Consequently, derivation of Bayesian prediction bounds was done by [8] for a new version of bathtub shape failure rate life time model of doubly type-II censored samples. Their research was of crucial importance, taking into consideration the functions of life time distribution founded upon the Bayesian intervals for prediction [9]. E-Bayesian estimation of Chen distribution based on type-I censoring scheme has also investigated [10].

Application of prediction techniques with two samples was done with the use of the parameters obtained previously for a single problem of statistical significance. The key problem of the outlier presence was identified with several approximation approaches combined, and it produces an impact on the estimations of previous density with implementation of unknown parameters. Illustration of the significance of the presented problem from a statistical perspective is done with numerical estimation and gamma conjugate joint prior and prior cases. The present research studies Bayesian prediction intervals from the Chen  $(\lambda, \beta)$  model (1) for future observations under two plans of sampling:  $x_{(1)}, x_{(2)}, \dots, x_{(n)}$  is an ordered random sample from model (2) (size  $n$ ) where  $x_{(r)}, x_{(r+1)}, \dots, x_{(n)}$ , the  $(n-r)$  are the sample ordered observations of the largest size. Statistical analysis applies the ordered observations which remained, i.e.,  $x = (x_{(1)}, \dots, x_{(r)})$ . It is evident that the sampling has its special cases like a complete ordered sample ( $r = n$ ).

In the first case, the observed sample is presented as follows:

\*Corresponding Author: Ayed R. A. Alanzi, Mathematics Department, Faculty of Science and Human Studies of Hotat Sudair, Majmaah University, Majmaah 11952, Saudi Arabia., E-mail: a.alanzi@mu.edu.sa

$x_{(1)} < \dots < x_{(r)}$  is Type II censored sample with  $r < n$ . An unobserved sample is presented with  $x_{(r+1)} < x_{(r+2)} < \dots < x_{(n)}$  as remaining values. The observed sample is used for the study of the remaining  $n - r$  values.

In the second case, the previously mentioned Type II of sample is represented in the same way as  $x_{(1)} < \dots < x_{(r)}$ . The future unobserved sample is shown as  $z_1 < z_2 < \dots < z_m$  from the same population. The key aim for the selected configurations of sampling is identification of the prediction intervals with the applied previous observations for specification of the future observations.

Derivation of a random sample  $x_{(1)} < x_{(2)} < \dots < x_{(n)}$  is done with the probabilities specified in (1) and (2) from the  $a$  population. In addition, the life test uses the assigned  $x_1, x_2, \dots, x_n$ . Recording of the failure times is done from the timeframe of  $r^{th}$  failure with  $r < n$ . The following likelihood function estimate is used for the Type-II censored data with the mentioned observations, taking into account:

$$L(\lambda, \beta; \underline{x}) \propto [1 - F_X(x_{(r)}; \lambda, \beta)]^{n-r} \prod_{i=1}^r [f_X(x_{(i)}; \lambda, \beta)] \\ = (\lambda\beta)^r \exp\left\{\sum_{i=1}^r (\beta \ln x_{(i)} + x_{(i)}^\beta) - \lambda T_1(\beta; \underline{x})\right\}, x_{(k)} > 0 \quad (3)$$

where

$$\underline{x} = (x_{(1)}, \dots, x_{(r)}), \\ T_1(\beta; \underline{x}) = \sum_{i=1}^r (e^{x_{(i)}^\beta} - 1) + (n - r)(e^{x_{(r)}^\beta} - 1). \quad (4)$$

Assuming that the researcher uses the function of simple prior density for measuring previous hypotheses for  $\lambda$  and  $\beta$ , the following estimations can be done:

$$\pi(\lambda, \beta) = \pi_1(\lambda) \pi_2(\beta), \quad (5)$$

with  $\pi_1(\lambda)$  viewed as a conjugate previously presented as

$$\pi_1(\lambda) = \frac{b_1^{a_1}}{\Gamma(a_1)} \lambda^{a_1-1} e^{-b_1 \lambda}, \lambda > 0 (a_1, b_1 > 0), \quad (6)$$

$$\pi_2(\beta) = \frac{b_2^{a_2}}{\Gamma(a_2)} \beta^{a_2-1} e^{-b_2 \beta}, \beta > 0 (a_2, b_2 > 0). \quad (7)$$

According to the suggestion of [11], the function presented in (6) and (7) can be used to identify the bathtub-shaped distribution with the help of two parameters. Summarization of the function of joint posterior density that  $\lambda$  and  $\beta$  parameters have can be done as follows with the help of the joint prior density function presented correspondingly in (14) and (19) as well as the likelihood function:

$$\pi_2^*(\lambda, \beta, \underline{x}) \propto \lambda^{r+a_1-1} \beta^{r+a_2-1} \\ \exp\left\{\sum_{i=1}^r (\beta \ln x_{(i)} + x_{(i)}^\beta) - b_2 \beta - \lambda [T_1(\beta; \underline{x}) + b_1]\right\} \quad (8)$$

Thus,  $\lambda$  and  $\beta$  can have their posterior density function presented as follows:

$$\pi^*(\lambda, \beta | \underline{x}) \propto g_1(\lambda | data) g_2(\beta | data) g_3(\lambda, \beta | data), \quad (9)$$

with  $g_1(\lambda | \beta, data)$  as a gamma density with the constraints of the  $r$  shape and  $T_1(\beta; \underline{x})$  scale, while  $g_2(\beta | data)$  can be viewed as a function of proper density estimated as follows:

$$g_2(\beta | data) \propto \frac{1}{[T_1(\beta; \underline{x})]^r} \beta^{r-1} \exp\left\{\sum_{i=1}^r \beta (\ln x_{(i)} - b_2)\right\}, \quad (10)$$

with  $g_3(\lambda, \beta | data)$  presented as

$$g_3(\lambda, \beta | data) = \lambda^{a_1} \beta^{a_2} e^{-\lambda b_1 + \lambda \beta}. \quad (11)$$

Hence, the following equation presents the Bayesian estimation of  $\lambda$  and  $\beta$  in any function, for instance  $g(\lambda, \beta)$  under the function of squared error loss:

$$\hat{g}(\lambda, \beta) = \frac{\int_0^\infty \int_0^\infty g(\lambda, \beta) g_1(\lambda | data) g_2(\beta | \lambda, data) g_3(\lambda, \beta | data) d\lambda d\beta}{\int_0^\infty \int_0^\infty g_1(\lambda | data) g_2(\beta | \lambda, data) g_3(\lambda, \beta | data) d\lambda d\beta} \quad (12)$$

As for the equation (21), conversion is evidently impossible with the output of a simple closed form; that implies that estimations of the Bayesian predications for the inputs of  $\lambda$  and  $\beta$  specified before are impossible in this form as well. Consequently, it can be suggested to follow the recommendation in [12] and use the approach of importance sampling technique for approximation of (21) with reference to the restriction of a simple closed form.

### 1.1 Importance sampling technique

The methodology of an importance sampling technique is applied for computing and validating the Bayes estimates on  $\lambda$  and  $\beta$  as well as  $g(\lambda, \beta)$  and other constructed relevant functions. Algorithm presents the function of posterior density with the peculiar features of the process of approximation.

---

#### Algorithm 1: Sampling Technique Algorithm

---

**Step 1: Estimation of  $\beta$**  on the basis of  $g_1(\beta | data)$ .

**Step 2: Estimation of  $\lambda$**  on the basis  $g_2(\lambda | \beta, data)$ .

**Step 3: Repetition** of the described stages following the principle of consecutiveness aimed at generating  $(\lambda_1, \beta_1), (\lambda_2, \beta_2), \dots, (\lambda_M, \beta_M)$

---

The following equation presents the process of approximation using the restrictions of Bayesian estimates for  $g(\lambda, \beta)$  and squared error loss control in the framework of the procedure of importance sampling:

$$\hat{g}_{BS}(\lambda, \beta) = \frac{\sum_{i=1}^M g(\lambda_i, \beta_i) g_3(\lambda_i, \beta_i | data)}{\sum_{i=1}^M g_3(\lambda_i, \beta_i | data)}. \quad (13)$$

## 2 Bayesian prediction under the condition of the presence of outliers

This part of research is devoted to predicting of future outputs of the studies in the presence of outliers. All the formal definitions used in this section are the same as in (1) and  $x_1, x_2, \dots, x_n$  is used as a random sample formed for the given function of the population



density from Chen  $(\lambda, \beta)$ . Furthermore,  $y_1, y_2, \dots, y_m$  is referred to as independent unobserved sample generated from the same data points, further denoted as a future sample. Furthermore, we probe the limits of Bayesian prediction for  $s^{th}$  in the framework of the future estimates that  $y_s, s = 1, 2, \dots, m$  has under the conditions of using single outliers. The following equation presents the  $y_s$  density function under the mentioned conditions for  $\theta$ :

$$h_1(y_s|\theta) = D(s) [(s-1)F^{s-2}(1-F)^{m-s}F^*f + F^{s-1}(1-F)^{m-s}f^* + (m-s)F^{s-1}(1-F)^{m-s-1}(1-F^*)f], \quad (14)$$

with

$$D(s) = \binom{m-1}{s-1}. \quad (15)$$

The functions of density and cumulative distribution are presented as  $f = f(y|\theta)$  and  $F = F(y|\theta)$  for any  $y^s$  provided that they cannot be referred as outliers with outliers as  $f^* = f^*(y|\theta)$  and  $F^* = F^*(y|\theta)$  [13]. The change of  $\lambda$  parameter by  $\lambda\lambda_0$  or  $\lambda$  depending on the classification of the outliers allows obtaining the  $f^*$  and  $F^*$  functions for the Chen  $(\lambda, \beta)$  model.

### 3 Predicting the minimum

It is possible to present distribution in the minimum in  $m$ -sized future sample with the single outlier of type  $\lambda\lambda_0$  given by using  $s$  as 1 in (14) as follows:

$$h_1(y_1|\theta) = (1-F)^{m-1}f^* + (m-1)(1-F)^{m-2}(1-F^*)f. \quad (16)$$

It is possible to obtain the density function of  $Y_1$  in the case of Chen  $(\lambda, \beta)$  in the presence of a single type  $\lambda\lambda_0$  outlier via substitution of (1) for  $f$  and (2) for  $F$  correspondingly in (16).  $f^*$  and  $F^*$  have values given by (1) and (2) further after  $\lambda$  is replaced by  $\lambda\lambda_0$ . Simplification of this density function is done in the following form:

$$h_1(y_1|\lambda, \beta) = f(y_1; \lambda(m + \lambda_0 - 1), \beta), \quad (17)$$

where  $f(\cdot, \cdot)$  is given by (1). The cdf of  $h_1(y_1|\lambda, \beta)$  can be presented as

$$H_1(y_1|\lambda, \beta) = F(y_1; \lambda(m + \lambda_0 - 1), \beta). \quad (18)$$

The predictive density of  $y_1$ , with  $x, \lambda$  and  $\beta$  is given by

$$h_1^*(y|x) = \int_0^\infty \int_0^\infty h_1(y_1|\lambda, \beta) \pi^*(\lambda, \beta|x) d\lambda d\beta. \quad (19)$$

The predictive distribution function of  $y_1$ , with  $x, \lambda$  and  $\beta$  is

$$H_1^*(y|x) = \int_0^\infty \int_0^\infty H_1(y_1|x, \lambda, \beta) \pi^*(\lambda, \beta|x) d\lambda d\beta. \quad (20)$$

It can be assumed that  $\{(\lambda_i, \beta_i); i = 1, 2, \dots, M\}$  should be viewed as MCMC samples obtained from  $\pi^*(\lambda, \beta|x)$ . The equation for corresponding estimation parameters that consistency of  $h_1^*(y_1|x, \lambda, \beta)$  and  $H_1^*(y_1|x, \lambda, \beta)$  has is as follows:

$$\hat{h}_1^*(y|x) = \sum_{i=1}^M h_1(y_1|\lambda_i, \beta_i) h_i, \quad (21)$$

$$\hat{H}_1^*(y|x) = \sum_{i=1}^M H_1(y_1|\lambda_i, \beta_i) h_i, \quad (22)$$

with

$$h_i = \frac{g_3(\lambda_i, \beta_i|data)}{\sum_{i=1}^M g_3(\lambda_i, \beta_i|data)}; \quad i = 1, 2, \dots, M. \quad (23)$$

All the mentioned equations serve as the basis for the a  $(1 - \tau)$  100 % Bayesian estimation for  $Y_1$ , which implies having  $P[L(x) \leq Y_1 \leq U(x)] = 1 - \tau$ , with  $L(x)$  as the highest and limit for  $y_1$  and  $U(x)$  as the lowest limit for  $y_1$ . On the basis of the prior estimates for (22),  $1 - \frac{\tau}{2}$  and  $\frac{\tau}{2}$ , we can claim that

$$P[Y \geq L(x)|x] = 1 - \frac{\tau}{2} \Rightarrow \hat{H}_1^*(L(x)|x) = \frac{\tau}{2}, \quad (24)$$

and

$$P[Y \leq U(x)|x] = \frac{\tau}{2} \Rightarrow \hat{H}_1^*(U(x)|x) = 1 - \frac{\tau}{2}. \quad (25)$$

Computing of the prediction bounds of  $y_1$  is done on the basis of the (24) and (25) equations.

### 4 Predicting the maximum

Obtaining of distribution of the maximum in a  $m$ -sized sample with the presence of a single outlier is done via using  $s = m$  in (14). The equation for  $Y_m$  density function for a provided  $\theta$  in the presence of a single outlier is as follows:

$$h(y_m|\theta) = (m-1)[F]^{m-2}[F^*]f + [F]^{m-1}f^*. \quad (26)$$

It is possible to obtain the density function of  $Y_m$  in the case of Chen  $(\lambda, \beta)$  in the presence of a single type  $\lambda\lambda_0$  outlier via substitution of (1) for  $f$  and (2) for  $F$  correspondingly in (26).  $f^*$  and  $F^*$  have values given by (1) and (2) further after  $\lambda$  is replaced by  $\lambda\lambda_0$ . Simplification of this density function is done in the following form:

$$h_1(y_s|\lambda, \beta) = \lambda_0 \sum_{j=0}^{m-1} A_{1j}(y_m) + (m-1) \sum_{j=0}^{m-2} A_{2j}(y_m), \quad y_m > 0, \quad (27)$$

with

$$\begin{aligned} A_{1j}(y_m) &= a_{1j}(m)f(y_m; \lambda(\lambda_0 + j), \beta), \\ A_{2j}(y_m) &= a_{2j}(m)[f(y_m; \lambda(j+1), \beta) - f(y_m; \lambda(\lambda_0 + j + 1), \beta)], \end{aligned} \quad (28)$$

and for  $\ell = 1, 2$ ,

$$a_{\ell j}(m) = (-1)^j \binom{m-\ell}{j}. \quad (29)$$

The cdf of  $y_m$  is presented with

$$H_1(y_s|\lambda, \beta) = \lambda_0 \sum_{j=0}^{m-1} A_{1j}^*(y_m) + (m-1) \sum_{j=0}^{m-2} A_{2j}^*(y_s), \quad y_s > 0, \quad (30)$$

with

$$A_{1j}^*(y_m) = \frac{a_{1j}(m)}{(\lambda_0 + j)} F(y_m; \lambda(\lambda_0 + j), \beta),$$

$$A_{2j}^*(y_m) = \frac{a_{2j}(m)}{(j + 1)} F(y_m; \lambda(j + 1), \beta)$$

$$- \frac{a_{2j}(m)}{(\lambda_0 + j + 1)} F(y_m; \lambda(\lambda_0 + j + 1), \beta). \quad (31)$$

and

$$\hat{H}_1^*(y|\underline{x}) = \sum_{i=1}^M H_1(y_{(m)}|\lambda_i, \beta_i) h_i. \quad (35)$$

with

$$h_i = \frac{g_3(\lambda_i, \beta_i|data)}{\sum_{i=1}^M g_3(\lambda_i, \beta_i|data)}; \quad i = 1, 2, \dots, M. \quad (36)$$

Table 1: Bayesian prediction intervals (95%) for  $y_1$  and  $y_{10}$  in the presence of type  $\lambda$  single outlier,  $r = 20, n = 30$ . Note: PP is Point Predictor, L is Lower bound, U is Upper bound, and CP is Coverage Percentage

$\lambda_0$	Observations	$y_1$	$y_{10}$
1	PP	0.219311	1.03915
	L	0.037503	0.773876
	U	0.470244	1.30324
	Length	0.432741	0.529363
	CP	95.45 %	94.85 %
2	PP	0.209173	1.02768
	L	0.035691	0.757975
	U	0.449826	1.29739
	Length	0.414135	0.539419
	CP	94.6 %	95.16 %
3	PP	0.200296	1.02638
	L	0.034113	0.085392
	U	0.431811	1.19415
	Length	0.397698	1.10876
	CP	93.77%	88.02%
4	PP	0.192439	1.02614
	L	0.032723	0.751878
	U	0.415762	1.29738
	Length	0.383039	0.545498
	CP	92.59%	95.36%
5	PP	0.185422	1.02609
	L	0.031487	0.751374
	U	0.401347	1.29738
	Length	0.36986	0.546003
	CP	91.54%	95.4%

Table 2: Bayesian prediction intervals (95%) for  $y_1$  and  $y_{10}$  in the presence of type  $\lambda$  single outlier,  $r = 25, n = 30$ . Note: PP is Point Predictor, L is Lower bound, U is Upper bound, and CP is Coverage Percentage

$\lambda_0$	Observations	$y_1$	$y_{10}$
1	PP	0.225826	1.04406
	L	0.039754	0.782087
	U	0.479422	1.30373
	Length	0.439668	0.521645
	CP	95.4 %	94.46%
2	PP	0.215545	1.03274
	L	0.037862	0.766312
	U	0.458946	1.298
	Length	0.421084	0.531691
	CP	94.75%	94.91%
3	PP	0.206536	1.03146
	L	0.036213	0.761754
	U	0.440865	1.29799
	Length	0.404652	0.536232
	CP	94.01 %	95.06 %
4	PP	0.198558	1.03123
	L	0.034759	0.760261
	U	0.424745	1.29799
	Length	0.389986	0.537724
	CP	93.18 %	95.1%
5	PP	0.191427	1.03117
	L	0.033465	0.759761
	U	0.410255	1.29799
	Length	0.37679	0.538224
	CP	92.12 %	95.11 %

$Y = y_m$  has predictive density of presented  $x$ ,  $\lambda$  and  $\beta$  as follows:

$$h_1^*(y|\underline{x}) = \int_0^\infty \int_0^\infty h_1(y_{(m)}|\lambda, \beta) \pi^*(\lambda, \beta|\underline{x}) d\lambda d\beta, \quad (32)$$

$Y = y_m$  has the function of predictive distribution of presented  $x$ ,  $\lambda$  and  $\beta$  as follows:

$$H_1^*(y|\underline{x}) = \int_0^\infty \int_0^\infty H_1(y_{(m)}|\underline{x}, \lambda, \beta) \pi^*(\lambda, \beta|\underline{x}) d\lambda d\beta. \quad (33)$$

Obtaining the  $h_1^*(y_{(m)}|\underline{x}, \lambda, \beta)$  and  $H^*(y_{(m)}|\underline{x}, \lambda, \beta)$  as simulation consistent estimators can be done in the following manner assuming that  $(\lambda_i, \beta_i); i = 1, 2, \dots, M$  can be viewed as MCMC samples generated from  $\pi^*(\lambda, \beta|\underline{x})$ :

$$\hat{h}_1^*(y|\underline{x}) = \sum_{i=1}^M h_1(y_{(m)}|\lambda_i, \beta_i) h_i, \quad (34)$$

A 100% Bayesian prediction interval  $(1 - \tau)$  for  $Y_m$  can be presented as  $P[L(x) \leq Y_m \leq U(x)] = 1 - \tau$ , with  $L(x)$  as the lower bound for  $y_m$  and  $U(x)$  as the upper bound for  $y_m$ . Thus equating (35)  $1 - \frac{\tau}{2}$  and  $\frac{\tau}{2}$ , it is possible to get the following equations:

$$P[Y \geq L(x)|\underline{x}] = 1 - \frac{\tau}{2} \Rightarrow \hat{H}_1^*(L(x)|\underline{x}) = \frac{\tau}{2} \quad (37)$$

and

$$P[Y \leq U(x)|\underline{x}] = \frac{\tau}{2} \Rightarrow \hat{H}_1^*(U(x)|\underline{x}) = 1 - \frac{\tau}{2}. \quad (38)$$

Solutions to (37) and (38) as nonlinear equations are found with the use of an iterative method aimed at evaluating the Bayesian prediction and obtaining its lower and upper bounds for the future sample  $y_m$  at its maximum point.

**Example.** I can use the previous (6) density equation to generate  $\lambda = 1.41617$  with the provided previous parameters  $a_1 = 2.2$  and  $b_1 = 1.8$  and then we use the previous (7) density equation to generate  $\beta = 1.91297$  for the previous parameters  $a_2 = 3.2$  and  $b_2 = 1.5$ . Chen distribution ( $\lambda = 1.41617, \beta = 1.91297$ ) with  $r$  in its deferent value is used to generate a random sample of size  $n = 30$ . Let us consider a case with one more sample of size  $m = 10$  in the presence of type  $\lambda \lambda_0$  single outlier. Our aim is to obtain the percentage of 95% as prediction bounds for  $Y_1$  and  $Y_{10}$  for the provided value of  $\lambda_0$  as minimum and maximum that the future sample has. Table 1, Table 2, and Table 3 present the mentioned bounds with  $\lambda_0$  in corresponding values.

Table 3: Bayesian prediction intervals (95%) for  $y_1$  and  $y_{10}$  in the presence of type  $\lambda \lambda_0$  single outlier,  $r = n = 30$ . Note: PP is Point Predictor, L is Lower bound, U is Upper bound, and CP is Coverage Percentage

$\lambda_0$	Observations	$y_1$	$y_{10}$
1	PP	0.225842	1.04428
	L	0.039745	0.782243
	U	0.479499	1.304
	Length	0.439754	0.521755
	CP	95.4%	94.44%
2	PP	0.21556	1.03295
	L	0.037853	0.133706
	U	0.459018	1.19737
	Length	0.421165	1.06367
	CP	94.75%	88.47%
3	PP	0.206549	1.03168
	L	0.036204	0.761905
	U	0.440932	1.29825
	Length	0.404728	0.536345
	CP	94.02 %	95.07%
4	PP	0.198569	1.03144
	L	0.03475	0.760412
	U	0.424807	1.29825
	Length	0.390057	0.537837
	CP	93.18%	95.1%
5	PP	0.191436	1.03138
	L	0.033457	0.035391
	U	0.410314	1.09987
	Length	0.376857	1.06448
	CP	92.12%	68.35%

## 5 Conclusion

The present study analyzes a single type  $\lambda \lambda_0$  outlier as the multiple outliers are more complicated. In case of no outliers in the homo-

geneous case, future observations can obtain Bayesian prediction bounds at the minimum and maximum level via assigning  $\lambda_0 = 1$  for the (22) and (35) equations. Table 1 demonstrates that future observations have their limits dependent on  $\lambda$  value.

**Conflict of Interest** The authors declare no conflict of interest.

**Acknowledgment** The author thanks to the Deanship of Scientific Research at Majmaah University for support and encouragement.

## References

- [1] I. R. Dunsmor, "The Bayesian predictive distribution in life testing models." *Technometrics*, **16**(3), 455-460, 1974. <https://doi.org/10.1080/00401706.1974.10489216>
- [2] A. Aitchison, and I.R. Dunsmore, "Statistical Prediction Analysis", Cambridge University Press, 1975.
- [3] S. Geisser, *Predictive inference* **55** CRC press, 1993.
- [4] Z. Chen, "A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function". *Statistics and Probability Letters* **49**, 155-162, 2000. [https://doi.org/10.1016/S0167-7152\(00\)00044-4](https://doi.org/10.1016/S0167-7152(00)00044-4)
- [5] M.A. Selim, "Bayesian Estimations from the Two-Parameter Bathtub-Shaped Lifetime Distribution Based on Record Values". *Pakistan Journal of Statistics and Operation Research* **8**(2):155-165, 2012. <https://doi.org/10.18187/pjsor.v8i2.328>
- [6] S. F. Niazi and A. M. Abd-Elrahman "Bayesian prediction bounds for a two-parameter lifetime distribution with bathtub-shaped failure rate function". *SYL-WAN* **159**(6), 34-50, (2015).
- [7] A. M. Sarhan, & J. Apaloo, "Inferences for a two-parameter lifetime distribution with bathtub shaped hazard based on censored data". *International Journal of Statistics and Probability*, **4**(4), 77-92, 2015.
- [8] S. F. Niazi and A. M. Abd-Elrahman, "Bayesian prediction bounds of doubly type-II censored samples for a new bathtub shape failure rate life time model". *Journal of Applied Mathematics and Statistical Sciences* **1**(2), 21-30, 2016.
- [9] A. R. Alanzi and S. F. Niazi, "Bayesian Prediction Intervals For A New Bathtub Shape Failure Rate Lifetime Model In The Presence Of Outliers". *Advances and Applications in Statistics*, **63**, Pages 1-22, 2020.
- [10] A. Algarni, A. M. Almarashi, H. Okasha, & H. K. T. Ng, "E-Bayesian Estimation of Chen Distribution Based on Type-I Censoring Scheme". *Entropy*, **22**(6), 636, 2020. <https://doi.org/10.3390/e22060636>
- [11] A. M. Sarhan, D. C. Hamilton, and B. Smith, "Parameter estimation for a twoparameter bathtub-shaped lifetime distribution" *Applied Mathematical Modelling*, (36), 5380-5392, 2012. <https://doi.org/10.1016/j.apm.2011.12.054>
- [12] M. H. Chen, & Q. M. Shao, "Monte Carlo estimation of Bayesian credible and HPD intervals". *Journal of Computational and Graphical Statistics*, **8**(1), 69-92, 1999. <https://doi.org/10.1080/10618600.1999.10474802>
- [13] N. Balakrishnan, & R. S. Ambagasptiya, "Relationships among moments of order statistics in samples from two related outlier models and some applications". *Communications in Statistics-Theory and Methods*, **17**(7), 2327-2341, 1988. <https://doi.org/10.1080/03610928808829749>

# Interpolatory Projection Techniques for $\mathcal{H}_2$ Optimal Structure-Preserving Model Order Reduction of Second-Order Systems

Md. Motlubar Rahman<sup>1</sup>, Mohammad Monir Uddin<sup>2</sup>, Laek Sazzad Andallah<sup>1</sup>, Mahtab Uddin<sup>3,\*</sup>

<sup>1</sup>Department of Mathematics, Jahangirnagar University, Savar, 1342, Bangladesh

<sup>2</sup>Department of Mathematics and Physics, North south University, Dhaka, 1229, Bangladesh

<sup>3</sup>Institute of Natural Sciences, United International University, Dhaka, 1212, Bangladesh

## ARTICLE INFO

Article history:

Received: 08 July, 2020

Accepted: 11 August, 2020

Online: 28 August, 2020

Keywords:

Interpolatory projections

IRKA

Second-order systems

Structure-Preserving MOR

$\mathcal{H}_2$  optimal condition

## ABSTRACT

This paper focuses on exploring efficient ways to find  $\mathcal{H}_2$  optimal Structure-Preserving Model Order Reduction (SPMOR) of the second-order systems via interpolatory projection-based method Iterative Rational Krylov Algorithm (IRKA). To get the reduced models of the second-order systems, the classical IRKA deals with the equivalent first-order converted forms and estimates the first-order reduced models. The drawbacks of that of the technique are failure of structure preservation and abolishing the properties of the original models, which are the key factors for some of the physical applications. To surpass those issues, we introduce IRKA based techniques that enable us to approximate the second-order systems through the reduced models implicitly without forming the first-order forms. On the other hand, there are very challenging tasks to the Model Order Reduction (MOR) of the large-scale second-order systems with the optimal  $\mathcal{H}_2$  error norm and attain the rapid rate of convergence. For the convenient computations, we discuss competent techniques to determine the optimal  $\mathcal{H}_2$  error norms efficiently for the second-order systems. The applicability and efficiency of the proposed techniques are validated by applying them to some large-scale systems extracted from engineering applications. The computations are done numerically using MATLAB simulation and the achieved results are discussed in both tabular and graphical approaches.

## 1 Introduction

A Linear Time-Invariant (LTI) continuous-time second-order system can be formed as

$$\begin{aligned} M\ddot{\xi}(t) + D\dot{\xi}(t) + K\xi(t) &= Hu(t), \\ y(t) &= L\xi(t), \end{aligned} \quad (1)$$

where the matrices  $M, D, K \in \mathbb{R}^{n \times n}$  are time-invariant, the input matrix  $H \in \mathbb{R}^{n \times p}$  describing the external access to the system and the output matrix  $L \in \mathbb{R}^{m \times n}$  represents the output of the measurement of the system. For  $M$  is an identity or invertible matrix, the system (1) is called a standard state-space system or can be converted into the standard state-space system. The system is  $n$ -dimensional, that contains the state vectors  $\xi(t) \in \mathbb{R}^n$ , input (control) vectors  $u(t) \in \mathbb{R}^p$ , and output vectors  $y(t) \in \mathbb{R}^m$ . It is said to be continuous-time system as the input-output relations are established in the continuous-time interval  $[0, \infty)$ . The system (1) is specified as the Single-Input Single-Output (SISO) system for  $p = m = 1$ , and otherwise

specified as the Multi-Input Multi-Output (MIMO) system, where  $p, m \ll n$ , i.e., the number of input and output of the system is much less than the number of states. If all the finite-eigenvalues of the matrix pencil of the system (1) lie in the left half-plane ( $\mathbb{C}^-$ ), it is called asymptotically stable. This kind of dynamical systems emerge in several engineering disciplines such as automation & control systems, structural mechanics or multi-body dynamics, and electrical networks [1, 2, 3]. Applying the Laplace-transformation in the system (1), the transfer function that defines input-output mappings can be express as

$$G(s) = L(s^2M + sD + K)^{-1}H; \quad s \in \mathbb{C}. \quad (2)$$

Because of the arrangement of a large number of disparate devices and their complex structures, in the engineering applications most of the systems turn into very large-scale. The large-scale systems are infeasible for the simulation, controller design, and optimization within the computing tools for reasonable computational time and the memory allocation. To avoid the adversities,

\*Corresponding Author: Mahtab Uddin, +8801550605560 & mahtab@ins.uju.ac.bd



Model Order Reduction (MOR) can be applied to well-approximate the original models. So that, the aim of the automation & control systems can be fulfilled by reducing the work-load, feasible time-management and efficient use of system components. That is the system (1) needs to be replaced by the  $r$ -dimensional Reduced-Order Model (ROM) as

$$\begin{aligned} \hat{M}\ddot{\hat{\xi}}(t) + \hat{D}\dot{\hat{\xi}}(t) + \hat{K}\hat{\xi}(t) &= \hat{H}u(t), \\ \hat{y}(t) &= \hat{L}\hat{\xi}(t). \end{aligned} \quad (3)$$

where  $\hat{M}, \hat{D}, \hat{K} \in \mathbb{R}^{r \times r}$ ,  $\hat{H} \in \mathbb{R}^{r \times p}$ ,  $\hat{L} \in \mathbb{R}^{m \times r}$  and  $r \ll n$ . The target is to minimize the value of  $r$  by the trial and error basis such that system (3) approximates the system (1) keeping the system properties invariant. The transfer-function corresponding to the ROM (3) is denoted by  $\hat{G}$  and defined as

$$\hat{G}(s) = \hat{L}(s^2\hat{M} + s\hat{D} + \hat{K})^{-1}\hat{H}; \quad s \in \mathbb{C}. \quad (4)$$

Some system oriented approximation requirements need to be fulfilled for the ROM (3), for instance the error norm  $\|y(t) - \hat{y}(t)\|$  or equivalently  $\|G(\cdot) - \hat{G}(\cdot)\|$  needs to be optimized under appropriate norm estimation, for example  $\mathcal{H}_2$ ,  $\mathcal{H}_\infty$  norms (see [4] and references therein).

Initially, the concept of interpolatory projection technique for MOR was developed in [5], and later updated in [6]. A modified approach was introduced in [7], in which the rational Krylov approach was implemented. In the MOR techniques for large-scale dynamical systems, Krylov based approaches can implicitly satisfy moment matching conditions avoiding the explicit computations to ignore the ill-conditioned simulations [8, 9]. A number of strategies have been flourished over the years for selecting the interpolation points and the tangential directions, which are the leading ingredients for finding the ROMs to get the best approximation of full models [10]. Nowadays selecting the suitable set of interpolation points is associated with the seeking of  $\mathcal{H}_2$  optimality for MOR techniques [11]. The Iterative Rational Krylov Algorithm (IRKA) was developed that guarantees the  $\mathcal{H}_2$  optimality of the investigation of interpolation points for MOR [12].

**Definition-1:** For the system (1), the  $\mathcal{H}_2$  norm can be defined as

$$\|G(s)\|_{\mathcal{H}_2}^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}(G(j\omega)^H G(j\omega)) d\omega, \quad (5)$$

where  $G(j\omega)^H$  is the harmonic conjugate of the frequency response  $G(j\omega)$  of the system and  $\omega \in \mathbb{R}$  is the frequency on the imaginary axis of the transfer function  $G(s)$  defined in (2).  $\mathcal{H}_2$  norm is the most appropriate way to investigate the performance of the ROMs achieved by IRKA. Estimating the  $\mathcal{H}_2$  norm by the improper integral (5) is impossible in practice.

**Definition-2:** The ROM (3) is said to be  $\mathcal{H}_2$  optimal if  $\hat{G}(s)$  is stable and minimize the error system under certain norm as

$$\|G_{err}\|_{\mathcal{H}_2}^2 = \|G(s) - \hat{G}(s)\|_{\mathcal{H}_2}^2; \quad s \in \mathbb{C}. \quad (6)$$

The details of the  $\mathcal{H}_2$  norm of the error system defined in (6) will be discussed later in Section 4.

Usually in the present MOR approaches for the second-order systems (1) includes the conversion of second-order systems into equivalent first-order forms and imply some suitable MOR techniques to get the first-order ROMs, which is one of the most eye-catching drawbacks of the MOR technique as they are structure destroyer. Here we propose the MOR techniques for the second-order system by applying the interpolatory techniques via IRKA, where the systems need not convert into the first-order form, that's why the structure of the system remains invariant. Also, we estimate the minimized  $\mathcal{H}_2$  error norm of the ROMs under a certain level of tolerance, by simulating the matrix equations governed from the error system.

## 2 IRKA for Model Order Reduction of First-Order Systems

To find the ROM of second-order systems through equivalent first-order forms, there are a handful number of MOR techniques available implementing IRKA [13]. The interpolation-based techniques for the first-order generalized system has been reviewed in this section in brief. Let us consider the following first-order form

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + D_a u(t), \end{aligned} \quad (7)$$

where the system matrices  $E, A \in \mathbb{R}^{k \times k}$ ,  $B \in \mathbb{R}^{k \times p}$ ,  $C \in \mathbb{R}^{m \times k}$ , and  $D_a \in \mathbb{R}^{m \times p}$  with  $E$  is non-singular. Assuming above conditions, the corresponding ROM can be formed as

$$\begin{aligned} \hat{E}\hat{x}(t) &= \hat{A}\hat{x}(t) + \hat{B}u(t), \\ \hat{y}(t) &= \hat{C}\hat{x}(t) + \hat{D}_a u(t), \end{aligned} \quad (8)$$

To find the ROM (8), there needs to generate two projection matrices  $W$  and  $V$ . Then the reduced matrices are formed as

$$\begin{aligned} \hat{E} &:= W^T E V, & \hat{A} &:= W^T A V, \\ \hat{B} &:= W^T B, & \hat{C} &:= C V, & \hat{D}_a &:= D_a. \end{aligned} \quad (9)$$

For the MIMO systems, interpolatory projection technique IRKA has been developed in [14], where the authors discussed that the interpolation points and the tangential directions need to be essentially updated until attain the  $\mathcal{H}_2$  optimality. The distinct interpolation points,  $\{\alpha_i\}_{i=1}^r \subset \mathbb{C}$  and  $\{\beta_i\}_{i=1}^r \subset \mathbb{C}$  are required to form right and left projector matrices  $W$  and  $V$  with the tangential directions  $b_i \in \mathbb{C}^m$  and  $c_i \in \mathbb{C}^p$ , respectively. Then these projector matrices need to be formed as

$$\begin{aligned} \text{Range}(V) &= \text{span}\{(\alpha_1 E - A)^{-1} B b_1, \dots, (\alpha_r E - A)^{-1} B b_r\}, \\ \text{Range}(W) &= \text{span}\{(\beta_1 E^T - A^T)^{-1} C^T c_1, \dots, (\beta_r E^T - A^T)^{-1} C^T c_r\}, \end{aligned} \quad (10)$$

Here  $\hat{G}(s)$  approximately interpolates  $G(s)$  satisfying the conditions of MOR, such that

$$\begin{aligned} G(\alpha_i) b_i &= \hat{G}(\alpha_i) b_i, & c_i^T G(\beta_i) &= c_i^T \hat{G}(\beta_i), & \text{and} \\ c_i^T G(\alpha_i) b_i &= c_i^T \hat{G}(\alpha_i) b_i & \text{when } \alpha_i &= \beta_i, & \text{for } i = 1, \dots, r. \end{aligned} \quad (11)$$

For a non-negative finite integer  $j$ , the following condition will be satisfied

$$c_i^T G^{(j)}(\alpha_i) b_i = c_i^T \hat{G}^{(j)}(\alpha_i) b_i, \\ c_i^T C[(\alpha_i E - A)^{-1} E]^j (\alpha_i E - A)^{-1} B b_i = c_i^T \hat{C}[(\alpha_i \hat{E} - \hat{A})^{-1} \hat{E}]^j (\alpha_i \hat{E} - \hat{A})^{-1} \hat{B} b_i, \quad (12)$$

where the  $j$ -th moment of  $G(\cdot)$  is defined as  $C[(\alpha_i E - A)^{-1} E]^j (\alpha_i E - A)^{-1} B$ , which evaluated at  $\alpha_i$  and represents the  $j$ -th derivative of  $G(\cdot)$ .

For the SISO systems, there need to consider the interpolation points but not the tangential directions. The complete procedures of IRKA for the SISO systems has been provided in [15].

The basis for the Krylov based MOR techniques require a adjustable interpolation points accomplishing the Petrov-Galerkin conditions [16]. The summary of the first-order IRKA procedures are provided in Algorithm 1 and Algorithm 2.

#### Algorithm 1: IRKA for First-Order MIMO Systems.

**Input** :  $E, A, B, C, D_a$ .  
**Output** :  $\hat{E}, \hat{A}, \hat{B}, \hat{C}, \hat{D}_a := D_a$ .

- 1 Consider the initial assumptions for the the interpolation points  $\{\alpha_i\}_{i=1}^r$  and the tangential directions  $\{b_i\}_{i=1}^r$  &  $\{c_i\}_{i=1}^r$ .
- 2 Construct  $V = [(\alpha_1 E - A)^{-1} B b_1, \dots, (\alpha_r E - A)^{-1} B b_r]$ ,  
 $W = [(\alpha_1 E^T - A^T)^{-1} C^T c_1, \dots, (\alpha_r E^T - A^T)^{-1} C^T c_r]$ .
- 3 **while** (not converged) **do**
- 4 Find  $\hat{E} = W^T E V$ ,  $\hat{A} = W^T A V$ ,  $\hat{B} = W^T B$ ,  $\hat{C} = C V$ .
- 5 **for**  $i = 1, \dots, r$ . **do**
- 6 Evaluate  $\hat{A} z_i = \lambda_i \hat{E} z_i$  and  $y_i^* \hat{A} = \lambda_i y_i^* \hat{E}$  to find  $\alpha_i \leftarrow -\lambda_i$ ,  
 $b_i^* \leftarrow -y_i^* \hat{B}$  and  $c_i^* \leftarrow \hat{C} z_i^*$ .
- 7 **end for**
- 8 Repeat Step-2.
- 9  $i = i + 1$ ;
- 10 **end while**
- 11 Find the reduced-order matrices by repeating Step-4.

#### Algorithm 2: IRKA for First-Order SISO Systems.

**Input** :  $E, A, B, C, D_a$ .  
**Output** :  $\hat{E}, \hat{A}, \hat{B}, \hat{C}, \hat{D}_a := D_a$ .

- 1 Consider the initial assumptions for the the interpolation points  $\{\alpha_i\}_{i=1}^r$ .
- 2 Construct  $V = [(\alpha_1 E - A)^{-1} B, \dots, (\alpha_r E - A)^{-1} B]$ ,  
 $W = [(\alpha_1 E^T - A^T)^{-1} C^T, \dots, (\alpha_r E^T - A^T)^{-1} C^T]$ .
- 3 **while** (not converged) **do**
- 4 Find  $\hat{E} = W^T E V$ ,  $\hat{A} = W^T A V$ ,  $\hat{B} = W^T B$ ,  $\hat{C} = C V$ .
- 5 **for**  $i = 1, \dots, r$ . **do**
- 6 Evaluate  $\hat{A} z_i = \lambda_i \hat{E} z_i$  and  $y_i^* \hat{A} = \lambda_i y_i^* \hat{E}$  to find  $\alpha_i \leftarrow -\lambda_i$ .
- 7 **end for**
- 8 Repeat Step-2.
- 9  $i = i + 1$ ;
- 10 **end while**
- 11 Find the reduced-order matrices by repeating Step-4.

### 3 IRKA for Structure-Preserving Model Order Reduction of Second-Order Systems

This section contributes the Structure-Preserving Model Order Reduction (SPMOR) of second-order system (1) via IRKA. One of the conversions of

second-order system (1) into first-order form (7) is as follows

$$x(t) = \begin{bmatrix} \xi(t) \\ \dot{\xi}(t) \end{bmatrix}, \quad E = \begin{bmatrix} I & 0 \\ 0 & M \end{bmatrix}, \quad A = \begin{bmatrix} 0 & I \\ -K & -D \end{bmatrix}, \quad (13) \\ B = \begin{bmatrix} 0 \\ H \end{bmatrix}, \quad C = [L \quad 0], \quad \text{and } D_a = 0.$$

In the conventional techniques, to obtain an efficient ROMs of the second-order systems, at first, it is to convert into (13) essentially [17]. Then converted first-order form (7) can be implemented substantially and by applying Algorithm 1 or Algorithm 2 the equivalent first-order reduced order model (8) can be achieved.

Sometimes, that explicit formulation of (13) is prohibitive as the original model structure is destroyed and we cant return back to the original system. The structure preservation is essential for the large-scale second-order systems to controller design, execute the computation, and optimization. SPMOR allows meaningful interpretation of the physical system and provides a more accurate approximation to the full model.

In the MIMO systems, to avoid the explicit conversion and derive the SPMOR approach, the  $i$ -th column of  $V$  and  $W$  utilizing the shifted linear systems need to be computed as

$$(\alpha_i E - A) v^{(i)} = B b_i, \quad \text{and} \\ (\alpha_i E^T - A^T) w^{(i)} = C^T c_i, \quad i = 1, \dots, r. \quad (14)$$

Inserting the assumptions defined in (13) and applying matrix algebra, the system of linear equations in (14) lead to the followings

$$\begin{bmatrix} \alpha_i I & -I \\ K & \alpha_i M + D \end{bmatrix} \begin{bmatrix} v_1^{(i)} \\ v_2^{(i)} \end{bmatrix} = \begin{bmatrix} 0 \\ H b_i \end{bmatrix}, \quad \text{and} \\ \begin{bmatrix} \alpha_i I & K^T \\ -I & \alpha_i M^T + D^T \end{bmatrix} \begin{bmatrix} w_1^{(i)} \\ w_2^{(i)} \end{bmatrix} = \begin{bmatrix} L^T c_i \\ 0 \end{bmatrix}. \quad (15)$$

Albeit the matrices in (15) has higher dimension  $2n$ , they are sparse and can be efficiently simulated by suitable direct or iterative solvers [18, 19]. After the elimination and simplification of the system of linear equations governed from (15) for  $v_1^{(i)}$ ,  $v_2^{(i)}$ ,  $w_1^{(i)}$  and  $w_2^{(i)}$ , we have the followings

$$v_1^{(i)} = (\alpha_i^2 M + \alpha_i D + K)^{-1} H b_i, \\ v_2^{(i)} = \alpha_i v_1^{(i)}, \\ w_2^{(i)} = (\alpha_i^2 M^T + \alpha_i D^T + K^T)^{-1} L^T c_i, \\ w_1^{(i)} = (\alpha_i M^T + D^T) w_2^{(i)}. \quad (16)$$

The concept of the formation (13) can be utilized in the system (7) and then by converting this to an equivalent sparse linear system, which ensures the structure preservation and rapid rate of convergence.

According to the techniques developed in [20], at each iteration, the first-order representation (13) will be considered and the Algorithm 1 will be applied to find desired set of interpolation points  $\{\alpha_i\}_{i=1}^r$  efficiently. Also, Algorithm 1 updates the tangential directions  $\{b_i\}_{i=1}^r$  &  $\{c_i\}_{i=1}^r$  accordingly.

Due to the structure of system, we can split the desired projectors  $V$  and  $W$  as position and velocity levels [21]. We can partition the projectors  $V$  and  $W$  as

$$V = \begin{bmatrix} V_p \\ V_v \end{bmatrix} \quad \& \quad W = \begin{bmatrix} W_p \\ W_v \end{bmatrix}, \quad (17)$$

where  $V_p$ , &  $W_p$  are the position levels and  $V_v$  &  $W_v$  are the velocity levels of  $V$  &  $W$  respectively. Partitioning  $V$  &  $W$  according to (17) and applying equation (16), we can write

$$V_p = [v_1^{(1)}, v_1^{(2)}, \dots, v_1^{(r)}]; \quad \& \quad V_v = [v_2^{(1)}, v_2^{(2)}, \dots, v_2^{(r)}]; \\ W_p = [w_1^{(1)}, w_1^{(2)}, \dots, w_1^{(r)}]; \quad \& \quad W_v = [w_2^{(1)}, w_2^{(2)}, \dots, w_2^{(r)}]. \quad (18)$$

Again, for the SISO systems, to find the projectors  $V$  and  $W$  we need to consider the first-order representation (13) and apply the Algorithm 2 to find desired set of interpolation points  $\{\alpha_i\}_{i=1}^r$  only. Then for the SISO systems, (16) turns into the followings

$$\begin{aligned} v_1^{(i)} &= (\alpha_i^2 M + \alpha_i D + K)^{-1} H, \\ v_2^{(i)} &= \alpha_i v_1^{(i)}, \\ w_2^{(i)} &= (\alpha_i^2 M^T + \alpha_i D^T + K^T)^{-1} L^T, \\ w_1^{(i)} &= (\alpha_i M^T + D^T) w_2^{(i)}. \end{aligned} \quad (19)$$

Since there are two sets of projectors as in (18), we will get two types of SPMOR for the system (1), position level and velocity level, respectively. We summarize the above ideas for computing SPMOR (3) in Algorithm 3 and Algorithm 4 for the second-order system (1).

#### Algorithm 3: IRKA for Second-Order MIMO Systems.

**Input :**  $M, D, K, H, L$ .  
**Output :**  $\hat{M}_p, \hat{D}_p, \hat{K}_p, \hat{H}_p, \hat{L}_p, \hat{M}_v, \hat{D}_v, \hat{K}_v, \hat{H}_v, \hat{L}_v$ .

- 1 Consider the initial assumptions for the the interpolation points  $\{\alpha_i\}_{i=1}^r$  and the tangential directions  $\{b_i\}_{i=1}^r$  &  $\{c_i\}_{i=1}^r$ .
- 2 Consider  $v_1^{(i)}, v_2^{(i)}, w_1^{(i)}$  and  $w_2^{(i)}$  are defined in (16) and construct  $V_p = [v_1^{(1)}, v_1^{(2)}, \dots, v_1^{(r)}]$  &  $V_v = [v_2^{(1)}, v_2^{(2)}, \dots, v_2^{(r)}]$ ,  
 $W_p = [w_1^{(1)}, w_1^{(2)}, \dots, w_1^{(r)}]$  &  $W_v = [w_2^{(1)}, w_2^{(2)}, \dots, w_2^{(r)}]$ .
- 3 **while (not converged) do**
- 4 Find  $\hat{M}_p = W_p^T M V_p$ ,  $\hat{D}_p = W_p^T D V_p$ ,  
 $\hat{K}_p = W_p^T K V_p$ ,  $\hat{H}_p = W_p^T H$ ,  $\hat{L}_p = L V_p$ .  
and  $\hat{M}_v = W_v^T M V_v$ ,  $\hat{D}_v = W_v^T D V_v$ ,  
 $\hat{K}_v = W_v^T K V_v$ ,  $\hat{H}_v = W_v^T H$ ,  $\hat{L}_v = L V_v$ .
- 5 Use the first-order representation (13) in Algorithm (1) to get  $\hat{E}, \hat{A}, \hat{B}$  and  $\hat{C}$ .
- 6 Evaluate  $\hat{A}z_i = \lambda_i \hat{E} z_i$  and  $y_i^* \hat{A} = \lambda_i y_i^* \hat{E}$  to find  $\alpha_i \leftarrow -\lambda_i$ ,  
 $b_i^* \leftarrow -y_i^* \hat{B}$  and  $c_i^* \leftarrow \hat{C} z_i^*$  for all  $i = 1, \dots, r$ .
- 7 Repeat Step-2.
- 8  $i = i + 1$ ;
- 9 **end while**
- 10 Find the reduced matrices by repeating Step-4 and Step-5.

#### Algorithm 4: IRKA for Second-Order SISO Systems.

**Input :**  $M, D, K, H, L$ .  
**Output :**  $\hat{M}_p, \hat{D}_p, \hat{K}_p, \hat{H}_p, \hat{L}_p, \hat{M}_v, \hat{D}_v, \hat{K}_v, \hat{H}_v, \hat{L}_v$ .

- 1 Consider the initial assumptions for the the interpolation points  $\{\alpha_i\}_{i=1}^r$ .
- 2 Consider  $v_1^{(i)}, v_2^{(i)}, w_1^{(i)}$  and  $w_2^{(i)}$  are defined in (19) and construct  $V_p = [v_1^{(1)}, v_1^{(2)}, \dots, v_1^{(r)}]$  &  $V_v = [v_2^{(1)}, v_2^{(2)}, \dots, v_2^{(r)}]$ ,  
 $W_p = [w_1^{(1)}, w_1^{(2)}, \dots, w_1^{(r)}]$  &  $W_v = [w_2^{(1)}, w_2^{(2)}, \dots, w_2^{(r)}]$ .
- 3 **while (not converged) do**
- 4 Find  $\hat{M}_p = W_p^T M V_p$ ,  $\hat{D}_p = W_p^T D V_p$ ,  
 $\hat{K}_p = W_p^T K V_p$ ,  $\hat{H}_p = W_p^T H$ ,  $\hat{L}_p = L V_p$ .  
and  $\hat{M}_v = W_v^T M V_v$ ,  $\hat{D}_v = W_v^T D V_v$ ,  
 $\hat{K}_v = W_v^T K V_v$ ,  $\hat{H}_v = W_v^T H$ ,  $\hat{L}_v = L V_v$ .
- 5 Use the first-order representation (13) in Algorithm (2) to get  $\hat{E}, \hat{A}, \hat{B}$  and  $\hat{C}$ .
- 6 Evaluate  $\hat{A}z_i = \lambda_i \hat{E} z_i$  and  $y_i^* \hat{A} = \lambda_i y_i^* \hat{E}$  to find  $\alpha_i \leftarrow -\lambda_i$  for all  $i = 1, \dots, r$ .
- 7 Repeat Step-2.
- 8  $i = i + 1$ ;
- 9 **end while**
- 10 Find the reduced matrices by repeating Step-4 and Step-5.

## 4 Estimation of $\mathcal{H}_2$ error norm of the Reduced Order Model

The error system of the ROM (3) of the system (1) using the first-order representation (13) has the form

$$G_{err} = G(s) - \hat{G}(s) = C_{err}(sE_{err} - A_{err})^{-1}B_{err}, \quad (20)$$

where  $G(s)$  and  $\hat{G}(s)$  are mentioned in (2) and (4), respectively. In the error system (20), we have considered

$$E_{err} = \begin{bmatrix} E & 0 \\ 0 & \hat{E} \end{bmatrix}, A_{err} = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix}, B_{err} = \begin{bmatrix} B \\ \hat{B} \end{bmatrix}, \text{ and } C_{err} = \begin{bmatrix} C & -\hat{C} \end{bmatrix}. \quad (21)$$

Here  $E, A, B$  and  $C$  are provided in the first-order representation (13) of the second-order system (1). Also,  $\hat{E}, \hat{A}, \hat{B}$  and  $\hat{C}$  are the reduced-order form of the first-order representation containing the reduced-order matrices defined in (3), which can be achieved by the Algorithm 3 for MIMO systems or Algorithm 4 for SISO systems.

Let us assume the controllability Gramian and observability Gramian for the error system (20) are  $P_{err}$  and  $Q_{err}$ , respectively, and they can be partitioned as

$$P_{err} = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix}, Q_{err} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix} \quad (22)$$

where  $P_{11}, Q_{11} \in \mathbb{R}^{n \times n}$ ,  $P_{22}, Q_{22} \in \mathbb{R}^{r \times r}$  and  $P_{12}, Q_{12} \in \mathbb{R}^{n \times r}$ . In terms of a Galerkin approach  $\hat{G}(s)$  can be defined by considering the projectors  $V = P_{12}P_{22}^{-1}$  and  $W = -Q_{12}Q_{22}^{-1}$ , and the achieved ROM satisfies the desired first-order conditions of the optimal  $\mathcal{H}_2$  norm condition with the property  $W^T V = I$  [22].

The Gramians  $P_{err}$  and  $Q_{err}$  can be attained by simulating the following Lyapunov equations

$$\begin{aligned} A_{err}P_{err}E_{err}^T + E_{err}P_{err}A_{err}^T + B_{err}B_{err}^T &= 0, \\ A_{err}^T Q_{err}E_{err} + E_{err}^T Q_{err}A_{err} + C_{err}^T C_{err} &= 0. \end{aligned} \quad (23)$$

Inserting (21) and (22) in (23), after simplification we have the following system of matrix equations

$$\begin{aligned} AP_{11}E^T + EP_{11}A^T + BB^T &= 0, \\ AP_{12}\hat{E}^T + EP_{12}\hat{A}^T + B\hat{B}^T &= 0, \\ \hat{A}P_{22}\hat{E}^T + \hat{E}P_{22}\hat{A}^T + \hat{B}\hat{B}^T &= 0, \end{aligned} \quad (24)$$

and

$$\begin{aligned} A^T Q_{11}E + E^T Q_{11}A + C^T C &= 0, \\ A^T Q_{12}\hat{E} + E^T Q_{12}\hat{A} - C^T \hat{C} &= 0, \\ \hat{A}^T Q_{22}\hat{E} + \hat{E}^T Q_{22}\hat{A} + \hat{C}^T \hat{C} &= 0. \end{aligned} \quad (25)$$

The efficient approaches to estimate the  $\mathcal{H}_2$  norm of the error system (20) suitably can be written as

$$\begin{aligned} \|G_{err}\|_{\mathcal{H}_2}^2 &= \text{trace}(C_{err}P_{err}C_{err}^T) \\ &= \text{trace}(CP_{11}C^T) + \text{trace}(\hat{C}P_{22}\hat{C}^T) + 2\text{trace}(CP_{12}\hat{C}^T) \\ &= \|G(s)\|_{\mathcal{H}_2}^2 + \|\hat{G}(s)\|_{\mathcal{H}_2}^2 + 2\text{trace}(CP_{12}\hat{C}^T), \end{aligned} \quad (26)$$

and

$$\begin{aligned} \|G_{err}\|_{\mathcal{H}_2}^2 &= \text{trace}(B_{err}^T Q_{err}B_{err}) \\ &= \text{trace}(B^T Q_{11}B) + \text{trace}(\hat{B}^T Q_{22}\hat{B}) + 2\text{trace}(B^T Q_{12}\hat{B}) \\ &= \|G(s)\|_{\mathcal{H}_2}^2 + \|\hat{G}(s)\|_{\mathcal{H}_2}^2 + 2\text{trace}(B^T Q_{12}\hat{B}). \end{aligned} \quad (27)$$

The  $\mathcal{H}_2$  norm of the full model needs to be estimated once using the corresponding Gramians, and that is infeasible for a large-scale system by solving corresponding Lyapunov equations using the direct solvers. In this

case, we can use the low-rank ADI based technique provided in Algorithm 1 of [23] or rational Krylov subspace method provided in Algorithm 2 of [24] to find the low-rank factors  $Z_p$  and  $Z_q$  of the Gramians  $P_{11} = Z_p Z_p^T$  and  $Q_{11} = Z_q Z_q^T$  defined in the first equations of (24) and (25), respectively. These approaches provide us feasible ways to find the  $\mathcal{H}_2$  norm of the full model. Then, instead using the Gramians of the full model, we can use the low-rank Gramian factors in the first part of (26) and (27) as given below.

$$\|G(s)\|_{\mathcal{H}_2}^2 = \text{trace}(CP_{11}C^T) = \text{trace}\{(CZ_p)(CZ_p)^T\} \quad (28)$$

$$\|G(s)\|_{\mathcal{H}_2}^2 = \text{trace}(B^T Q_{11} B) = \text{trace}\{(B^T Z_q)(B^T Z_q)^T\} \quad (29)$$

Since the Lyapunov equations defined by third equations in (24) and (25) corresponding to the ROMs are very small in size, they can be solved by any direct solver or MATLAB `lyap` command in every iterations to find the second part of (26) and (27), respectively. Moreover to find the third part of (26) and (27), we can solve the generalized Sylvester equations given in the second equations in (24) and (25), respectively, using the sparse-dense Sylvester equation approach provided in Algorithm-4 of [25].

## 5 Numerical Results

The proposed techniques will be validated by applying them to the data generated from the real-world models given below. The computations will be carried out with the MATLAB<sup>®</sup> R2015a (8.5.0.197613) on a board with Intel<sup>®</sup>Core<sup>™</sup>i5 6200U CPU with 2.30 GHz clock speed and 16 GB RAM.

**Example-1: [International Space Station Model (ISSM)]** International Space Station Model is a part of modern aeronautics satisfying the second-order system defined in (1). This model is essential to control the system and maintain the varieties of operational modes at the international space station. The control of the system ensures the dynamic interaction within the components including adjustable structure alignments and clarification of the technical legitimacy of the flight.

The physical model of the international space station being assembled part by part. The modeling and control of the vibrations of the docking of an incoming spaceship is the prime concern under system constraints. Two types of ISSM exist for fulfilling the requirements of the space stations, namely the Russian service Module (1R) and the Extended Service Module [26]. In our work, we have taken the Russian service Module (12A) with state dimension  $n = 270$  and the number of input-output is 3. The target model provides a sparse system derived from a complicated mechanical system.

**Example-2 [Clamped Beam Model (CBM)]** The Clamped Beam Model is derived from micro-actuators, which is mainly a non-linear electrodynamic system satisfying a set of partial differential equations with the boundary conditions [27]. A usual micro-actuator combines two parallel conducting electrodes, one of which is movable and the rest is fixed. The movable electrode is formed as the clamped beam.

By the discretization approach Finite Element Method (FEM), the Clamped Beam Model has been derived as (1). For the Clamped Beam Model the state dimension  $n = 348$  with single input-output. The force applied to the system at the movable end is taken as the input, whereas the resulting displacement is considered as the output.

**Example-3 [Scalable Oscillator Model (SOM)]** Scalable Oscillator Model is derived from a special class of mechanical systems consisting of three damping chains with all of the chains bonded to constant support by an extra damper at one end, whereas the other end of them bonded to a

heavy and inflexible mass [28]. There are some amount of spring units that united the heavy mass. All of the three individual chains composed of  $n_1$  number of masses and unit springs with uniform stiffness. The masses  $m_1, m_2, m_3$  with the stiffness's  $k_1, k_2, k_3$ , respectively, are the model parameters for the desired model. For the bonding mass  $m_0$  is the mass with stiffness  $k_0$ . Here  $\vartheta$  is the viscosity of extra supporting damper and all of the individual oscillator chains.

The mathematical formulation of the Scalable Oscillator Model can be written in the form of the system (1) having the state dimension  $n_\xi = 3n_1 + 1$ . The matrix  $M = \text{diag}[m_1 I_{n_1}, m_2 I_{n_1}, m_3 I_{n_1}, m_0]$  represents the combination of the masses of the system with the stiffness  $K$ . The damping matrix  $D$  consists the stiffness of the oscillator chains as the elements in the block diagram form and the bonding terms in the last row and column at  $n_1, 2n_1$  &  $3n_1$  positions of the diagonals. The model used in the numerical computation of single input-output with state dimension  $n = 9001$ . For practical purposes, we have considered the following assumptions.

$$m_1 = 1, m_2 = 2, m_3 = 3, m_0 = 10, \\ k_1 = 10, k_2 = 20, k_3 = 1, k_0 = 50, \quad \& \quad \vartheta = 5.$$

**Example-4 [Butterfly Gyro Model (BGM)]** For the applications of inertial navigation, vibrating micro-mechanical based Butterfly-Gyro is an essential engineering tool for its theoretical and practical contribution. The basic Butterfly-Gyro structure consists of a wafer stack of three layers with the body sensor in the middle. The model is named Butterfly-Gyro because of the existence of a pair of double wings attached with a common frame in the layout of the sensor [29].

The mathematical formulation employing the matrix equations has been governed from the physical model numerically using the multiphysics simulation software ANSYS7 and can be written as the system (1). The original derived model is a  $n_\xi = 17361$ -dimensional second-order system with single input but 12 outputs. To adjust the structure of our computation we have taken the transpose of the output matrix as the input matrix. Thus, in our work, the number of input-output is 12.

Above models discussed here are available in the web-page for the *Oberwolfach Benchmark Collection*<sup>1</sup> in detail.

Table 1 depicts the dimensions of the full models and their ROMs via IRKA based SPMOR achieved by the Algorithm 3 and Algorithm 4. Here the dimensions of the ROMs are taken appropriately such the ROMs approximate the corresponding full model sufficiently.

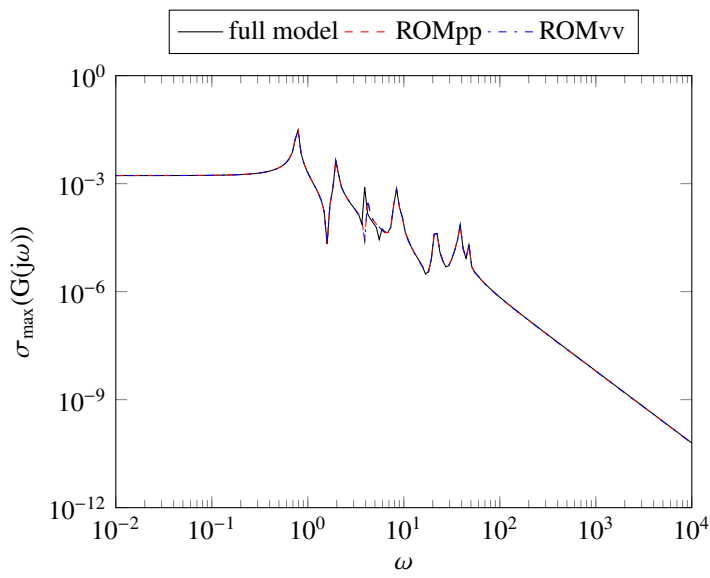
Table 1: Model examples and dimensions of full models and ROMs

Model	type	full model ( $n$ )	ROM ( $r$ )
ISSM	MIMO	270	20
CBM	SISO	348	30
SOM	SISO	9001	50
BGM	MIMO	17361	70

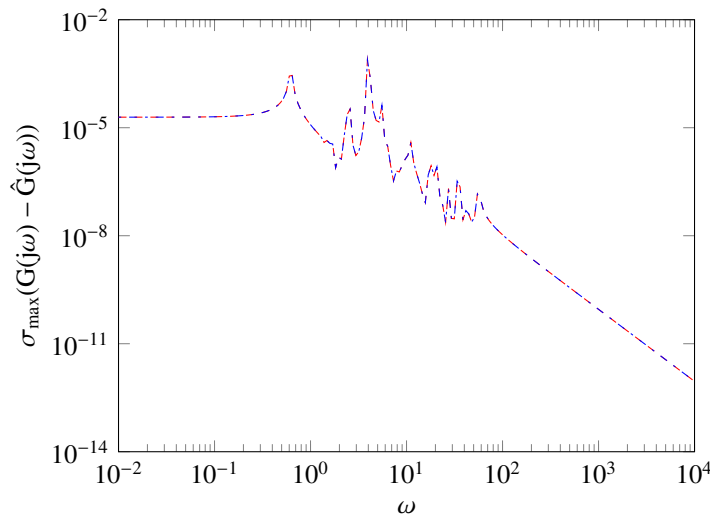
The transfer functions of the full models and corresponding ROMs in the sense of position-level and velocity-level with the desired dimensions are compared in the Figure 1, Figure 2, Figure 3 and Figure 4 for the models discussed above. In each of the figures, the first sub-figures display the comparisons of the transfer functions of the full model and its ROMs considering both of the position-level and velocity-level, whereas the second and third sub-figures depict their absolute errors and relative errors, respectively.

<sup>1</sup><https://sparse.tamu.edu/Oberwolfach>

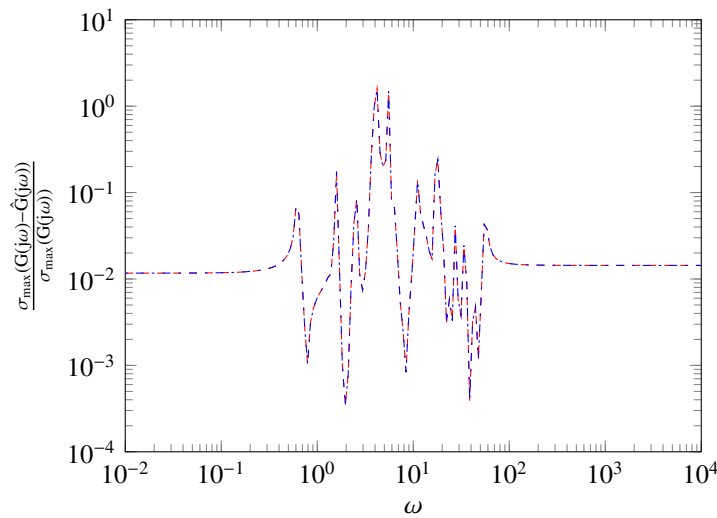




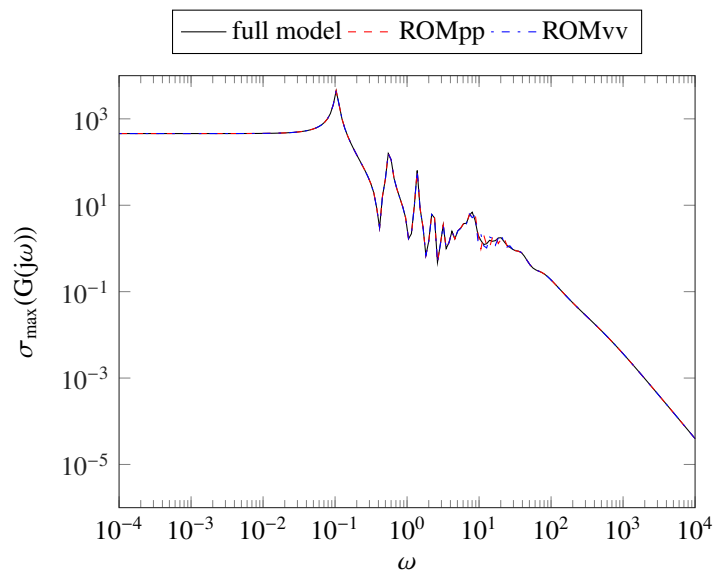
(a) Sigma plot



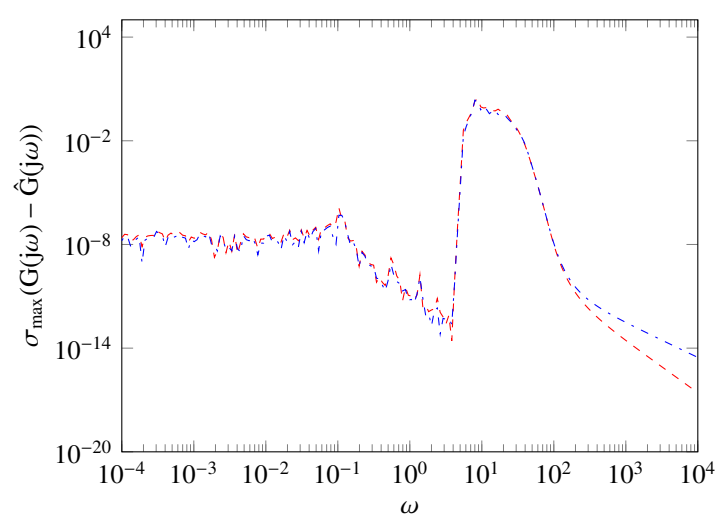
(b) Absolute error



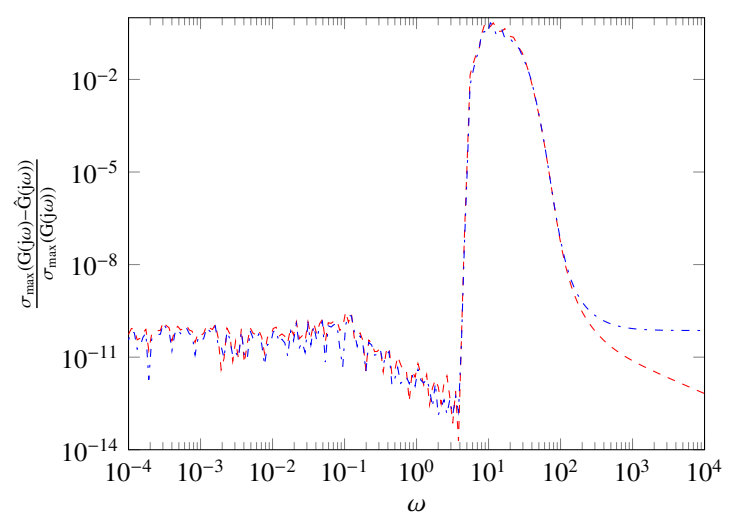
(c) Relative error



(a) Sigma plot



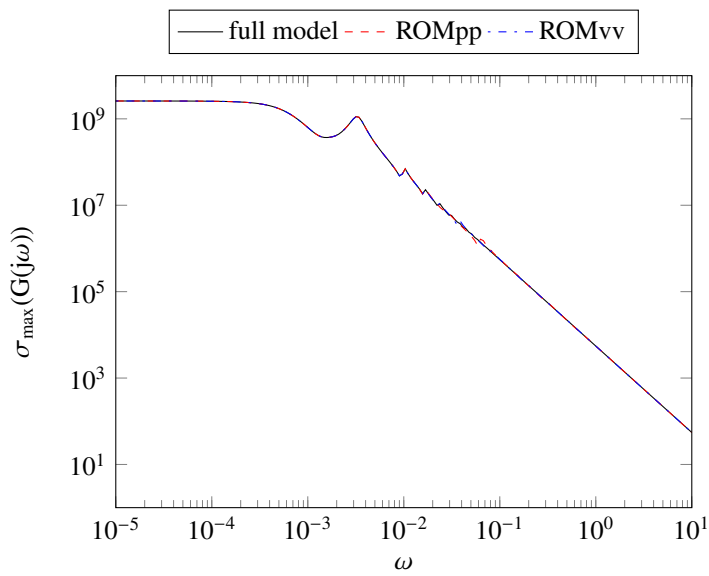
(b) Absolute error



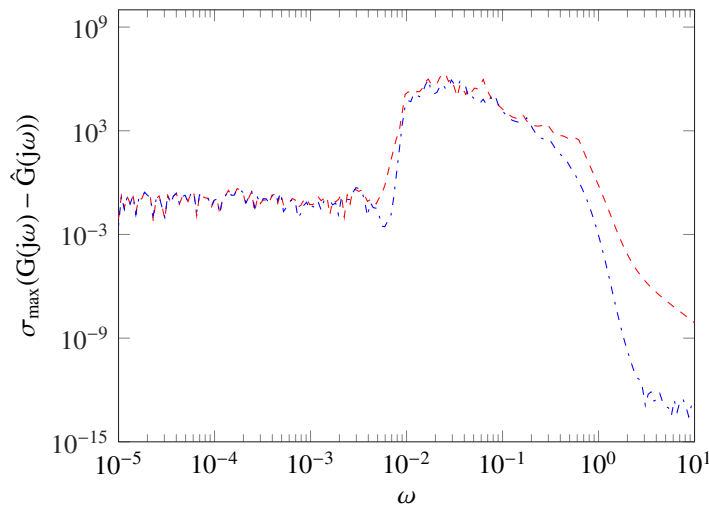
(c) Relative error

Figure 1: Comparison of full model and ROM computed by Algorithm 3 (for MIMO system) of the ISSM

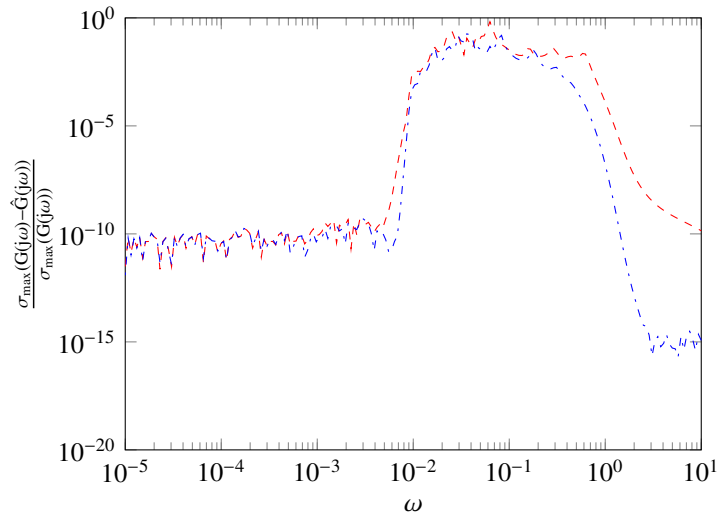
Figure 2: Comparison of full model and ROM computed by Algorithm 4 (for SISO system) of the CBM



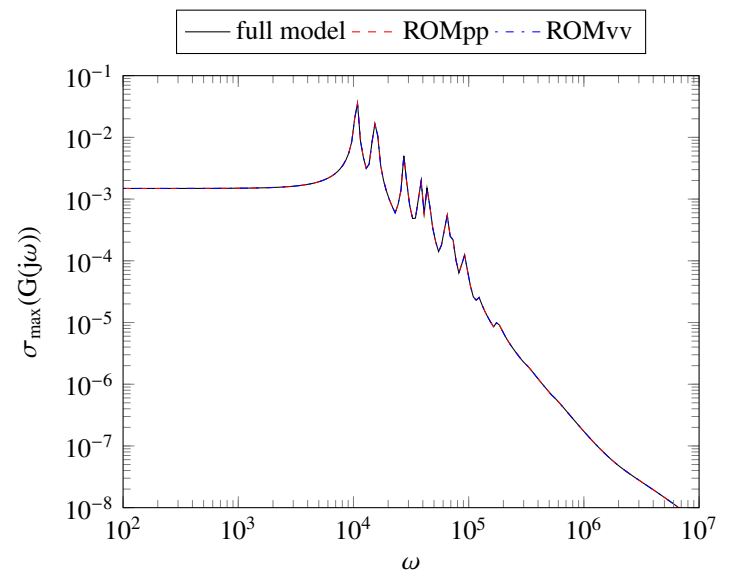
(a) Sigma plot



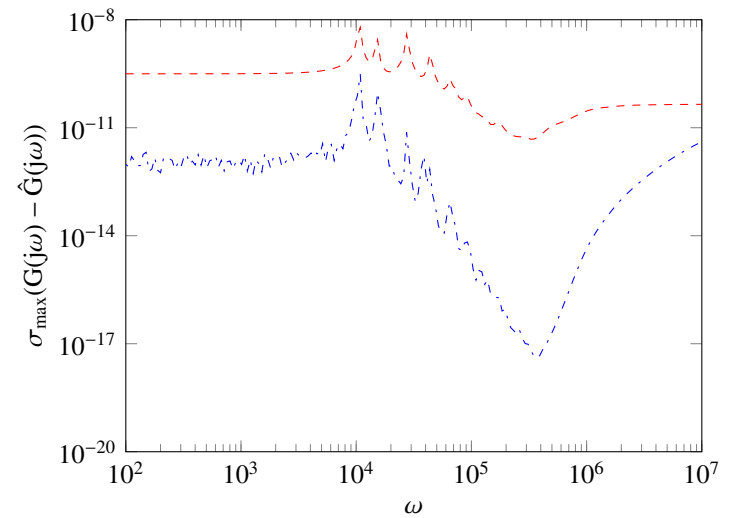
(b) Absolute error



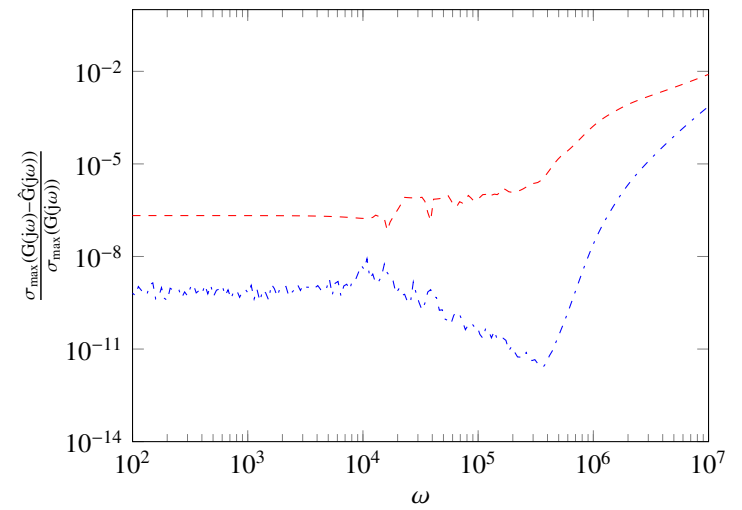
(c) Relative error



(a) Sigma plot



(b) Absolute error



(c) Relative error

Figure 3: Comparison of full model and ROM computed by Algorithm 4 (for SISO system) of the SOM

Figure 4: Comparison of full model and ROM computed by Algorithm 3 (for MIMO system) of the BGM

From the figures presented above, we have observed that the transfer functions of the full models and their ROMs of position-level and velocity-level are nicely matched. For this reason, the absolute errors and the relative errors for all cases are sufficiently small. Therefore it can be said that the proposed IRKA based SPMOR technique for finding ROMs of both of the levels of the second-order systems is efficient and robust for the target models.

Speed-up of the frequency-responses of the target models represents in the Table 2. For the time-convenient comparison, we have measured the execution time of the frequency-responses for a single cycle of the full models and corresponding ROMs. The speed-up of the system execution indicates the feasibility of the time-allocation which is the fundamental requirement for the system automation and controller design. It has been noticed that the proposed technique is more efficient based on time for higher-dimensional models. Noted that the ROMs of position-level and velocity-level have the same execution time and for this, we have omitted to show them individually.

Table 2: Speed-up comparisons for ROMs

	Model	dimension	execution time (sec)	speed-up
ISSM	full model	270	$5.67 \times 10^{-4}$	2.78
	ROM	20	$2.04 \times 10^{-4}$	
CBM	full model	348	$3.19 \times 10^{-3}$	12.97
	ROM	30	$2.46 \times 10^{-4}$	
SOM	full model	9001	$1.66 \times 10^{-2}$	55.54
	ROM	50	$2.98 \times 10^{-4}$	
BGM	full model	17361	$6.48 \times 10^{-1}$	609.13
	ROM	70	$1.07 \times 10^{-3}$	

In Table 3 the  $\mathcal{H}_2$  error norms of the ROMs for the target models for both position-level and velocity-level have been illustrated. Here, to find the  $\mathcal{H}_2$  norm of the full model we have used a low-rank ADI approach. It is observed that the  $\mathcal{H}_2$  error norm for CBM is small enough and for ISSM it is better. For BGM, the  $\mathcal{H}_2$  error norm is the best, whereas, for SOM, the  $\mathcal{H}_2$  error norm is less satisfactory in comparison to other target models but still not intolerable. Thus, it is evident that the proposed techniques are capable to estimate the  $\mathcal{H}_2$  optimal SPMOR for second-order systems with a few exceptions.

Table 3:  $\mathcal{H}_2$  error norm of the ROMs

Model	ROM	$\ G - \tilde{G}\ _{\mathcal{H}_2}$
ISSM	Position	$1.2 \times 10^{-6}$
	Velocity	$5.1 \times 10^{-6}$
CBM	Position	$1.1 \times 10^{-3}$
	Velocity	$3.6 \times 10^{-3}$
SOM	Position	$1.8 \times 10^1$
	Velocity	$3.6 \times 10^1$
BGM	Position	$6.6 \times 10^{-12}$
	Velocity	$7.8 \times 10^{-12}$

## 6 Conclusions

We have introduced the Structure-Preserving Model Order Reduction (SPMOR) techniques for second-order systems through the interpolatory

projection-based method Iterative Rational Krylov Algorithm (IRKA). In those techniques, the second-order systems need not convert into the first-order forms explicitly which are essential for preserving some important features of the systems. The SPMOR techniques enhance the rapid rate of convergence of the simulations and ensure the feasibility of further manipulations. We have derived the approaches to find the  $\mathcal{H}_2$  error norm for the optimality condition of the Reduced-Order Models (ROM) achieved from the second-order systems, that confirm the fidelity of the proposed techniques. We have investigated the applicability and efficiency of the proposed techniques numerically using MATLAB simulation and applied them to some real-world models derived from engineering applications. From the numerical investigations, it is evident that the proposed techniques are efficient to find ROMs of the second-order systems, that can approximate the full models with reasonable absolute errors and relative errors. The speed-up of the target systems provided the testimony of the aptness of the proposed techniques in the automation and control systems. Finally, it can be concluded that proposed techniques provide ROMs of the second-order systems with the  $\mathcal{H}_2$  optimal condition.

**Acknowledgment** This work is funded by the Bangladesh Bureau of Educational Information and Statistics (BANBEIS) under a project bearing ID MS20191055. Also, the first author is affiliated to the University Grants Commission (UGC) of Bangladesh.

## References

- [1] R. Riaza, "Differential-Algebraic Systems. Analytical Aspects and Circuit Applications," Singapore: World Scientific Publishing Co. Pte. Ltd., 2008.
- [2] M. M. Uddin, J. Saak, B. Kranz, and P. Benner, "Computation of a compact state space model for an adaptive spindle head configuration with piezo actuators using balanced truncation," *Production Engineering*, **6**, pp.577–586, 2012.
- [3] M. M. Uddin, "Gramian-based model-order reduction of constrained structural dynamic systems," *IET Control Theory & Applications*, 2018.
- [4] S. Gugercin, A. C. Antoulas, and C. A. Beattie, " $\mathcal{H}_2$  model reduction for large-scale dynamical systems," *SIAM J. Matrix Anal. Appl.*, **30**(2), pp. 609–638, 2008.
- [5] C. d. Villemagne and R. E. Skelton, "Model reductions using a projection formulation," *International Journal of Control*, **46**(6), pp.2141–2169, 1987.
- [6] E. Grimme, "Krylov projection methods for model reduction," Ph.D. Thesis, University of Illinois at Urbana Champaign, May 1997. [Online]. Available: <https://tel.archives-ouvertes.fr/tel-01711328>
- [7] A. Ruhe, "Rational Krylov algorithms for non-symmetric eigenvalue problems. ii. matrix pairs," *Linear algebra and its Applications*, **197**, pp.283–295, 1994.
- [8] M. M. Uddin, "Computational methods for model reduction of large-scale sparse structured descriptor systems," Ph.D. Thesis, Otto-von-Guericke-Universität, Magdeburg, Germany, 2015. [Online]. Available: <http://nbn-resolving.de/urn:nbn:de:gbv:ma9:1-6535>
- [9] M. Uddin, "Numerical study on continuous-time algebraic Riccati equations arising from large-scale sparse descriptor systems," Master's thesis, Bangladesh University of Engineering and Technology, 2020.
- [10] C. Lein, M. Beitelshmidt, and D. Bernstein, "Improvement of Krylov subspace reduced models by iterative mode-truncation," *IFAC-PapersOnLine*, **48**(1), pp. 178–183, 2015.
- [11] G. Flagg, C. Beattie, and S. Gugercin, "Convergence of the iterative rational Krylov algorithm," *Systems & Control Letters*, **61**(6), pp.688–691, 2012.
- [12] S. Gugercin, C. Beattie, and A. Antoulas, "Rational Krylov methods for optimal  $\mathcal{H}_2$  model reduction," submitted for publication, 2006.
- [13] M. M. Uddin, "Computational Methods for Approximation of Large-scale Dynamical Systems," Chapman and Hall/CRC, 2019.
- [14] Y. Xu and T. Zeng, "Optimal  $\mathcal{H}_2$  model reduction for large scale MIMO systems via tangential interpolation," *International Journal of Numerical Analysis and Modeling*, **8**(1), pp. 174–188, 2011.

- [15] K. Sinani, "Iterative rational Krylov algorithm for unstable dynamical systems and generalized co-prime factorizations," Ph.D. dissertation, Virginia Tech, 2016.
- [16] K. Carlberg, M. Barone, and H. Antil, "Galerkin v. least-squares Petrov–Galerkin projection in non-linear model reduction," *Journal of Computational Physics*, **330**, pp.693–734, 2017.
- [17] P. Benner, P. Kürschner, and J. Saak, "An improved numerical method for balanced truncation for symmetric second order systems," *Mathematical and Computer Modelling of Dynamical Systems*, **19**(6), pp. 593–615, 2013.
- [18] T. A. Davis, "Direct Methods for Sparse Linear Systems, Fundamentals of Algorithms," Philadelphia, PA, USA: SIAM, 2006.
- [19] Y. Saad, "Iterative Methods for Sparse Linear Systems," Philadelphia, PA, USA: SIAM, 2003.
- [20] S. A. Wyatt, "Issues in interpolatory model reduction: Inexact solves, second-order systems and DAEs," Ph.D. dissertation, Virginia Tech, 2012.
- [21] P. Benner and J. Saak, "Efficient balancing-based mor for large-scale second-order systems," *Mathematical and Computer Modelling of Dynamical Systems*, **17**(2), pp. 123–143, 2011.
- [22] D. A. Wilson, "Optimum solution of model-reduction problem," in *Proceedings of the Institution of Electrical Engineers*, **117**(6), pp.1161–1165, 1970.
- [23] S. Hasan, A. M. Fony, and M. M. Uddin, "Reduced model based feedback stabilization of large-scale sparse power system model," in *International Conference on Electrical, Computer and Communication Engineering (ECCE)*, IEEE, 2019, pp.1–6.
- [24] M. M. Uddin, M. S. Hossain, and M. F. Uddin, "Rational Krylov subspace method (RKSM) for solving the lyapunov equations of index-1 descriptor systems and application to balancing based model reduction," in *9th International Conference on Electrical and Computer Engineering (ICECE)*, IEEE, 2016, pp.451–454.
- [25] P. Benner, M. Köhler, and J. Saak, "Sparse-dense Sylvester equations in  $H_2$ -model order reduction," 2011.
- [26] S. Gugercin, A. Antoulas, and N. Bedrossian, "Approximation of the international space station 1R and 12A models," in *Proceedings of the 40th IEEE Conference on Decision and Control (Cat. No. 01CH37228)*, **2**, IEEE, 2001, pp. 1515–1516.
- [27] A. C. Antoulas, D. C. Sorensen, and S. Gugercin, "A survey of model reduction methods for large-scale systems," *Contemp. Math.*, **280**, pp.193–219, 2001.
- [28] N. Truhar and K. Veselić, "An efficient method for estimating the optimal dampers' viscosity for linear vibrating systems using Lyapunov equation," *SIAM J. Matrix Anal. Appl.*, **31**(1), pp.18–39, 2009.
- [29] J. Lienemann, D. Billger, E. B. Rudnyi, A. Greiner, and J. G. Korvink, "MEMS compact modeling meets model order reduction: Examples of the application of Arnoldi methods to microsystem devices," in the technical proceedings of the 2004 nanotechnology conference and trade show, *Nanotech*, **4**, 2004.



# Towards Classification of Shrimp Diseases Using Transferred Convolutional Neural Networks

Nghia Duong-Trung<sup>1</sup>, Luyi-Da Quach<sup>1</sup>, Chi-Ngon Nguyen<sup>\*2</sup>

<sup>1</sup>Software Engineering Department, FPT University, Can Tho city, 94000, Vietnam

<sup>2</sup>College of Engineering Technology, Can Tho University, Can Tho city, 94000, Vietnam

## ARTICLE INFO

Article history:

Received: 29 May, 2020

Accepted: 12 August, 2020

Online: 28 August, 2020

Keywords:

Shrimp diseases

Transfer learning

Convolutional neural networks

## ABSTRACT

Vietnam is one of the top 5 largest shrimp exporters globally, and Mekong delta of Vietnam contributes more than 80% of total national production. Along with intensive farming and growing shrimp farming area, diseases are a severe threat to productivity and sustainable development. Timely response to emerging shrimp diseases is critical. Early detection and treatment practices could help mitigate disease outbreaks, leading to on-site diagnostics, instant services recommendation, and front-line treatments. The authors establish a contribution hub for data collection in the ethnographic fieldwork of Mekong delta. Several deep convolutional neural networks are trained by applying the transfer learning technique. We have investigated six common reported shrimp diseases. The classification accuracy is achieved of **90.02%**, which is very useful in extremely non-standard images. Throughout the work, we raise the attention of shrimp experts, computer scientists, treatment agencies, and policymakers to develop preventive strategies against shrimp diseases.

## 1 Introduction

The shrimp farming industry in Vietnam has been expanding continuously for decades, especially in the Mekong Delta. Regarding the shrimp industry alone, the region contributes more than 80% of total production, mainly from Bac Lieu (BL), Ca Mau (CM), Ben Tre (BT), and Soc Trang (ST) [1], see Figure 1. In 2017, these provinces contributed more than 50% of Vietnam's total shrimp production. In 2018, the region's total shrimp farming area was about 720,000 hectares, with a total of 745,000 tons. Leading the production is the giant tiger shrimp (*Penaeus monodon*), whose output is around 300 ktons per year. In the first seven months of 2019, Vietnam's shrimp exports reached US\$ 1.8 billion (<https://tongcucchuysan.gov.vn/en-us/Fisheries-Trading/doc-tin/013419/2019-08-27/vietnam-shrimp-exports-started-to-reverse>). The leading shrimp importers of Vietnam include the EU, the USA, Japan, China, Korea, Canada, Australia, ASEAN countries, and Switzerland, accounting for 96% total export value (<http://seafood.vasep.com.vn/702/onecontent/certifications.html>). On January 18, 2018, Decision 79/QD-TTg of the Vietnamese Prime Minister issued a national plan for shrimp industry development to 2025. The goal is to export shrimp worth 10 billion USD from 2020 to 2025 (<https://customsnews.vn/small-production->

[http://agro.gov.vn/vn/tID25840\\_Nam-2020-xuat-khau-tom-dat-55-ty-USD.html](http://agro.gov.vn/vn/tID25840_Nam-2020-xuat-khau-tom-dat-55-ty-USD.html)) (i.e., an average growth of 12.7% per year) of which the value of brackish water shrimp exports will be USD 8.4 billion; the total area of brackish shrimp farming will reach 750,000 hectares; the centralized area for rearing giant freshwater prawn will be 50,000 hectares; lobster farming will reach 1,300,000  $m^3$  cages; total shrimp production will reach 1,153,000 tons (average growth of 6.73% per year); brackish shrimp will be archived by 1,100,000 tons; giant freshwater shrimp reaches 50,000 tons, and lobster reaches 3,000 tons ([http://agro.gov.vn/vn/tID25840\\_Nam-2020-xuat-khau-tom-dat-55-ty-USD.html](http://agro.gov.vn/vn/tID25840_Nam-2020-xuat-khau-tom-dat-55-ty-USD.html)). Although they occupy less than 10% of the land devoted to shrimp farming, they contribute about 80% of total production [?]. The growth of the shrimp industry is necessary to provide shrimp-based products for increasing global population demands. The need for such manageable growth is intensively recognized as an essential goal. However, strengthening the intensification of production can also make shrimp more susceptible to diseases. For example, in 2012, one-sixth of the total area was severely damaged by infectious diseases, acute hepatopancreatic necrosis disease (AHPND), or early mortality syndrome (EMS) [2]. Warnings on potentially sustainable intensification indicate the increasing understanding of the government and shrimp farmers of the environmental, economic, and social impacts of increasing

\*Corresponding Author: Chi-Ngon Nguyen, Can Tho University, Can Tho city, 94000, Vietnam. Contact No. (+84)0918.538.224. Email: ncngon@ctu.edu.vn

aquaculture production and its trade-off. The Vietnamese government has cooperated with international partners to find solutions to improve the sustainable and effective shrimp farming practices of the country and the Mekong Delta, focusing on small-scale shrimp farming (<https://psmag.com/environment/the-environmental-impacts-of-shrimp-farming-in-vietnam>), see Figure 2. Along with intensive farming and growing shrimp farming area, diseases are a severe threat to productivity and sustainable development of shrimp farming.

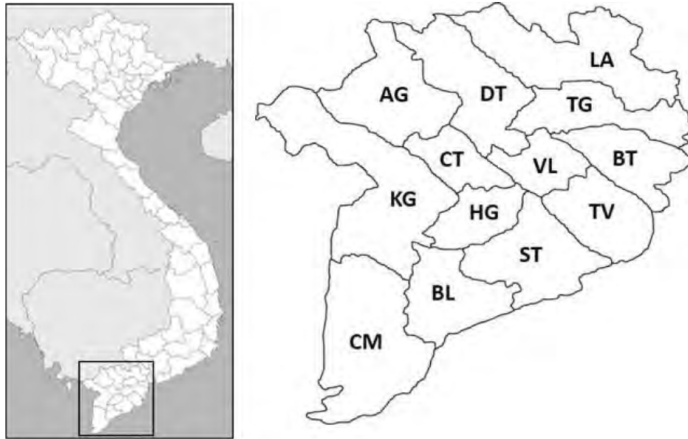


Figure 1: Mekong Delta of Vietnam and its provinces.

Although there are many policies and plans from the government and localities for the development of the shrimp industry in the Mekong Delta, shrimp production in many provinces continues to suffer significant economic losses due to the impact of shrimp diseases. Besides, the intensive farming of many shrimp crops per year is also affected because shrimp farmers have to spend time preparing new shrimp ponds after the shrimp diseases. According to an FAO survey conducted from 2013 to 2016, all respondents said shrimp diseases affected their farms (<http://www.fao.org/3/ca6702en/ca6702en.pdf>). Commonly reported diseases are AHPND, EMS, white spot syndrome virus (WSSV), white feces syndrome (WFS), and yellow-headed virus (YHV). Regarding WSSV, shrimp farmers are aware of an infection occurrence over an extended period, e.g., from 25 to 60 days after stocking. The delay in disease detection and treatment intervention can lead to the severe loss of the entire shrimp crop. Shrimp farmers reported that EMS, WSSV, and YHV diseases could not be controlled while WFS is manageable. The most regular exercise shrimp farmers apply to reduce disease risks is to buy quality food from reputable hatcheries and prepare and disinfect ponds. However, these reactions depend significantly on their own experiences without much scientific knowledge and standardization of treatment. These exercises make the shrimp industry in Mekong Delta vulnerable to the rapid and un-manageable propagation of emerging diseases. According to statistics on Asian Fisheries Science, 8.72% of *Penaeus monodon* ponds and 32.48% of *Penaeus vannamei* ponds in 2014 were reported to have been affected by shrimp disease [3]. In 2015, the monodon-affected ponds of AHPND were reported to be 5,875 hectares, while more than 5,509 hectares used for *Penaeus vannamei* culture were infected. WSSV remains the most important viral pathogen of farmed shrimp, with the ability to infect quickly

and often results in from 80% to 100% crop failure. Infections of *Penaeus monodon* and *Penaeus vannamei* ponds in the Mekong Delta in 2015 were estimated US\$ 11.02 million loss. When preparing to submit this article, the authors read the information in Quang Ngai province (Vietnam) and knew that the People's Committee of Binh Son district had just decided to handle the compulsory destruction of 21 diseased shrimp ponds (<http://vietlinh.vn/tin-tuc/2020/nuoi-trong-thuy-san-2020-s.asp?ID=470>). Accordingly, local authorities were forced to destroy all white leg shrimp in 21 shrimp ponds with an area of 49700m<sup>2</sup> of 11 shrimp farmers. Dead shrimps were diagnosed with WSSV and AHPND. Shrimp farmers did not know the identity of shrimp diseases and asked for information on shrimp forums. It causes devastating delays in detection and timely intervention.

Consequently, timely response to emerging shrimp diseases is critical. Early detection and treatment practices could help mitigate disease devastation, leading to on-site diagnostics, instant services recommendation, and front-line treatments.

In contrast with the abundance of research on shrimp diseases [4]- [7], computer science applications for them are uncommonly explored, especially identification and classification. Thanks to the ubiquity of mobile devices and social media networks, shrimp farmers have many online channels and forums to post images of shrimp diseases and ask for any detection, help on knowledge and treatments, see two examples in Figure 3. The volume of information related to shrimp diseases has been accumulated overtimes, and those being generated by shrimp farmers should be notified to experts, treatment agencies, and policymakers. There is a need to have a place to collect images for integrating legacy information with contemporary knowledge related to shrimp diseases. Accurate and early detection is preliminary to any intervention and recommendation of treatments. It is obvious to say that early detection and intervention are better than treatment when diseases get worse. Hence, leveraging technology in the easily automatic identification of these diseases has become essential.



Figure 2: An example of small-scale shrimp farming in Mekong Delta. Shrimp farmers dig a pond and establish several holding pens next to their house. Image courtesy to Roberto Schmidt/AFP/Getty Images.

To the best of our knowledge, the authors have made several contributions as follows.

- First, we establish a contribution hub where farmers can send images of emerging shrimp diseases. This kind of contribution hub or center has not existed before. Several images have been sent to shrimp experts at local universities for labeling.
- Then, a machine learning system is deployed to identification diseases. Although the number of images is small, the detection accuracy should not be sacrificed and less demand for computing power. More specifically, the authors apply convolutional neural networks applying the transfer learning technique. Addressing those requirements lead to our second contribution.
- Third, we have investigated six common reported shrimp diseases, which make the experiments the most intensive in the literature.



Figure 3: Shrimp farmers took pictures of infected samples, posted them to a Facebook forum and asked for help on the emerging diseases. The text above the pictures on the left means "What is the disease of this shrimp? How to cure?"

## 2 Related Work

The purpose of this work is neither to discuss existing research on shrimp diseases in aquaculture through the lens of biologists [8, 9] nor the effects of diseases over shrimp industry from an ecological perspective [10]. By the time we wrote a literature review, there is very little research on the classification or detection of shrimp diseases based on images. YHV recognition application was designed in [11]. They proposed an image processing algorithm based on the Niblack algorithm to detect and eliminate shrimp with only YHV from the gather lines. Another work devoted to WSSV detection in shrimp images using the K-means clustering technique (<https://www.semanticscholar.org/paper/White-Spot-Syndrome-Virus-Detection-in-Shrimp-using-Nagalakshmi-Jyothi/5d77b95d947adf375139a01662b6b308e2024811>). They equipped special spectrographic cameras, which are either not affordable to shrimp farmers or applicable in real scenarios. Another exciting research on the identification of softshell and sound shrimps, e.g., caused by physical illness or pathological disease, has been done in [12]. Several exciting shrimp image analysis

<sup>1</sup><https://sites.google.com/view/shrimp-image-collection/home>

applications but not disease detection can be referred to as shrimp freshness evaluation [13] and marine organism classification [14].

Many previous researchers have focused their attention on deploying a more accurate deep convolutional neural network regarding shrimp images. Zihao *et al.* developed ShrimpNet-3 [15], for recognizing shrimp, which is an architecture modification from LeNet-5 [16]. Similarly, the idea of modifying the architecture of a deep neural network like AlexNet, in this case, is investigated in [17]. That paper addressed the task of classification between a softshell and sound shrimp from an imaging perspective. The mentioned works examined several architectures, but the main point is that the networks need to be trained from scratch. These papers have not investigated the transfer learning approach which our work intends to.

## 3 System Components

In this Section, the authors briefly describe the proposed architecture to satisfy the paper's contributions. We illustrate our system components in Figure 4. We launch a data collection hub<sup>1</sup> where shrimp farmers across Mekong Delta region and other countries can send their images. It happens is step 1 in the system components. In step 2, shrimp images are sent to experts in local universities for ground-truth labeling. We can see in Figure 3, shrimp farmers send pictures with a diverse background, including fingers, catching equipment, soil, and floor. In step 3, we replace the background by white color. Steps 4, 5, and 6 are the necessary machine learning process. However, we do not train the models from scratch. Instead, we apply the transfer learning technique, which is further described in Section (4). At step 7, we contact the farmers who had sent shrimp images to ask for further details of the emerging diseases.

## 4 Technical Background

### 4.1 Transfer Learning

The traditional machine learning process emphasizes the entire process of independent learning from data collection, model selection and training, performance evaluation, and metadata tuning. Modern object identification and classification models with millions of parameters can take weeks to complete training. Therefore, the first need is that the database used for model training must be enormous. The second requirement is computer infrastructure and financial support, especially in high power-consumption algorithms such as neural networks.

**Definition 1** Given a training data set  $X_{train} \in \mathbb{R}^{n \times m}$  and equivalent label  $Y_{train} \in \mathbb{R}^n$  where  $n$ , and  $m$  are the number of samples and the number of attributes, respectively. Traditional machine learning methods seek to build a  $f$  prediction model by training on pairs  $(X_{train}, Y_{train}) = \{(x_i, y_i), \dots, (x_n, y_n)\}$ .

A principal assumption in applying machine learning algorithms is that the training and test data must share a similar feature space,



more specifically  $X_{train}$  and  $X_{test}$  are described using the same attributes. However, this assumption can be difficult to satisfy in many real-world applications. It reduces the generalization possibilities of machine learning models. The difference between transfer learning traditional machine learning is shown in Figure 5.

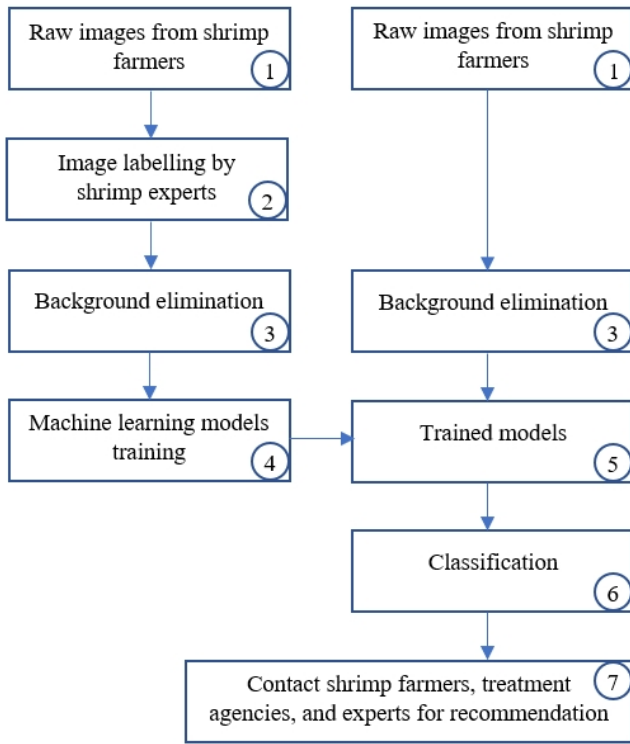


Figure 4: Our proposed system components.

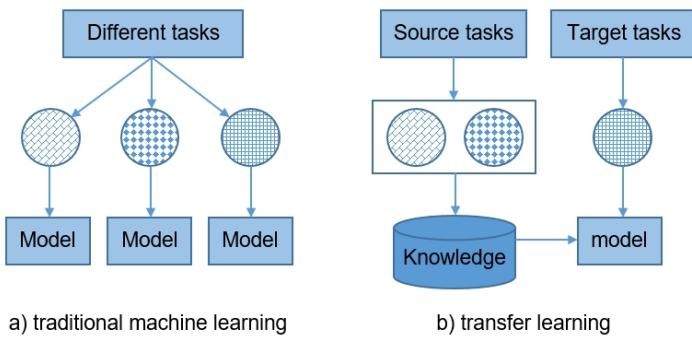


Figure 5: Difference between traditional machine learning and transfer learning.

The idea of transferring knowledge is a new approach to machine learning practices. It develops a mechanism of knowledge transferability in one or more source tasks and uses it to improve the prediction capacity in a new task [18]- [21]. It is like the propagation of knowledge from a well-developed domain with a lot of learning data to a less-developed domain that is limited due to insufficient data [22, 23]. The method allows machine learning models to be applied to new data drawn from distributions that are entirely different from the original data sources. Machine learning algorithms avoid suffering from the cold start problems by leveraging fully trained models on predefined large datasets [24, 25], Transfer learning aims

to improve generalization ability in the target task by leveraging knowledge from the source task(s). We can define transfer learning as follows.

**Definition 2** Given a source domain  $\mathcal{D}^{source}$  where we define a learning task  $\mathcal{T}^{source}$ . We denote  $\mathcal{D}^{target}$  and  $\mathcal{T}^{target}$  are a target domain and a target learning task, respectively. Transfer learning technique aims to help improve the learning of the  $f$  prediction model for  $\mathcal{T}^{target}$  utilizing knowledge in  $\mathcal{D}^{source}$  and  $\mathcal{T}^{source}$ , where  $\mathcal{D}^{source} \neq \mathcal{D}^{target}$  or  $\mathcal{T}^{source} \neq \mathcal{T}^{target}$ .

By re-weighting the observations in  $\mathcal{D}^{source}$ , the effects of the various samples are reduced. Conversely, similar instances will contribute more to  $\mathcal{T}^{target}$  and may lead to a more accurate prediction.

### 4.2 Convolutional Neural Networks (CNNs)

Convolutional neural networks (CNNs), initially attending several large-scale image classification challenges [24]-[26], is a class of deep, feed-forward artificial neural networks that are applied to the vast majority of machine learning problems due to the outstanding performance. The architecture of CNNs is usually a composition of layers that can be grouped by their functionalities. The high performance of CNNs is achieved by (i) their ability to learn rich level image representations and (ii) leveraging a tremendous amount of data. It can take millions of estimated parameters to characterize the network. However, in the situation of limited data sources, training CNNs from scratch is not very useful. There is a phenomenon in the learning of level image that each layer represents from color blobs, lines, and basic shapes to a high-level mixture of them. Several first-layer patterns appear to be generic to any dataset, Hence, its knowledge can be transferred and re-used. CNNs can be trained on large-scale datasets, and then re-weighted on a particular target dataset. The idea of training deep neural networks applying transfer learning has been explored by previous research [27]-[30].

#### 4.2.1 Inception-based models

The Inception deep convolutional architecture was introduced as GoogLeNet or Inception-v1 in [31]. Later architecture improvement was the introduction of batch normalization or Inception-v2 [32] and additional factorization or Inception-v3 [33].

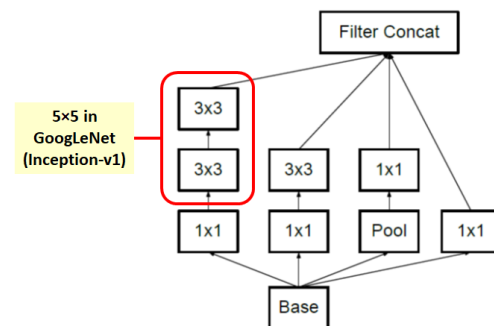


Figure 6: Inception module A.

In image representation, the prominent patterns can exist in different size variations. Therefore, the problem is to choose the



appropriate filter size that slides through the image. The idea of inception effectively addresses the mentioned problem by introducing several Inception modules where multiple-size filters operate on the same level. See Figures 6, 7, and 8 or Inception Module A, B, and C respectively. Consequently, the computational expense is significantly reduced because of the reduction of connections/parameters. The outline of our adapted implementation architecture is described in Table 10.

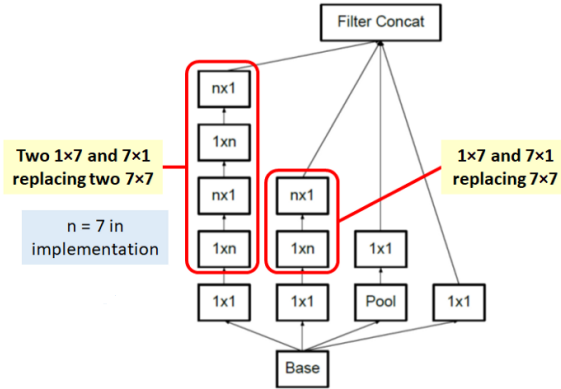


Figure 7: Inception module B.

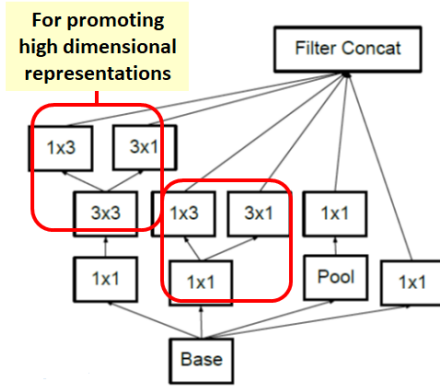


Figure 8: Inception module C.

#### 4.2.2 Depthwise separable convolution-based model

MobileNets is a family of mobile-first computer vision models designed by Google [34]. Its primary design is to help develop robust applications on mobile devices where restricted resources exist. Similar to how other popular CNNs models, it can be developed for addressing classification, detection, and segmentation tasks. The main point in Mobilenets is that two separated depthwise and pointwise convolutions perform the filtering and combination processes. More precisely, the model splits the convolution into a  $3 \times 3$  and a  $1 \times 1$  depthwise and pointwise, respectively. The depth-wise convolution applies a single filter to each input channel; the pointwise convolution combines the outputs of the depth-wise ones. These two different convolutions are illustrated in Figure 9. The depthwise separable convolution architecture is described in Table 1. The comparison between the inception-v3 model and mobilenets is presented in Table 2.

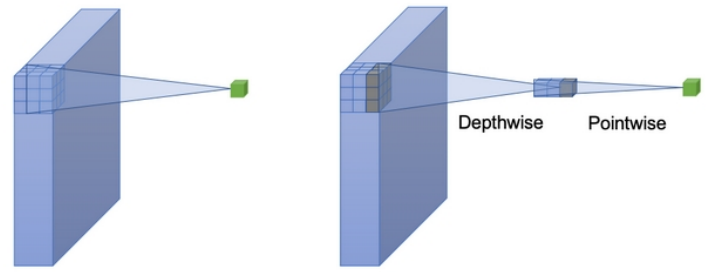


Figure 9: A depthwise and a pointwise convolutions.

Table 1: The MobileNets architecture [34]. At the softmax layer,  $c$  is the number of labels. dw means depthwise.

Layer (Stride)	Filter size	Input size
conv2d (s2)	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
conv2d dw (s1)	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
conv2d (s1)	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
conv2d dw (s2)	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
conv2d (s1)	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
conv2d dw (s1)	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
conv2d (s1)	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
conv2d dw (s2)	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
conv2d (s1)	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
conv2d dw (s1)	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
conv2d (s1)	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
conv2d dw (s2)	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
conv2d (s1)	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
$5 \times$ conv2d dw (s1)	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
conv2d (s1)	$1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
conv2d dw (s2)	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
conv2d (s1)	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
conv2d dw (s2)	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$
conv2d (s1)	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
avg_pool (s1)	$7 \times 7$	$7 \times 7 \times 1024$
fully connected (s1)	$1024 \times 1000$	$1 \times 1 \times 1024$
softmax (s1)	predictions	$1 \times 1 \times c$

Table 2: The comparison between the inception-based and MobileNets models.

Experimented models	# Mult-Adds (Million)	# Parameters (Million)
Inception	5000	23.6
MobileNets 1.0-244	569	4.24

## 5 Experiments

### 5.1 Datasets

In this work, the authors have evaluated the machine learning models on different data sources. We collected images from a wide range of shrimp forums<sup>2 3 4</sup>, where farmers posted the shrimp images and asked for help. We also launch a data collection hub as

<sup>2</sup><https://www.facebook.com/ShrimpCulture/>

<sup>3</sup><https://www.facebook.com/shrimp.nursery.tech/>

<sup>4</sup><https://www.facebook.com/ShrimpDisease/>

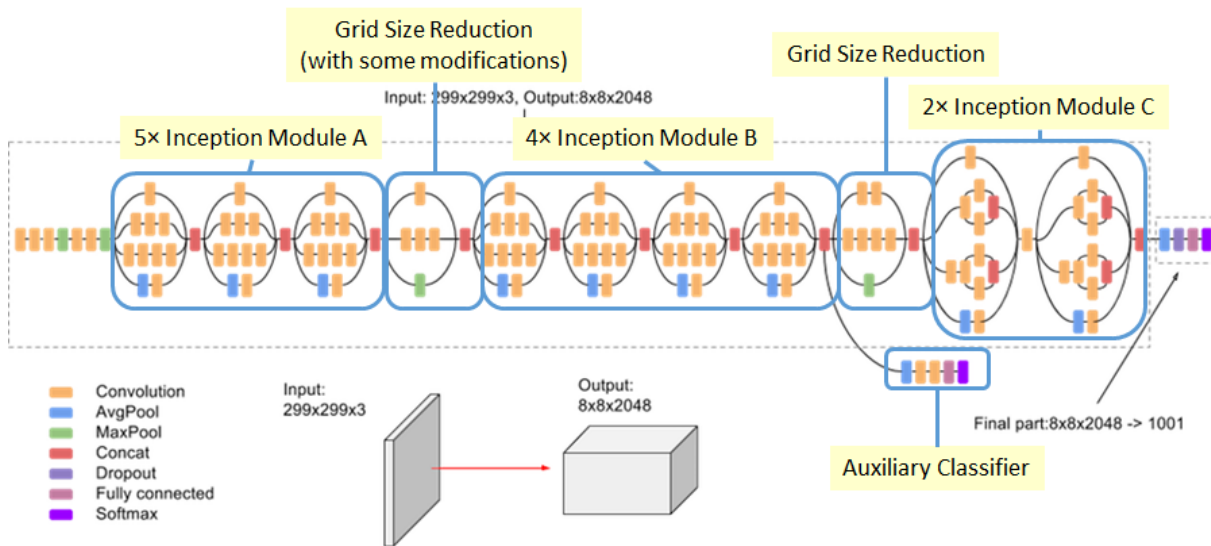


Figure 10: The outline of inception-based architecture. At the softmax layer, 1001 is replaced by the number of predicted labels in our dataset.

discussed in Section 3. The hub has been advertised to provinces administrations and shrimp farmers in Mekong Delta. The images were collected under various conditions, such as lighting, angle, distance, and background. Then, they were fed into learning models without manually resizing.

We sent raw images to experts at the College of Aquaculture & Fisheries<sup>5</sup> and experts in our project for labeling. The image background was manually removed using Adobe Photoshop by several technical assistants. Then the images were rechecked by the experts to be accepted as training images. Since the number of images we have collected weekly is not that much, the proposed manual pre-processing works acceptably.

Table 3: Shrimp diseases dataset used in the experiments.

#	Label	# Observations
1	Black gill	37
2	Black spot	52
3	WSSV	88
4	IMNV	45
5	NHB	47
6	YHV	35
7	Healthy shrimp	44
	<i>Total images</i>	348

At the time of conducting this paper, the authors receive images of 6 most asked shrimp diseases such as YHV [35], black gill [36], black spot [37], WSSV [38], infectious myonecrosis virus (IMNV) [39], and necrotizing hepatopancreatitis bacterium (NHB) [40]. We also add healthy shrimps as the seventh class. Example of photos collected by the authors presented in Figure 11. We show the distribution of samples into classes in Table 3.

<sup>5</sup><https://caf.ctu.edu.vn/en/>

## 5.2 Experimental Results

The implementation process is executed using Anaconda Python Distribution for Python 3.6 installed in Windows 10, TensorFlow version 1.7, and NVIDIA GPU Computing Toolkit 10.2. Our experiments were conducted on a regular machine equipped with a Core i7 with 2.5GHz clock speed, GPU NVIDIA GeForce 940MX, and 16GB of RAM.

In our experiments, we grid-search several hyperparameters as follows. The learning rate (lr) is {0.1, 0.01}. The number of epoch (ep) is {1000, 2000, 3000}. We randomly shuffle the dataset into an 80% training set, a 10% validation set, and a 10% test set, which is a regular choice in machine learning practices. We refer to it as the 80/10/10 splitting schema. The models are tuned by the best score on the validation set. The tuning process is to find the best combination of hyperparameters of the learning rate and the number of epoch. After that, we run the train models on the test set for final classification accuracy. We run the experiments five times and take the average scores.

The input size of the MobileNets model is  $224 \times 224 \times 3$  for height, width, and channel, respectively. Regarding the Inception model, the input size is  $299 \times 299 \times 3$ . The performance of models is reported in Table 4 and the identification accuracy is highlighted in Figure 12. The reported running time includes (i) shuffling images into 80/10/10 schema, (ii) training the model on the training/validation sets, and (iii) performing prediction capacity on the test set. Note that the best hyperparameters combination is 0.01 and 2000 for the learning rate and the number of epoch, respectively.

## 6 Remarks

The experimentation has been conducted on shrimp images of 7 categories with a total of 348 samples. Removing the image's background is the only step for the pre-processing procedure on collected



Figure 11: Shrimp diseases images. From left to right and the first row to bottom row, the labels are black gill, black spot, WSSV, IMNV, NHB, YHV. The last image illustrates the result of background elimination. The red circles indicate the disease expression.

images. The authors also tried the experiments without the step of background elimination, but the results were quite low. It is clearly stated in Table 4 that the combination of transfer learning and several convolutional neural networks works well in the shrimp images classification task. The inception-v3 model achieves the best classification accuracy score at 90.02%.

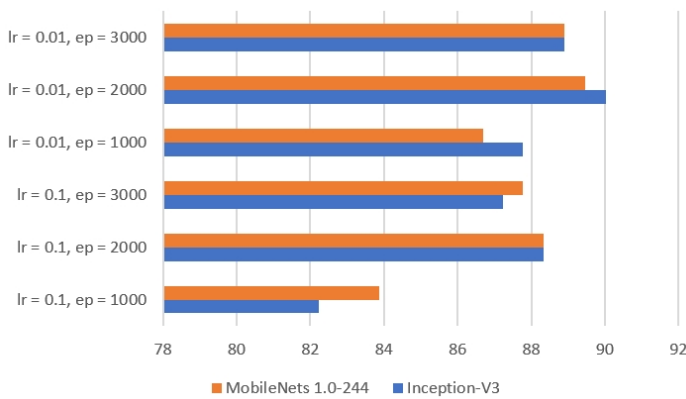


Figure 12: Accuracy comparison between two models.

One thing to note is that the shrimp' images themselves have not appeared in the ImageNet dataset, on which the models have been pre-trained. The experiment results have proved the remarkable ability of knowledge transferability even if the dataset distribution is different. The best score of each version of the models is reached when the learning rate is 0.01, and the number of epoch is 2000.

By the time of conducting this paper, there are newer versions of inception and mobilenets. However, the authors choose

inception-v3 and mobilenets-v1 because they are the standard transfer learning models, successfully integrated into many machine learning libraries, and can be adequately trained on a wide range of machines such as regular laptops, Google cloud TPU (<https://cloud.google.com/tpu/docs/inception-v3-advanced>), smartphones, and small IoT devices (<https://software.intel.com/en-us/articles/inception-v3-deep-convolutional-architecture-for-classifying-acute-myeloidlymphoblastic>). Newer versions need to be configured in different working environments.

## 7 Conclusion and Outlook

The growth of the shrimp industry is essential to provide shrimp-based products for global population demands, and the need for such manageable growth is intensively recognized as a principal goal. In the shrimp industry of Mekong delta of Vietnam, the farming sector is highly fragmented over regions. Shrimp farmers are vulnerable to epidemic diseases once they occur. Although various sources and agencies exist which aim to support services and treatments, there is not a single place where shrimp experts, farmers, computer scientists, treatment agencies, and policymakers can effectively connect. Shrimp farmers have to rely on neighbors and social networks for help on emerging diseases. By looking at the problems' insight, the authors, throughout this work, pave several first stones of a complete solution. We have launched a data collection center that shrimp farmers assure to send images and questions. We apply the transfer learning technique to leverage the effectiveness of trained deep convolutional neural networks on different non-standard datasets. A good accuracy score is archived, confirming this research direction is worth investigating. We have investigated six common reported

Table 4: The performance of two models regarding hyperparameters combination. The best scores are in **bold**.

Hyperparameters combination	Models			
	Inception-V3		MobileNets 1.0-244	
	Classification accuracy (%)	Running time (s)	Classification accuracy (%)	Running time (s)
lr = 0.1, ep = 1000	82.22 ± 1.47	97.75 ± 0.43	83.88 ± 2.30	<b>25.24 ± 2.26</b>
lr = 0.1, ep = 2000	88.34 ± 1.25	150.61 ± 5.29	88.34 ± 2.34	53.51 ± 6.95
lr = 0.1, ep = 3000	87.22 ± 1.53	146.99 ± 2.86	87.78 ± 1.53	75.09 ± 2.79
lr = 0.01, ep = 1000	87.78 ± 2.50	104.21 ± 3.08	86.68 ± 5.34	41.12 ± 0.57
lr = 0.01, ep = 2000	<b>90.02 ± 1.53</b>	128.97 ± 9.41	89.46 ± 2.34	58.47 ± 1.41
lr = 0.01, ep = 3000	88.90 ± 2.80	157.96 ± 14.17	88.90 ± 1.97	75.67 ± 1.07

shrimp diseases and one class of healthy shrimp.

**Conflict of Interest** The authors declare no conflict of interest.

**Acknowledgment** This study was funded in part by the Can Tho University - Improvement Project VN14-P6, supported by a Japanese ODA. We thank our colleagues from the College of Aquaculture and Fisheries for advising on the fishery knowledge. The authors thank the farmers and anonymous contributors who participated and sent images used in the experiments. The discussions expressed are those of the authors and not necessarily those of the funding agency.

## References

- [1] N. T. P. Lan, "Social and ecological challenges of market-oriented shrimp farming in Vietnam," SpringerPlus, **2**(1), 1–10, 2013, doi:10.1186/2193-1801-2-675.
- [2] T. B. T. Nguyen, "Good aquaculture practices (VietGAP) and sustainable aquaculture development in Viet Nam," in Resource Enhancement and Sustainable Aquaculture Practices in Southeast Asia: Challenges in Responsible Production of Aquatic Species: Proceedings of the International Workshop on Resource Enhancement and Sustainable Aquaculture Practices in Southeast Asia 2014 (RESA), 85–92, Aquaculture Department, Southeast Asian Fisheries Development Center, 2015.
- [3] A. Shinn, J. Pratoomyot, D. Griffiths, T. Trong, N. Vu, P. Jiravanichpaisal, M. Briggs, "Asian shrimp production and the economic costs of disease," Asian Fish. Sci. S, **31**, 29–58, 2018, doi:10.33997/j.afs.2018.31.S1.003.
- [4] V. Boonyawiwat, N. T. V. Nga, M. G. Bondadreantoso, "Risk Factors Associated with Acute Hepatopancreatic Necrosis Disease (AHPND) Outbreak in the Mekong Delta, Viet Nam," Asian Fisheries Society, **31**, 226–241, 2018, doi:10.33997/j.afs.2018.31.S1.016.
- [5] S. Thitamadee, A. Prachumwat, J. Srisala, P. Jaroenlak, P. V. Salachan, K. Sritunyalucksana, T. W. Flegel, O. Itsathitphaisarn, "Review of current disease threats for cultivated penaeid shrimp in Asia," Aquaculture, **452**, 69–87, 2016, doi:10.1016/j.aquaculture.2015.10.028.
- [6] R. W. Doyle, "Inbreeding and disease in tropical shrimp aquaculture: a reappraisal and caution," Aquaculture research, **47**(1), 21–35, 2016, doi:10.1111/are.12472.
- [7] G. Braun, M. Braun, J. Kruse, W. Amelung, F. G. Renaud, C. M. Khoi, M. Duong, Z. Sebesvari, "Pesticides and antibiotics in permanent rice, alternating rice-shrimp and permanent shrimp systems of the coastal Mekong Delta, Vietnam," Environment international, **127**, 442–451, 2019, doi:10.1016/j.envint.2019.03.038.
- [8] B. K. Dey, G. H. Dugassa, S. M. Hinzano, P. Bossier, "Causative agent, diagnosis and management of white spot disease in shrimp: A review," Reviews in Aquaculture, 2019, doi:10.1111/raq.12352.
- [9] T. J. Sullivan, A. K. Dhar, R. Cruz-Flores, A. G. Bodnar, "Rapid, CRISPR-Based, Field-Deployable Detection Of White Spot Syndrome Virus In Shrimp," Scientific Reports, **9**(1), 1–7, 2019, doi:10.1038/s41598-019-56170-y.
- [10] J. Xiong, W. Dai, C. Li, "Advances, challenges, and directions in shrimp disease control: the guidelines from an ecological perspective," Applied microbiology and biotechnology, **100**(16), 6947–6954, 2016, doi:10.1007/s00253-016-7679-1.
- [11] T. Q. Bao, T. C. Cuong, N. D. Tu, L. T. Hieu, et al., "Designing the Yellow Head Virus Syndrome Recognition Application for Shrimp on an Embedded System," Exchanges: The Interdisciplinary Research Journal, **6**(2), 48–63, 2019, doi:10.31273/eirj.v6i2.309.
- [12] Z. Liu, F. Cheng, W. Zhang, "Identification of soft shell shrimp based on deep learning," in 2016 ASABE Annual International Meeting, 1, American Society of Agricultural and Biological Engineers, 2016, doi:10.13031/aim.20162455470.
- [13] M. Ghasemi-Varnamkhashti, R. Goli, M. Forina, S. S. Mohtasebi, S. Shafiee, M. Naderi-Boldaji, "Application of image analysis combined with computational expert approaches for shrimp freshness evaluation," International Journal of Food Properties, **19**(10), 2202–2222, 2016, doi:10.1080/10942912.2015.1118386.
- [14] H. Lu, Y. Li, T. Uemura, Z. Ge, X. Xu, L. He, S. Serikawa, H. Kim, "FD-CNet: filtering deep convolutional network for marine organism classification," Multimedia tools and applications, **77**(17), 21847–21860, 2018, doi:10.1007/s11042-017-4585-1.
- [15] Z. Liu, X. Jia, X. Xu, "Study of shrimp recognition methods using smart networks," Computers and Electronics in Agriculture, **165**, 104926, 2019, doi:10.1016/j.compag.2019.104926.
- [16] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, **86**(11), 2278–2324, 1998, doi:10.1109/5.726791.
- [17] Z. Liu, "Soft-shell Shrimp Recognition Based on an Improved AlexNet for Quality Evaluations," Journal of Food Engineering, **266**, 109698, 2020, doi:10.1016/j.jfoodeng.2019.109698.
- [18] L. Torrey, J. Shavlik, "Transfer learning," in Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques, 242–264, IGI Global, 2010, doi:10.4018/978-1-60566-766-9.ch011.
- [19] K. Weiss, T. M. Khoshgoftaar, D. Wang, "A survey of transfer learning," Journal of Big Data, **3**(1), 9, 2016, doi:10.1186/s40537-016-0043-6.
- [20] M. Christopher, A. Belghith, C. Bowd, J. A. Proudfoot, M. H. Goldbaum, R. N. Weinreb, C. A. Girkin, J. M. Liebmann, L. M. Zangwill, "Performance of Deep Learning Architectures and Transfer Learning for Detecting Glaucomatous Optic Neuropathy in Fundus Photographs," Scientific reports, **8**(1), 16685, 2018, doi:10.1038/s41598-018-35044-9.
- [21] N. Duong-Trung, L.-D. Quach, C.-N. Nguyen, "Learning Deep Transferability for Several Agricultural Classification Problems," International Journal of Advanced Computer Science and Applications, **10**(1), 2019, doi:10.14569/IJACSA.2019.0100107.
- [22] N. Duong-Trung, L.-D. Quach, M.-H. Nguyen, C.-N. Nguyen, "Classification of Grain Discoloration via Transfer Learning and Convolutional Neural Networks," in Proceedings of the 3rd International Conference on Machine Learning and Soft Computing, 2732, Association for Computing Machinery, New York, NY, USA, 2019, doi:10.1145/3310986.3310997.



- [23] N. Duong-Trung, L.-D. Quach, M.-H. Nguyen, C.-N. Nguyen, "A Combination of Transfer Learning and Deep Learning for Medicinal Plant Classification," in Proceedings of the 2019 4th International Conference on Intelligent Information Technology, 8390, Association for Computing Machinery, New York, NY, USA, 2019, doi:10.1145/3321454.3321464.
- [24] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," 2009, doi:10.1109/CVPR.2009.5206848.
- [25] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., "Imagenet large scale visual recognition challenge," *International journal of computer vision*, **115**(3), 211–252, 2015, doi:10.1007/s11263-015-0816-y.
- [26] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, A. Zisserman, "The pascal visual object classes (voc) challenge," *International journal of computer vision*, **88**(2), 303–338, 2010, doi:10.1007/s11263-009-0275-4.
- [27] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, "How Transferable Are Features in Deep Neural Networks?" in Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, NIPS14, 33203328, MIT Press, Cambridge, MA, USA, 2014, doi:10.5555/2969033.2969197.
- [28] O. Russakovsky, J. Deng, Z. Huang, A. C. Berg, L. Fei-Fei, "Detecting avocados to zucchinis: what have we done, and where are we going?" in Proceedings of the IEEE International Conference on Computer Vision, 2064–2071, 2013, doi:10.1109/ICCV.2013.258.
- [29] M. Oquab, L. Bottou, I. Laptev, J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 1717–1724, 2014, doi:10.1109/CVPR.2014.222.
- [30] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogue, J. Yao, D. Molura, R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE transactions on medical imaging*, **35**(5), 1285–1298, 2016, doi:10.1109/TMI.2016.2528162.
- [31] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition, 1–9, 2015, doi:10.1109/CVPR.2015.7298594.
- [32] S. Ioffe, C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167, 2015.
- [33] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, "Rethinking the inception architecture for computer vision," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2818–2826, 2016, doi:10.1109/CVPR.2016.308.
- [34] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.
- [35] P. Srisapoo, K. Hamano, I. Tsutsui, K. Iiyama, "Immunostimulation and yellow head virus (YHV) disease resistance induced by a lignin-based pulping by-product in black tiger shrimp (*Penaeus monodon* Linn.)," *Fish & shellfish immunology*, **72**, 494–501, 2018, doi:10.1016/j.fsi.2017.11.037.
- [36] M. E. Frischer, R. F. Lee, A. R. Price, T. L. Walters, M. A. Bassette, R. Verdiyev, M. C. Torris, K. Bulski, P. J. Geer, S. A. Powell, et al., "Causes, diagnostics, and distribution of an ongoing penaeid shrimp black gill epidemic in the US South Atlantic Bight," *Journal of Shellfish Research*, **36**(2), 487–500, 2017, doi:10.2983/035.036.0220.
- [37] S. Senapati, G. P. Kumar, C. B. Singh, K. M. Xavier, M. Chouksey, B. Nayak, A. K. Balange, "Melanosis and quality attributes of chill stored farm raised whiteleg shrimp (*Litopenaeus vannamei*)," *Journal of Applied and Natural Science*, **9**(1), 626–631, 2017, doi:10.31018/jans.v9i1.1242.
- [38] B. Verbruggen, L. K. Bickley, R. Van Aerle, K. S. Bateman, G. D. Stentiford, E. M. Santos, C. R. Tyler, "Molecular mechanisms of white spot syndrome virus infection and perspectives on treatments," *Viruses*, **8**(1), 23, 2016, doi:10.3390/v8010023.
- [39] K. P. Prasad, K. Shyam, H. Banu, K. Jeena, R. Krishnan, "Infectious Myonecrosis Virus (IMNV)—An alarming viral pathogen to Penaeid shrimps," *Aquaculture*, **477**, 99–105, 2017, doi:10.1016/j.aquaculture.2016.12.021.
- [40] J. M. Leyva, M. Martinez-Porchas, J. Hernandez-Lopez, F. Vargas-Albores, T. Gollas-Galvan, "Identifying the causal agent of necrotizing hepatopancreatitis in shrimp: Multilocus sequence analysis approach," *Aquaculture Research*, **49**(5), 1795–1802, 2018, doi:10.1111/are.13633.

## Determination of ERP Readiness Assessment using Agile Parameters: A Case Study

Santo Fernandi Wijaya<sup>\*1</sup>, Harjanto Prabowo<sup>2</sup>, Ford Lumban Gaol<sup>3</sup>, Meyliana<sup>1</sup>

<sup>1</sup>Information Systems Department, School of Information System, Bina Nusantara University, Jakarta, 11480, Indonesia

<sup>2</sup>Management Department, BINUS Business School Undergraduate Program, Bina Nusantara University, Jakarta, 11480, Indonesia

<sup>3</sup>Computer Science Department, BINUS Graduate Program, Bina Nusantara University, Jakarta, 11480, Indonesia

### ARTICLE INFO

Article history:

Received: 02 July, 2020

Accepted: 05 August, 2020

Online: 28 August, 2020

Keywords:

Agile ERP

Implement

Readiness assessment

Textile industry

### ABSTRACT

At present, in the era of digitization the organizations need Enterprise Resource Planning (ERP) systems to have the ability adapt to changes with rapid response in order to increase the competitive advantage. The fact, many companies have failed to implement ERP which is proven to be not go live on time, so that the implementation value is to be very expensive. For this reason, it is important to consider other methods that can reduce failures in implementing ERP that are reviewed from the perspective of readiness assessment. Based on previous research, there are some challenging on the ERP readiness assessment. We also found that some areas on the ERP readiness that still not explore more serious such as ERP and agile method. Therefore, based on the challenging and open are on the ERP readiness, we will explore more development the framework of ERP readiness assessment using agile parameters to help the industry. The purpose of this research is to identify gaps and propose improvements which are the weaknesses in order to improve performance of the organization. The result of this research is developing the framework using an agile parameter for determination of the ERP readiness assessment with a case study in the textile industry in Indonesia. Hopefully, this research has contribution as a measurement tool for assessing the organizational readiness in order to increase agility in the industry. This research methodology uses the combination of qualitative and quantitative research methods using the NVivo software as a result of FGD data processing and using the Technique for Order Preference by Similarity (TOPSIS) for validation and verification tests.

## 1. Introduction

Now, more than ever before, technology must provide satisfying services and ensure that everyone can get it the information they need quickly and precisely for the decision making. Thus, technology is changing the way people work towards digitization. Information systems could manage, disseminate, and produce the information that accurate, real-time, and informative in order to the provide attractive services to managerial levels for the decision making. ERP is an integrated information system that allows people to monitor business activities. ERP implementation for the industry is an absolute necessity in improving effective, efficient, and improving organizational performance. With the use of ERP, it allows

management of the organization to apply the work method with the paperless principle in which business activities can be monitored online and real-time, so that it can be obtained easily obtain various information quickly, accurately and informally without knowing the time and place. Thus, the application of ERP is very helpful managerial level in making decisions appropriately and quickly. This is a reason for organization to increase technology support in managing business activities. Therefore, the use of ERP systems is fundamental of the organizations in order to improve the performance and have a competitive advantage. However, the reality in implement ERP for the industry tends to experience a failure risk level of up to 40% [1]. Based on the previous research that the critical factors which cause failure in ERP implementation is the readiness assessment

<sup>\*</sup>Corresponding Author Santo Fernandi Wijaya, Email: santofw@binus.ac.id

[2]. Thus, the adopt agile method in ERP readiness assessment is to be considered. The principle of agile methods that emphasize the collaboration and communication work and responding to changes [3]. For this reason, it is the important to assess the organizational readiness using agile parameters. The agile parameters for ERP can follow changes and increase business value. Even so, suggesting whether the agile method is suitable for ERP implementation in the industry, the authors will carry out further research. This research aims to solve the complexity in ERP implementation by focusing on evaluating ERP readiness using agile parameters to answer the research questions as follows:

- a. Whether the agile method is suitable for ERP implementation in the industry?
- b. What kind of the framework is proposed for ERP readiness assessment in the industry?

**2. Literature Review**

*2.1. ERP readiness assessment*

The dimensions and factors of agile method are organization, processes, people, and project [4]. The dimensions and factors of agile are shown in the Table 1.

Table 1: Dimensions and factors of agile [4]

Dimensions	Agile factors	
Organization	1	Collaborative work environment
	2	Top management support - involvement
	3	Adaptive view toward change
	4	Cooperative horizontal business culture
	5	People oriented culture
Process	6	Adaptive/iterative requirements management
	7	Early delivery of important features
	8	Regular and frequent communication
	9	Test driven environment
People	10	Co-location of staff & stakeholders
	11	Adaptive leadership style
	12	Self-organizing teams
Project	13	Close team customer relationship
	14	Rapid/early delivery of value
	15	Emergent requirements
	16	Fluid project schedule
	17	Customer involvement
	18	Continuous & incremental business value

The parameters of an agile method are focusing on adaptive and iterative, business needs, on time, collaborating work environment, accountability, responsiveness [5, 6]. The authors analyze and compares the framework of previous research from survey literature by comparing the following references, research focus, identification factors, identify the main components as measurement tools, validation methods, and model evaluations. The comparison framework of the classification models for ERP readiness assessment is shown in the Table 2.

Considering the result analyze of previous research which little research *provides about the framework of agile ERP readiness assessment in order to increasing the organization agility*. Thus, the authors propose the development of a

framework for the ERP readiness assessment using agile parameters to find out how effective the framework can reduce ERP complexity. The goal of the readiness assessment is to identify gaps and propose improvements that are weaknesses in order for increasing the organization performance.

Table 2: Comparison framework of the classification models for ERP readiness assessment (source: Survey Literature)

No	Research focus	Components	Measurement tool	Method used	ID Paper
1	Development of assessment tool for ERP readiness	Project Management, Organizational, Change management	Framework hierarchy of Critical Factors for ERP Readiness	AHP & FANP	[2]
2	Developing of organizational readiness assessment model	Organizational	Data gathering & analysing for organizational readiness	FCM, DEMATEL Cluster	[5]
3	Development of assessment tool	Technoware, Humanware, Infoware, Orgaware	Critical Factors for ERP Readiness	TOPSIS	[7]
4	Development of framework for ERP assessing readiness	Organizational	Scale measurement questionnaire	CFA, SEM	[8]
5	Development of ERP readiness assessment model	Top management, Project management, People, Change management, Technical requirement	Comparison Survey Literature & Case study	AHP, MOORA & TOPSIS	[9]
6	Development of framework for ERP readiness	Organizational context, Business processes, Perception of ERP, External	Data synthesis tool	SPSS software Cluster Analysis	[10]
7	Framework for ERP assessing readiness	Organizational	Rating 9S model based by agility factors criteria	TOPSIS, Fuzzy AHP	[11]

*2.2. ERP readiness factors*

Adopting the Literature Review Prisma [12], researchers conducted a search through database searching with the ERP readiness factor. The results of the filtering of the article after reading the abstract and the introduction of the selected articles, then the authors makes a mapping and comparison which is a factor of ERP readiness. From these factors, the authors define

and meta analyzes of the factors based on the writing in the article. Then the authors categories the ERP readiness factors into four main dimensions, namely processes, people, organizational, and technology. The four main dimensions are based on the Leavitt Diamond Model [13]. The main dimensions of the Leavitt Diamond Model consist of structure, technology, people, and tasks, with the following explanation:

- Structure approach. Is an effort that is applied to improve organizational behavior through increased task performance, such as changes in communication systems, authority systems, workflow systems, structural changes. This is categorized in the organizational.
- Technology approach. Respond to technological trends to support significant organizational changes, and problem solving such as work measurement techniques.
- Task approach. Refers to the process of producing goods or services, which are categorized as processes.
- People approach. Changing the behavior of people in organizations, and changing human behavior that will cause changes to complete tasks, achieve performance, and meet the qualifications criteria of people as needed

Based on the literature survey, the authors do a mapping of the ERP readiness factors that affect each of these components. The mapping results show that there are 18 factors that can be considered for these 4 components. The result mapping for the comparison of ERP readiness factors is shown in the Table 3.

Table 3: Comparison of ERP readiness factors (source: Survey literature)

Dimensions	No	Factors	ID Paper
Processes	1	Business process change	[14]
	2	Process business integration	[14, 15, 16]
	3	Development of business	[17, 14, 18, 19]
People	4	Management support	[14, 20, 21]
	5	Skill project team	[4, 9]
	6	Train and education	[13, 20]
Organizational	7	Organization agile	[1, 4, 9, 18, 20, 21]
	8	Communication & collaboration	[18, 22]
	9	Project management	[16, 18, 19, 22, 23]
	10	Industry characteristics	[24]
	11	Change management	[14, 18, 22, 25]
	12	Strategy planning	[15, 18, 23, 24]
	13	Organization culture	[14, 22]
Technology	14	ERP implementation	[14, 15, 20, 22, 23]
	15	Agile method	[14, 17, 20, 25]
	16	Technological	[14]
	17	System Integration	[14, 24]
	18	ERP system	[14, 20, 22, 23]

### 2.3. Agile for ERP

The agile factors are flexibility, speed, responsiveness, competency, agility, and integration [25]. The mapping result for definition of agile factors depicted in the Table 4.

Table 4: Definition of Agile factors [source: survey literature]

Factors	Definition	ID Paper
Flexibility	Ability to process different products and achieve different objectives with the same facilities	[17, 20, 26]
Speed	Ability to carry out tasks and operations in shortest possible time	[17, 20, 26]
Responsive ness	Ability to identify changes and respond quickly	[20, 26]
Competency	Extensive set of abilities that provide a basis for productivity, efficiency and effectiveness of a company's activities	[17, 20, 26]
Account ability	Proves to be the most important catalytic agent in guiding the drivers of organization agility	[17, 20]
Integration	Relations between the individual system components, easy & effortless flow of the materials, information & communication between the system components, organizational structures, people, & technology	[20]

The collaborative & iterative function in the agile method can adapt to changes, simplify processes, move quickly, and work teams. [27]. Besides that, Agile approaches to supply all effectively sequence, quick response, & manage the manufacturing process in order to reduce lead times [27]. Therefore, the agile for ERP can be considered for analysis and carried out further research as a way for reducing the complexity of ERP implementation. Iterative and collaborative of each stage of implementation, thereby making project work more effective and efficient. Based on the explanation of the definition of agile factors, it can be considered to help the stages of the processes for implement ERP for achieving the organization agility.

### 2.4 ERP in textile industry

One of the important roles of ERP for the industry is system integration, which make it possible to reduce repetition of work, thereby increasing work efficiency. This also applies to the textile industry. Supply chain issues in textile activities are a customization of system, typical product, marketplace demand, product variety, product life cycle, customer drivers, profit margins, dominant costs, stockout penalties, information enrichment, and forecasting mechanism [27]. Effective change management scheme including enough staff training is necessary, and clear business vision and understand the scope of installation complexity [28]. ERP for textiles to be able to control the quality of raw materials better, so as to produce finished products according to customer demand, easily obtain various information in real-time, precise, and fast, including in the financial reporting process [6, 20, 28]. The essence of agile methods is collaborative and effective communication, adaptation & iterative response to



change, to overcome the complexity of ERP [29]. The main process in the textile process is routine sequential production processes and each process requires strict control, especially in fabric processing to produce finished fabrics. This is to increase efficiency in the process of textile production. It really requires the role of ERP to monitoring production activities automatically and paperless, so that it can improve the way work more effectively and efficiently.

### 3. Methodology

#### 3.1. Research Methodology

This research methodology uses NVivo to analyze qualitative data processing by using the Focus Group Discussion (FGD) method and TOPSIS method is used in order to analyze and rank the weight of each criterion. TOPSIS is a multi-criteria decision-making method. TOPSIS determines the ideal solution and the negative-ideal solution, selects the alternative with the shortest distance from the ideal solution and the farthest distance from the negative-ideal solution as the best alternative [29]. FGD is one method for conducting interviews with participating industry experts and professionals, so that they can obtain information and feedback based on their experience and knowledge. TOPSIS method can determine the ranking of factors of each criterion. The results of Nvivo and TOPSIS data processing are expected to better understand the results of data analysis for the answer of the research questions.

TOPSIS is a decision-making method for ranking and prioritizing Multi Criteria Decision Making (MCDM). MCDM was first introduced by [29]. TOPSIS aims to rank using the principle that the chosen alternative has the shortest distance from the positive ideal solution and the farthest distance from the negative ideal solution from a geometric point of view by using the distance between two points to determine the relative proximity of an alternative to the optimal solution. The positive ideal solution is defined as the sum of all the best values that can be achieved by each attribute, while the negative-ideal solution consists of all the worst values achieved for each attribute. The stages of weighting the TOPSIS method are as follows [29]:

1. Building a normalized weighted matrix using the following formula:

$$Y_{ij} = \frac{X_{ij}}{\sqrt{\sum_{i=1}^m X^2_{ij}}}$$

2. Building a weighted normalized matrix by multiplying the normalized matrix with the weighting value of entropy weighting, with the normalization formula for the weighting matrix v as follows:

$$V = \begin{bmatrix} V_{11}V_{12} \dots V_{1n} \\ V_{21}V_{22} \dots V_{2n} \\ \dots \\ V_{m1}V_{m2} \dots V_{mn} \end{bmatrix} = \begin{bmatrix} W_{1r11}W_{2r12} \dots W_{nr1n} \\ W_{1r21}W_{2r22} \dots W_{nr1n} \\ \dots \\ W_{1rm1}W_{2rm2} \dots W_{nrmn} \end{bmatrix}$$

3. Determine the matrix for positive and negative ideal solutions. The positive ideal solution is denoted by A +, while the negative ideal solution is denoted by A-. With the following formula:

$$A^* = \left\{ \left[ \max_i V_{ij} \mid j \in J \right], \left[ \min_i V_{ij} \mid j \in J^1 \right] \right\} \quad i = 1, 2, \dots, m$$

$$= \{V_1^*, V_2^*, \dots, V_j^* \dots, V_n^*\}$$

$$A^- = \left\{ \left[ \min_i V_{ij} \mid j \in J \right], \left[ \max_i V_{ij} \mid j \in J^1 \right] \right\} \quad i = 1, 2, \dots, m$$

$$= \{V_1^-, V_2^-, \dots, V_j^- \dots, V_n^-\}$$

4. Calculate the distance between the value of each alternative solution with a positive ideal solution matrix and a negative ideal solution matrix. By using the following formula:

$$S_i^* = \sqrt{\sum_{j=1}^n (V_{ij} - V_i^*)^2}, \quad i = 1, 2, \dots, m$$

$$S_i^- = \sqrt{\sum_{j=1}^n (V_{ij} - V_i^-)^2}, \quad i = 1, 2, \dots, m$$

5. Calculating the preference value to an alternative ideal solution, with the following formula:

$$C_i^* = \frac{S_i^-}{S_i^* + S_i^-}$$

6. Ranking the value of Ci +. The best solution is to have the shortest distance to the ideal solution and the farthest distance to the ideal negative solution.

#### 3.2. Data Collection

The authors conducted interviews with respondents with the following criteria such as the respondents understanding business processes of the textile industry in Indonesia, the respondents have managerial positions, and understand the process business of ERP Systems. The authors also distribute the questionnaires to users in the industry as validation and further testing. It is hoped that the results of FGD and Questionnaire process will get meaningful feedback and comments and provide solutions to the problems of this research. The characteristic of respondents is shown in the Table 5.

Table 5: Respondents characteristic

Respondent profiles	FGD		Questionnaire	
	Frequen cy	%	Frequency	%
<b>Gender</b>				
Male	11	85%	69	70%
Female	2	15%	29	30%
<b>Age</b>				
20-30	5	38%	50	51%
31-40	3	23%	36	37%
41-50	4	31%	11	11%
>51	1	8%	1	1%

<b>Educational Level</b>				
Diploma	0	0%	31	32%
strata-1	7	54%	51	52%
strata-2	6	46%	16	16%
<b>Tenure</b>				
1-2 years	0	0%	12	12%
2-3 years	0	0%	6	6%
4-5 years	0	0%	0	0%
>5 years	13	100%	26	27%
<b>Background of education</b>				
IT / IS	13	100%	55	56%
Finance/Accounting	0	0%	28	29%
Marketing	0	0%	15	15%
<b>Role of respondents</b>				
Staff	0	0%	47	48%
Manager/General Manager	10	77%	41	42%
Director	3	23%	10	10%
<b>Industry type</b>				
Industry-Other	0	0%	14	14%
Industry-Textile	10	77%	72	73%
IT Consultant	3	23%	12	12%

3.3. Data Analysis

The Data processing this research begins with the data analysis from the node that to identify problems in dimensions and provide the description based on the list of FGD questionnaires. A list of questionnaires for FGD is shown in Table 6.

Table 6: List of questionnaires for Focus Group Discussion

No	Questionnaire for Focus Group Discussion
1	To improve the performance of textile companies, strategies are needed. One strategy is the use of an ERP system. What do you think?
2	ERP implementation is complex system. In your experience, what critical factor causes the successful implementation of ERP for the industry
3	ERP projects are usually coordinated by the IT team. Actually, the ERP project is not an IT project. What do you think of the ERP project statement as a project management?
4	Before deciding on ERP implementation, it is necessary to consider readiness factors such as organizational readiness, project management readiness, & change management readiness. Which factors are dominant in achieving implementation ERP success?
5	Organizational agility is a critical factor that determines effectiveness in implementing ERP. The following are organizational readiness factors, namely: Project management, Training & education, Business Process Reengineering, System Integration. In your experience, which are the most dominant factors of organizational readiness to support the success of ERP implementation?
6	The following are the project management sub-factors, namely having a project management plan, having a formal project team, holding project status meetings regularly, setting realistic time targets, defining the scope of the project in detail, conducting effective communication & strict supervision of implementation schedule & costs. In your opinion, which are the dominant factors of project management factors that influence the success of ERP implementation?
7	The following are the training and education sub-factors, namely Training as needed in sufficient detail, Training substantially to increase the level of user understanding, Giving confidence to users about how to work using the new system, Training is handled by knowledgeable & competent trainers according to characteristics industry. In your opinion, which are the dominant factors of project management factors that influence the success of ERP implementation?

8	The following are the Business Process Reengineering (BPR) sub-factors in ERP implementation, namely: Redesigning business processes before configuring ERP, Standardizing current business processes. Having sufficient experience in running business processes from integrated systems, Organizations running computerized systems with good infrastructure, organizations that have a culture of sharing information, organizations that have adequate financial resources. In your opinion, which is the dominant factor of the BPR factor that affects the success of ERP implementation?
9	The following are sub-factors of System Integration factors in ERP implementation, namely: Organizations that integrate the way ERP works in supporting operations become more effective, Organizations that are able to integrate ERP with other information systems in the organization. In your opinion, which are the dominant factors of the Integration System factor that can influence the success of ERP implementation?
10	One of the success factors of ERP implementation is the implementation methodology. In an effort to reduce the failure of ERP implementation, it is necessary to consider the agile methodology of ERP implementation. The agile methodology has effective feedback characteristics in each iteration & focuses on system integration. What do you think about developing an agile model as an alternative to improving the quality of ERP implementation for an industry?
11	Management support undertakes a change management process in an effort to overcome the complexity of project management. What is your opinion, what is the role of management support in the process of change management in overcoming the complexity of ERP project management?
12	Reducing the risk of failure of ERP implementation for the industry, it is necessary to develop the Agile ERP method by emphasizing the incremental & iterative process in implementing ERP for the industry. What do you think of the proposal to develop an ERP agile method in ERP implementation for an industry?

In conducting interviews, the authors record the interview activities as evidence and for further data processing. then the recording will be transcribed in Microsoft word. Then the authors do data processing using the NVivo application. Based on the factors are identified from the results of data processing, the authors make comparisons with factors from the results of literature studies. After that, the authors build a hierarchy of research models.

Similarly, for a list of questionnaire statements distributed to users for further data management is shown in Table 7.

Table 7: List questionnaire statement

No	Factors	Questionnaire Statement
1	Business process change	I have a positive belief that evaluating the readiness of business process change is the critical factor in supporting implementation ERP success for the industry
2	Process business integration	I have to be positive before implementing ERP, it is necessary to standardize business processes taken from the ERP system as an effective first step
3	Development of business	I have a positive belief that one way to manage business processes is to be more effective for the industry, so it is necessary to develop technological innovations
4	Management support	I have positive beliefs that the support and active involvement of management levels, ERP implementation can be successful and timely.
5	Skill project team	I have positive beliefs even though the ERP project is coordinated by the IT team, but actually the ERP project is not an IT project but a project management that involves the management of an organization that is committed and supports the change process

6	Training & education	I have a positive belief that the intensive training process is important to ensure users better understand how ERP systems work effectively
7	Organization agile	I have a positive belief, even though ERP implementation has done a lot for the Industry and the failure rate is still relatively high. However, an evaluation of the implementation readiness can reduce ERP implementation failure.
8	Communication & collaborative	I have a positive belief that effective and fast communication of information and collaboration is important in achieving successful ERP system implementation
9	Project management	I have a positive belief that the skills of the project team will provide feedback to provide solutions to difficulties in ERP implementation.
10	Industry characteristics	I have a positive belief that one of the advantages of the agile ERP model is to iterate at each stage to ensure the stages of the ERP implementation process run well.
11	Change management	I have a positive belief that the readiness to carry out the stages of change management by following the workings of the selected ERP system is one of the critical factors in determining the success of ERP implementation.
12	Strategic planning	I have a positive belief that the use of an ERP system is an industrial adaptation requirement to improve work efficiency, information integration, productivity, and the organizational performance.
13	Organization culture	I have a positive belief that corporate culture is a cultural value that is rooted and difficult to change. However, organizational readiness needs to make changes that follow the business processes of the selected ERP system
14	ERP implementation	I have a positive belief that getting effective opinions and feedback from users in every iteration is important to ensure ERP implementation runs effectively
15	Agile method	I have a positive belief that increasing the success of ERP implementation, so before implementing ERP for an industry, it is necessary to assess the organization's readiness for the change process
16	Technological	I have a positive belief that one of the impacts of using an ERP system is to get various information as needed and to support decision making at the managerial level of the organization
17	System integration	I have a positive belief that ERP is an integrated system and needs technological innovation for optimal use, so as to reduce duplication of work and increase organizational productivity.
18	ERP Systems	I like the use of ERP systems in completing work, because it will give effect to the work to be more efficient and effective

4. Discussion of the Results

4.1 The agile method for ERP implementation.

The result of FGD processing by NVivo that rank the factors for each dimension that the processes dimension is 20%, people dimension is 28%, organizational dimension is 37%, and the technology dimension is 14 %. Thus, the organizational readiness is the most important. The results of FGD by NVivo application is shown in the Table 8.

Table 8: The result indicators of managing data analysis

No	Factors	#	Respondent													
			1	2	3	4	5	6	7	8	9	10	11	12	13	
1	Business process change	66		5	1	6	1		1	1	3	1	7	6	4	
2	Process business integration	22	1	2	1	3	1			2	7	2	2	1		

3	Development of business	6							1	1		3		1		
4	Management support	73		2	1	1	2	5	1	4	1	2	1	6	9	10
5	Skill project team	41			1		2	1	8	1	0		1	0	1	8
6	Training & education	15	1	2	1	2	2		1		2	3			1	
7	Organization agile	78	4			1	1	5	2	1	7	4	2	9	8	3
8	Communication & collaborative	40	1			5	3	2	3		9	2	1	0	4	1
9	Project management	16								1	4	1	4	3	3	
10	Industry characteristics	14			1	1			3	3	2		3	1		
11	Change management	10		1					1			2	1	5		
12	Strategic planning	8		1					1	1				4		1
13	Organization culture	6		2									2			2
14	ERP implementation	15							1	1			3	1	5	1
15	Agile method	19	2						1	2			8	2		2
16	Technological	10						1				3	2		1	2
17	System integration	10	1							1			5			3
18	ERP systems	10							1	1	4		3		1	

Based on the calculated processing follow the stages of weighting the TOPSIS method are as follows:

1. Normalized weighted matrix is 31710
2. Weighted normalized matrix by multiplying the normalized matrix with the weighting value min is 0,61 and max is 1,30
3. Matrix for positive is 0,47 and matrix for negative is 1,51
4. Value of each alternative solution with a positive ideal solution matrix is 2,76 and a negative ideal solution matrix is 5,08
5. Preferences value is 11,67
6. Ranking the value of each factors are weights sub-factors shown in the Table 9.

The summary result of data questionnaire processing by TOPSIS approach that rank the factors for each dimension that the processes dimension is 17%, people dimension is 15%, organizational dimension is 40%, and the technology dimension is 28 %. Thus, the organizational readiness is the most important also. The results weight analysis of TOPSIS approach is shown in the Table 9.

Table 9: The result weight analysis of TOPSIS

No	Factors	Squared	Min	Max	(-) addition	(+) addition	(-) score	(+) score	Preference score	Final weights
1	Business process change	1753	0,03	0,07	0,09	0,02	0,30	0,14	0,68	34%
2	Process business integration	1838	0,03	0,08	0,12	0,02	0,35	0,16	0,70	34%

3	Development of business	1667	0,03	0,07	0,09	0,02	0,29	0,15	0,66	32%
4	Management support	1744	0,05	0,08	0,04	0,02	0,19	0,15	0,57	33%
5	Project team	1433	0,04	0,11	0,18	0,13	0,42	0,35	0,54	32%
6	Training & education	1836	0,04	0,07	0,04	0,02	0,20	0,13	0,60	35%
7	Organization agility	1760	0,03	0,07	0,10	0,02	0,31	0,14	0,68	15%
8	Communication & collaborative	1797	0,04	0,07	0,03	0,02	0,18	0,13	0,58	12%
9	Project management	1865	0,03	0,07	0,10	0,02	0,31	0,12	0,72	15%
10	Change management	1751	0,03	0,07	0,10	0,02	0,32	0,15	0,68	15%
11	Industry characteristics	1674	0,03	0,07	0,09	0,03	0,30	0,16	0,66	14%
12	Strategic planning	1747	0,03	0,06	0,08	0,09	0,29	0,13	0,69	15%
13	Organization culture	1681	0,03	0,08	0,11	0,03	0,33	0,17	0,66	14%
14	ERP implementation	1786	0,04	0,07	0,03	0,09	0,18	0,13	0,58	18%
15	Agile method	1738	0,03	0,08	0,14	0,03	0,37	0,17	0,68	21%
16	Technological	1938	0,05	0,08	0,05	0,02	0,23	0,12	0,65	20%
17	System integration	2008	0,04	0,07	0,05	0,01	0,22	0,11	0,68	21%
18	ERP systems	1694	0,03	0,07	0,09	0,02	0,30	0,15	0,66	20%
		31710	0,61	1,30	1,51	0,47	5,08	2,76	11,67	400%

From the comparison of the results of data processing it can be said that the organization dimension is the important that must be considered in readiness assessment before implementing ERP for the industry. While one dominant factor of the organizational dimension is the organization agile. So that, the agile factors can be considered. Thus, it can be said that the agile method is suitable in ERP implementation for the industry.

4.2 ERP readiness assessment framework in the industry.

The measurement the organization's readiness to make changes is wise, so that the goal of implementing ERP can be realized. This assessment considers the business process integration factors of the selected ERP system that the people and project teams involved, the technology tools used, and the organization's readiness to respond to change. So, four main components such as processes, people, organizations and technology can work using agile parameters, such as flexibility, responsiveness, speed, competency, accountability, and integration. The results of the assessment of the four components with collaborative agile parameters in order to increase agility in the industry. For this

reason, the authors propose the ERP readiness assessment framework in the industry that can be shown in Figure 1.

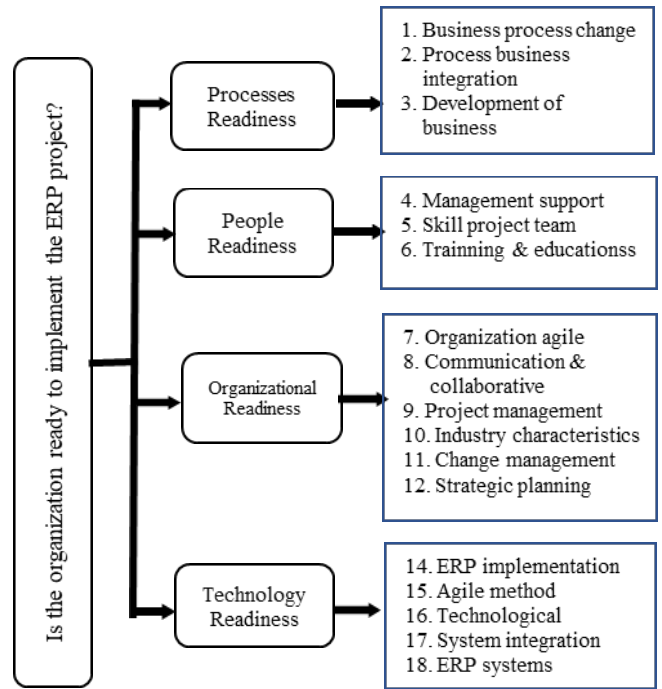


Figure 1: Framework hierarchy of ERP readiness assessment

5. Conclusion

The focus of integrated information systems is to provide solutions by aligning information technology and business processes to meet business needs. The readiness assessment is considered as a fundamental factor in order to increase agility in the industry before the organization's management decides to implement the ERP project. The agile method approach is to be considered as one of the suitable methods for assessing organizational readiness to support organizational activities with rapid adaptation and simplification of processes with focusing on quality and time. The agile factors activities such as flexibility, speed, responsiveness, competency, accountability, integration.

The result of this research indicate that the readiness factors are mapped with agile factors can contribute significantly in order to improve the organization agility. Before the organization decision to implement the ERP project, it is necessary to consider the readiness factors of the main components such as processes, people, organizational, and technology. The readiness assessment model for ERP implementation as one solution to reduce the complexity of ERP implementation can be considered. The authors realize that the research has limitations in analyzing the results of data processing and case study. Therefore, a challenge for future research is how effective to transfer this method practically and necessary to conduct further research as validation agile factors and maturity in ERP readiness assessment with case studies in other industries that can be used as benchmarks for the agile ERP readiness assessment framework.



**References**

- [1] Z. Zhang, et al., "Framework of ERP systems implementation success in China: An empirical study". *International Journal of Production Economics*, 98(1), 56-8., 2005. DOI: 10.1016/j.ijpe.2004.09.004.
- [2] J. Razmi., R. Ghodsi, M.S. Sangari, "A fuzzy ANP model to assess the state of organizational readiness for ERP implementation". In 4th International Conference on Information and Automation for Sustainability. 13. 2009. DOI: 10.1109/ICIAFS.2008.4783989
- [3] R.O. Yusuff, J.A. Sonibare, "Characterization of textile industries' effluents in Kaduna, Nigeria and pollution implications". *Global Nest: Int. J.*, 6(3), 212-221, 2004.
- [4] R. Avazpour, E. Ebrahimi, M.R. Fathi, "Prioritizing agility enablers based on agility attributes using fuzzy prioritization method and similarity-based approach". *International Journal of Economy, Management and Social Sciences*, 3(1), 143-153, 2014.
- [5] S. Ahmadi, E. Papageorgiou, C.H. Yeh, R. Martin, "Managing readiness-relevant activities for the organizational dimension of ERP implementation". *Computers in Industry*, 68, 89-104. 2015.
- [6] H.N. Gandhi, "Analisis Tingkat Kematangan Proses Bisnis Perusahaan Garmen Kelas Menengah". (Doctoral dissertation, Institut Teknologi Sepuluh Nopember). repository.its.ac.id/id/eprint/42365. 2017.
- [7] D. Soysa, J. Nanayakkara, "Readiness for ERP implementation in an organization: Development of an assessment model". *International Conference on Information and Automation* (pp.27-32). IEEE. 2.2006. DOI: 10.1109/ICINFA.2006.374147
- [8] P. Hanafizadeh, A.Z. Ravasan, "A McKinsey 7S model-based framework for ERP readiness assessment". *International Journal of Enterprise Information Systems (IJEIS)*, 7(4), 23-63.2011.
- [9] M. Kirmizi, B. Kocaoglu, "The key for success in enterprise information systems projects: development of a novel ERP readiness assessment method and a case study". *Enterprise Information Systems*, 14(1), 1-37.2020. DOI: 10.1080/17517575.2019.1686656
- [10] L. Raymond, S. Rivard, D. Jutras, "Evaluating readiness for ERP adoption in manufacturing SMEs". *International Journal of Enterprise Information Systems (IJEIS)*, 2(4), 1-17.2006. DOI: 10.4018/jeis.2006100101
- [11] S. Shiri, A. Anvari, H. Soltani, "Identifying and prioritizing of readiness factors for implementing ERP based on agility (extension of McKinsey 7S model)". *European Online Journal of Natural and Social Sciences: Proceedings*, 4(1), 56, 2015.
- [12] A. Liberati, D.G. Altman, J. Tetzlaff, C. Mulrow. "The PRISMA statement for reporting systematic reviews and metaanalyses of studies that evaluate health care interventions: explanation and elaboration. *J. Clin.Epidemiol.*". 62(10). 2009. DOI: 10.7326/0003-4819-151-4-200908180-00136
- [13] H.J. Leavitt, "Applied Organizational Change in Industry, Structural, Technological and Humanistic Approaches. *Handbook of organizations*. 1965.
- [14] L.M. Liao, C.J. Huang, X.Y. Lin, "Applying Project Management Perspective for ERP Implementation: A Case Study". *Proceedings of Engineering and Technology Innovation*, 8, 40-45.2018.
- [15] A.N. Hidayanto, M.A. Hasibuan, P.W. Handayani, Y.G. Sucahyo, "Framework for Measuring ERP Implementation Readiness in Small and Medium Enterprise (SME): A Case Study in Software Developer Company". *JCP*, 8(7), 1777-1782. 2013.
- [16] A. Kraljic, T. Kraljic, "Agile Software Engineering Practices and ERP: Is a sprint too fast for ERP Implementation?". 2017.
- [17] Campanelli & Parreiras. "Agile methods tailoring–A systematic literature review". *Journal of Systems and Software*, 110,85-100. 2015. DOI: /10.1016/j.jss.2015.08.035
- [18] M. Mariouli, J. Laassiri, "Applying Agile Procedure Model to Improve ERP Systems Implementation Strategy". In *International Conference Europe Middle East & North Africa Information Systems and Technologies to Support Learning* (pp.471-481). Springer, Cham. 111.2019. DOI: 10.1007/978-3-030-03577-8\_52
- [19] J. Ram, D. Corkindale, M.L. Wu, "Examining the role of organizational readiness in ERP project delivery". *Journal of Computer Information Systems*, 55(2), 29-39. 2015. DOI: 10.1080/08874417.2015.11645754
- [20] J.J.A. Baig et al., "Evaluation of agile methods for quality assurance and quality control in ERP implementation". In *Intelligent Computing and Information Systems (ICICIS)*, Eighth International Conference, 252-257, 2017. DOI: 10.1109/INTELCIS.2017.8260055
- [21] P.C. Hui et al., "Enterprise resource planning systems for the textiles and clothing industry". *Innovative Quick Response Programs in Logistics Supply Chain Management* (pp. 279-295). Springer Berlin Heidelberg. 28.2010. DOI: 10.1007/978-3-642-04313-0\_14
- [22] A.A. Fetouh et al., "Applying Agile Approach in ERP Implementation". *IJCSNS*, 11(8), 2011.
- [23] M.I.M. Nofal et al., "Conceptual model of enterprise resource planning and business intelligence systems usage". *International Journal of Business Information Systems*, 21(2), 178-194.2016.
- [24] G. Lee, W. Xia, "Toward agile: an integrated analysis of quantitative and qualitative field data on software development agility". *MIS Quarterly*, 34(1), 87-114. 2010. DOI: 10.2307/20721416
- [25] J. Azar et al., "Using the agile implementation model to reduce central line-associated bloodstream infections". *American journal of infection control*. 47(1). 2018 DOI: 10.1016/j.ajic.2018.07.008
- [26] B. Sherehly et al., "A review of enterprise agility: Concepts, frameworks, and attributes". *International Journal of industrial ergonomics*, 37(5), 445-460. 2007. DOI: 10.1016/j.ergon.2007.01.007
- [27] M. Bruce et al., "Lean or agile: a solution for supply chain management in the textiles and clothing industry?". *International journal of operations & production management*, 24(2), 151-170. 2004. DOI: 10.1108/01443570410514867
- [28] R. Rambola et al., "An Effective Synchronization of ERP in Textile Industries". *Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 969-973, 2018. DOI: 10.1109/ICECA.2018.8474686
- [29] C.L. Hwang, K. Yoon, "Methods for multiple attribute decision making". In *Multiple attribute decision making* (pp.58-191). Springer, Berlin, Heidelberg. 186.1981. DOI:10.1007/978-3-642-48318-9\_3.

## A Review on Cross-Layer Design Approach in WSN by Different Techniques

Sultana Parween\*, Syed Zeeshan Hussain

Department of Computer Science, Jamia Millia Islamia, New Delhi, 110025, India

---

### ARTICLE INFO

Article history:

Received: 12 June, 2020

Accepted: 07 August, 2020

Online: 28 August, 2020

---

Keywords:

Cross-Layer Design (CLD)

Cross-layer Framework (CLF)

Data management Framework

Energy efficiency

Quality of Service (QoS)

TinyCubus

Wireless Sensor Network (WSN)

---

---

### ABSTRACT

Wireless Sensor Networks (WSN) include a large number of sensor nodes that are connected to each other with the limitations in energy sources, battery life, memory, mobility and computational capacity. Since the traditional layered architecture was appropriate only for the wired network. It works within a strict boundary that leads to more energy usage as well as more communication traffic. It also impacts on the overall network lifespan and performance of the system. Energy efficiency and network lifespan are the primary concern of WSN due to the fact that each node in the network operates with extremely limited energy. Recent studies have shown that the Open System Interconnection (OSI) model cannot meet the demands of the wireless sensor network.

To overcome such limitations, the cross-layer design has been introduced. It allows direct interactions between protocol at non-adjacent layers. In this paper, we present different types of cross-layer design techniques in Wireless Sensor Network (WSN) and discusses several cross-layer proposals given by researchers. At the end, the paper highlights some challenges faced in implementing CLD in Wireless Sensor Networks.

---

### 1. Introduction

Nowadays WSN has evolved into one of the “on-demand” wireless systems that can be generally used in many fields like environmental monitoring, habitat monitoring, military surveillance, smart spaces and medical systems, and so on [1]. The key purpose of the WSN is the detection, collection, computation and transfer of the data to the sink node.

Normally, sensor nodes are powered by batteries, that are difficult to replace or recharge. Power limitation is to be considered one of the most significant factors while designing a sensor network. However, as interference, fading and noise are common in wireless network communication. To overcome the error caused by the above reason during the data transmission, sensor nodes need to use more energy [2]. There are many techniques and protocols used by several researchers to reduce energy consumption. There are still many open challenges to developing an appropriate energy-saving technique for sensor nodes in WSN. Increasing the lifespan of nodes in the wireless sensor network is important as the usage of this network is unreliable for much of the time and it is difficult to repair the batteries or faulty nodes. Thus, several researchers have carried out various solutions for an energy-efficient model in different situations. The traditional layered

architecture might not be suitable and efficient in the wireless environment due to the basic features of wireless communication [3]. In this point of view, cross-layer design for cross-layer optimization being an optimal approach, that is not only to improve but also optimize the overall network performance. Different cross-layer interaction and optimization techniques used by the WSN designer and developers to provide QoS in WSNs. The major contribution of this paper is consequently that of supplying researchers with a summarized comparison of existing CLD approaches by the different techniques in WSN. As Jurdak defines it “Cross-layer design techniques with respect to the reference layered model are the design of protocols, architectures, or algorithms, that not only exploit but also offer a set of inter-layer interaction which is a superset of the standard interfaces provided by the overall reference layered architecture” [4].

The focus of this paper is to explore the cross-layer optimization techniques in the context of wireless sensor networks. The remaining portion of the paper is organized in the following sections: Section 1 describes the brief overview of CLD and its objective and significance. It also explains the need for cross-layer design in WSN. Section 2 presents the research gap. Section 3 briefs the cross-layer design framework in WSN. Section 4 describes the cross-layer design implementation method. Section 5 covers different cross-layer design techniques. Section 6 briefs

---

\*Corresponding Author: Sultana Parween, New Delhi, India | Email: [sultana.tech@gmail.com](mailto:sultana.tech@gmail.com)

the categorization of recent cross-layer design by a different technique. The challenges of cross-layer design in wireless sensor networks are discussed in section 7. Finally, Section 8 gives the conclusion of this paper.

1.1. A brief explanation of Cross-Layer Design

The OSI layered architecture comprised of basically seven layers and these layering structures brings many benefits: it breaks the complex task of networking into separate parts simplifies design; it makes protocols easier to optimize, manage, and maintain; the lower layers can be changed without affecting the upper layers [5]. Although the traditional layering structure is beneficial for wired networks because it provides modularity, transparency, standardization, and abstract functionality. It means that this structure is not suitable for a wireless network. The main purpose behind this is that, there are strict boundaries in the OSI model, all the protocol is proposed for a single layer and each layer cannot communicate directly with the non-adjacent layer. And this characteristic is opposite to energy limitation, node mobility, and the dynamic nature of wireless sensor networks [6]. An approach that could exchange information across many layers is needed to solve this problem. Therefore, the cross-layer technique has affected increased attention. The basic idea of a cross-layer is to use the close interaction between different layers for improving the performance of the entire network. For many domains, this technique has been widely used.

1.2. The objective of Cross-Layer Design in WSN

The major goals that needs to be considered while designing a cross layer architecture are mentioned below.

Table 1: Objective of CLD

Goals	Explanation
1.Network lifetime maximization	It is significant in the context of enhancing the flawless operation of battery-constrained WSNs.
2.Energy-efficient routing	The CLD is responsible for creating a perfect routing algorithm to calculate the best routing path, which could not only consume minimum power but also maintain the network connectivity.
3.Reliable event detection and transfer	The main request for an event-based sensor network is to detect an event and transport the data accurately.
4.Optimization among multiple, conflicting objectives	To find the optimal scheme, which concerning the conflicting goals, to accomplish the best performance of the sensor network
5.Bringing flexibility into the application-specific design of WSNs	The CLD should be flexible enough to satisfy the dynamic requirement of different application.

1.3. Why do we need a Cross-Layer Design?

The traditional layered method avoids communication among non-adjacent layers in the protocol stack, which limits its use in

wireless sensor networks (WSN) [7]. CLD method helps in breaking the traditional waterfall idea that was considered in traditional network protocols. CLD doesn't destroy the traditional layered structure of the networks [8]. It additionally presents the inter-layer communication between non-adjacent layers. With the help of CLD, the internal status and parameters of every layer can be revealed to the different layers ensuing in the removal of redundancy. CLD also helps in figuring out the behaviour of different layers by way of retrieving and receiving the facts from them [9]. CLD helps in increasing the functionality of the wireless network, quality of service, mobility, and security in sharing information amongst all layers. To get the maximum energy consumption in WSNs, a cross-layer can be useful.

1.4. Significance of Cross-Layer Design

The motivating points that trigger to apply the cross-layer design approach were numbered as follows.

- It improves network efficiency and satisfy the quality of service requirements of various applications [10].
- It increases network performance and optimizes network throughput [11].
- It is used to enhance the performance of sensor nodes.
- It allows communication among various non-adjacent layers and the exchange of information amongst them.
- It helps to reduce the transfer load through exchanging information across different layers.
- It maintains the features and functions related to the layers, but to allow for coordination, communication, and optimization of protocols across various layers.
- It is an appropriate technology for overcoming some of the current TCP/IP stack limitations, particularly for wireless networks.

1.5. Assessment among Conventional layered approach and CLD approach

After a review of existing work, we find Comparisons between CLD Approach and the Conventional layered approach [11].

Table 2: Comparison table of Conventional Layered approach and CLD approach

CLD Approach	Conventional Layered Approach
CLD approach minimizes the communication traffic efficiently.	Communication traffic is more in a conventional layered approach.
Nodes ingest less energy.	Energy consumption by the nodes is more likely to be in this approach.
Appropriate for both wired as well as wireless networks.	More appropriate for a wired network.
We can interact with one layer to any other non-adjacent layer in a cross-layered approach.	We cannot interact with one layer to the other non-adjacent layer in the conventional layered approach.

2. Research Gap

After a review of several works the gaps in existing work are mentioned below:-

- It is not clear how when and where the cross-layer proposals should be implemented in CLD.
- It is hard to review and redesign because changes in subsystem may cause changes in other parts of the system [12].
- There is a lack of standardization in cross-layer design that leads to many problems. The network performance should be decreased due to this reason.
- Parameters that are similar across all layers need to be identified.

### 3. Cross-Layer Design Framework

#### 3.1. TinyCubus

The architecture of TinyCubus provides a new conventional model that can satisfy many different application specifications. It is flexible and adaptive by nature and applied to the top of TinyOS with the help of the necessary programming language [13]. It has many components like a Data management framework (DMF), a Cross-Layer Framework (CLF), and a Configuration Engine (CE). Tiny DMF as stated in [14]. It offers a set of data management as well as the system components that are selected dynamically. It is also responsible for selecting the proper implementation based on system information. The cube of fig 1 named ‘Cubus’ integrates all the optimization parameters like energy, communication latency, etc.

CLF helps in cross-layer information sharing & management. E.g. if some of the application and the link-layer component requires information regarding the actual network neighbourhood, then one of its components in the framework collected that information and provide it to all others. A tiny cross-layer framework is not sufficient to give proper parameterization for most of the cases. This problem is solved by the Tiny Configuration Engine. It enables reliable and efficient code delivery, by taking into consideration sensor topology and its assigned functionality [15]. Figure 1 shows the key components in TinyCubus.

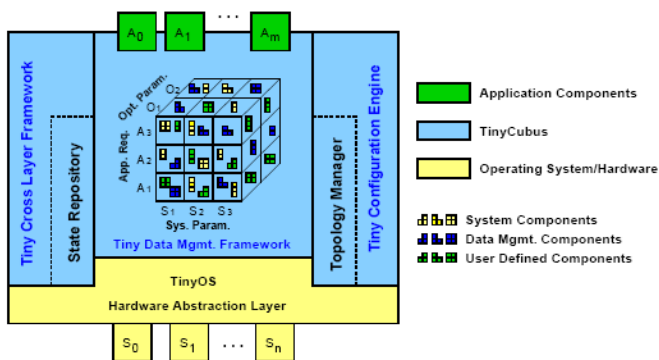


Figure 1: The architecture of TinyCubus [15]

#### 3.2. DMA-CLD Framework

DMA-CLD refers to the Dynamic Multi-Attribute Cross-Layer Design framework [16]. In this framework, various objectives are met. The main purpose of DMA-CLD is to enable cross-layer connections with wireless ad hoc networks & WSN to achieve various as well as probably overlapping goals (networking, nodal, single layer, and cross-layer). DMA-CLD framework recognizes a

set of routes chosen to maximize the network performance according to given parameters like "residual battery," "reliable packet transmission", etc. as input [17]. Figure 2 showing the main idea for DMA-CLD.

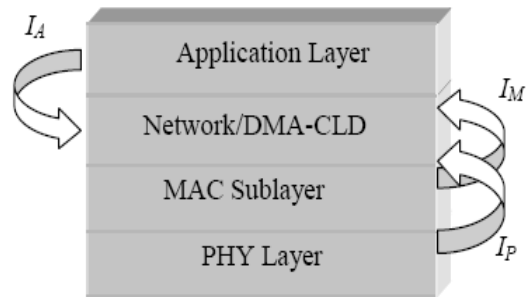


Figure 2: The architecture of DMA-CLD [16]

#### 3.3. Optimization Agent-Based Framework

The main aim of this framework is to emphasize wireless channel & physical layer properties. It is an extended version of DMA\_CLD [17]. It utilizes an optimizing agent to transmit the significant changes in the physical layer as well as a wireless channel to other layers. The basic interactions among different layers are intra-layer interaction or inter-layer interaction. The directions of interaction were maybe bottom-up or top-down. Figure 3 shows the OA based framework.

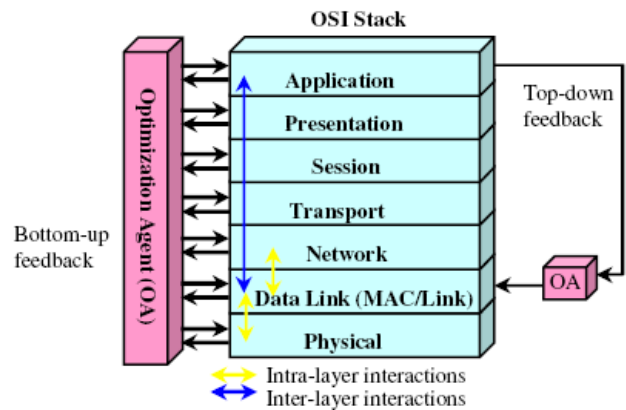


Figure 3: The architecture of the OA based framework [17]

#### 3.4. Horizontal Framework

This framework comprises of the basic entities: cross-layer management entity, application, and protocol stack as shown in figure 4. The main reason for using the Horizontal model rather than a vertical model framework is to optimize the protocol stack height and to reduce resource use [17]. The CLM provides a shared data structure and needs to take care of different features of the sensor network, such as topology control and energy savings whereas APS is accountable for the data transmission of only application-specific. It also includes several extra features for the application and other layers that can be used in the protocol stack. The main concept for using this framework is just to streamline the



protocol stack and isolate such activities like CLM entity modules, thereby enabling application programming easier [14].

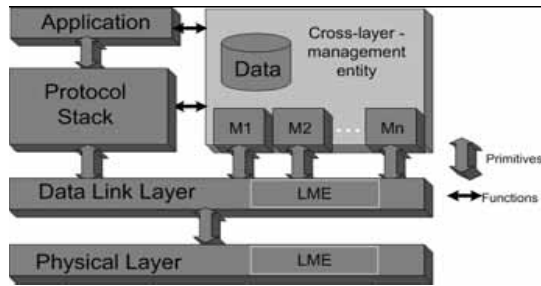


Figure 4: The architecture of Horizontal framework[14]

### 3.5. XLM Framework

XLM is a unified cross-layer framework that was created to achieve reliable and effective communication in WSN with minimal energy costs. It incorporates the protocol layer's most essential feature into a single cross-layer system that is used to provide all the requirements for effective communication [14]. XLM's architecture is built upon an initiative concept that gives full authority to each and every sensor node to determine, when and how to take part in the process of communication [18].

## 4. Cross-Layer Design Implementation Method

After an analysis of existing work, three Implementation strategies for the cross-layer design approach are identified:

### 4.1. Straightforward communication between layers

In straightforward communication among the layers, allow the variables at one layer are accessible to all the other layers during runtime. Straightforward communication among layers is the easiest way of exchange information among layers in runtime. It allows direct interaction among layers [19]. It is shown in Figure 5:

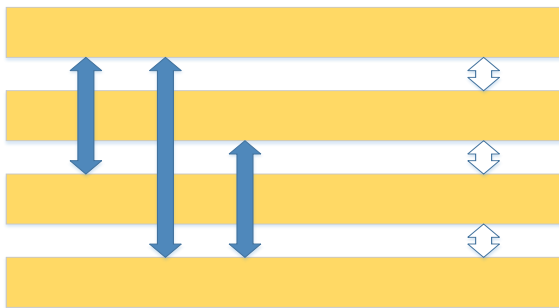


Figure 5: straightforward communication between different layers[19]

### 4.2. A Common database between the layers

As shown in Figure 6, there is a common database among the layers, which can be shared by all the layers. Just like a new layer, it provides a resource-share service to each and every layer [20]. The common database method is ideal for common resource sharing. The common database can connect to different layers at once by an optimization program [21].

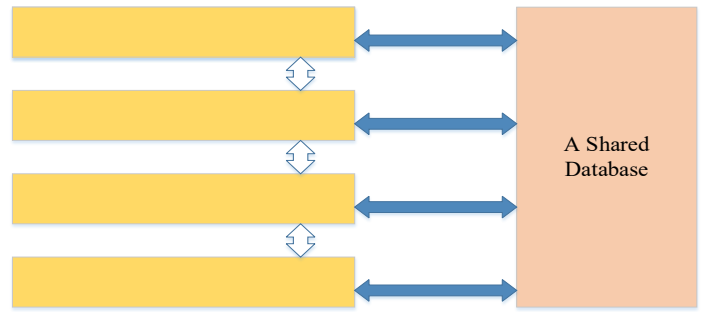


Figure 6: A Common Database across layers [19]

### 4.3. Completely new abstractions

Each layer can communicate from each and every other layer in a completely new abstraction. A completely new abstraction approach allows strong communications between different layers and it provides more flexibility during runtime [21]. It is shown in Figure 7.

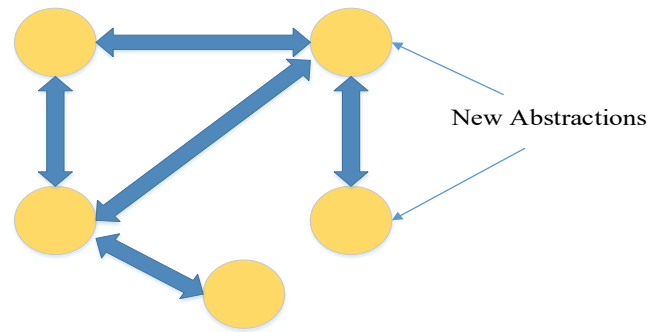


Figure 7: Completely new abstractions [21]

## 5. Cross-Layer Design Optimization Techniques

Each layer contributes to different functionalities in the protocol stack of WSNs. Crossing these layers can accomplish performance gain in different scenarios. The most common of these performance enhancements are discussed in this section. Energy efficiency is a broad category as the nature of WSN is resource-constrained with respect to energy [22]. The key focus of each similar contribution is the energy preservation of such networks. Other areas such as QoS and communication, reliability, error correction, network resource management along with security are also discussed. Cross-Layer approach is one of the most significant emerging technique that communicates across layers for upgrading the complete performance of the network, increases energy efficiency, reduces power consumption, and deliver real-time communications for achieving quality of service (QoS) [23].

### 5.1. Energy Efficiency

An energy-efficient usage is one of the main and critical constraints in the WSN. For prolonging the lifespan of the network, it is essential to save energy at every node. The major challenges of WSN are to design an energy-efficient algorithm that significantly reduces the usage of energy during the transmission

of data. Many researchers are working to improve energy efficiency through a different technique [24].

Cross-layer methods proved that it is better than layered equivalent and it offers much better optimization outcomes. Though, the overall performance of WSNs is greatly enhanced by layer cooperation in cross-layer based systems [25].

The following are several other works that make a significant contribution to energy efficiency by using cross-layer design approaches.

The authors in [7] proposed different modulation techniques to extend network lifetime and use the energy of WSNs in the best possible way. They suggested that when the sensor nodes within a diameter of 35 to 40 meters be run on BPSK or 8PSK modulation technique. When the sensor nodes and their CH (cluster head) and border nodes above 40 meters be run on 16 PSK or 4 QAM modulation technique.

The authors in [26] presented a very popular technique to solve energy consumption issues i.e. LEACH (Low-energy adaptive clustering hierarchy). It is appropriate for large-scale networks and problems of the wireless channel are identified throughout the packet forwarding. They are taking into consideration the residual power and SNR rate. LEACH has significantly improved with multi-hop communications by determining the right possible route from the source cluster head to the destination cluster head. The simulation outcome clearly indicates that it outperforms the traditional LEACH in terms of several parameters like network lifespan, power consumption & PDR.

In a typical LEACH algorithm, it is difficult to share information among each layer, so the communication overhead increased. But the authors in [27] proposed a novel technique to improve energy efficiency and reduce overhead. The algorithm of routing for Leach-CLO (Cross-Layer Optimization) is built upon the typical Leach algorithm along with a cross-layer optimization model. The WSN should also communicate with the bottom three layers to save energy. Routes are determined by the network layer through residual energy information, thereby, reducing energy usage. The experimental findings indicate that this algorithm can balance energy usage in wireless sensing networks effectively & efficiently in comparison to other related algorithms.

In [28], the authors proposed a cross-layer optimization technique that considers both the routing layer and the MAC layer. The routing layer balances the traffic flow through multiple paths that can enhance energy conservation. The MAC layer controls the retry limit of retransmissions for each wireless link, further energy conservation can be achieved, thus improving the network lifetime.

The authors in [29] proposed a cross-layer optimization technique that improves Dynamic Source Routing (DSR) to reduce the frequency of recomputed paths. It enables DSR to initiate route discovery when link failure is occurring. This technique can improve the DSR's routing energy efficiency.

In [30], the authors proposed an energy-aware routing technique based on a cross-layer approach for WSNs. A cross-layer routing algorithm uses the information from different layers to help the routing protocol in deciding the next hop. A fuzzy logic-based approach used to prolong the sensor node's lifetime.

The authors in [31] presented a cross-layer technique HECMAC. The technique was based on the integrated procedure of a duty cycling MAC protocol and a reconfigurable beam-steering antenna. The authors have used information from MAC and network layers to handle the scheduling process of MAC to reduce power consumption. Although the proposed technique has improved the performance but, achieving an accurate antenna adaptive approach is an extremely challenging task.

On the other hand, the author in [1] proposed an energy-efficient technique i.e. EECL (Energy-efficient cross layer) approach. The EECL approach utilizes X-MAC protocol to maintain an asynchronous duty cycle process as they consume less energy. It provides short preliminary instructions which shift to get up/sensing mode just for the routing node whereas the other nodes are in sleeping mode. Therefore, X-MAC protocol overcomes the issues of low power listening: overhearing, unnecessary preamble, and mismatch with packetizing wireless channels.

The authors in [32] suggested a new technique named CLEEP designed for the improvement of energy efficiency through WSN. It utilizes the cross-layer method by taking into consideration only three layers. Throughout the physical layer, it handles the transmission power among the nodes and periodically maintains the neighbour's table. By using the neighbour's table, each node makes a routing table in the network layer. Eventually, the MAC layer uses the information of the routing table to find the duty-cycle of the node. The findings of simulation reveal that the new protocol is more energy-efficient and leads to achieving better performance enhancement.

Based on the above method, the authors in [33] suggested another technique for the improvement of energy efficiency i.e. EECP. It is a type of CLEEP protocol and used essentially to improve and optimized the performance of the network. It utilizes those data that are produced by the layers. It is mainly designed to facilitate effective data communication by reducing energy consumption. In this protocol data routes towards the node that has the highest energy rate and that is also nearest to the sink. But on the other hand, it takes into account the MAC layer information to calculate the node's service cycle and to increase the sleep' state. A comparative analysis reveals that EECP may boost system performance when compared with CLEEP.

The authors in [34] developed a hybrid technique named ATEER. They proposed a WSN adaptive threshold routing protocol for the energy-efficient cross-layer network. In this paper, the algorithm presents the initialization, deployment, selection of CH, the formation of a cluster, and the transmission of gathered information to the base station. The simulation results show that

ATEER has enhanced network lifetime as compared to old techniques

In [35], the authors applied a metaheuristic technique like SA (simulated annealing) for cross-layer energy-aware topology control problem. They suggested optimizing problems to reduce energy consumption through energy control, packet transfer via ECC, and topology regulation in WSNs with power-aware. SA algorithm should be useful in the evaluation of the modified topology. When they compared their results it indicates the SA provides a much better solution than the random sampling.

The authors in [36] presented an energy-efficient transmission technique named EAP-CMAC. This protocol employed physical layer network coding (PNC) that could minimize transmission time slots and achieving better performance and power efficiency than conventional network coding. Direct communication, conventional communication, and PNC-based communication are selected by the EAP-CMAC based on correct communication mode and taking into account the target queue as well as the link quality of the source-destination. However, a combined relay function and energy algorithm were measured in the context with position and residual energy node that not only increases the network lifespan but also saving power. The simulation results demonstrate that EAP-CMAC's performance in terms of network

life is better than IEEE 802.11 and CoopMAC. Besides, the suggested optimal level of power allocation tends to increase the performance of EAP-CMAC by 7 % particularly in comparison with the same power allocation for network lifespan.

In [37], the authors implemented a cross-layer architecture with a different timescale technique that included constraints such as congestion and power control, and processing delay. A primary aim is just to develop a cross-layer architecture that improves overall efficiency as well as reduces the overall communication latency and energy usage.

The authors in [38] presented a solution for 6LoWPAN Wireless Sensor Network to boosts the efficiency of mobility handover via different extents. Such approaches emphasize the convergence of the two separate processes i.e. Layer 2 & Layer 3 handover. It provides an effective combination of these handovers to improve the overall handover performance.

In [39], the authors have proposed a framework trying to find the optimal packet size for transmission with decreased energy consumption. And based on the calculation the authors find the packet size and energy consumption are related to three parameters, that are the condition of channel, medium access and routing.

Table 3: Comparative analysis of different CLD approaches for WSN by energy efficiency

Reference Paper	CLD approach	Layers contributed in CLD approach	Result of CLD approach	Future scope
[1]	Proposed CLD approach for energy-efficient communication i.e. EECL for cloud wireless green communications	Network, MAC and Physical layer	The short preamble implemented in X-MAC only with target node address and minimizes delay and overhearing issues.	Future work would be to implement the suggested EECL approach with various network characteristics.
[7]	The CLD approach for the optimization of layers and framework for WSN has been proposed.	Network, DLL and Physical layer	Uniform clustering can be applied to minimize interference, path loss, and preserve energy.	Future work may include testing of the hardware analysis and evaluating the findings.
[26]	Proposed cross-layer approach for deriving analytical expression of the energy consumption along with the small and large scale channel fading.	Physical layer, Network layer	Improving the existing cross-layer clustering approach concerning network lifespan and power efficiency, without penalties for the PDR.	Evaluate the optimum residual energy values as well as the SNR thresholds allowing for PDR energy minimization and delay constraints.
[27]	A ring cross-layer optimization model is suggested to improve energy efficiency	Network, DLL and Physical layer	LEACH cross-layer optimization is used for energy saving.	-----
[34]	Energy efficient CLD approach which is based on adaptive threshold sensitive distributed routing protocol	Transport layer, MAC layer, Physical layer	ATEER is tested and simulated by previously established routing protocols. It has increased the lifespan of the network in contrast with the old techniques.	-----

[35]	Adopted Cross-layer method for energy-aware optimization of power control and technique of topology control using Simulated Annealing	Physical, Data link, and Network layer	Simulated Annealing leads to better solutions as compared to Random Sampling.	To derive conclusions on energy consumption and topology control comparing Customized Metaheuristic Probabilistic to SA.
[33]	Proposed a novel protocol named EECP for routing in WSN	Network layer, MAC layer, and Physical layer	EECP reduces power consumption and increases the lifespan of the wireless network	Plan to extend the work by enhancing the protocol in terms of the selection of the next-hop and allowing the case of equal costs and eliminating loops.
[36]	Proposed CLD approach for new cooperative communication protocol	Physical and Network layer	EAP-CMAC performs better than the existing protocol and improves energy efficiency.	Performance of EAP-CMAC will be considered in vehicular networks in terms of high mobility.
[38]	Cross-layer mobility support scheme	Network layer and link layer	Evaluated and analyzed the protocol's performance parameters such as latency handover, cost of handover, and packet loss rate.	-----
[39]	CLD approach to get the optimum size of data packets in WSN	MAC and Physical layer	The packet length is optimized to lower energy usage per bit and enhance the network lifespan.	-----
[40]	Proposed CLD for magnetic induction based WUSNs	Transport layer, Network layer, MAC layer	DEAP provides high system throughput and extremely low power consumption within just an obligatory delay.	-----
[41]	CLD approach for energy-efficient and reliable routing to be achieved in WSN	MAC, Network and PHY layer	DRPC can save more energy when compared with DPC and ETX-based routing.	-----
[42]	CLD approach to improve energy efficiency	Network layer, Physical layer, MAC layer	To improve interactions in CLD, the interface among the various layers needs to be updated.	Introduce the new protocol for energy conservation.

### 5.2. Quality of Service

A CLD has now become a better way to increase Quality-of-Service in WSN. Implementation of QoS techniques in WSNs is considerably challenging because WSNs were originally developed for non-real-time applications with low data-rate. WSN designers and developers, therefore, lead to various cross-layer communication and optimization techniques to provide QoS in WSNs. A few cross-layer design techniques are described in the following section.

The authors in [43] presented the cross-layer QoS techniques as delay-aware and reliability-aware techniques. They highlighted the main QoS challenges in WSNs that are related to delay and reliability-aware applications.

The authors discussed a cross-layer technique in [44] are IQOS that optimizes the quality of service factors of the sensor nodes at the bottom three layers of the wireless protocol stack because QOS factors exist in these layers. Another cross-layer technique presented in [45] is XLCP (cross-layer-communication protocol) based on QOMOS architecture. It is capable to offer soft and scalable service differentiation in QOMOS architecture.

A cross-layer approach is an important technique to optimize the QoS requirement of different types of data. In [46], the authors utilized a cross-layer design between the physical layer and MAC layers. Analyze and compare the performance of the QoS control scheme and standard scheme. The results show that the proposed scheme performed better than the standard scheme.

The authors in [47] proposed a novel cross-layer framework for QoS support in WMSNs. The objective is to maximize the capacity of the deployed network to enhance the number of video sources given that the QoS constraint of each source is also preserved. It is accomplished by implementing Wyner-Ziv lossy distributed source coding at the sensor node with a variable group of pictures (GOP) size.

### 5.3. Reliable Communication

The cross-layered framework performs better than the layered solution. Many experimental simulations show that the average transmission reliability of the cross-layered design has an improvement when compared with the layered counterpart. Congestion is always caused by communications waste and minimization in energy efficiency. The authors in [49] presented a cross-layer approach that takes the benefit of communication



among MAC and network layer to detect and control congestion. The MAC layer detects congestion by querying the network layer before transmitting any packet. When congestion is observed, the MAC layer begins its priority scheduling function using priority-based packet dispatching and packet drop policy to manage congestion, thus increasing real-time transmission in WSN. A Novel congestion mitigation technique is used in [50]. This

technique works well in the application, MAC, and network layer. SUIT (Sensor Fuzzy-based Image Transmission) is a new cross-layer transport protocol that provides a fuzzy logic method for estimating and mitigating the congestion efficiently. It uses two different techniques that are adapting video frame rate at the source node and new congestion mitigation technique which can adapt the quality of the image on the fly.

Table 4: Comparative analysis of different CLD approaches for WSN by QoS.

Reference Paper	CLD approach	Layers contributed in CLD approach	Result of CLD approach	Future scope
[43]	CLD approach to QoS in WSNs to meet the optimum latency and reliability	MAC layer and Physical layer	Qos challenges in WSN that are associated to delay and reliability-aware applications.	-----
[44]	CLD approach for IQOS control scheme in WSN.	Network, DLL and Physical layer	IQOS improve energy effectiveness, quality of service and system throughput in WSN.	-----
[45]	CLD approach for QoS in WMSNs	Network layer and Data link layer	XLCP helps to provide and distinguish QoS to various levels of service class.	Evaluate the performance of the current protocol in terms of node mobility and loss.
[46]	CLD approach for QoS in WBAN	Physical layer and MAC layer	Performance Analysis and optimization of the quality of service (QoS) control system in WBANs.	Evaluate and analyse the other error control schemes and access protocols.
[47]	A framework for CLD in support of QoS under WMNs	Application, Network, and DLL	QoS constraint can be achieved by implementing Wyner-Ziv lossy distributed video coding.	-----
[48]	CLD approach for QoS in WMSNs	Network and MAC layer	Clustered multipath routing in WMSNs as well as QoS aware scheduling for various traffic classes.	Emphasis on maximizing threshold values based on Network specification in mathematical representation and Simulation.

Table 5: Comparative analysis of different CLD approaches for WSN by reliable communication

Reference Paper	CLD approach	Layers contributed in CLD approach	Result of CLD approach	Future scope
[49]	A framework of Cross-Layer to improve transmission reliability	Network layer and MAC layer	Congestion detection and congestion control to accomplish real-time transmission in WSN.	Plans for more network parameters to be tested.
[50]	SUIT is a cross layer-based image transport protocol	Application, Transport, MAC and Network layer	The fuzzy-logic method is used for evaluating and minimizing the congestion effectively.	Plan to apply prioritization of packets based on fuzzy-logic technique.
[51]	CLD scheme to enhance video quality in WMSNs	Physical and Transport layer	Compressed sensing video transmission to provide high-quality video to the receiver.	Measure the system performance and complexity with other video encoders and channel encoding.
[52]	Forwarding scheme based on CLD for video sensor network(VSN)	MAC and Network layer	LBRF and DLBS introduce significant improvement as regards to the reliable data delivery in VSN	-----
[53]	CLD approach for reliable video transmission	Transport, Network and Physical layer	RRA method is used to improve data gathering performance in WMSNs.	-----

5.4. Security

Security is always the main concern of WSNs. The idea of cross layering has improved the efficiency of WSN communication. It also reduces modularity by the exchange of parameters between layers that enhance wireless sensor network communication capacity, reliability, and efficiency. Notwithstanding these many successes, there have been security threats in a cross-layering model.

The authors in [54] suggested a new cross-layer technique i.e. T-XLM, which considers a comprehensive solution to securing a sensor-based network. This model uses the trusted method to initiate and coordinate interactions between entities in an intimidating routing process. A cross layering framework of T-XLM builds on the XLM. They also introduced TruFiX, a T-XLM inspired protocol of configurable cross-layer. The security efficiency of TruFiX was compared with RBSS protocols through comprehensive simulation experiments. The results indicate that the TruFiX performs better than the RBSS-based protocols to packet delivery.

XLM security features are not strong in [54], so another authors in [55] suggested the fuzzy logic calculation method which is used to provide defence against attacks. It is mostly based on trust and used to calculate the reputation of nodes to allow safe data

transmission and accurate data delivery through a cross-layer architecture. It shares interlayer information by using different parameters to reduce security attacks on WSN.

In [57], the authors have proposed a cross-layer optimization technique to secure the delivery of images through wireless channels. The desirable parameters–BER, ARQ retry limit, and transmission rate are optimized across the three layers (i.e. PHY, MAC, and APP) with respect to image distortion reduction required and energy-efficient communication.

In [58] the authors have presented an Efficient Dynamic Selective Encryption Framework (EDES) to ensure the security of multimedia traffic in WMSNs. EDES proposed three security levels and the selection depends upon the energy and QoS parameters. A cross-layer approach is used for EDES to take different parameters from the physical, MAC, and upper layers. The capacity metric is used for the evaluation to increase or decrease the level of security. The EDES is a two-step framework wherein in the first step, network performance parameters are determined by combining the QoS parameters and residual energy to capacity function CAP. This requires parameters from the MAC, network, and upper layers. The second step selects the security level based on CAP.

Table 6: Comparative analysis of different CLD approaches for WSN by Security

Reference Paper	CLD approach	Layers contributed in CLD approach	Result of CLD approach	Future scope
[54]	CLD framework introduces TruFiX mechanism	Physical and Network layer	T-XLM specific protocol TruFiX performs better than RBSS(resource bound security solution) based protocols with regard to security performance.	Both SVM and MCDA are being used to evaluate the effectiveness and efficiency of each routing estimation technique.
[56]	CLD Based security Framework	The Network layer, MAC and Physical layer	CLBCSF( cross-layer based comprehensive security framework) is an appropriate model used for the hierarchical clustering.	CLBCSF gives a new direction towards protecting the communication links and extending the lifespan of the whole network.
[57]	Cross-layer resource allocation technique	Application layer, MAC layer, and Physical layer	The images that are moving across the network must be protected.	-----
[58]	EDES- based on a cross-layer approach	Application layer, Network, and MAC layer	Defining the security level for the various scenario can improve energy efficiency and QoS.	Plan to extend the study to dynamic MWSNs by using mobility.

5.5. Error Control

An efficient cross-layer erasure coding technique in [59] determines the redundancy rate of erasure coding. It keeps trying to adapt the channel condition with the traffic load in the application layer. Another technique to detect redundancy in the wireless link layer is the hybrid FEC/ARQ method. Afterward, the objective function is proposed for FEC/ARQ approach. It controls the parameters through packet loss Information. At last, the UEP approach was used to improve the efficiency and consistency of the transmission of video packets. The result of the simulation verified the performance of the cross-layer error control method in contrast to other approaches with respect to different parameters like energy efficiency, PSNR, and delayed-restricted PSNR.

5.6. Network Resource Management

The authors in [63] represented a new cross-layer technique namely cross-layer collaborative communication (CL-CC). The sensor nodes must be adjusted to the changes in their environment to achieve successful network management. At any time, CL-CC uses temporal correlation to correctly estimate the node’s sleeping time. It assists the node is transmitting data of higher quality and also reducing redundancy of data and energy usage. Spatial correlation is also used by sensor nodes to change their sensing activities. The simulation result indicates that the performance of CL-CC is better than the CC-MAC, LSWCC & Anycast to improve network performance.

Table 7: CLD approaches for WSN by error control

Reference Paper	CLD approach	Layers contributed in CLD approach	Result of CLD approach	Future scope
[59]	Novel adaptive cross-layer protocol intended for error control across WMSN	Application layer and Link layer	NAC offers improved performance as a contrast to the other error control methods in real-time multimedia streaming across networks.	-----
[60]	CLD approach for error control scheme in 3-D UWSNs	Network, MAC and Physical layer	HARQ-I and FEC block codes have high error correction feature and a small n/k ratio is the most appropriate for error control schemes.	Initial tests with an underwater research platform, and assessment and comparison of results by using Fountain Codes.
[61]	CLD approach for error control in WMSNs	Application layer, Transport layer, and Physical layer	FEC approach capable to deliver strong bit error correction to solve packet loss in multi-hop transfer.	-----
[62]	Cross-layer method for error control mechanism in WSN	Network, MAC and Physical layer	Hybrid ARQ and FEC methods significantly enhance error resilience Communication in contrast to ARQ.	FEC scheme can be used in multimedia applications.

Table 8: CLD approaches for WSN by Network Resource Management

Reference Paper	CLD approach	Layers contributed in CLD approach	Result of CLD approach	Future scope
[63]	Correlation among nodes to improve the effectiveness of the network by CLD approach	MAC layer	CL-CC determines the node's sleep time to transfer high - quality information while reducing data redundancy and power consumption.	Aim to expand the current work on mobile WSN.
[64]	CLD for network lifetime optimization	Network and MAC layer	Optimize the network life of WNSs by evaluating the impact of both the physical layer parameters and SPP.	-----

Table 9: Abbreviations used in this work

Abbreviations	Description	Abbreviations	Description
CLD	Cross-Layer Design	WMSNs	Wireless Multimedia Sensor Networks
WSN	Wireless Sensor Networks	SUIT	Sensor Fuzzy-based Image Transmission
DMF	Data Management Framework	LBRF	Load Balanced Reliable Forwarding
CLF	Cross-Layer Framework	DLBS	Directional Load Balanced Spreading
CE	Configuration Engine	RRA	Radius and data generation Rate Adjustment
CLM	Cross-Layer Management	T-XLM	Trust-based Cross-Layer Framework
QoS	Quality of Service	CLBCSF	Cross-Layer Based Comprehensive Security Framework
LEACH	Low-Energy Adaptive Clustering Hierarchy	SVM	Support Vector Machine
PDR	Packet Delay Ratio	MCDA	Multi-Criteria Decision System
BPSK	Binary Phase-Shift Keying	RBSS	Resource Bound Security Solution
CH	Cluster Head	NAC	New Adaptive Cross-layer Error Control Protocol
DEAP	Distributed-Environment-Aware Protocol	UEP	Unequal Error Protection
EECP	Energy Efficient Cross-layer Protocol	ARQ	Automatic Repeat Request
EAP-CMAC	Energy-Aware-Physical-layer Network Coding Cooperative MAC	HARQ	Hybrid ARQ
DRPC	Distributed Routing Algorithm with Transmission Power Control	FEC	Forward Error Correction
DPC	Distributed Power Control	PSNR	Peak Signal-to-Noise Ratio
WUSNs	Wireless Underground Sensor Networks	CL-CC	Cross-Layer Collaborative Communication
IQOS	Improved Quality of Service	MAC	Media Access Control
XLCP	Cross-Layer Communication Protocol	CC-MAC	Correlation-based Collaborative MAC
WBANs	Wireless Body Area Networks	LSWCC	Localised Slepian-Wolf Coding-based Correlation approach

## 6. Categorization of Recent Cross-Layer Design by different techniques

Except for the classification based on the number of layers participate in the optimization. Another categorization for cross-layer design in WSN could be technique. I summarize the table 10 according to the recent work of CLD in the WSN.



Table 10: Categorization of CLD by different techniques

Author et. al./Year	Energy efficiency	Reliable communication	Quality of service	Security	Error correction	Network resource
Zhang et al. (2018)	√					
Babber et al. (2017)	√					
Saini et al. (2017)	√		√			
Sarvi et al. (2017)					√	
Umar et al. (2017)				√		
Sadiq et al. (2017)	√					
Singh et al. (2017)	√					
Xenakis et al. (2016)	√					
Ammar et al. (2016)	√					
Yan et al. (2016)	√					
Anbagi et al. (2016)			√			
Ranga et al. (2013)		√				
Das et al. (2015)	√					√
Lin et al. (2015)	√					
Fouzi et al. (2015)	√					
Demir et al. (2014)			√			
Gawdan et al. (2011)				√		

## 7. Cross-Layer Design Challenges

Although there are many advantages of implementing the cross-layer architecture, there are still some new challenges faced by the system. Some of them are discussed in table 11.

Table 11: Challenges of CLD

Challenges of CLD	Main Reason	Possible Solution
Coexistence problem in CLD[3]	Each CLD has its individual standard for the interaction among different layers.	The coexistence of CLD is determined by the standardization of interfaces.
Signalling problem in CLD	There is also no standardized format or way of exchange of cross-layer information in the wireless network.	To make use of ICMP Messages and packet headers to avoid implementing a signalling protocol.
Extra workload induced through cross-layer signalling	It results in an additional burden while sharing information among the nodes.	The layered abstraction method and distributed mechanism may be used for reducing overhead.
The nonexistence of universal CLD	CLD for any particular application might not be appropriate for any other application.	Universal CLD is unlikely to exist for all of the purposes.
Destruction of layered architecture[65]	Cross-layer may disrupt the very well-organized layered design into a disorganized design.	TCP/IP model is used to enhance the features and performance of the layer without destroying its layered design.

## 8. Conclusion

In this paper, we have presented the goals and significance of CLD and its implementation method in WSN. It demonstrates the several frameworks for cross-layer design within the WSN. We have discussed in detail various cross-layer optimization techniques for wireless sensor networks. To bring wireless sensor networks more secure, energy-efficient, and optimized, the different aspects of the cross-layer design approach have been discussed.

In this paper, we have put forward many such techniques that not only improve the overall interactions but also improve the Quality-of-service(QoS). Furthermore, the challenges of cross-layer design have been highlighted and identified their possible solutions. This could be a reason for the future improvement in cross-layer design to overcome existing work limitations.

## References

- [1]. A.S. Sadiq, T.Z. Almohammad, R.A.B.M. Khadri, A.A. Ahmed, J. Lloret, "An Energy-Efficient Cross-Layer approach for cloud wireless green communications," 2017 2nd International Conference on Fog and Mobile Edge Computing, FMEC 2017, 230–234, 2017, doi:10.1109/FMEC.2017.7946436.
- [2]. Mahmoud Q, editor. *Cognitive networks: towards self-aware networks*. John Wiley & Sons; 2007.
- [3]. B. Fu, Y. Xiao, H.J. Deng, H. Zeng, "A survey of cross-layer designs in wireless networks," *IEEE Communications Surveys and Tutorials*, **16**(1), 110–126, 2014, doi:10.1109/SURV.2013.081313.00231.
- [4]. Jurdak R. *Wireless ad hoc and sensor networks: A cross-layer design perspective*. Springer Science & Business Media; 2007.
- [5]. F. Wang, "A Survey of Recent Cross-layer Designs in Sensor Networks," (December), 2016.
- [6]. Z. Hamid, F.B. Hussain, "QoS in wireless multimedia sensor networks: A layered and cross-layered approach," *Wireless Personal Communications*, **75**(1), 729–757, 2014, doi:10.1007/s11277-013-1389-0.
- [7]. K. Babber, R. Randhawa, "A Cross-Layer Optimization Framework for Energy Efficiency in Wireless Sensor Networks," *Wireless Sensor Network*, **09**(06), 189–203, 2017, doi:10.4236/wsn.2017.96011.
- [8]. S. Networks, "Cross-layer design and optimisation for wireless sensor networks Weilian Su\* and Tat L. Lim," **6**(1), 2009.
- [9]. L. Song, "Cross Layer Design in Wireless Sensor Networks: A Systematic Approach," (December), 2008.
- [10]. D. Wu, S. Ci, H. Luo, H.F. Guo, "A theoretical framework for interaction measure and sensitivity analysis in cross-layer design," *ACM Transactions on Modeling and Computer Simulation*, **21**(1), 2010, doi:10.1145/1870085.1870091.
- [11]. Devi, C. Dhivya, and K. Vidya. "A survey on cross-layer design approach for secure wireless sensor networks." In *International Conference on Innovative Computing and Communications*, pp. 43-59. Springer, Singapore, 2019.
- [12]. R. Ranjan, S. Varma, "Challenges and Implementation on Cross Layer Design for Wireless Sensor Networks," *Wireless Personal Communications*, **86**(2), 1037–1060, 2016, doi:10.1007/s11277-015-2972-3.
- [13]. V. Srivastava, M. Motani, "The road ahead for cross-layer design," 2nd International Conference on Broadband Networks, BROADNETS 2005, 2005, 593–602, 2005, doi:10.1109/ICBN.2005.1589660.
- [14]. O. Karaca, R. Sokullu, "Comparative study of cross layer frameworks for wireless sensor networks," *Proceedings of the 2009 1st International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace and Electronic Systems Technology, Wireless VITAE 2009*, (June 2009), 896–900, 2009, doi:10.1109/WIRELESSVITAE.2009.5172568.
- [15]. P.J. Marron, A. Lachenmann, D. Minder, J. Hähner, R. Sauter, K. Rothermel, "TinyCubus: A flexible and adaptive framework for sensor networks," *Proceedings of the Second European Workshop on Wireless Sensor Networks, EWSN 2005*, 2005, 278–289, 2005, doi:10.1109/EWSN.2005.1462020.
- [16]. A.M. Safwat, "A novel framework for cross-layer design in wireless ad hoc and sensor networks," *GLOBECOM - IEEE Global Telecommunications Conference*, 130–135, 2004, doi:10.1109/glocomw.2004.1417561.
- [17]. N. Zhao, L. Sun, "Research on cross-layer frameworks design in wireless sensor networks," *Third International Conference on Wireless and Mobile Communications 2007, ICWMC '07*, 5–7, 2007, doi:10.1109/ICWMC.2007.75.
- [18]. M.S. Sumalatha, V. Nandalal, "An intelligent cross layer security based fuzzy trust calculation mechanism (CLS-FTCM) for securing wireless sensor network (WSN)," *Journal of Ambient Intelligence and Human Computing*, (0123456789), 2020, doi:10.1007/s12652-020-01834-1.
- [19]. Saini, Rakesh Kumar, and Sandip Vijay Ritika. "A Survey on Cross Layer Design implementation in Wireless Sensor Network." *Int J Adv Inf Sci Technol IJAIST* **5** (2016): 101-107.
- [20]. Sarkar, Subir Kumar, Tiptur Gangaraju Basavaraju, and C. Puttamadappa. *Ad hoc mobile wireless networks: principles, protocols and applications*. CRC Press, 2007.
- [21]. D. Lin, S. Li, "TCLA: A triangular cross-layer architecture for wireless sensor networks," 4th International Conference on Frontier of Computer Science and Technology, FCST 2009, 272–278, 2009, doi:10.1109/FCST.2009.50.
- [22]. H.A. Khattak, Z. Ameer, I.U. Din, M.K. Khan, "Cross-layer design and optimization techniques in wireless multimedia sensor networks for smart cities," *Computer Science and Information Systems*, **16**(1), 1–17, 2019, doi:10.2298/CSIS181115004K.
- [23]. P. Sarwesh, N.S. V. Shet, K. Chandrasekaran, "ETRT – Cross layer model for optimizing transmission range of nodes in low power wireless networks – An Internet of Things Perspective," *Physical Communication*, **29**, 307–318, 2018, doi:10.1016/j.phycom.2018.06.005.
- [24]. R. Singh, A.K. Verma, "Efficient image transfer over WSN using cross layer architecture," *Optik*, **130**, 499–504, 2017, doi:10.1016/j.ijleo.2016.10.143.
- [25]. L.D.P. Mendes, J.J.P.C. Rodrigues, "A survey on cross-layer solutions for wireless sensor networks," *Journal of Network and Computer Applications*, **34**(2), 523–534, 2011, doi:10.1016/j.jnca.2010.11.009.
- [26]. A. Ben Ammar, A. Dziri, M. Terre, H. Youssef, "Multi-hop LEACH based cross-layer design for large scale wireless sensor networks," 2016 International Wireless Communications and Mobile Computing Conference, IWCMC 2016, 763–768, 2016, doi:10.1109/IWCMC.2016.7577153.
- [27]. W. Zhang, X. Wei, G. Han, "An Energy-Efficient Ring Cross-Layer Optimization Algorithm for Wireless Sensor Networks," *IEEE Access*, **6**, 16588–16598, 2018, doi:10.1109/ACCESS.2018.2809663.
- [28]. F. Bouabdallah, N. Bouabdallah, R. Boutaba, "Cross-layer design for energy conservation in wireless sensor networks," *IEEE International Conference on Communications*, 0–5, 2009, doi:10.1109/ICC.2009.5198872.
- [29]. N. Chilamkurti, S. Zeadally, A. Vasilakos, V. Sharma, "Cross-layer support for energy efficient routing in wireless sensor networks," *Journal of Sensors*, **2009**, 2009, doi:10.1155/2009/134165.
- [30]. T. Jaradat, D. Benhaddou, M. Balakrishnan, A. Al-Fuqaha, "Energy efficient cross-layer routing protocol in Wireless Sensor Networks based on fuzzy logic," 2013 9th International Wireless Communications and Mobile Computing Conference, IWCMC 2013, 177–182, 2013, doi:10.1109/IWCMC.2013.6583555.
- [31]. M. Luca, M. Vincenzo, L. Patrono, "HEC-MAC: A hybrid energy-aware cross-layer MAC protocol for wireless sensor networks," *International Journal of Distributed Sensor Networks*, **11**(4), 2015, doi:10.1155/2015/536794.
- [32]. S. Liu, Y. Bai, M. Sha, Q. Deng, D. Qian, "CLEEP: A Novel Cross-Layer Energy-Efficient Protocol for Wireless Sensor Networks," 2–5, 2008.
- [33]. F. Semchedine, W. Oukachbi, N. Zaichi, L. Bouallouche-Medjkoune, "EECP: A New Cross-layer Protocol for Routing in Wireless Sensor Networks," *Procedia Computer Science*, **73**(Awict), 336–341, 2015, doi:10.1016/j.procs.2015.12.001.
- [34]. R. Singh, A.K. Verma, "Energy efficient cross layer based adaptive threshold routing protocol for WSN," *AEU - International Journal of Electronics and Communications*, **72**, 166–173, 2017, doi:10.1016/j.aeue.2016.12.001.
- [35]. A. Xenakis, F. Foukalas, G. Stamoulis, "Cross-layer energy-aware topology control through Simulated Annealing for WSNs," *Computers and Electrical Engineering*, **56**(March), 576–590, 2016, doi:10.1016/j.compeleceng.2016.02.015.
- [36]. M. Sami, N.K. Noordin, F. Hashim, S. Subramaniam, A. Akbari-Moghanjoughi, "An Energy-Aware Cross-Layer Cooperative MAC Protocol for Wireless Ad Hoc Networks," *Journal of Network and Computer Applications*, **58**, 227–240, 2015, doi:10.1016/j.jnca.2015.08.009.
- [37]. Q.V. Pham, H.L. To, W.J. Hwang, "A multi-timescale cross-layer approach

- for wireless ad hoc networks,” *Computer Networks*, **91**, 471–482, 2015, doi:10.1016/j.comnet.2015.08.007.
- [38]. X. Wang, Q. Sun, Y. Yang, “A cross-layer mobility support protocol for wireless sensor networks,” *Computers and Electrical Engineering*, **48**, 330–342, 2015, doi:10.1016/j.compeleceng.2015.06.017.
- [39]. S.R. Chudasama, S.D. Trapasiya, “Packet size optimization in wireless sensor network using cross-layer design approach,” *Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2014*, 2506–2511, 2014, doi:10.1109/ICACCI.2014.6968357.
- [40]. S.C. Lin, I.F. Akyildiz, P. Wang, Z. Sun, “Distributed Cross-Layer Protocol Design for Magnetic Induction Communication in Wireless Underground Sensor Networks,” *IEEE Transactions on Wireless Communications*, **14**(7), 4006–4019, 2015, doi:10.1109/TWC.2015.2415812.
- [41]. Y. Chen, L. Gao, Y. Xing, W. Yi, “Cross-Layer Design for Energy-Efficient Reliable Routing in Wireless Sensor Networks,” *Proceedings - 11th International Conference on Mobile Ad-Hoc and Sensor Networks, MSN 2015*, 31–36, 2016, doi:10.1109/MSN.2015.44.
- [42]. B. Jain, G.S. Brar, J. Malhotra, “A Survey on Design and Development of Energy Efficient Cross Layer Approaches in Wireless Sensor Networks,” **2**(4), 273–279, 2014.
- [43]. I. Al-Anbagi, M. Erol-Kantarci, H.T. Mouftah, “A survey on cross-layer quality-of-service approaches in WSNs for delay and reliability-aware applications,” *IEEE Communications Surveys and Tutorials*, **18**(1), 525–552, 2016, doi:10.1109/COMST.2014.2363950.
- [44]. R.K. Saini, . R., “IQOS: A Cross Layer Design Approach for Improve Quality of Service in Wireless Sensor Networks,” *International Journal of Advanced Science and Technology*, **103**, 1–8, 2017, doi:10.14257/ijast.2017.103.01.
- [45]. A.K. Demir, H.E. Demiray, S. Baydere, “QoS-MOS: Cross-layer QoS architecture for wireless multimedia sensor networks,” *Wireless Networks*, **20**(4), 655–670, 2014, doi:10.1007/s11276-013-0628-3.
- [46]. K. Takabayashi, H. Tanaka, C. Sugimoto, K. Sakakibara, R. Kohno, “Cross-Layer Design and Performance Analysis of Quality of Service Control Scheme for Wireless Body Area Networks,” *IEEE Access*, **5**, 22462–22470, 2017, doi:10.1109/ACCESS.2017.2762078.
- [47]. G.A. Shah, W. Liang, O.B. Akan, “Cross-layer framework for QoS support in wireless multimedia sensor networks,” *IEEE Transactions on Multimedia*, **14**(5), 1442–1455, 2012, doi:10.1109/TMM.2012.2196510.
- [48]. I.T. Almkaw, M. Guerrero Zapata, J.N. Al-Karaki, “A cross-layer-based clustered multipath routing with QoS-aware scheduling for wireless multimedia sensor networks,” *International Journal of Distributed Sensor Networks*, **2012**(2), 2012, doi:10.1155/2012/392515.
- [49]. V. Ranga, R. Gupta, “Cross-Layered Framework in WSNs to Enhance Transmission Reliability,” 2013.
- [50]. C. Sonmez, S. Isik, M.Y. Donmez, O.D. Incel, C. Ersoy, “SUIT: A cross layer image transport protocol with fuzzy logic based congestion control for wireless multimedia sensor networks,” *2012 5th International Conference on New Technologies, Mobility and Security - Proceedings of NTMS 2012 Conference and Workshops*, (May), 2012, doi:10.1109/NTMS.2012.6208750.
- [51]. S. Pudlewski, T. Melodia, “DMRC: Distortion-minimizing rate control for wireless multimedia sensor networks,” *2009 IEEE 6th International Conference on Mobile Adhoc and Sensor Systems, MASS '09*, 563–572, 2009, doi:10.1109/MOBHOC.2009.5336954.
- [52]. S. Isik, M.Y. Donmez, C. Ersoy, “Cross layer load balanced forwarding schemes for video sensor networks,” *Ad Hoc Networks*, **9**(3), 265–284, 2011, doi:10.1016/j.adhoc.2010.07.002.
- [53]. L. Shu, M. Hauswirth, Y. Zhang, J. Ma, G. Min, Y. Wang, “Cross layer optimization for data gathering in wireless multimedia sensor networks within expected network lifetime,” *Journal of Universal Computer Science*, **16**(10), 1343–1367, 2010, doi:10.3217/jucs-016-10-1343.
- [54]. I.A. Umar, Z.M. Hanapi, A. Sali, Z.A. Zulkarnain, “TruFiX: A Configurable Trust-Based Cross-Layer Protocol for Wireless Sensor Networks,” *IEEE Access*, **5**, 2550–2562, 2017, doi:10.1109/ACCESS.2017.2672827.
- [55]. N.K. Goud, S.G. Assistant, I.T. Depart, “An Efficient Trust Based Cross Layer Protocol for Wireless Sensor Networks,” **118**(20), 343–357, 2018.
- [56]. I.S. Gawdan, C.O. Chow, T.A. Zia, Q.I. Gawdan, “Cross-layer based security solutions for wireless sensor networks,” *International Journal of Physical Sciences*, **6**(17), 4245–4254, 2011, doi:10.5897/IJPS11.542.
- [57]. W. Wang, D. Peng, H. Wang, H. Sharif, “A cross layer resource allocation scheme for secure image delivery in wireless sensor networks,” *IWCMC 2007: Proceedings of the 2007 International Wireless Communications and Mobile Computing Conference*, 152–157, 2007, doi:10.1145/1280940.1280973.
- [58]. A. Rachedi, L. Kaddar, A. Mehaoua, “EDES - Efficient dynamic selective encryption framework to secure multimedia traffic in Wireless Sensor Networks,” *IEEE International Conference on Communications*, 1026–1030, 2012, doi:10.1109/ICC.2012.6364221.
- [59]. B. Sarvi, H.R. Rabiee, K. Mizanian, “An adaptive cross-layer error control protocol for wireless multimedia sensor networks,” *Ad Hoc Networks*, **56**, 173–185, 2017, doi:10.1016/j.adhoc.2016.12.008.
- [60]. M.C. Domingo, M.C. Vuran, “Cross-layer analysis of error control in underwater wireless sensor networks,” *Computer Communications*, **35**(17), 2162–2172, 2012, doi:10.1016/j.comcom.2012.07.010.
- [61]. Y. Yang, Y. Chen, W. Yi, “Cross-layer forward error control for reliable transfer in wireless multimedia sensor networks,” *2010 7th IEEE Consumer Communications and Networking Conference, CCNC 2010*, 0–1, 2010, doi:10.1109/CCNC.2010.5421716.
- [62]. M.C. Vuran, I.F. Akyildiz, “Error control in wireless sensor networks: A cross layer analysis,” *IEEE/ACM Transactions on Networking*, **17**(4), 1186–1199, 2009, doi:10.1109/TNET.2008.2009971.
- [63]. S.N. Das, S. Misra, “Correlation-aware cross-layer design for network management of wireless sensor networks,” *IET Wireless Sensor Systems*, **5**(6), 263–270, 2015, doi:10.1049/iet-wss.2014.0110.
- [64]. H. Yetgin, K.T.K. Cheung, M. El-Hajjar, L. Hanzo, “A Survey of Network Lifetime Maximization Techniques in Wireless Sensor Networks,” *IEEE Communications Surveys and Tutorials*, **19**(2), 828–854, 2017, doi:10.1109/COMST.2017.2650979.
- [65]. D.K. Sah, T. Amgoth, “Parametric survey on cross-layer designs for wireless sensor networks,” *Computer Science Review*, **27**, 112–134, 2018, doi:10.1016/j.cosrev.2017.12.002.

## Design and Implementation of a PLC Trainer Workstation

Matthew Oluwole Arowolo\*, Adefemi Adeyemi Adekunle, Martins Oluwaseun Opeyemi

*Department of Mechatronics Engineering, Federal University Oye – Ekiti, 371104, Nigeria*

---

### ARTICLE INFO

*Article history:*

*Received: 28 May, 2020*

*Accepted: 05 August, 2020*

*Online: 28 August, 2020*

---

*Keywords:*

PLC

Ladder logic

PLC Workstation

Programming

Software

---

### ABSTRACT

*The Programmable Logic Controller (PLC) is an important component for industrial automatic engineering operation. Hence, the need to comprehend its basis of operation becomes an inevitable task. Some of the problems is industrial PLC is an expensive, pre-built hardware kit, acquisition of programming software and its requisite programming competence is a challenge. Thus, this paper present's the design steps for a desktop PLC trainer workstation for industrial automatic engineering operation. Although researchers have proposed and reported several PLC trainers but they fail to discuss the hardware connection of the input/output components neither is the basic PLC automatic - operation nor PLC component symbols with description discussed. These are the areas discussed in this paper to train participant on PLC programming knowledge. The develop PLC workstation consist of push buttons and switches for input signals and for output signal buzzer, indicator lights and blower. The control aspect comprises the PLC, timer and relay. The PLC module is the MITSUBISHI FX 1S-30MR-001, the simulation software is the MITSUBISHI FXTRN-BEG-EL and the programming software is the MITSUBISHI GX Developer version 8. Authors presented two automatic control application scenarios to train participants and evaluate how the trainer is applicable to real-world situation. We conducted a survey after training to measure the impact of our approach for PLC programming knowledge for participants and result show enhanced knowledge in design step module and LL programming module significantly with our approach.*

---

### 1. Introduction

The PLC replaced the antique relay control logic in the late 1970s in the control of machines and processes [1]-[3]. Some of the benefits of the PLC over the relay control logic are flexibility, higher reliability, communication possibilities, faster response time, and easier troubleshooting [1]. Hence, it has become a vital component in the industry for engineering operation control. The PLC is a microprocessor-based controller; it receiver's analogue and digital signal input from input component such as switches and sensors and apply instructions stored in it programmable memory to control outputs to output components such as motors, pneumatic devices and status indicator. It implement's functions such as logic and sequence [1,4]. The rapid pace of technological development with new model and innovation of PLC technology and its flexibility has encouraged its application beyond industrial control spectrum. Therefore, the development of competence through training in the cabling, programming of PLC and its

application become imperative for student and person with interest in the field of industrial automation. Nevertheless, some of the problems is that industrial PLC is an expensive, prebuilt hardware kit also, to acquire programming software and its requisite programming competence is a challenge [2].

The programming languages defined by IEC 6-1131 for PLC is the Ladder Logic (LL), Structure Text (ST), Function Block (FB) and Instruction List (IL) [3]. The PLC programming device can be a handheld device or the personal computer (PC). However, the PC is commonly used for PLC programming because it is readily available and portable. The LL is the most used programming language because it is simple to comprehend and implement [5]. Although researchers have proposed and reported several PLC trainers but they fail to discuss the hardware connection of the input/output components neither is the basic PLC - automatic operation nor PLC component symbols with description discussed. These are the areas discussed in this paper to train participant on PLC programming knowledge.

Authors in [2] described the development of an embedded PLC for teaching students. Authors combined LabVIEW software

\*Corresponding Author: Matthew. O. Arowolo , Department of Mechatronics Engineering, Federal University Oye – Ekiti, 371104, Nigeria, +2348038066827 & [arowolo.oluwole@fuoye.edu.ng](mailto:arowolo.oluwole@fuoye.edu.ng)



and the AVR Microcontroller with the VB modules to achieve the embedded PLC built bottle filling plant for it application. The programming language used for the embedded PLC is the FB. Although the embedded PLC setup is flexible, relatively easy and affordable to teach the basic principle for PLCs, they did not present or discuss FB program for the bottle filling application. Also, survey report of their application shows moderate performance in stability and reliability. In [6] the development of a PLC Trainer Kit Simulator Automation Lab at the Polytechnic of Sultan Abdul Halim Mu'adzam Shah (POLIMAS) was described. The training kit comprises the Omron PLC CPU unit with 12 inputs and 8 outputs control. Input and output devices are bank of switches and light indicators respectively. They used CX-Programmer for CP1E version 1.0 to program the PLC using ladder diagram and instruction list PLC programming languages. The fabrication of a multiple input/output (I/O) PLC module for educational purpose to enhance the learner's theoretical comprehension and hands-on skill especially for programming, cabling, circuit design and problem solving is presented in [7]. Their module consists of I/O devices such as push buttons (normally open), DC motor (24V), DC relay (24V), DC solenoid piston cylinder (24V) and DC light (24V) capable of interfacing with PLC controller produced by Matsushita, Omron, Siemens. Survey report from the trainees show that 95.70 % attest to the enhancement in their theoretical comprehension and hands-on skill competence in their learning process. However, the programming aspect of the PLC is not discussed. Authors in [8] presented a PLC Based Electrical Machine Trainer Kit developed for Electrical Engineering Practices in the Department of Electrical Engineering Education at Faculty of Engineering, Universitas Negeri Yogyakarta. Their approach is research and development with reference to the ADDIE model from RobertMaribe Branch. The installed PLC is Zelio SR2.201FU and the console dimension is 44.1 cm × 100 cm and 92.7 cm × 100 cm with a front tilt angle of 80 °. Authors examined the performance of trainer kit on 8 practical experiment; rotation control of DC motors; rotation control of three phase induction motor; rotation control of one phase induction motor; starting DC motor; starting 3 phase induction motor using auto-transformer; dynamic DC motor braking; DC motor braking by plugging; and braking 3-phase induction motor by DC injection. Result show the trainer kit has a good performance, indicated by the electrical components and the practical work description can function appropriately as planned. The development of an affordable and portable PLC trainer for industrial control process application is described in [9]. They used the Delta DVP14SS2 PLC, WPLSoft software and switches as inputs and pilot lamps as outputs. Instructors trained student on cabling and programming of PLC with hands-on training on Traffic light automation application. Authors carried out pre and post training evaluation for trainee and result show significant improvement of about 45.8% in the trainees' capacity to wire and program a PLC for automation control. Authors in [10] presented a PLC educational training platform using the TM221CE16R and TM241CEC24R PLCs with TM3AM6/G analogue input/output module XX918A3C2M12 X-ray detector, XUB4APBNL2 photoelectric detector and XS8D1A1PBL5 inductive detector as sensors. They programmed HMISTU855 machine interface using the Vijeo Designer 6.2 Software. Authors presented a comparative study of the PLC programme using SoMachine Basic and SoMachine V4.3

software. A PLC training kit compact and portable for learning on industrial automation practice is presented in [11]. The hardware includes: PLC Omron CP1E-N20R, plug and play I/O interface, DC motor module, electro-pneumatic module integrated in a portable box with dimension 41 × 23 × 11.8 cm.

The objective of this paper is to present design steps for a desktop PLC trainer workstation for industrial automatic engineering operation with emphasis on hardware connection of the input/output components. We present basic PLC automatic operation such as latching and PLC component symbols with description. In order to train students and person's with interest in the field of automatic control for industrial automation. The design and implementation comprise the cabling, programming of PLC using ladder logic with the PC, downloading and uploading the program to the PLC through USB cable and testing on the workstation.

## 2. Design Step

The materials for the development of the proposed PLC training workstation are: MITSUBISHI FX 1S-30MR-001 PLC, simulation software; MITSUBISHI FXTRN-BEG-EL and programming software; MITSUBISHI GX Developer version 8, Power Supply Unit, USB programming cable, switches, push buttons, blower, relay and buzzer.

### 2.1. MITSUBISHI FX 1S-30MR-001 PLC

Mitsubishi PLC is applied in this work to achieve the desired control specification written in LL. We used the MITSUBISHI GX Developer version 8 programming software to write the LL program. The PLC has 28 number of I/O (discrete, I = 14, O = 14), 24 V DC input signal and power supply, with relay and transistor output type. Figure 1 shows the PLC diagram.



Figure 1: Mitsubishi FX 1S-30MR-001 PLC

### 2.2. Power supply unit

The PLC operate on 24 V DV converted from a 220 V AC source voltage. The 24 V DC is distributed through the I/O peripherals as shown in Figure 2.

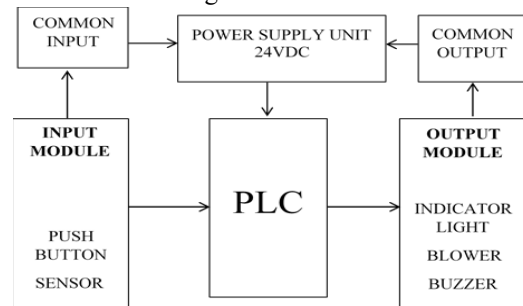






Figure 2: System architecture for the PLC trainer workstation

Table 1 presents other materials used in the setup.

Table 1: Hardware Components

S/N	Item	Specification	Model
1	Buzzer	Active buzzer with predefined frequency (2300 ±300 Hz)	
2	Blower	Rated Voltage: 5V-12V- 24V, Operation Voltage: 3.5-27.6V, Rated Current: 0.08-0.45A, Rated Speed:3000-5000RPM, Air Flow: 23.72-51.23CFM, Static Pressure:3.1-9.63H2O, Noise Level: 24.84-45.32dBA, Weight: 63g	
3	Relay	G2R-1-E Omron relays, Rated coil voltage: 24VDC / 110VAC. Rated switching current: 16A at 250VAC / 16A at 30VDC.	
4	Timer	Time Delay Relay Solid State Timer, ST3PA-B 0-10S Power ON Delay Timer Relay Knob Control Time Relay with Base AC 220V	

5	Push Buttons	Siemens Flush Push Button Buttons 3SB5000-OACO1	
6	Indicator light	16mm 12V Round Panel Mount LED Light CNGAD NXD-213	

2.3. Means of Programming

The means of programming is an HP core i5 laptop computer with windows 10 operating system installed with MITSUBISHI GX Developer version 8 to create/edit the LL diagram, download/upload a control program into or from the PLC. The laptop and the PLC are interfaced using a USB communication cable.

2.4. Hardwired circuit

Figure 2 present’s the system architecture for the PLC trainer workstation. The hand wiring schematic of the I/O component and AC source is shown in Figure 3. Input components are connected to the PLC on the X block and output components are connected to the PLC on the Y block. The AC power source is connected to the neutral (N) and life (L) terminal on the PLC. The common terminal of the Y block of the PLC is inter-connected. Input component terminal is connected to the common and a X address on the X block of the PLC. Output component terminal is connected to 24 V DV and a Y address on the Y block of the PLC.

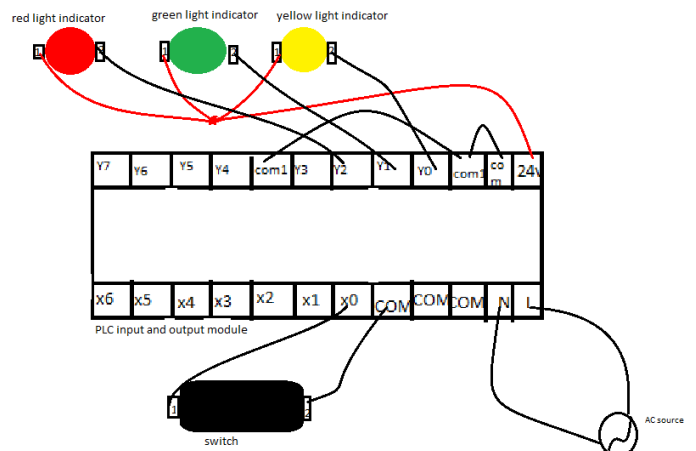


Figure 3: Hand wiring schematic of the I/O component

The hardware component and PLC are mounted on a plywood panel board with aluminium frame shown in Figure 4 with the dimension (120× 60× 5) cm.

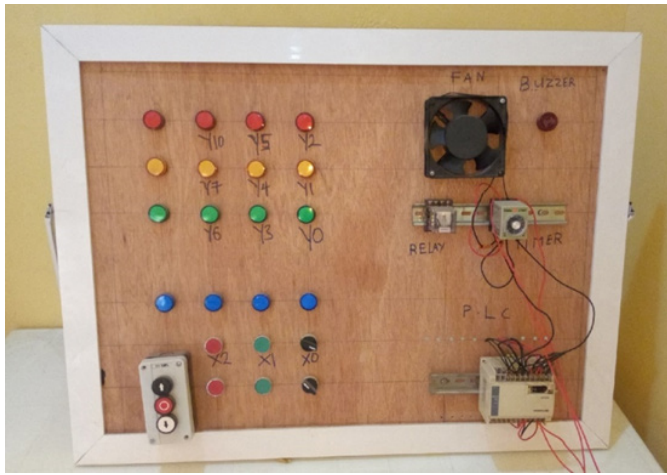


Figure 4: PLC Training workstation

2.5. Programming Language

The LL defined by IEC 6-1131 is the programming language used in this work. The LL programming structure is similar to ladder with horizontal rails and vertical rungs. In the LL structure; power flows from left horizontal rail to right horizontal rail through the connecting vertical rung. Circuit components are shown on the vertical rung in their normal condition. Every rung must specify a control operation with at least one input and output component on each rung. Input component is shown on the left and output component on the right of the rung. Same component can appear multiple time on a LL. I/O component are addressed as specified by the PLC manufacturer. The PLC reads the LL left to right, top to bottom with the last rung specified as END or RET. Table 2 below presents LL symbols with description.

Table 2: LL Symbols with description

S/N	Symbol	Description
1		Input component, normally open contact
2		Input component, normally closed contact
3		Input component, rising edge detection (contact is energized when signal switch from 0 to 1)
4		Input component, falling edge detection (contact is

		energized when signal switch from 1 to 0)
5		
6		instructed out component; SET means set output Y001 energized. And RST means reset output.

The addresses used in Mitsubishi are; X, Y, M, T and C for input component, output component, internal relay, timer and counter. Internal relay number ranges from 000 – 999 and special function relay M8000 is energized by the run button of the PLC. Timer and counter have the K number that specifies the time in seconds and unit to be counted. However, conversion for timer K number to second is 0.1sec to 1 K number.

3. Methodology

Participants are taken through the step by step wiring of the PLC trainer workstation and LL programming discussed above. Under lecturer supervision participant hardwired the PLC trainer workstations as shown in figure 3. To train participant on automatic control they are taken through LL programming for PLC with LL symbols with description explained above. Participant are trained on writing LL program for automatic control and are given assignment on latching procedure to develop their PLC programming skill. The participants are required to write two latching procedure with LL program using [PB1] (X20) for output ON, [PB2] (X21) for output OFF and signal lamp (Y20). Figure 5 and 6 are two random selected LL program by participant.

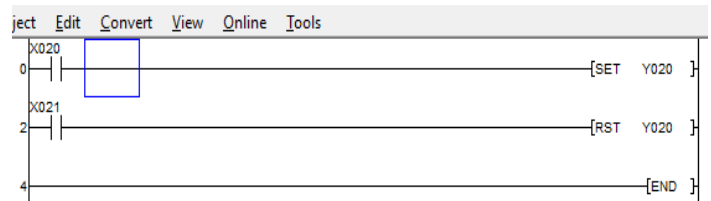


Figure 5 Instructed Output Latching Procedure

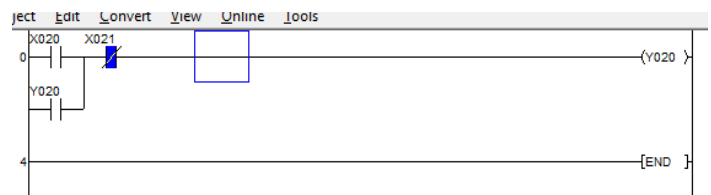


Figure 6 OR Logic with Precedence Latching Procedure

To show that the develop trainer workstation can be applied to train participants on real-world automatic control, two real-world scenarios are presented in table 3. Participants attempt the scenarios and the LL program is shown in figure 7 and 8.

Table 3: Participant Test Modules

No	Scenario1		Scenario 2
i.	When push button X0 is pressed blower Y1 is switched ON and runs for 20 seconds	i.	When switch X2 is ON buzzer Y2 is ON for 5 seconds
ii.	After 20 seconds the blower goes OFF and the buzzer Y2 is ON for 10 seconds	ii.	After buzzer goes OFF LED light Y4 (green) is ON for 10 seconds
iii.	Step i and ii above is repeated	iii.	After 10 seconds Y4 goes OFF and Y5 (red) is ON for 5 seconds
		iv.	The process i – iii is repeated

To measure the impact of the trainer on PLC knowledge of participants a survey was conducted after training. An observation sheet was given to participant for comment on the module of PLC technology tutored. This allowed participant to give their evaluation on each module tutored. To fill the observation sheet participant must partake in all the 3-module tutored and summary presented in section 4.

4. Results and discussion

The summary of the observation sheet filled by the 23 participant taking the course; Automation and Robotics (MTE502) in the Department of Mechatronics where PLC is one of its course content in the laboratory exercise is presented in table 4. The 3module tutored are reflected.

Table 4: Summary of observation sheet

S/N	Topic	Point			
		5	3	1	0
Design step module					
1	Improved knowledge about PLC input and output components	8	12	3	
2	I comprehend PLC system architecture	4	11	5	3
Design step module					
3	I comprehend the description of a PLC unit	2	14	5	2
4	I comprehend PLC system hardwiring	1	17	2	3
LL programming module					
5	I gained more knowledge about PLC input and output components symbol with description	9	12	2	
6	Improved PLC programming using LL	6	11	4	2
Automatic control module					
7	Improved knowledge on automatic control concept such as latching and logic with PLC using LL	2	11	7	3
8	I comprehend the use of timer and relay with LL on PLC	1	8	6	8
9	I can independently write LL program for automatic control with PLC	2	9	8	4

Point: strongly agreed = 5; agreed = 3; disagreed = 1; strongly disagreed = 0

To categorise the participant in two categories; strongly agreed enhance knowledge ( $\geq 3$ ) and strongly disagreed enhance knowledge ( $\leq 1$ ) on PLC using the trainer. Table 5 present the percentage of  $\geq 3$  and  $\leq 1$  grade point on the topics.

Table 4, 5 and Figure 9 present the summary of the observation sheet for the 23 participants. This result show the participants agreed that the PLC trainer workstation enhance their knowledge on design step module and LL programming module significantly. However, about 50 % of participant did not show

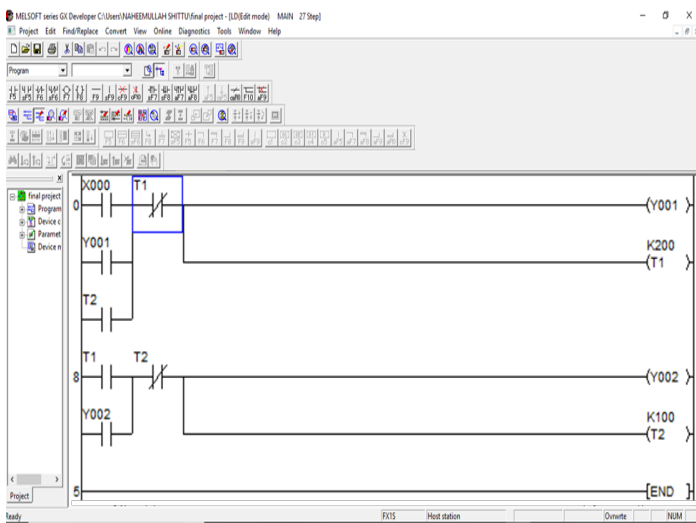


Figure 7: LL Scenario 1

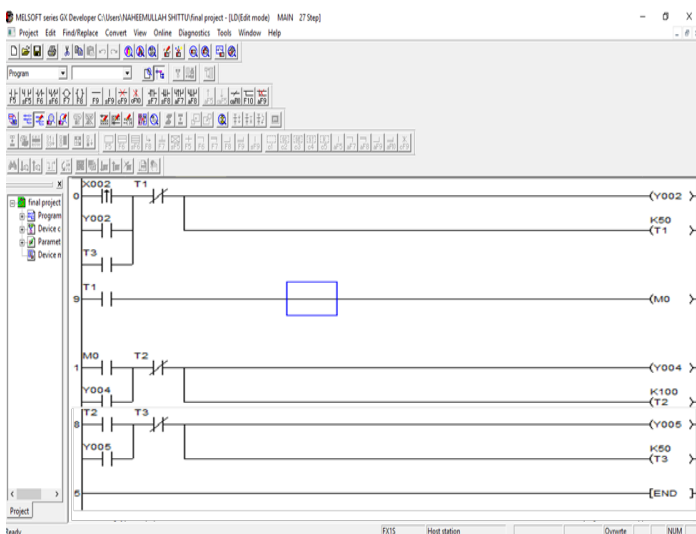


Figure 8: LL Scenario2



confidence writing LL program for automatic control independently and is evident on the automatic control module survey. This can be enhanced with further practice with Human Machine Interface (HMI) for real-time control visualization.

Table 5: Percentage of  $\geq 3$  and  $\leq 1$  grade point

S/N	Topic	% Point	
		$\geq 3$	$\leq 1$
Design step module			
1	Improved knowledge about PLC input and output components	86.96	13.04
2	I comprehend PLC system architecture	65.22	34.78
3	I comprehend the description of a PLC unit	69.57	30.43
4	I comprehend PLC system hardwiring	78.26	21.74
Average point		75.00	25.00
LL programming module			
5	I gained more knowledge about PLC input and output components symbol with description	91.30	8.69
6	Improved PLC programming using LL	73.91	26.09
Average point		82.61	17.39
Design step module			
7	Improved knowledge on automatic control concept such as latching and logic with PLC using LL	56.52	43.48
8	I comprehend the use of timer and relay with LL on PLC	39.13	60.87
9	I can independently write LL program for automatic control with PLC	47.83	52.17
Average point		47.83	52.17

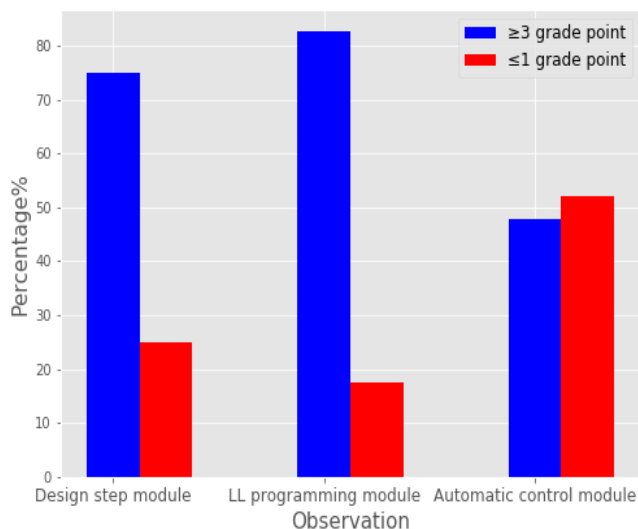


Figure: 9 Percentage comparison of  $\geq 3$  and  $\leq 1$  grade point

## 5. Conclusion

This paper present's the design steps for a desktopPLC trainer workstation for industrial automatic engineering operation. Although researchers have proposed and reported several PLC trainers but they fail to discuss the hardware connection of the input/output components neither is the basic PLC automatic operation nor PLC component symbols with description discussed. These are the areas discussed in this paper to train participant on PLC programming knowledge. A survey of the 23 participant taking the course; Automation and Robotics (MTE502) in the Department of Mechatronics where PLC is one of its course content presented show that 68.43 % of participant strongly agree to improvement in PLC operation knowledge. This percentage is superior tothe45.8 % presented in [9]. Hence, our approach; PLC hardware connection for input/output components, basic PLC automatic operation and PLC component symbols with description discussion with the participant improved their performance in the 3 modules evaluated. However, about 50 % of participant did not show confidence writing LL program for automatic control independently and is evident on the automatic control module survey. This can be enhanced with further practice and Human Machine Interface (HMI) for real-time control visualization.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

Authors like to acknowledge the staff and student of the department of Mechatronics Engineering, Federal University, Oye-Ekit for their support during this work.

## References

- [1] M. Barrett, "The Design of a Portable Programmable Logic Controller (PLC) Training System for Use Outside of the Automation Laboratory" in 2008 International Symposium for Engineering Education, Dublin City University, Ireland, 2008.
- [2] K. Bhise, S. S. Amte, "Embedded PLC Trainer Kit with Industry Application" Int. J. Eng. Sci. Innovative Tech (IJESIT), 4(3), 1-9, 2015.
- [3] C. D. Johnson, Process Control Instrumentation Technology (8th ed.), Upper Saddle River, New Jersey, Prentice Hall, 2006.
- [4] W. Bolton, Programmable Logic Controllers, (4th ed.), UK, ELSEVIER, 2006.
- [5] D. H. Gawali, V. K. Sharma, "FPGA Based Micro-PLC design Approach" Advances in Computing, Control, and Telecommunication Technologies" in 2009 International Conference on Communication, Computing and Electronics Systems, 2009.https://doi.org/10.1109/ACT.2009.167
- [6] M. Mahadi, N. A. Mohd Amin, M. Ab-Rahim, M. S. Abdul Majid, "PLC Trainer Kit Simulator: An Improvement for Automation Study in Polimas" Appl. Mechanics and Mater., 786, 367-371, 2015. https://doi.org/10.4028/www.scientific.net/AMM.786.367
- [7] B. Ibrahim, A. A. Ahmad, T. Saharuddin, "Multiple Input/Outputs Programmable Logic Controller (PLC) Module for Educational Applications" in 2015 Innovation & Commercialization of Medical Electronic Technology Conference (ICMET), Shah Ala, Malaysia, 2015.https://doi.org/10.1109/ICMETC.2015.7449570
- [8] S. Sukir, A. S. J. Wardhana, "Performance of A Programmable Logic Controller Based Electrical Machine Trainer Kit" J. Phys.: Conf. Ser., 1413 (2019) 012011, IOP Publishing, 2019.https://doi.org/10.1088/1742-6596/1413/1/012011

- [9] R. Awingot, A. Albert, F. A. Oliver, "Development of a Programmable Logic Controller Training Platform for the Industrial Control of Processes" *Am. Sci. Res. J. Eng., Tech. Sci. (ASRJETS)*, 186-196, 2016.
- [10] C. Sărăcin, L. G. Deaconu, "Educational Platform Dedicated to the Study of Programmable Logic Controllers and the Human Machine Interface" in 2019 The 11th International Symposium on Advanced Topics in Electrical Engineering, Bucharest, Romania, 2019.  
<https://doi.org/10.1109/ATEE.2019.8725021>
- [11] E. S. Maarif, S. Suhartinah, "Compact Portable Industrial Automation Kit for Vocational School and Industrial Training" in 2017 International Symposium on Materials and Electrical Engineering (ISMEE), 2017.  
<https://doi.org/10.1088/1757-899X/384/1/012011>

## An Explanatory Review on Cybersecurity Capability Maturity Models

Adamu Abdullahi Garba<sup>1,\*</sup>, Maheyazah Muhamad Siraj<sup>2</sup>, Siti Hajar Othman<sup>2</sup>

<sup>1</sup>Yobe State University Damaturu, Computer Science, Yobe State University Damaturu, 1144, Nigeria

<sup>2</sup>School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, 81310, Malaysia

### ARTICLE INFO

*Article history:*

*Received: 26 June, 2020*

*Accepted: 14 August, 2020*

*Online: 28 August, 2020*

*Keywords:*

*Cybersecurity Maturity Model*

*Capability Maturity Model*

*Security Models*

*Confidentiality*

### ABSTRACT

Cybersecurity is growing exponentially day by day in both the public and private sectors. This growth also comes with a new and dynamic cyber-threats risk that causes both sectors' performance to halt. These sectors must update their cybersecurity measures and must understand the capability and maturity of their organization's cybersecurity preparedness. Cybersecurity maturity models are widely used to measure how ready an organization is when it comes to cybersecurity. The main aim of this article is to conduct a comprehensive review of the current cybersecurity capability maturity models using a systematic review of published articles from 2011 to 2019. A comparative study was conducted based on Halvorsen and Conradi's taxonomy. The review indicated almost all the cybersecurity maturity model consists of similar elements like maturity levels and processes but significantly lacks the validation process, it was observed each of the models were predominantly designed for a specific purpose and also for different organization size and application domain.

## 1. Introduction

Cybersecurity is a method of protecting organization assets, through the identification of threats that can compromise the critical information stored in the organization systems, it also involves the protection, identification, and responding to threats. However, cybersecurity evolves from computer security which means securing the physical components of a computer system from any damage, to information security which means securing the stored information in a computer system from unauthorized access by maintaining its Confidentiality, integrity and availability (CIA) then to cybersecurity which includes both computer security and information security and also adding the security of information being transferred across a different medium (wired & wireless) and also access from anywhere. The advancement of the cybersecurity domain is dated back to the 1950s. The field of cybersecurity emerged as a result of Robert Morris testing the worlds' network vulnerability in 1980 when he uses a virus he created to test the size of the internet, to protect organization assets, an organization needs to improve their cybersecurity practices. The knowledge of cybersecurity has also been used negatively, the Russian in the 1980s attacked around 400 military computers in the US which include the pentagon computers[1-3]. Therefore, knowledge of cybersecurity capability maturity models is essential

as the research area is new and growing exponentially, critical review in the existing models and their applications is important to know, to fill this gap this paper intends to answer the following objective:

- To identify currently available cybersecurity capability maturity models available for this study from 2011 to 2019 using systematic review (SR).
- To identify the main difference between the cybersecurity capability maturity model and their levels
- To understand the application of the cybersecurity capability maturity models.

This paper consists of eight sections, section 1 is the introduction, the second section 2 discusses on the evolution of cybersecurity capability maturity models, section 3 discusses on the method used in conducting the research, section 4 explains the Review on cybersecurity capability maturity models, section 5 explains the results and discussion of the comparative analysis of the identified models, it further discusses the importance of the research and explained how the objective was achieved, section 6 explains the future direction of the research and from where other authors can continue to explain the research direction and lastly, section 7 is the conclusion.

\*Corresponding Author: Adamu Abdullahi Garba, adamugaidam@gmail.com

1.1. Review Method and Protocol

The systematic literature review is defined as “ a well-defined study or methodology for identifying, analyzing, and also interpreting all available evidence related to a specific research question [4,5]. This method was mostly used in medicine [6], but not now it has been adopted by many fields of studies like social science, information system, and computer science, software engineering [5]. In this study systematic review, “which aims for exhaustive searching, quality assessment, inclusion, and exclusion criteria which are typically narrative with tabular form was adopted ”[7]. The aim or analysis of this method is to explain what is known for practices, what remains unknown, and recommendation for future research.

This systematic search started with a well-developed review protocol based on the procedures of the SR review. The protocol includes: background study as evolution, review method, research objectives, and data extraction criteria, and for this study. This section helps to increase the accuracy of the review and also reduces bias in conducting the research. Table 1 describes the review criteria.

Table 1: Inclusion and exclusion criteria

Included article	Excluded article
Full text and available	Full text but not available
Year of publication from 2011 – 2019	Outside range of the year
Published in English	Non-English
Only focus on the domain (cybersecurity)	Were outside domain
Must be a model	Not related to objectives
Found in the selected database.	Duplicated studies

1.2. Inclusion and exclusion criteria

This section is mainly to set up the criteria of inclusion and exclusion for the researchers to follow when doing the study. This research considers the following articles (emerald, IEEE explore, Scopus, the web of science and science direct) published in English, also published from 20011 to 2019 in the digital database. Table 1 shows the steps used in conducting the research.

As part of Step1, we searched articles that have the phrase “cybersecurity”, “security model” AND “maturity”, from different databases. After following the protocol mention in table 1, we collected 220 articles relevant to your objective

Step 2 is where we used the inclusion and exclusion criteria to determine article very close to your objective by reading abstract [5], we removed all the papers that do not have the word “cybersecurity model” and “cybersecurity capability maturity model”. At the end of this phase, we only obtained 30 articles.

Step 3, the article obtained in step 2 was critically analyzed and read fully with more depth analysis. Based on the full- text, the previous criteria were applied to identify the actual articles that are related to our objective. The articles used are fully cybersecurity oriented. At the end of this phase, only 7 articles were obtained. These seven articles are selected based on table 1 criteria.

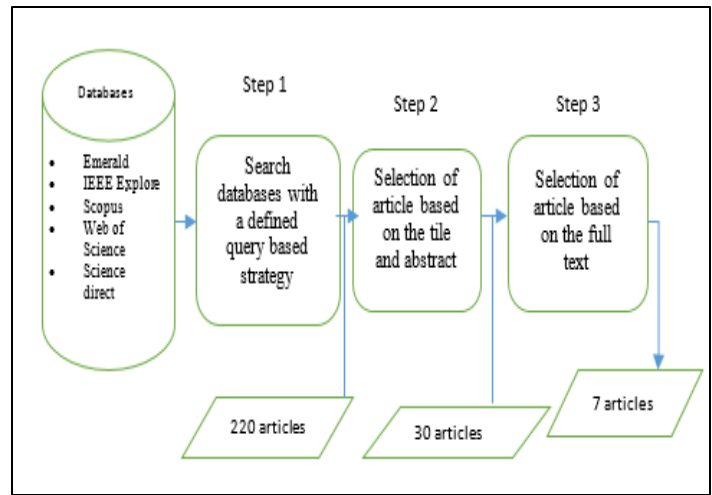


Figure 1. Search and selection of articles considered for this study.

2. Cybersecurity Capability Maturity Models

The cybersecurity Capability maturity model has emerged from the capability maturity model been design from the quality management field in the 1930s. it becomes popular in the 1990s when it was first developed by software engineering institutions [8]. Today all these models are based on this basic model, the model has a set of a structured set of operations and activities that improve over time [9]. the model is later being adapted into many fields of studies to identify or measure the maturity level of an organization or process or even product quality as they are widely known. The capability maturity model (CMM) which was for software industries has some key elements for an effective software development process[10] the model has 5 basic process maturity levels called, initial, repeatable, defined, managed, and optimizing [11]. A best- practices and process efficiency is provided for every five levels of each process for evaluating the maturity [12]. Also [13] conducted a study regarding the capability maturity models, in his research he identified and compares many maturity models for software domain and product quality, while [14] surveyed maturity models specifically for knowledge management to find out how far it contributed to the measurement of knowledge management, but the study only emphases on one special type of maturity model, therefore it is not suitable for general mapping of the maturity model research, the model was designed for software products as guidance as well as for management excellence in producing quality software”[15].

The cybersecurity maturity model offers a framework for assessing the maturity of a security program and guidance on how to reach the next level [16]. The cybersecurity maturity model provides a pathway that enables the organization to measure where they are along that path. This can be a valuable tool not only for improving Cybersecurity efforts but also for collaborating with upper management and getting the support needed to enhance Cybersecurity culture in organizations. There are various Cybersecurity Maturity Models from which to choose, Based on the systematic review performed regarding the currently available Cybersecurity models published to the knowledge of the author from 2011 to 2019 are; Cybersecurity Capability Maturity Model (C2M2), Electricity Subsector Cybersecurity Capability Maturity Model (ES-C2M2), Oil and Natural Gas Subsector



Cybersecurity Capability Maturity Model (ONG-C2M2), National Initiative for Cybersecurity Education-Cybersecurity Capability Maturity Model (NICE-C2M2), Community Cyber Security Maturity Model (CCSMM), African union maturity model for cybersecurity (AUMMCS) and Federal Financial Institute of Examination Council Capability Maturity Model (FFIEC- CMM). These identified models are selected because they focus on cybersecurity, other models were found during the SR but were not fully focusing on cybersecurity, like the National Institute of Standards and Technology Cybersecurity Framework (NIST CSF), The NIST CSF differs from the C2M2, as NIST doesn't consider the CSF a maturity model, Instead of 10 domains, NIST represents five Cybersecurity functions: identity, protect, detect, respond and recover Models identified section 5 shows Comparisons Cybersecurity Capability Maturity Models results then followed by a discussion on the Comparisons Cybersecurity Capability Maturity Models.

### 3. Methodology

Based on the systematic review performed regarding the currently available cybersecurity capability maturity models published, the author has identified the following: C2M2, ES-C2M2, ONG-C2M2, NICE-C2M2, CCSMM, FFIEC- CMM, and AUMMCS as explained in the previous section. The author adapted Halverson and Conradi taxonomy of software process improvement (2001), this taxonomy consists of 21 features peculiar to software process and is grouped into 5 categories: **general, process, organization, quality, and result**. Each category refers to:

- **General:** the features that describe the overall attribute of improvement.
- **Process:** the feature that explains the way the organization uses the features.
- **Organization:** this explains the relationship between the features and organization and how they work simultaneously.
- **Quality:** this explains the feature related to the quality dimension.
- **Result:** this explains the feature of the results as the result of using the environment, the cost of achieving the result. In this paper, only **general, process, organization, and results** are adapted as the other one has no relation to Cybersecurity Capability Maturity Models. The feature that falls under each category is modified to suit Cybersecurity terms as shown in Table 2 below.

Table 1: Halverson and Conradi taxonomy

Category	Feature
General	Cybersecurity oriented
	Origin
	Purpose
	Prescriptive/ descriptive
	Maturity level
Process	Field Applicable
	Define role
	Depth of assessment

	Assessment
	Assessor
Organization	Actors
	Organization size
	Level of documentation
	Organization Environment
Result	Implementation cost

The features related to **General** group are defined below:

- **Cybersecurity Oriented:** this feature depicts which model was purposely designed for Cybersecurity maturity.
- **Origin:** this criterion is used to know the country, lab, organization that created or design the model e.g. the US.
- **Prescriptive/Descriptive:** the criteria used here is either Prescriptive: if the model is enforcing a rule to be used if the model is adapted, while descriptive: if a model is describing or classifying its objectives and how to follow it, not enforcing rules.
- **Maturity Level:** the criteria are used to understand the level of maturity for each model number 1- 5 are used, the more level a mode is the more level of the maturity increases.
- **Field Applicable:** the criteria is used to know the area where the model is applicable criteria include: organization, paper lab. University.
- **Define Role:** this feature explains how well the roles and functions are evaluated using “ Yes” if a role is defined and else “ No” is used.
- **Depth of Assessment:** the criteria are used here is either “General” if the assessment is not in-depth and “specific” if the assessment is in-depth that is more than one level.
- **Actors:** the criteria used here are “ management, staff, communities or states “ to know who is using the model directly.
- **Organization Size:** this criterion is used to know the size of the organization for appropriate adaption, criteria used here are: large, medium, small, or all.
- **Level of Documentation:** criteria use are either “high” when a model has implementation guide and other supporting documents that will help adaptor to implement the model, “medium “is when no more details are available on the implementation guide but there are white papers and other supporting documents, “low” in both implementation and white paper are not available but other introductory documents are available.
- **Organization Environment:** criteria “Overall” is used if the model focuses on the entire organization while “ Explicit” if the model focuses on a specific unit or department in the organization.
- **Assessment:** the feature is explained by the name of a process to be assessed in the organization e.g. risk, maturity, customer, employee, organization.
- **Assessor:** the criteria use in this feature are “ internal” if the assessor is from the environment the model is implemented, “ external” if the assessor comes from outside the workplace, and “internal and external” where the assessor can be both.

- **Validation Method:** this criteria is use to know the method of validation includes: survey, case study. Experiment.
- **Implementation Cost:** this criteria is use to know how much to spend when implementing the model.

#### 4. Review

This section explains the main structure and domain found in the identified maturity models based on their focus on cybersecurity. Based on the SR the identified from scientific articles are C2M2, ES-C2M2, ONG-C2M2, NICE-C2M2, CCSMM, FFIEC- CMM, and AUMMCS models.

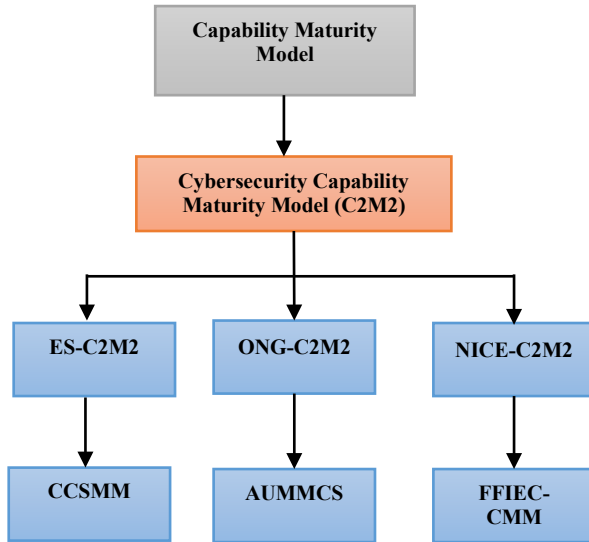


Figure 2: Descriptive diagram of maturity models

Figure 2 above elaborates more on the identified maturity models, it indicates C2M2 is the origin or the first model to be designed in the cybersecurity domain, The difference from the models is either from the number of levels like C2M2 has 4 levels while CCSMM has 5 levels, application area, also C2M2 can be assessed both internally and externally while CCMM can only be asses externally. However, the most important concept of all is that they are only design for cybersecurity specifications, i.e cybersecurity orientation.

##### 4.1. Cybersecurity Capability Maturity Model (C2M2)

The Cybersecurity Capability Maturity Model was designed by Carnegie Mellon University in collaboration with the US Department of Energy in 2014 [17]. The model has ten domains and each domain is a grouping of cybersecurity practices. Also, many objectives are grouped to be in one domain which represents achievements the model contains ten domains with grouped objectives and Maturity levels. (Appendix A shows the domain and also the maturity level)

##### 4.2. Electricity Subsector Cybersecurity Capability Maturity Model (ES-C2M2)

The Electricity Subsector Cybersecurity Capability Maturity Model was designed by the department of energy USA to protect the electricity subsector from any form of cybersecurity attacks. [18]. This model was designed as a subsector of the C2M2 i.e.

Independent guidance. Both models' general purpose is almost the same, which is to improve cybersecurity capabilities by allowing continuous benchmarking. This model also has ten domain and are the same as the C2M2. The model was developed with the main four primary sector functions as listed below.

- Generation
- Transmission
- Distribution
- Markets

There is a difference between the models in reporting of incidents, C2M2 mentioned ISACs in general in the DOE form while ES-C2M2 threat and vulnerability incident are reported to electricity sector information sharing and analysis center specifically [18]. The model is purposely for electricity sector organizations.

##### 4.3. Oil and Natural Gas Subsector Cybersecurity Capability Maturity Model (ONG-C2M2)

The Oil and Natural Gas Subsector Cybersecurity Capability Maturity Model was developed or derived from ES-C2M2 first version, the main reason for this model was to address the threats and vulnerabilities uniquely characterized by the Oil and Gas subsector. This model can be used to support cybersecurity capabilities in the ONG subsector, it enables organizations to evaluate and benchmark their capabilities. The ONG-C2M2 was designed to address problems in the oil and gas sector only. The development process was extensively cantered with public and private sector experts through pilot facilitation, working sessions. This model uniquely includes the exploration, gathering, production, processing storage, and transportation of petroleum liquids and natural gas. The critical areas where threats can occur, from exploration to storage as technology is used to do all the processes, where security has been not tightened well, an attacker from nowhere can hinder the process or even stops the organizational activities. The ONG-C2M2 and the ES-C2M2 are derived from the C2M2, therefore they share the same domain and maturity levels only place of applications, and the purpose of design differs. This is the reason the author did not repeat the same tables to avoid repetitions, see Appendix A shows the domain and also the maturity level)

##### 4.4. National Initiative for Cybersecurity Education Capability Maturity Model (NICE)

The NICE model was designed by then US President George Bush under the directive of national security in 2008, [19]. The model was designed purposely to select the staff with cybersecurity background and skills. The model comprises three key components focusing on staff security structure at the management and the role of staff, the model was officially published in 2014 [19]. The model consists of three domains: Process and Analytics, Integrated Governance, and Trained Professionals and Enabling Technology, which also three maturity levels (*limited level, progressing level, and optimized level*).

#### 4.5. Federal Financial Institute of Examination Council Capability Maturity Model (FFIEC- CMM)

The Federal Financial Institute of Examination Council Capability Maturity Model was published to guide or navigate the increasing complexity of the cyber risk landscape. This assessment tool was designed to help managers assess their institution's cybersecurity preparedness, evaluate its risk, and determine what risk management practices and controls are needed to attain the desired state. This tool has two-part as shown below:

- **The Inherent Risk Profile:** these are risks posed to the organization by technologies and connection types, delivery channels, online and mobile products, and other external threats.
- **Cybersecurity Maturity:** this helps the organization to measure the level of risk and corresponding controls. The level starts from baseline to innovation. The model contains five domains and some assessment factors

#### 4.6. African Union Maturity Model for Cybersecurity (AUMMCS)

The African union maturity model for cybersecurity was made available in 2014 by the center for the cybersecurity at the University of Johannesburg on security and protection of personal data at the convention of African member states, this model covers three sections: *electronic transactions, personal data protection and promoting cybersecurity and combating cybercrime* [20]. The model was intended to help member states of the African Union to evaluate their cybersecurity status against a specific part of the convention. This model can be utilized in two ways: one as a self-assessment by a specific country against the specification of the convention, two as a comparison by the AU between different member states in order to see how they can compare themselves as requirements are concerned. The model only covers the promotion of cybersecurity and combating cybercrime[20]. The model has four objectives:

- A national culture of cybersecurity does exist.
- A national Cybersecurity policy does exist.
- Public-private partnerships, initiated by the government, do exist.
- Cybersecurity capacity building on all levels, driven by the government, does exist.

The model also has four MLS maturity levels:

- ML0: Nothing Exists At All.
- ML1: Very Basic Position.
- ML2: Progressed Position.
- ML 3: Stable Position.

#### 4.7. The Community Cybersecurity Maturity Model (CCM2)

The community cybersecurity maturity model was developed in San Antonio Texas by the Centre for Infrastructure Assurance and Security (CIAS) [21]. this model was designed to the needs of state and community to the development of a practical and sustainable program of cybersecurity in Texas United States. This model identified the character of community and state as their cyber-security program mature, such aspect includes knowledge, security policies, procedures, information exchange, and

cybersecurity training and education. The main importance of this model is to respond to the linkage that exists among states since more communities made up a state. Also, the model is shown in a three-dimensional way [21]. This model is made up of five maturity levels with the lowest initial level showing characteristics for communities that do not share a minimal level of cyber preparedness in the four key areas known as *Awareness, information sharing, processes and procedures, and integration*. [21]. An example, like top managers at level one, would have little or even no awareness of any cybersecurity threat and its damage, also have little or no information sharing on the cyber event between entities within organization cities or states. Also, few processes or procedures would be in place in the community to handle any cyber threat and lastly, lack of or no mechanism in place (security exercise) to evaluate the level of preparedness of the community or its capacity to respond to any threats. Initially, the model focus on designing the roadmap for the community to follow than later was identified that it is not yet robust enough to adequately represented what is needed for a community to be secure individual community must have a certain level of security as well program necessary to address prevention and detection of cyber threats. The model shows how important information sharing is with other communities so that current threats picture can be obtained and to be able to alert other communities that might be affected and share measures are taken. The model was then expended to include three-dimension to include a third axis that will indicate characteristics and activities at an individual organizational level as well as at a state level.

#### 4.8. Comparative Evaluation of Cybersecurity Capability

The section shows a well and detailed explanation of the identified maturity models, a descriptive diagram of the maturity models, also table 3 shows a summary of the comparisons among the models using the adopted taxonomy features from Halverson and Conradi's taxonomy of software process improvement.

#### 4.9. Maturity Model

**Note:** 1 yet to be determined

Table 3 indicated how Halverson and Conradi's taxonomy features were used in explaining the identified cybersecurity capability maturity models. The features give a full description of all the models, such as their origin, reasons for creating the model, number of maturity levels, where it is applicable, who can use the model in the organization, how depth the implementation guidelines, etc. This description will help other organizations to see the features of each model and where it can be applied. Furthermore, the table can be an inside for top management of an organization that has less knowledge of cybersecurity to decide or decide on what type of model would suit their organization. The research aims to identify models from 2011 to 2019, but mostly relevant material from 2011 to 2019 was used, this is because only a few models are specifically cybersecurity oriented, those what were identified but did not fit the inclusion criteria includes like Control Objectives for Information and Related Technology Organization (COBIT), Project Management Maturity Model (OPMM) and Siemens Knowledge Management Maturity Model (KMMM) were not used in this research.

Table 3: Comparative review on cybersecurity capability maturity model

<b>Model</b> <b>Features</b>	<b>C2M2</b>	<b>ES-C2M2</b>	<b>ONG-C2M2</b>	<b>NICE-C2M2</b>	<b>CCSMM</b>	<b>FFIEC-CMM</b>	<b>AUMMCS</b>
Cybersecurity Oriented	Yes	Yes (derived from vC2M2)	Yes (derived from ES-C2M2)	Yes	Yes	Yes	Yes
Origin	The US.Dept of Energy	The US.Dept of Energy	The US.Dept of Energy	The US.Dept of Energy	CIAS	US Federal Financial Institute Of Examination Council	Centre For Cyber Security At The University of Johannesburg
Maturity level	4	3	4	3	5	5	4
Purpose	Assessment of cybersecurity capabilities for any organization comprises of a maturity model evaluating a tool	Tailored to energy subsector	Tailored to the oil and natural gas subsector	Tailored to three areas: process and analytics, integrated governance, skilled practitioners and technology for work development	Tailored to communities yardstick to know the security posture	Tailored to as assessment tools to identify organizational risk and determine their cybersecurity maturity	Tailored to ensuring citizens and government and business are protected African member states
Actors	Management	Management	Management	Staff	Communicates	Management/ Employees	States
Organization Size	large	large	large	large	Medium	large	All
Level of Documentation	Medium	Medium	Medium	Medium	Low	High	Medium
Organization Environment	Overall	Overall	Overall	Explicit	Explicit	Explicit	Overall
Define role	Yes	Yes	Yes	Yes	No	No	No
Depth of Assessment	Specific	Specific	Specific	General	Specific	Specific	General
Field Applicable	Organization	Electricity	Oil and Natural Gas	Workforce	Communities	Financial Organization	African states
Prescriptive/ descriptive	Both	Both	Both	Both	Descriptive	Both	Both
Assessment	Organization maturity	Electricity grid protection	Oil and gas protection	Organization maturity	Community protection	Organization maturity	Data protection
Assessor	Internal and external	Internal and external	Internal and external	Internal and external	external	external	external
Implementation Cost	1	1	1	1	1	1	1i



## 5. Result and Discussion

Cybersecurity has been growing exponentially day by day both in private and public sectors, so also cyber threats. These threats are dynamic and organizations need to be updated on the measures they should adopt to secure the critical assets. In this article, a review was carried out to identify the most recent mechanism used in protecting and also identifying the maturity of cybersecurity in an organization. The research is limited to models from 2011 to 2019 and also those designed specifically for cybersecurity as specified in the first objective. Table 3 elaborated on the comparisons of the identified models, the table shows most of the models have basic similarities, such as domains and levels, but also differ in some areas which include the level of implementation and guidelines, the actor's role, the field of application, and also assessment. Objective two of this article was to identify the maturity and the level of the models, the models have almost similar maturity description. Some models use levels like C2M2, while others use the baseline to innovation and others use initial to vanguard to describe how maturity increase from one level to the next. Overall they have basic similarities.

However, some models were derived from other models like in the case of the ES-C2M2, and the ONG-C2M2 models are derived from C2M2. Most models are more specific than generic. The last objective was to understand the application domain of the identified models. Most of the models' application domain includes organization, oil and gas section, communities, banking sectors, and even continent as a form of guidelines as shown in table 3. Certain models are designed to be used for the entire organization like C2M2 while other are not like NICE-C2M2. The complete adoption of a model seems to be impossible as the most model is designed for a specific purpose as in NICE-C2M2 which was designed for skilled staff. This discussion further shows that organization can assess their needs before selecting an appropriate model to measure their cybersecurity maturity level.

## 6. Research Direction

This paper explains the cybersecurity maturity models properties and their similarities and Applications domain, based on the reviews of all the available models, no any author explains the validation process of the proposed model before implementation, therefore, a future research can focus on how cybersecurity capability maturity models are evaluated and also cost of implementation of the model in an organization as no model explains the financial standpoint.

## 7. Conclusion

In conclusion, cybersecurity measures is an essential entity to be known by all organizations, identifying organizational maturity level and knowledge on cybersecurity is a most, also knowing what model to be used in identifying the maturity level is important. There is limited research on cybersecurity capability maturity models and their application as the research area is new and growing exponentially. This research will serve as the first step in knowing the relevant cybersecurity capability maturity models available and also areas of application. However, all the identified models are fully based on cybersecurity but adopting

can be impossible, however, the models can be adapted and customized. Tables 3 give a clear view of all the models and how to choose a suitable model for any organization based on the features used. Furthermore, only C2M2 focuses on the entire organization while others focused on cybersecurity. Lastly, all models found after the SR lacks cost implementation, therefore, to know how much to spend for implementing any model depends highly on the size of the organization and the number of critical assets to be protected.

## Conflict of Interest

The author declares no conflict of interest

## Acknowledgment

This journal would not have been possible without the exceptional support and guidance of my supervisors Dr. Maheyzah Muhamad and Dr. Siraj Siti Hajar Othman. Their enthusiasm, knowledge, and exacting attention to details have been an inspiration and kept this journal work on track. I would also like to thank all the reviewers that gave their comments to make this paper acceptable to the community of knowledge.

## References

- [1]. M. Dunn Caveltly, "Breaking the Cyber-Security Dilemma: Aligning Security Needs and Removing Vulnerabilities," *Science and Engineering Ethics*, **20**(3), 701–715, 2014, doi:10.1007/s11948-014-9551-y.
- [2]. Nicholas R., *The Cybersecurity Dilemma*, Duke University, 2011.
- [3]. R.W. Taylor, *Cyber Crime and Cyber Terrorism*, 4th ed., Pearson Education, Inc., United State, 2019.
- [4]. Z. Soltani, N.J. Navimipour, "Customer relationship management mechanisms: A systematic review of the state of the art literature and recommendations for future research," *Computers in Human Behavior*, **61**, 667–688, 2016, doi:10.1016/j.chb.2016.03.008.
- [5]. Kitchenham, *Guidelines for performing Systematic Literature Reviews in Software Engineering*, Durham Durham, UK, 2007, doi:10.1145/1134285.1134500.
- [6]. A.D. Oxman, "Systematic Reviews: Checklists for review articles," *BMJ*, **309**(6955), 648–651, 1994, doi:10.1136/bmj.309.6955.648.
- [7]. M.J. Grant, A. Booth, "A typology of reviews: An analysis of 14 review types and associated methodologies," *Health Information and Libraries Journal*, **26**(2), 91–108, 2009, doi:10.1111/j.1471-1842.2009.00848.x.
- [8]. M.C. Paulk, "A History of the Capability Maturity Model for Software," *The Software Quality Profile*, **1**(1), 5–19, 2009.
- [9]. C. V Weber, S.M. Garcia, M. Bush, "Key Practices of the Capability Maturity Model," 1993.
- [10]. Y. Goksen, E. Cevik, H. Avunduk, "A Case Analysis on the Focus on the Maturity Models and Information Technologies," *Procedia Economics and Finance*, **19**(15), 208–216, 2015, doi:10.1016/s2212-5671(15)00022-2.
- [11]. R.M. Adler, "A dynamic capability maturity model for improving cyber security," 2013 IEEE International Conference on Technologies for Homeland Security, HST 2013, 230–235, 2013, doi:10.1109/THS.2013.6699005.
- [12]. P. Byrnes, M. Phillips, "-- ~ 47L ~ Software Engineering Method Description ESC-TR-96-002," (April), 1996.
- [13]. D. Budgen, M. Turner, P. Brereton, B. Kitchenham, "Using Mapping Studies in Software Engineering," *Ppigi*, **2**, 195–204, 2008.
- [14]. N. Khatibian, T. Hasan gholoi pour, H. Abedi Jafari, "Measurement of knowledge management maturity level within organizations," *Business Strategy Series*, **11**(1), 54–70, 2010, doi:10.1108/17515631011013113.
- [15]. G.B. White, "The Community Cyber Security Maturity Model The Center for Infrastructure Assurance and Security," *Proceedings of the 40th Hawaii International Conference on System Sciences*, (June), 1–8, 2007, doi:10.1109/HICSS.2007.522.
- [16]. L. Johnson, *Cybersecurity framework*, 2020, doi:10.1016/b978-0-12-818427-1.00012-4.
- [17]. W. Miron, K. Muita, "Technology Innovation Management Review

Cybersecurity Capability Maturity Models for Providers of Critical Infrastructure,” Technology Innovation Management Review, 4(October), 33–39, 2014, doi:10.22215/timreview/837.

- [18]. A. Sorini, E. Staroswiecki, 8. Cybersecurity for the Smart Grid, Elsevier Ltd, 2017, doi:10.1016/B978-0-12-805321-8.00008-2.
- [19]. J.A.C.-M. and I.D.S.-G. Angel Marcelo Rea-Guaman, Tomás San Feliu, “Comparative Study of Cybersecurity Capability Maturity Models Angel,” Computer Standards and Interfaces Software Process Improvement and Capability Determination Conference 2017, 60, 1–2, 2018, doi:10.1016/j.csi.2018.05.002.
- [20]. S.H.B. Von Solms, “A maturity model for part of the African Union Convention on Cyber Security,” Proceedings of the 2015 Science and Information Conference, SAI 2015, 1316–1320, 2015, doi:10.1109/SAI.2015.7237313.
- [21]. G.B. White, “The community cyber security maturity model,” 2011 IEEE International Conference on Technologies for Homeland Security, HST 2011, 173–178, 2011, doi:10.1109/THS.2011.6107866.

**Appendix A**

Table 4: C2M2 Domain description

Domains	Grouped Objectives
Asset, Change and Configuration Management	Manage Asset inventory Manage Asset configuration Manage changes to Asset Management Activities
Cybersecurity Program Management	Established Cybersecurity Program Strategy Sponsor Cybersecurity Program Established And Maintain Cybersecurity Architecture Perform Secure Software Development Management Activities
Event and Incident Response, Continuity of Operation	Detect Cybersecurity Events Escalate Cybersecurity Events And Declare Incidents Respond To Incident And Escalated Cybersecurity Events Plan Continuity Management Activities
Identify and Access Management	Established And Maintain Identities Control Assess Management Activities
Information Sharing and Communications	Share Cybersecurity Information Management Activities
Risk Management	Established Cybersecurity Risk Management Strategy Manage Cybersecurity Risk Management Activities
Situational Awareness	Perform Logging Perform Monitoring Established And Maintain A Common Operating Picture Management Activities
Supply Chain and External Dependencies Management	Identify Dependencies Manage Dependency Management Activities
Threat And Vulnerability Management	Identify And Respond To Threats Reduce Cybersecurity Vulnerabilities Management Activities
Workforce Management	Assign Cybersecurity Responsibilities Control The Workforce Life Cycle Develop a Cybersecurity Workforce Increase Cybersecurity Awareness Management Activities

Table 5: C2M2 Maturity level description

Maturity level MIL	indicator	Level description
Level 0		This level has no practices or processes defined. There is no stable environment for activities. MIL 0 is given as a result of the domain objective not achieved.
Level 1		This level contains a set of initial practices. This level activities are usually ad hoc and chaotic. MIL 1 is scored if there is an initial practice performed
Level 2		This level has more stable practice compared to MIL, more confidence is achieved at this level as the result of the performance and is sustained over time.
Level 3		At MIL 3 policy is applied to the practices to further stabilize the operations in the organization and is guided by top-management directives. Also, the staff s’ are fully trained and fully funded.

Table 6: ES-C2M2 and ONG-C2M2 domain description

Domain	Practices
Risk	Risk Assessment
Assets	Asset, Change, and Configuration Management
Access	Identity and Access Management
Threat	Threat and Vulnerability Management
Situation	Situational Awareness
Sharing	Information Sharing And Communication
Response	Event And Incident Response, Continuity Of Operations
Dependencies	Supply Chain And External Dependencies Response Management
Workforce	Workforce Management
Cyber	Cybersecurity Program Management

Table 7: ES-C2M2 and ONG-C2M2 maturity level

Maturity Level	Description
MIL 0 “ Not Performed”	This level describes the domain has achieved nothing.
MIL 1 “ Initial”	This level shows only initial practices are performed
MIL 2 “Performed”	The level is characterized by having well-documented practices, stakeholders’ involvement, and provision of standards or guidelines for practice implementation.
Mil 3 “Managed “	This level shows all practices and activities are fully guided by policy, also practice is only assigned to adequate skills personal. The formed policy are periodically evaluated for improvement

## A Method for Detecting Human Presence and Movement Using Impulse Radar

Young-Jin Park\*, Hui-Sup Cho

Division of Electronics & Information System, DGIST, Daegu, 42988, South Korea

---

### ARTICLE INFO

*Article history:*

*Received: 01 July, 2020*

*Accepted: 24 August, 2020*

*Online: 28 August, 2020*

---

*Keywords:*

*UWB Impulse Radar*

*Machine Learning*

*Human Detecting*

*Noncontact measurement*

---

---

### ABSTRACT

*Using non-invasive and non-contact sensors to measure a person's presence or movement helps improve the quality of life for both healthy people and patients. In this paper, a method of measuring the presence and motion of a person is proposed by utilizing UWB Impulse Radar, which is low power consumption and safe to radiate to the human body. The experimental stage of this study is divided into the stage of extracting features by signal processing from radar signals, the stage of generating datasets with 3~6 kinds of labels, and the stage of performing and verifying machine learning by imaging. In this experiment, a small number of images were used because only good quality signals were selected and used by radiating radar signals to the human body. The experiment result show high accuracy when using neural networks such as GoogLeNet and SqueezeNet. Experiments in this study confirmed that radar signals could be used to detect human presence and motion as a result of studies using the proposed method.*

---

### 1. Introduction

This paper is an extension of the work originally presented in the 2019 International Conference on Information and Communication Technology Convergence (ICTC) [1].

Non-contact measurements are useful for health care, human activity, security, surveillance, etc., and the prevalence of COVID-19 in 2020 is a constant increase in demand for non-invasive, contactless measurements of physiological functions in modern society. Especially, non-contact sensors such as radar technology can be used to identify health conditions and movements without limiting human activities. In commercial applications, radar sensors are widely used for LED control and alarm monitoring in indoor and outdoor environments, as well as in smart homes and cities. A radar sensor can be applied in areas where there are restrictions on the use of sound, infrared, vibration, and camera sensors. For example, acoustic and vibration sensors are very vulnerable to acoustic noise, and infrared sensors frequently generate false alarms in outdoor environments. Moreover, camera sensors are relatively expensive, require high signal processing, have low performance at night, and have lens contamination problems. On the other hand, radar sensors are robust against weather conditions and their performance does not decrease at night. In addition, since signal processing is relatively easy and effective for detecting a target, it is possible to effectively utilize radar sensors in various indoor and outdoor environments [2].

Compared with continuous-wave radar systems, ultra-wideband (UWB) radars have localization capability, consume less power, and can monitor multiple subjects [3]. Radar technology is aimed at detecting targets in aviation and military areas. Recently, studies have been conducted to detect human bodies at close range, and to detect heart rate and respiration. UWB impulse radars are used to obtain biological information due to their low risk of exposure to electromagnetic waves and low power consumption. It has become an emerging technology for indoor localization and tracking. UWB radar has many advantages, including high spatial resolution, ability to mitigate interference, through-wall visibility, simple transceiver, and low cost [4].

In this study, radar signal processing and machine learning are applied in a system to detect human presence and movement. In addition, the machine learning data set utilizing most contactless sensors is generated as a data set by data processing techniques of various statistical calculation methods. However, this study does not use a statistical calculation method, but converts the signal-processed result into images to generate a data set, machine learning, and shows the result. As a method of research, data is received from radar in chronological order, and features are extracted through signal processing. The extracted feature data is divided into six labels, and machine learning is used to determine the existence and movement of the current person through different experiments for each label. As a result of the research, a total of three experiments were conducted: changing the epoch sizes and changing the composition of the label. In addition, even though

\*Corresponding Author: Young-Jin Park, [yjpark@dgist.ac.kr](mailto:yjpark@dgist.ac.kr)

this is the result of research on data sets using fewer images, the result is satisfactory in the research of converting radar signals into images and detecting the presence or absence of subject and movement.

In Section 3, the machine learning process is discussed by extracting features with a dataset configured through the preprocessing of radar signals. Section 4 presents the experimental results using several test sets. Finally, the paper is concluded in Section 5.

## 2. Related Works

The most common method is to use image sensors to detect the presence or movement of a subject. However, there are scenarios when this method cannot be used, such as conditions involving personal privacy infringement. In order to overcome problems, the UWB impulse radar signal was used in this study. Machine learning algorithms is used to determine and classify subject presence and movement. The features for the machine learning model are extracted using signals from the UWB impulse radar module, as shown in Figure 1. The impulse radar module has one transmitter (Tx) and receiver (Rx). The Tx sends very narrow pulses, and the Rx receives the reflected pulses. The received signal passes through several signal processing steps to extract the target signal. However, this target signal is generally perturbed by clutter, noise, and attenuation. Therefore, the removal of unwanted signals and signal compensation are crucial tasks for improving the detectability of a target [5]. The UWB impulse radar emits short pulses through antennas, and a radar transceiver [6] that digitizes pulses returning from the target using sampling methods is used in the experiment. A pulse with a Gaussian envelope and sine wave less than 0.4 ns in width has a 6.8 GHz frequency and a 2.3 GHz bandwidth. The sampler collection read from the radar chip in particular is called the frame, and the delay between individual samplers yields the equivalent sampling rate of the frame [7].

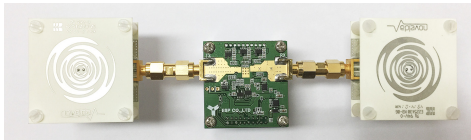


Figure 1: UWB impulse radar hardware module

In this study, the experiments receive frames sequentially from the impulse radar and the process of performing basic signal processing is the same. In the previous work [1], a feature set was created using the signal distribution by calculating standard deviation, root mean square (RMS), etc. However, in this study, the signal is processed in the frequency domain to generate an image, and the dataset used as input for machine learning is converted into an image. Furthermore, support vector machine (SVM) was used in the previous experiment to classify patterns. The major strength of SVM is that training is relatively easy and it has no local optimum, unlike neural networks. It scales relatively well to high-dimensional data, and the tradeoff between classifier complexity and error can be explicitly controlled. This weakness is mitigated by a good kernel function [8-13]. However, in this study, deep learning (GoogLeNet) was used for image classification rather than SVM. In 2014, Google published its network, GoogLeNet, to the imageNet large scale visual

recognition challenge. Its performance (6.7%) is slightly better than that of VGGNet (7.3%). The main attractive feature of GoogLeNet is that it runs very fast due to the introduction of a new concept called the inception module, thus reducing the number of parameters to only 5 million, which is 12 times less than that of AlexNet. It also uses lower memory and power [14].

## 3. Method

### 3.1. Preprocessing

Figure 2 shows the overall flowchart of the proposed algorithm. First, a raw signal from the radar is collected to the frameset. A feature set is then made by extracting the signal characteristics through signal processing. Next, a dataset is created by extracting the characteristics based on each action label—no one, in front of the radar, moving in front of the radar, and so on—and then converting them into images to be used in machine learning.

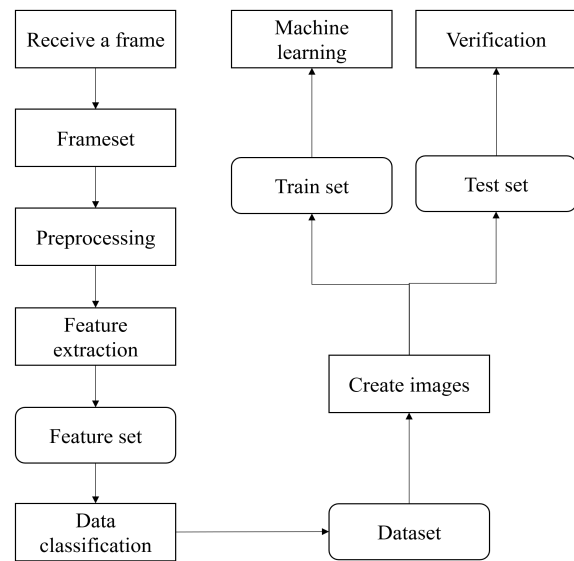


Figure 2: Flowchart

The created images are stored for each label and divided into train and test sets for machine learning and verification, respectively. The train and test sets are required when constructing models in machine learning. The training set is used to construct a model that gives a result that is close to the expected actual value, and the test set is used to check if the model constructed is reasonable. In other words, if a suitable prediction coefficient for the machine learning algorithm is found using the training set, the performance of the model can be verified with the test set. The results of the experiment suggest that the model can classify subject presence and movement using images created with radar signals.

In this study, one frame with 512 samplers is received every 50 ms and its appearance is shown in Figure 3(a). The frame set is a collection of frames accumulated in chronological order, and a single frame with a total of 48,000 frames (40 min.) used in this study is shown in Figure 3(b). For your information, three frames can be used for breathing and heart rate extraction, and can be used in various ways, such as identifying a person's path of travel.



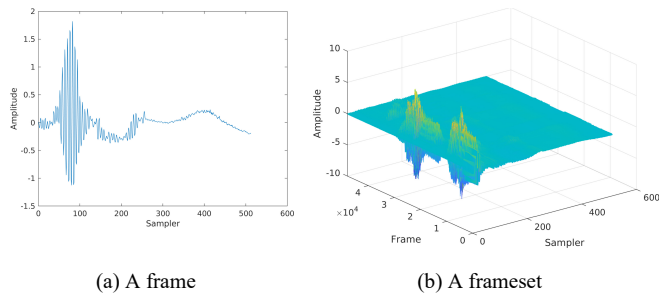


Figure 3: Chronologically ordered a feature set

Note that Figure 4 shows the feature extraction process. The raw signal is a frame that is received directly from the radar. Remove the background, etc. of the signal and store it in chronological order in the frameset. In this experiment, frameset is used to extract characteristic information of signals by means of signal processing, such as frequency analysis, digital filtering, etc., and moreover, because the subject's position is fixed, all areas of frameset are not used.

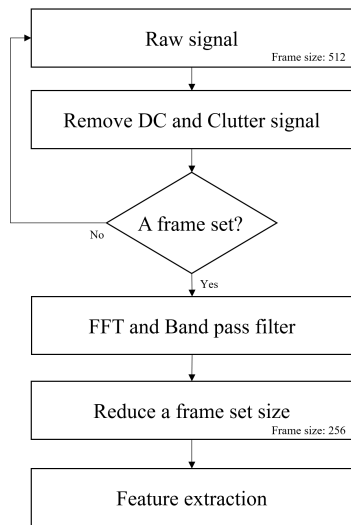


Figure 4: Chronologically ordered a feature set

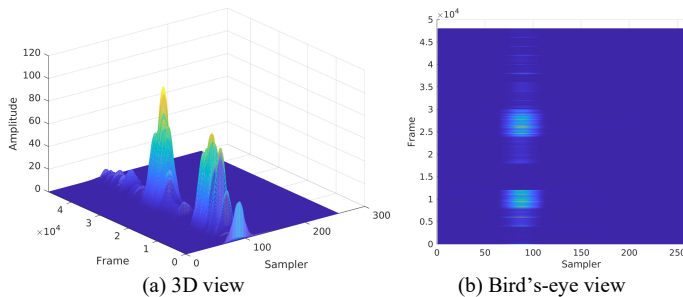


Figure 5: Chronologically ordered a feature set

The original radar signal contains noise and unstable signals, so accurate machine learning results cannot be obtained. Therefore, a feature set that can be used can be obtained only after the signal processing process. For reference, as shown in Figure 5, the feature set was used in half the size of a frameset that make up the frame, using only 256 samplers, or half of the total, to remove unnecessary data information. Figure 5(b) presents the strength of the overall signal when the signal is seen from a bird's eye view.

The signal strength appears in six forms, which are mapped to the six labels to be used in the experiment. The x-axis represents samplers within the dataset, and the y-axis represents collected each frame.

Each axis in Figures 5 represents the x-axis representing the sampler, the y-axis representing the frame set and the z-axis representing the amplitude of the signal.

### 3.2. Machine Learning

The total size of the frame set is 24,576,000 double types. In this experiment, the subject is always in a fixed position (within one meter) and is located approximately 70 to 120. A region of interest (ROI) was set to be used to reduce the size and create a data set in the Figure 6, using only the data in a specific area. Therefore, only those areas are set as areas of interest, and the rest of the data is not needed. The data set for creating images using feature sets uses only the area of interest, reducing the amount of data to be used.

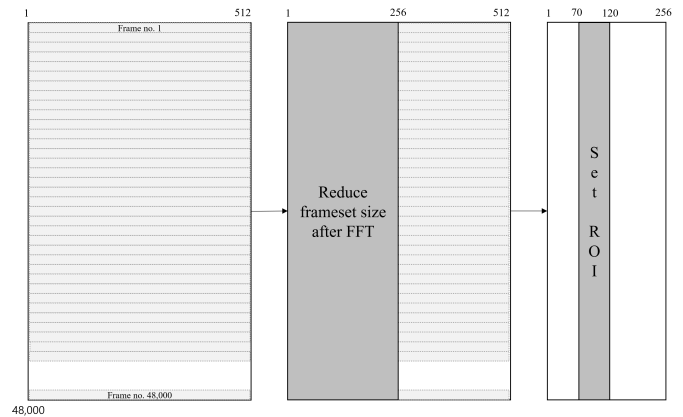


Figure 6: Reduce process of image size

In addition, the labels used to predict the condition of the subjects were divided into six, as shown in Table 1 below.

Table 1: Classification labels

Label	Amount of Data	Location of the subject
0	10,000	No one in front of the radar
1	10,000	Fix in front of the radar
2	10,000	Moving in front of the radar
3	6,000	Back
4	6,000	Left 30°
5	6,000	Right 30°

The dataset was created to generate images for machine learning by classifying the data of the same labels from scattered data within feature set. Due to the characteristics of radar signals, various behaviors (each label) are mixed, such that it was difficult to generate images for machine learning. Hence, a block consisting of data of the same label was created, and the size of the data for each label was compared, as shown in Figure 7. The x-axis represents six labels, and the y-axis represents the number of data within each label. In Figure 7, the 48,000-frame dataset has a complete configuration for each label, and six blocks are created with each label. This is because the data need to be classified under the same label to generate images for machine learning.

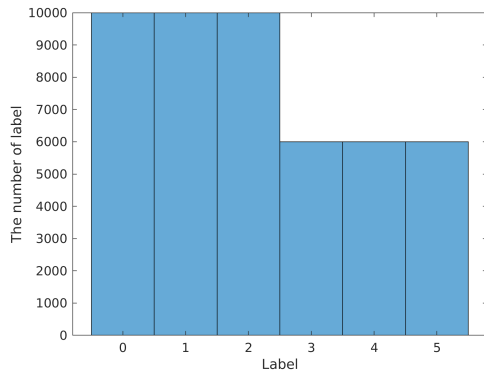


Figure 7: Composition of each label of a data set

Figure 8 shows the composition of the data shape for each label. The x-axis represents samplers of a specific area within the dataset with a corresponding label, and the y-axis represents the amplitude. This section describes the process of creating a dataset to generate an image required for machine learning. Particularly, the quality of the image depends on signal processing. This is the most important aspect because machine learning results vary depending on the quality of the image. The training set was used to implement the machine learning model, and the test set were used to test the model accuracy. If only the front part of the image, which represents the position of the actual subject, was converted into an image, a slightly faster performance is possible, but it is difficult to achieve noticeably better results. However, when experimenting with more images or with embedded equipment, it is necessary to reduce the size of the basic information of the images, even if they are damaged.

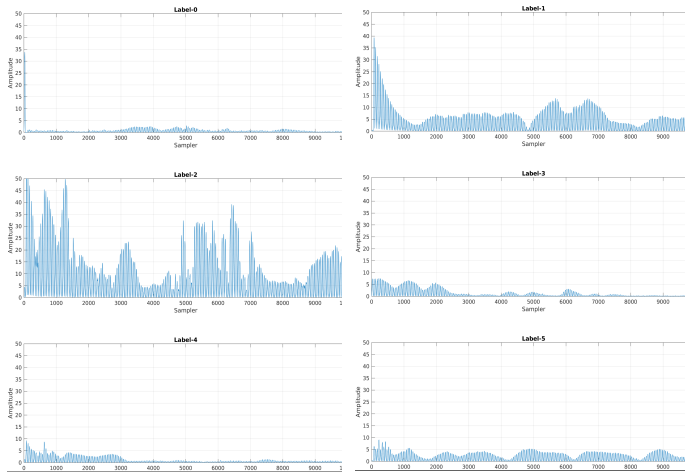


Figure 8: Reconfiguration dataset using labels

The experiments show that machine learning can determine the presence and movement of a person using images generated by radar signals. The current experimental setting and results are better than expected at the beginning of the experiment. However, if the situation changes, such as the distance and location of the radar, person, and different physical conditions of several subjects, the results can be significantly different. Therefore, it is necessary to collect more experimental data and to improve pretreatment algorithms to obtain cleaner images.

#### 4. Experiment

In the preprocessing stage, the radar signals were processed to produce a basic feature set. This is the most important process because machine learning results may vary depending on the feature set produced in the preprocessing stage. Next, we created a dataset from the feature set using classification labels and then created images for machine learning. The images generated were divided into a training set for model construction and test set for verification. The generated images were also used as input data for machine learning. The experiments conducted in this study showed a high level of classification accuracy. However, datasets pretreated only by signal classification seemed to be quite sensitive to the quality of the data when used as features for machine learning. The algorithm was implemented in hardware and software environments, as summarized in Table 2.

Table 2: System environment

Environment	Specification and Version
Processor	Intel i7-8700k
Memory	32 GB
GPU	Single Nvidia Titan Xp
Operating System	Ubuntu 18.04 LTS
Language	Matlab (2019b)
Radar module	NVA6201 (Novelda)

The algorithm proposed was used to generate an image using the dataset created in the previous step. The image was then stored separately by label. During the machine learning process, the images were loaded and learned, and the results were shown. The experiment uses frames received for 40 minutes using a total of 48,000 signals received every 50 milliseconds. The total number of images created using the proposed algorithm is 240 and the train set and test set are used in machine learning at a ratio of 8:2. The images for machine learning were converted into a spectrogram, which combines the waveform and spectrum features to visualize the wave, using the hamming window method. Furthermore, an RGB (Red, Green, Blue) image with an actual size of 224×224×3 was created and stored by label. The created image appeared in the form shown in Figure 12, and features obtained by signing the radar’s signal at the front of the image exist using spectrogram that is a visual tool of representing the signal length. Meanwhile, features that appear in front of the radar are required. However, the images were not randomly cut or resized for the experiment. When more than one person is present in future studies or multiple radars are used, new targets appear in the blank area as they are currently being signaled. Therefore, the original image was used because the accuracy could vary depending on whether signal processing was performed or not.

Table 3 summarize experimental environment and accuracy. For experiment- 1, it is a measure of how well six labels measured using radar can be distinguished. In experiment-2, label-4 and label-5 are experiments on subjects in fixed positions, such as Label 1. However, since only the angle of measurement is different, label-4 and label-5 can be integrated with label-1. In other words, it is an experiment that combines data set where similarities exist into a single family of labels. In experiment-3, the simplest experiment was conducted, and although the actual result

is 100%, the accuracy will be decreased as the number of images increases. Depending on the experimental conditions, it can be confirmed that the results of the three experiments vary depending on the composition of the label. As mentioned above, label-1, label-4, and label-5 in experiment-1, label-4, and label-5 in experiment-2 are datasets for subject who have different angles but are not moving. Therefore, in experiment-2, we used the image of having three labels on label-1.

Table 3: Experiment condition and accuracy

	Experiment-1	Experiment-2	Experiment-3
Network	GoogLeNet		
Image size	224×224×3		
Train set	192		
Test set	48		
Epochs	32		
Used label Types	0, 1, 2, 3, 4, 5	0, 1(include 4 and 5), 2, 3	0, 1, 2
Accuracy	91.7%	97.6%	100%

The Figure 9, Figure 10 and Figure 11 show a graph that monitors train and validation for machine learning, and the graph at the top represents accuracy, and the graph at the bottom represents loss. The x-axis represents iteration of machine learning, and the y-axis represents the accuracy (%) and loss rate. The black dotted line is the validation result for the actual train set. Figure 7 shows result of the first experiment with GoogLeNet using six labels.

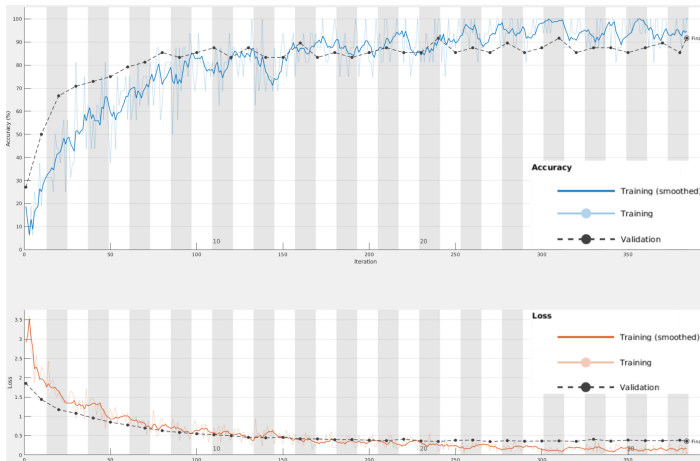


Figure 9: Accuracy and loss function for the experiment-1

Figure 10 shows the result of second experiment performed in a similar manner as the experiment-1. In the experiment, classification by label was slightly simplified. In the case of label-0 and label-2, no change was observed. For label-4 and label-5, radiated radar pulses from the left and right sides of the subject were integrated into the radiated radar pulses from the front of the subject (label-1). Label-3, which radiates radar signals from the back of the subject, including poor results from previous experiments, was excluded from the experiment and machine learning was conducted.

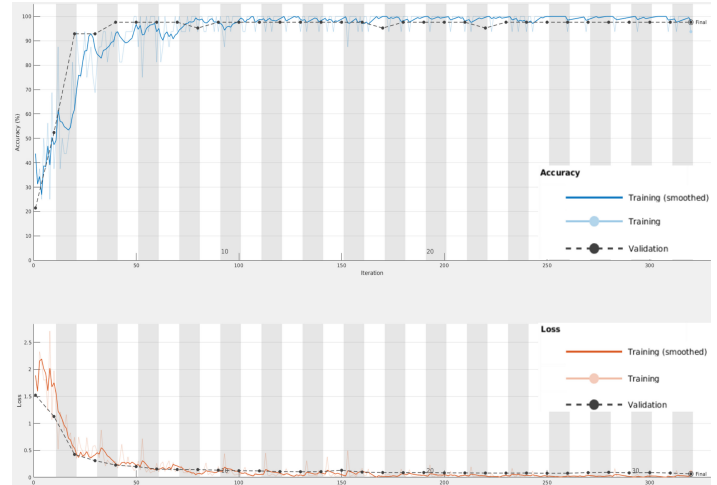


Figure 10: Accuracy and loss function for the experiment-2

Figure 11 shows the machine learning result using only label-0, label-1, and label-2. The results show that label-3 is difficult to recognize. Furthermore, label-4 and label-5 had high predicted rates when integrated into label-1, indicating that these three labels may have similar features.

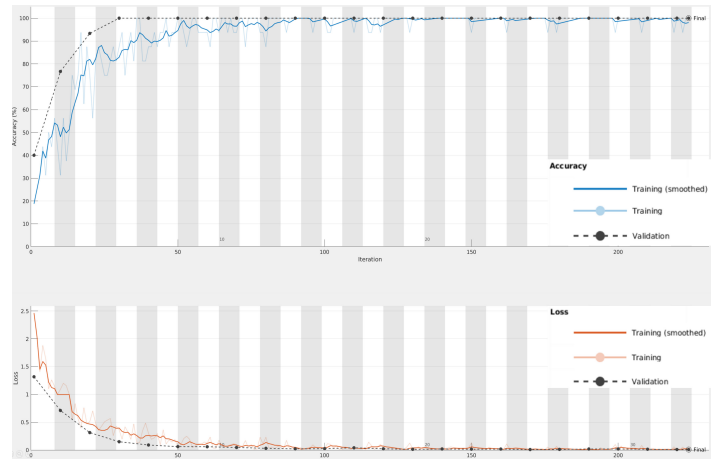


Figure 11: Accuracy and loss function for the experiment-3

Figure 12 shows the result of some validated images for the machine learning model, with high predictive results for label-0 indicating when no one is present and for label-1 and label-2, when a person are located in front of the radar. For label-3, the radar pulse was emitted by the back of the subject, but this was insufficient to judge the presence of a subject. Label 4 and label-5 is the data that is changed only from left to right under the same conditions as label 1, and high predictions can be seen. Information about each label can be found in Table 1. In fact, the use of images representing signal strength within frequencies conducted in this study alone cannot determine whether the subject is human. Therefore, in the future, research will be needed to process biometric signals such as respiration at the same time to determine if they are human or not.

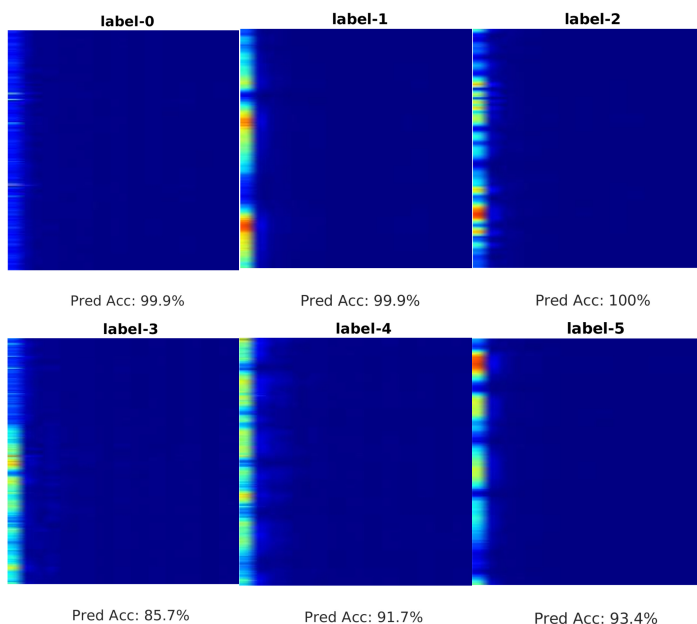


Figure 12: Result images created by the experiment-1

In addition, experiments using other neural networks like Resnet series and SqueezeNet were conducted. In fact, in the case of SqueezeNet, it has the advantage of learning speed, similar to GoogLeNet, and the results were similar. However, experiments using the Resnet series (Resnet18, Resnet50 and Resnet101) of neural networks with more network layers to improve accuracy showed that the results were rather lower at around 80 percent. The reason is thought to be the result of poor learning, passing through many layers because of the small number of images used in the experiment. In this experiment, a total of 240 images (train set: 192 test set: 48) are used. Although much more data was collected than was used in this experiment, many data were discarded due to errors and inaccuracies in the data. Therefore, it is not easy to filter out quality data by analyzing radar signals obtained by emitting them to living things. In order to have higher accuracy, it is necessary to have the integrity of the radar signal and to obtain at least five times the image used in this experiment.

## 5. Conclusion

Research and development using machine learning is ongoing worldwide, particularly in the prediction of subject activities. In this study, subject presence and movement were determined via machine learning, using an ultra-wideband impulse radar signal. Using a set of data imaged with electromagnetic signals in machine learning, experimentation using several neuronal networks resulted in high classification prediction results. An analysis of human-to-electromagnetic interactions shows that UWB impulse radars have low risk associated with human exposure. Therefore, in this paper, the UWB Impulse Radar signal was released to the human body and processed and imaged the signal. In addition, a new method was proposed to detect subject presence and motion by utilizing the created images as data sets of machine learning. Experiments have shown that subject presence and motion are sufficiently possible to be detected. Based on the results, it is judged that if the proposed method is further developed, it can be used in a variety of applications, including medical care and crime prevention.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

This work was supported by the DGIST R&D Program of the Ministry of Science, ICT, and Technology of Korea (20-IT-09).

## References

- [1] Y. J. Park, H. S. Cho, "An experiment of human presence and movement using impulse radar and machine learning", 2019 International Conference on Information and Communication Technology Convergence (ICTC), 2019, doi:10.1109/ICTC46691.2019.8939858.
- [2] J. Kwon, S. Lee, N. Kwak, "Human detection by deep neural networks recognizing micro-doppler signals of radar", in 15th European Radar Conference (EuRAD), 26-28, 2018, doi:10.23919/EuRAD.2018.8546605.
- [3] L. Ren, Y. S. Koo, Y. Wang, A. E. Fathy, "Noncontact heartbeat detection using UWB impulse Doppler radar", in Proceedings of the IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing System, 14-16, 2015, doi:10.1109/BIOWIRELESS.2015.7152128.
- [4] R. J. Fontana, "Recent system applications of short-pulse ultra-wideband (UWB) technology", IEEE Trans. Microw. Theory Tech., 52, 2087-2104, 2004, doi:10.1109/TMTT.2004.834186.
- [5] V. H. Nguyen, J. Y. Pyun, "Location detection and tracking of moving targets by a 2D IR-UWB radar system", Sensors 15(3):6740-62, 2015, doi:10.3390/s150306740.
- [6] Novelda, NVA620x Preliminary Datasheet, Novelda AS, Oslo, Norway, 2013.
- [7] Y. J. Park, H. S. Cho, "Detecting human movement during vital sign measurement using heterogeneous sensors", ITC-CSCC2019, 2019, doi:10.1109/ITC-CSCC.2019.8793461.
- [8] V. Jakkula, "Tutorial on Support Vector Machine (SVM)", School of EECS, Washington State University, Pullman 99164.
- [9] Tutorial slides by Andrew Moore. <http://www.cs.cmu.edu/~awm>
- [10] C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery, Kluwer Academic Publishers, 2, 121-167, Boston, 1998, doi: 10.1023/A:1009715923555.0.
- [11] N. Cristianini, J. Shawe-Taylor, "An Introduction to Support Vector Machines and other Kernel-based Learning Methods", Cambridge University Press, 2000, doi:10.1017/CBO9780511801389.
- [12] J. P. Lewis, Tutorial on SVM, CGIT Lab, USC, 2004.
- [13] Tutorial slides by Andrew Moore. <http://www.cs.cmu.edu/~awm>
- [14] M. Sewak, Md. R. Karim, P. Pujari, "Practical Convolutional Neural Networks - GoogLeNet architecture", 2018, <https://subscription.packtpub.com/book/data/9781788392303/4/ch04lvl1sec31/googlenet-architecture>.



## Ethics as a Motivation Indicator in Second Language Vocational Digital Teaching

Natalya Viktorovna Matveeva\*, Ludmila Vladimirovna Makar

The Linguodidactics Department, Institute of Humanities, Russian University of Transport (MIIT), 9, bld. 9, Obrazcova Street, Moscow, 127994, Russia

### ARTICLE INFO

Article history:

Received: 31 March, 2020

Accepted: 12 May, 2020

Online: 28 August, 2020

Keywords:

College Students

Computer Games Development

Ethical Teaching Skills

Foreign Language

Out-Of-Classroom Autonomous Activity

Teacher-Student Rapport

### ABSTRACT

*The non-selective second language course at vocational colleges and universities makes teachers strive at fostering students' motivation to learn by choosing from a variety of enhancing factors. Teacher's personality and skills if they comply with pedagogical ethics are considered to be inspiring for students to learn. The aim of this piloting study was to collect empirical data and ascertain the actual motivational power of ethical teaching skills when organizing and conducting autonomous learning activities for college students on the example of educational computer games.*

*The ethical teaching skills applied when creating and using computer games as out-of-classroom voluntary activities were conceptualized. Their possible influence on learning motivation of college students was analysed with the use of oral and written feedback procedures, and statistical data retrieved from Metrika.Yandex and students' certificates submitted.*

*This allowed to determine five specific groups of students valuing the activity, which in their turn were united into two larger formations: students with distinct intrinsic motivation in the activities-related areas (60%) and socially and psychologically responsive students (40%). A possible method to evaluate statistically the efficiency of some ethical skills specified is suggested, whereas the other skills need further research under modified parameters.*

*Thus, ethical teaching skills are effective in the case the college students possess an intrinsic motivation in an area actualized by an activity suggested. Therefore, they are unlikely to be regarded as an independent motivator, but the indicator of a particular intrinsic motivation characteristic to a student.*

### 1. Introduction

This paper is an extension of work originally presented in *Proceedings 2019 International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS)*. It stated that there existed the gap between college and university syllabus in the number of professionally oriented themes (not in favour of college students); individual words learnt ( $\approx 800$  at college vs 600 at university); and in learning methods. In addition, it has been revealed that college students underestimate the role of searching information online, critical thinking, making quick decisions and some other basic competences and skills as essential for their future career. The language computer games to overcome the differences were

developed in Adobe Animate, and their plot and the interface design were described [1].

The piloting research was initiated at Serpukhov college (Moscow Region, Russia). The college is comprehensive with approximately 1400 students from the south of Moscow region and adjacent areas of Tula and Kaluga regions mastering a wide range of specialities. All of them were invited to take part in the games. However, due to the geographically dispersed location of the college faculties, for the purpose of adequate data collection in part of classroom discussions we managed to keep in close touch only with the students of engineering specialities who study in one of the premises (450). This is why we consider this group of students the core of our research and the research itself piloting.

The first invitation to bridge the gap through the games was placed on the college official Internet-site at the beginning of the academic year. It offered a brief thematic review of the games and

\*Corresponding Author: Natalya Matveeva, Russian University of Transport (MIIT), +79262460543, [matveevanv@mail.ru](mailto:matveevanv@mail.ru)

description of test conditions. Students were informed that we would be glad if they shared the results with us sending screenshots of their game certificates to the contact e-mail. We also organized a public page in a social network in case the students wanted to discuss gameplay. The English teachers of Serpuukhov college duplicated information in classes. Even a month later there was no great activity surge, and we found out that this new type of learning activity was not welcome by students with their arms outstretched.

The question followed: as information technology itself and the intriguing form of presenting educational materials such as educational computer game couldn't be the dominating motives for college students in their independent work in learning English as a second language, a teacher should ethically motivate them to study.

Ethics in teaching profession attracts attention of the professional community. The quality of a teacher is directly connected to setting a good moral example for students and the skills of acting in accord with professional ethical standards [2, 3]. In Russia researchers challenge the task of developing pedagogical ethics in order to make the educational process more effective [4]. A need for continuous teacher training to develop ethical competency is highly emphasized [5 – 7]. The teacher's ethical competency is thought to comprise, firstly, the knowledge of ethics and ethical aspects in teaching, secondly, skills of pedagogical interaction and communication, thirdly, the ability to implement the knowledge of ethics in specific pedagogical situations. Among ethically sound personal qualities in teaching, tolerance, poise, goodwill, caring attitude, and the ability to empathize appear to be the most representative ones [8]. Teaching ethically in a digital environment is of great importance as well. Exactly this aspect of teaching ethics is in the focus of our research.

## 2. Problem Statement

It has been found out that the teacher's personality can have a major long-term impact on students' motivation to learn a second language, including autonomous, or as Lamb & Wedell said, "in their own time, of their own volition beyond the classroom" [9] learning. Melyokhina & Bobohodzhaeva highlighted that learners are particularly inspired by teachers who are content competent and able to build good teacher-student rapport, demonstrate academic literacy and interest in teaching [10]. Thus, the issue of teaching ethics as a factor motivating students to study is becoming a research agenda.

In college and university educational context, where the second language is a mandatory, non-elective job-oriented course and teachers have to face mixed ability and mixed motivation classes, the issue of motivating students is of the utmost importance. A great number of motivators having been scrutinized by the teaching community worldwide [11], still potentially effective ones must be revealed and investigated.

## 3. Research Questions

Without minimizing the significance of ethical knowledge and ethically sound personal qualities of a teacher in developing students' learning motivation, we tend to attribute a greater deal of success in this field to the practical skills of applying these two in

such fundamental pedagogical situations as 1) designing and implementing language learning courses, classes, materials and activities and 2) building relationships with students. As to the designing tasks, it is generally accepted that informing learners about the aims, providing them with information meaningful to them, using facilitating instructional techniques to perform and achieve goals, etc. is ethical. When building rapport with students, it is ethical to respect their rights, dignity and freedom, their autonomy and privacy, to avoid conflicts, provide choice, timely and objective assessment of works, give informational feedback.

Using this line of thinking, we raised the question as to whether these teaching skills with strongly felt ethical background can play into the college teachers' hands to engage the learners into autonomous training aimed to compensate an existing gap between college and university syllabus in the number of 1) professionally oriented themes (not in favour of college students); 2) individual words learnt ( $\approx 800$  at college vs 600 at university); 3) in learning methods. In addition, it has been revealed that college students underestimate the role of searching information online, critical thinking, making quick decisions and some other basic competences and skills as essential for their future career [12].

## 4. Purpose of the Study

The objective of this research is to collect empirical data and ascertain the degree of motivational power of ethical teaching skills in conducting out-of-classroom learning activities at college. For the purpose, we 1) conceptualize the ethical teaching skills of primary importance when creating and using an autonomous learning activity in the form of a computer game; 2) analyse their possible influence on second language learning motivation of college students.

## 5. Research Methods

Below we outline two sets of ethical teaching skills corresponding to the pedagogical situations mentioned which are of high value in digital environment.

### 5.1. Ethical Skills Important for Designing and Implementing Learning Materials and Activities

*Using extra-curricular interests, skills and competences.* When designing materials and activities for autonomous language learning, it is essential to provide for bringing different interests and skills students may have a good command of, these may be drawing, love for sport or IT. While choosing the ones we could rely upon, what we were quite sure of is that the absolute majority of our students are quite competent in computers and digital technologies. The rich virtual environment offers what young people appreciate: simulations, games, digital objects, communication tools. It is natural then to choose students' IT competence for designing autonomous English learning activities.

*Providing information meaningful for students.* Due to the characteristics of their age group college students are not inclined to work for the long-term perspective. Their life and professional trajectory are far from being clear and well defined. In their free time, they are more fascinated with "having fun", which in fact means communicating with friends and trying new or pleasurable things to do.

Our games are quite in this vein. The topics for the games “The Era of Great Geographical Discoveries: Meeting of Civilizations” and “Arctic Convoys of the Second World War” concern the watershed periods in human history. Being new, they suggest a challenge to be engaged in indirect communication with antecedent generations and civilizations through conceptual thinking and problem solving, acquire and share knowledge of them. Moreover, remembering Macmurray [13], who pleaded for online educational systems based on “learning to be human”, we can say the games implement the humanistic approach based on human perception, communication and interaction, which are also the key to the field of learning languages.

*Instructional context* providing, following Walters’ words [14], “choice, optimal challenge, informational feedback, interpersonal involvement, and acknowledgement of feelings” can play a great part in motivating L2 learners. The skill to provide the instruction of this kind is an ethical challenge for a teacher.

We provide a full-range instruction set on the splash screen of each game. It briefly describes the theme, plot and gameplay, states the main aim of each game, informs about helpful Internet resources and the systems of controls for successful playing. The instruction gives a clear rationale that allows participants to fully grasp the specific meaning and importance of the activity.

*Utility value* of the games designed proceeds from the research task to bridge the gap between college and university foreign language curricula. The themes of the games raise issues on sociology, psychology, culture and religion, communication, economic systems, which are not included in the syllabus at colleges, and also concern civilized values in general – war and peace, independence, non-intervention, friendship, cooperation, poverty and prosperity. Besides language and learning skills, the games promote college students’ digital literacy, life, research and professional skills such as navigating, interacting, searching information online, quick reading and making decisions. All kinds of skills gained will benefit college graduates both if they decide to continue education at university and if they go straight to work.

## 5.2. Ethical Skills Important for Teacher-Student Relationship

*Respect for autonomy.* Autonomy, according to Noels [15] and some other researcher such as Lozovoy & Zashchitina [16], refers to “a sense that one can freely choose to engage in activities that they find personally meaningful. Teacher’s autonomy enhancing behavior includes fostering relevance, allowing criticism and providing choice [17]. We followed the advice of Jang, Reeve & Deci [18] to use informational language that provides options for students and identifies the potential meaningfulness of the activity suggested. We highlighted that the games comprised professionally oriented tasks, and the knowledge and skills students had acquired during their studies at college would help them as players to survive in the game worlds currently and as specialists in the future. The tasks were specifically designed for players to form good generalizations about what will work well when they face more complex professional and life challenges.

*Commitment to success* means a student’s willingness to engage in and persist at a task (Wolters, 2003). In order to sustain engagement and persistence in the games we foresaw the possibility for the participants to play through the games again if

the score does not satisfy them, simultaneously drilling the words and memorising the answers thus climbing to game and academic success.

*Feedback.* In-class short discussions were organized to see how much interest the games provoked, find out the difficulties and drawbacks, and promote learners’ sense of competence. Students were motivated to express their opinions about the games. Some reactions just needed explanations or encouragement from the teachers, while others resulted in a forthcoming hot round-table discussion on the topics the games offered. Some of the students demonstrated involvement, immediacy and interest. They provided help to those who experienced difficulties whether IT or English course based. The feedback then was threefold: from students to the teachers, from the teachers to the students, and between students.

*Evaluation* brings up an issue of rewards. To be effective, rewards need to be valued by their potential recipients and timely administered. Making voluntary participation and confidentiality the corner-stones of the activity, we avoided the necessity of extrinsic rewards, which have been found to interfere with the process and quality of learning (Reeve, 2005). Two kinds of evaluation were ensured, namely, direct individual clear-eyed assessment and indirect, implied self-evaluation. The former was provided by scoring and timing programmed for the whole game. The score and time spent in a game are shown in the last frame, which is also the learner’s certificate. The latter was supposed to arouse during feedback in-class sessions. Although these sessions pursued a different goal (see above), it is difficult to argue the possible outbreak of latent self-assessment thoughts as a result of comparing one’s own actions with those of groupmates. In both cases, students did not have to share their results with anyone keeping them private if timid or unsatisfied.

*Confidentiality.* Some students felt comfortable and were not shy to share their scores without being forced. Nevertheless, we strictly followed this principle regarding those students who showed no inclination to discuss anything related to the activity. We did not experience personal data collection, we did not pursue students who refused to take part in the activities, just we were happy for those who had passed the games and shared the score with us.

## 6. Procedure

We were allowed to deliver information about the computer games developed to fill the gap in learning functional lexis during the buzz session with Sergey Chekmaev, the Russian science-fiction writer and a participant of literary adaptation of numerous computer games, *TimeZero*, *Destiny*, *StarQuake* being some of them. The audience of the session, as we saw it, was the students who were “computer game players” and we hoped exactly these ones were the vanguard open to innovations.

Students could launch the games any time they wanted during five-month period starting from the first announcement.

Short classroom feedback took place once a fortnight at the lessons, their duration depending on the number of questions on procedural and rule matters the students had.

The issues raised by the topics of the games were discussed during final round table talks when the students discussed and argued the topics-related issues. During these talks students' L2 and human progress was reviewed.

After this debriefing the site was not closed and had more visitors in the spring and during the summer holidays.

Having conceptualised ethical issues of primary importance when creating and using educational computer games as learning activities we monitored their possible influence on students' L2 learning motivation by means of quantitative and qualitative methods.

To do statistics we processed data from 1) Metrika.Yandex, which is the analytical widget installed on the site to count unique IPs and the number of visits, to fix the non-personalised data regarding the traffic, geography, age of the visitors and time spent on the site available and 2) the students' certificates submitted. Only results proved by certificates sent by emails or VKontakte were subject to the analysis. The certificates shown but not sent were neglected.

The qualitative analysis is based on the students' questions and comments expressed orally during classroom feedback sessions or committed in writing (mostly in English!) via e-mail, VKontakte or directly to the teachers and sorting them out in accordance with the ethical aspects listed above. During these sessions the students could express their main motives why they played the games explicitly, and some of them did. This gave us an opportunity to disclose the motivation these students had. Data retrieved are presented in the tables below.

**7. Findings**

This section presents data retrieved through the investigation: the total number of unique IPs and the total number of visits; students' scores.

Table 1: Statistical data retrieved from Metrika.Yandex

The number of unique IPs	The number of views
275	692

Table 2: Non-personalized students' scores

Certificate	Score, %	Motivation	Time spent online, min.
1	92	Positive attitude to life	6
2	90	Computer game player	1
3	88	Computer game player	1
4	80	Entering a university	-
5	80	Computer game player	16
6	80	Investigation for cheating	-
7	77	Computer game player	-
8	74	Computer game player	1
9	71	Interest to the topics	9
10	68	Interest to the topics	5
11	67	Inclination to success	9
12	61	Computer game player	17

13	60	High interest to learning English	17
14	51	Social activist	5
15	50	General interest to IT as a speciality	12
16	50	High interest to learning English	19
17	49	Computer game player	9
18	48	Computer game player	12
19	48	General interest to IT as a speciality	20
20	47	Investigation for cheating	20
21	47	General interest to IT as a speciality	9
22	44	Curiosity	4
23	40	General interest to IT as a speciality	16
24	31	Interest to the topics	-
25	28	Interest to the topics	-
26	28	Social activist	15
27	26	Interest to the topics	6
28	26	General interest to IT as a speciality	2
29	26	Computer game player	6
30	25	Curiosity	-
31	23	Computer game player	6
32	23	Interest to the topics	5
33	20	General interest to IT as a speciality	4
34	20	Knock-on effect	-
35	16	Social activist	9

Table 3. Students' Questions and Comments Grouped by Ethical Skills Involved

No	Ethical Skill	Students' question and comments
<i>Ethical Skills Important for Designing and Implementing Learning Materials and Activities</i>		
1	Using extra-curriculum (in this research: IT) interests, skills and competences	How should I activate the game? Can I play the game using a mobile phone? What should I do to get rid of advertising banners on the site? What should I do to start playing? I type in the correct word but the score doesn't improve.
2	Providing information meaningful for students	Do you think a peaceful meeting of civilization possible? Is it possible to help the Arctic convoy survive in a game?
3	Instructional context	How to pass the game through? What is it necessary to do there? Where should I



		get information on the topic? Should I know everything included into the game?
4	Utility value	Do we help you (a teacher) in anything? Do you (a teacher) want us to play? Are you (a teacher) preparing a report?
<i>Ethical Skills Important for Teacher-Student Relationship</i>		
5	Respect for autonomy	The control system is too complicated. I didn't manage to catch the point. Is it a compulsory or voluntary activity? No desire to be measured again.
6	Commitment to success	I've passed through the game several times (12) and didn't get a score I can share.
7	Evaluation	Is it possible to cheat?
8	Confidentiality	Have you passed through the game, my groupmate? Do you, teacher, need help?

## 8. Discussion

Below we suggest our interpretation of the statistical data and students' reactions to the activities received with a glance to their personalities and conclusions we have come to.

*Using extra-curriculum interests, skills and competences.* Initially we supposed the students' high IT competence in general and in computer games particularly. However, their questions (table 3) displayed that some IT areas were underdeveloped, among them 1) operating some software programmes, namely, Adobe Flash Player; 2) understanding that the Internet is an unrestricted advertising media; 3) unskillful situational manipulating with a keyboard and a mouse.

*Providing information meaningful for students.* The games content proved to be topical. The possibility of a peaceful meeting of civilization caused a really hot classroom discussion as well as the alternative for the Arctic convoy PQ-17 to do the 'the Russian run' and survive. The students' willingness to put their opinions on the lessons the humankind learnt made us certain that we had chosen information which was of interest to our students provoking critical thinking and arguing the point of view.

Questions dealing with *instructional context* testify to its particular importance and indicate the necessity of its provision more than once and in different forms. In addition, the technological difficulties students experienced with the software, advertising, etc. made us think of the necessity to cover them in a written form, too, thus providing sufficient guidance and support throughout the entire process of participation.

The in-lead about the *utility value* of the games brought the students to speculations about their future jobs and responsibilities, the difference between college and university educational routes and subsequent careers being clearly understood. In their turn, students showed their sincere interest to the utility value of the games for us, teachers. They tried to discover if we had any personal motivation for the work done other than declared. Here we felt the interdependence of teaching and learning motivations, the so-called knock-on effect described in a number of research [19, 20].

As we have repeatedly stressed, we did not press the students to play, it's them who decided. This *respect for autonomy* was a novelty for them used to compulsory activities. Students returned to this aspect, expressing no desire to be measured and assessed again. This is quite in line with Huang, Han, & Schnapp's conclusions [21] that state standardised testing regimes cause a strong concern. Assured that the activity was voluntary, some of such students participated, not sharing the score though.

*Commitment to success* ranged from very high (some students passed through the game up to twelve times) to extremely low when the students refused to finish the games alleging technical difficulties, for example, complicated control system. It is interesting to note that it took only one minute to make two of the best scores (90% and 88%), while the winner (92%) spent 6 minutes online (table 2). The winners evidently had passed through the game several times, memorizing the answers and drilling the words. The six of the worst results (20 - 26%) were also received with little time spent in a game (2 - 6 minutes), but this possibly shows they passed through the test formally just to see what it looked like.

Some students were concerned with possible cheating, which was quite unexpected, as we planned these activities to be self-regulated. Such remarks undeliberately reminded us that being placed on the web, all teaching materials should be treated as objects of protection [22]. We fixed the vulnerability, gave praise for valuable help, and thanked for passing through the game with a good (probably winded up) score (certificates 6, 20) and doing additional work of investigation. Now the score a learner gets is always correct and can be relied on in *evaluation* procedures.

Teachers strictly followed the principle of *confidentiality* and did not divulge the scores or any personal game information, but students asked each other directly about passing through the game. They were happy to see the like-minded persons, provided help to those who asked for, withholding comments in respect to those students who did not play or did not want to share the score or discuss anything regarding the game.

## 9. Conclusion

1. Students' independent activity before we started to apply ethical principles in stimulating them to play our games was exceptionally low. It gave us the understanding that neither information technology be itself or educational computer games as an intriguing format of presenting the education materials motivated students to strive for the best results in learning English as a second language. Teachers should ethically motivate students to learn English autonomously.

2. While ethically motivating students we retrieved statistical data of high participation of college students of Engineering (275 out of 450) – more than 50 % of enrollees, even higher than we expected when starting.

3. The games stimulated active behavior from the specific groups of students, presented in Table 4 below in accordance with their strength.

Table 4. Specific groups of students

№	Groups of students	Number	Percentage
1.	students interested in IT and computer games including those two investigating for vulnerabilities	18	51.4 %
2.	students driven by personality underlying motives	8	22.8 %
3.	interested in the topics	6	17 %
4.	with positive attitude to the language course	2	5.8 %
5.	planning to study at university	1	2.9 %

4. Since the activity was voluntary, it was evidently valued by the students with distinct intrinsic motivation in our games-related areas – IT, language, entering university (groups 1, 4, 5 – 60%) and students energized with their personalities (groups 2, 3 – 40%). Hence, the ethical principles are not likely to possess the motivational power by themselves but provide support in developing motivated learners, on the one hand, and on the other hand, provoke an interest from socially and psychologically responsive students contributing to their human and professional growth. And from this standpoint the student’s reaction to them can serve as a kind of indicator whether they possess any sort of intrinsic motivation or do not.

5. The impetus with the least driving power appears to be orientation to higher education – only one four-year student openly declared his intention to continue education at university. The others though keeping this possibility in mind are not so definite about their professional plans for the future. This is quite in line with higher educational establishment enrolment quality monitoring data, which highlight that the amount of college graduates annually entering Russian universities is only 6-10% of the total number of enrollees.

6. The ethical skills of *respect for autonomy* and *confidentiality* provide psychological comfort to learners and cannot be measured. Still, the influence of the others on students’ motivation can roughly be evaluated. Thus, the statistics and efficiency of *using extra-curriculum interests, skills and competences* may climb to 100 % covering all the participants, *providing information meaningful for students* – 17 % embracing students fascinated by the topics. *Utility value* can be estimated on the basis of such drivers as entering a university, interest to IT as a speciality, knock-on effect giving 20-25%.

7. The research also proves that students do need teachers’ personal assessments and interest in their scores and in their personalities in general. Communication with others and possibility to be a member of a community motivate students much more than information technology itself and even educational games as a means of delivering knowledge.

### Conflict of Interest

The authors declare no conflict of interest.

### Acknowledgment

The authors thank the staff of the Chair of Linguodidactics of the Russian University of Transport for their support in their work.

Great thanks to the students of Serpukhov college (Moscow Region, Russia) who gave their feedback, asked questions and kindly shared their scores.

### References

- [1] N.V. Matveeva, “Computer Games to Fill the Gap in Learning Functional Lexis at Russian Colleges and Universities” Proceedings of 2019 International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS), 639-643. 2019. <http://doi.org/10.1109/ITQMIS.2019.8928440>.
- [2] R.R. Puhan, L. Malla, S.K. Behera, “Current Ethical Issues in Teacher Education: A Critical Analysis on Pre-Service and In-service Emerging Teachers”. American Journal of Educational Research, 2(12A), 1-7, 2014. <http://doi.org/10.12691/education-2-12A-1>.
- [3] M. Sultana, “Ethics in Teaching Profession” ABC Journal of Advanced Research, 3(1), 44-50. 2014. URL: <http://oaji.net/articles/2015/814-1431431772.pdf>
- [4] O.S. Tsaregorodtseva, “Razvitiye pedagogicheskoi etiki uchitelya kak faktor povysheniya effektivnosti obrazovatel'nogo protsesssa v shole”. Ph.D Thesis, Murmansk, 2009. (In Russian).
- [5] V.M. Grebennikova, V.V. Doroshenko “Ethics of Teaching in the Context of New Educational Paradigm”, Modern Problems of Science and Education, 1(1), 1056, 2015. (In Russian).
- [6] I.M. Shadrina, “Model of Building the Moral Culture of a Student – Future Teacher”, The Volga district pedagogical bulletin, 6(3), 117-125, 2018. (In Russian).
- [7] N.E. Skripova, “The professional ethics for teachers: perspective directions of the professional development system”, Professional development system scientific bases, 2(23), pp. 28-35, 2015. (In Russian).
- [8] E.V. Pyaterikova, “Pedagogical Foundations of Ethical Competence for Future Teachers”, Nizhny Novgorod State Linguistic University Bulletin. Issue 21. Intercultural communication and teaching foreign languages. Challenges in Education, 181-188, 2013. (In Russian).
- [9] M.V. Lamb & M. Wedell, “Cultural contrasts and commonalities in inspiring language teaching”, Language Teaching Research, 19(2), 207–224, 2015. <http://dx.doi.org/10.1177/1362168814541716>.
- [10] E.A. Melyokhina, Ya.R. Bobohodzaeva, “Motivation to Learning a Foreign Language: Pedagogical Aspect”, Teacher in Higher Education in XXI century, 1, 89-93, 2016. (In Russian)
- [11] M. Lamb, “The motivational dimension of language teaching”, Language Teaching, 50(3), 301-346, 2017. <http://doi.org/10.1017/S0261444817000088>.
- [12] N.V. Matveeva, “Professionally Oriented Vocabulary in English Classes is Truly Fascinating”, Secondary Vocational Education, 12, 10-15, 2018. (In Russian)
- [13] J. Macmurray, “Learning to be Human”, Oxford Review of Education, 38, 2012 - Issue 6, 2012: Special issue: Learning to be human: the educational legacy of John Macmurray <https://doi.org/10.1080/03054985.2012.745958>
- [14] C.A. Wolters, “Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning”, Educational Psychologist, 38(4), 189–205, 2003.
- [15] K.A. Noels, “Learning Japanese, learning English: Promoting motivation through autonomy, competence and relatedness”, In M. Apple, D. Da Silva & T. Fellner (Eds.), Language Learning Motivation in Japan, 15–34, 2013. <http://doi.org/10.21832/9781783090518-004>

- [16] A.Y. Lozovoy, E.K. Zashchitina, "Online Education: Pros and Cons", Proceedings of 2019 International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS), 631–633, 2019. <http://doi.org/10.1109/ITQMIS.2019.8928455>.
- [17] A. Assor, H. Kaplan, and C. Roth, "Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing teaching behaviors predict students' engagement in schoolwork", *British Journal of Educational Psychology*, **72**, 261-278, 2002.
- [18] H. Jang, J. Reeve, & E.L. Deci, "Engaging students in learning activities: It is not autonomy-support or structure but autonomy-support and structure", *Journal of Educational Psychology*, **102**, 588-600, 2010.
- [19] S. Aydin, "Factors causing demotivation in EFL teaching process: A case study", *Qualitative Report*, **17**(51), 1–13, 2012.
- [20] M. Bernaus, A. Wilson & R. C., "Gardner Teachers' motivation, classroom strategy use, students' motivation and second language achievement", *Porta Linguarum*, **12**, 25–36, 2009.
- [21] J. Huang, T. Han, & K. Schnapp, "Do High-Stakes Test Really Address English Learners' Learning Needs? – A Discussion of Issues, Concerns, and Implications", *International Journal of Learning and Development*, **2**(1), 499-508, 2012.
- [22] V.V. Zolotarev, M.N. Zhukova, "Role Model Features in Educational Serious Games", Proceedings of 2019 International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS), 583–586. 2019. <http://doi.org/10.1109/ITQMIS.2019.8928292>

## The Effects of Transmission Power and Modulation Schemes on the Performance of WBANs in on-Body Medical Applications

Marwa Boumaiz<sup>\*1</sup>, Mohammed El Ghazi<sup>1</sup>, Mohammed Fattah<sup>2</sup>, Anas Bouayad<sup>1</sup>, Moulhime El bekkali<sup>1</sup>

<sup>1</sup>LIASSE-ENSA, Sidi Mohammed Ben Abdellah University, Fes, 30050, Morocco

<sup>2</sup>EST, Moulay Ismail University, Meknes, 50050, Morocco

### ARTICLE INFO

Article history:

Received: 13 July, 2020

Accepted: 02 August, 2020

Online: 28 August, 2020

Keywords:

Wireless Body Area Networks

IEEE 802.15.6

On-body medical applications

Packet Loss Rate

Throughput

Latency

### ABSTRACT

Wireless Body Area Networks support the operation within multiple frequency bands. Thus, they can be integrated in several applications, one of which is on-body medical monitoring applications, as concerned in this paper. Therefore, the purpose of this study is to present the impact of transmission power and both of Differential Binary Phase Shift Keying and Differential Quadrature Phase Shift Keying modulation schemes, on the performance of a WBAN model based on the IEEE 802.15.6 2.4 GHz narrow-band, dedicated to on-body medical applications. This involves identifying the modulation scheme(s) and transmission power level(s) to be adopted for these applications, that can be classified into three types depending on their data rate (low, medium and high data rate medical applications), in order to meet Packet Loss Rate and latency requirements. The numerical study has confirmed that the adoption of DBPSK modulation and low transmission powers provides good performance for low data rate monitoring applications. At medium data rates, a relatively increased transmit power was required. However, at high data rates, DQPSK modulation with a 0 dBm transmission power seemed to be the right choice to be made in terms of the mentioned performance indicators.

### 1. Introduction

The actual improved life expectancy all over the world has led to an increased rate of population ageing [1], which may overburden conventional healthcare infrastructures and consequently increase healthcare costs. In addition, each year millions of people die due to late diagnosis of certain fatal diseases such as cancer, cardiovascular diseases, Parkinson's disease and many others [1]. These losses could have been largely avoided if those illnesses were detected in time.

The massive growth of electronic systems [2] resulted in the appearance of small biosensors operating at the human body scale as Wireless Body Area Networks. When used in healthcare systems, WBANs can simplify the monitoring of simple parameters such as body temperature, blood pressure, ECG, etc., or be integrated in more complicated processes like changing programs for implantable pacemakers and defibrillators [3],

retrieving biokinetic information [4–6], or adjusting body limbs movements when damaged by car accidents, for example.

In medical applications [7], WBANs are considered to be the key solution to prevent a variety of cardiovascular diseases [8–9], such as myocardial infarction, which are often related to intermittent rather than permanent anomalies [10]. In addition, as the number of diabetic people worldwide is expected to reach 380 million by 2025, WBANs can also be adopted in diabetes monitoring [11–12]. Therefore, it will be necessary to integrate medical monitoring systems based on WBANs in our day to day life, in order to correctly and timely dose medicines, and thus reduce the risk of many complications. Similarly, miniaturized sensors can further be used to detect cancer cells, allowing the doctor to first diagnose tumors without the need for a biopsy [10]. As for people suffering from asthma, WBANs may also ensure the detection of allergenic agents in the environment surrounding the patient. In [13] for example, The authors proposed an alert system triggered by the detection of an environment the patient is allergic

\*Corresponding Author: Marwa Boumaiz, [marwaboumaiz@gmail.com](mailto:marwaboumaiz@gmail.com)



to. Moreover, they can be integrated into the rehabilitation process to bring patients' functional capacity back to normal [5].

For all the above mentioned medical applications, data traffic can be classified into three categories [10]: on-demand, urgent and normal traffic. For diagnostic purposes, on-demand traffic is initiated by the coordinator in order to obtain information about the patient's state of health. As for urgent traffic, it is initiated by sensor nodes when the data collected exceeds a predefined threshold. This type of data is not time-dependent and is generated in an unpredictable manner. As for normal traffic, it generally concerns normal non-critical data, that include discrete and systematic health monitoring.

In non-medical applications, WBANs take place in fitness, performance and wellness monitoring, in cognitive biometry [14] for a secure authentication to information systems, in serious gaming [15] for educational purposes, and in sport.

Given this variety of applications, the actual wireless short range technologies do not fully comply with the specific requirements and technical challenges of WBANs [16]. For this reason, IEEE 802.15.6 was created in November 2007 to deliver an international standard for highly reliable, short-range, and bodywide communications, while providing a broad range of data rates going from 75.9 Kbps (narrowband) to 15.6 Mbps (ultra-broadband); to cover a variety of applications. The standard provides an advanced MAC layer that serves three physical layers: The Narrowband (NB), the Ultra Wide Band (UWB) and the Human Body Communication (HBC) physical layers. These are chosen according to the intended application. We are interested in the narrowband physical layer, which alone offers 7 different frequency bands, one of which is the 2.4 GHz band (2400-2483.5 MHz) and is often favored over the other ones for its worldwide availability [17]. It is also the most mature band [18] which offers a larger bandwidth. Furthermore, it uses familiar PHY modules, that are already extensively applied in WiFi and Bluetooth [19], and adopts smaller antennas, which makes it ideal for the majority of on body WBAN applications. All these features contribute to make the 2.4 GHz ISM band of IEEE 802.15.6, the most widely used band.

According to the specifications of the IEEE 802.15.6 standard [20–21], The 2.4 GHz narrowband provides two modulation schemes: DQPSK and DBPSK. In addition, the nodes in a WBAN must be able to transmit their data using a power level ranging from -10 dBm (0.1 mW) to 0 dBm (1 mW) [1], so as not to exceed the specific absorption rate fixed at 1.6 W / kg in 1g of body tissue. However, the choice of one of these two modulations and the appropriate level of node transmission power, may be related to the intended application and its technical requirements, such as data rate, duty cycle, sensitivity to data loss and latency, etc.

Therefore, the aim of this paper is to study the adequacy of DBPSK and DQPSK modulations, as well as different transmission power levels in the 2.4 GHz narrowband, for a realistic WBAN model based on the IEEE 802.15.6 standard, with the requirements of normal traffic on-body medical monitoring applications, such as electrocardiogram, pulse oximetry, body temperature measurement etc., in terms of PLR, throughput and latency.

The proposed contributions of the present paper are based on two detailed numerical studies dealing with the impact of transmission power on one hand, and the two modulation schemes DBPSK and DQPSK on the other, on the performances of a WBAN model based on the IEEE 802.15.6 standard. This evaluation is done by means of three parameters: PLR, throughput and latency.

The remainder of this paper is presented as follows:

Section 2 gives an overview of the IEEE 802.15.6 standard, addressing its technical requirements for the MAC layer, with a special highlight on the beacon mode with superframe boundaries, and for the Narrow Band physical layer. Section 3 presents the PHY and MAC layer configuration of the studied WBAN model. The emphasis in this section is put on the adopted on-body channel model and the settings of MAC and radio modules. Section 4 discusses the obtained results from the performed study. and Section 5 concludes the paper.

## 2. The IEEE 802.15.6 Standard

### 2.1. Technical requirements

To implement a WBAN compliant with the IEEE 802.15.6 standard, several requirements should be considered as listed below [1]:

- Bit rate: Inter-sensor communications within WBANs should be carried out at data rates ranging from 10 kbps to 10 Mbps.
- Packet Error Rate (PER): For the majority (~95%) of the best performing links, the PER must be below 10% for a reference payload of 256 bytes.
- Easy integration and removal of nodes: Adding, removing or replacing nodes in a WBAN, should be done easily and in less than 3 seconds approximately.
- Maximum number of nodes: the size of a WBAN must not exceed 256 nodes.
- Communication reliability in case of mobility: The performance of a WBAN network must not be degraded to the point of causing data loss if the patient is mobile. Furthermore, WBANs must not cause any discomfort to the patient while moving.
- Latency and jitter: Latency and jitter are two important metrics for evaluating the performance of a WBAN. Latency should not be greater than 125 ms in medical applications and 250 ms in non-medical applications. For the jitter, it should be lower than 50 ms.
- On-body and in-body WBANs should be able of coexisting.
- Coexistence between adjacent BANs: up to 10 WBANs should be able to coexist in a space of 6m<sup>3</sup> approximately.
- Transmission power: All nodes in a WBAN must be capable of transmitting data in the range [-10 dBm, 0 dBm].
- Specific Absorption Rate (SAR): according to the Federal Communications Commissions (FCC) SAR is 1.6 W / kg in 1 g of body tissue.
- Cross-interference: WBANs must be able to provide acceptable functional performances in an environment of heterogeneous coexistence with other wireless network technologies (wifi, IEEE 802.15.14, etc.).

- Quality of Service (QoS) management must be able to provide security mechanisms, and priority features for services in a WBAN.

## 2.2. MAC layer of the standard

In order to control channel access in a WBAN, the IEEE 802.15.6 standard has defined a sophisticated MAC layer serving 3 physical layers that are designated according to the desired application.

To ensure time-referenced resources allocations in its WBAN, the coordinator (hub) must divide its transmission time axis (channel), into a series of superframes while having the choice of delimiting them or not with beacons (frames transmitted by the hub to simplify network management, such as coordination in the access to the media and power management for nodes, or to synchronize time in a WBAN, etc). However, in some cases, the coordinator may not need to time reference its allocations in the WBAN, so it can operate without using a time base, therefore, without having to transmit beacons. This means that a hub in a WBAN is able to function in one of the three access modes:

- The beacon mode with superframe boundaries.
- The non-beacon mode with superframe boundaries.
- The non-beacon mode without superframe boundaries

In this paper we only focus on Beacon mode with superframe boundaries, in which the hub sends a beacon (B) in the beginning of every active superframe to delimit it. unlike non-bacon modes, where the WBAN node can transmit its data to the coordinator using CSMA/CA or poll it to receive the data. In the beacon mode with superframe boundaries, the access phases that constitute each superframe are:

- Exclusive access phases (EAP1 and EAP2): These are time slots allocated by the hub to the rest of the WBAN nodes to transmit urgent and high-priority traffic.
- Random access phases (RAP 1 and RAP2): Dedicated to ensure random access to the medium, using CSMA/CA or Slotted ALOHA protocols (but not both at once).
- Managed Access Phases (MAP): During these access phases, the coordinator can guarantee different methods of access to the medium by the nodes of the WBAN, namely: improvised access, scheduled access and unscheduled access.
  - Improvised access: where polled (uplink allocation interval, adapted to unexpected or additional ordinary traffic service) or posted (downlink allocation interval, adapted to "unexpected" or "additional" traffic service on the downlink) allocations are generally assigned outside the planned uplink or downlink allocations.
  - Scheduled access: which is based on a pre-reservation and a validated scheduling, in such a way that nodes and the coordinator obtain scheduled recurrent time intervals to initiate frame transactions.
  - Unscheduled access: which is a combination of scheduled access and polling.

- Contention access phase (CAP): with this access phase, a time interval is defined by the coordinator and announced via a previous non-beacon frame (B2) for contention access.

## 2.3. Narrowband physical layer (NB)

As noted earlier in section I, due to the multitude of applications supported by WBANs, The IEEE 802.15.6 standard has defined three physical layers (NB, UWB and HBC). These are responsible for activating/deactivating the radio transceivers of sensor nodes, evaluating the RF channel by listening to the RF transmissions of the different sensor nodes (this mission is known as Clear Channel Assessment or CCA) and controlling the transmission/reception of data on the channel. In this work we are interested in the NarrowBand physical layer (NB), specific to medical applications like the capture of physiological signals such as temperature, blood pressure, ECG, EEG, etc.

### 2.3.1. Physical frame structure of NB-PHY

As shown in figure 1, the Physical Protocol Data Unit (PPDU) for the NB PHY layer includes:

- A PLCP (Physical Layer Convergence Procedure) preamble: 90-bit coded and used to help the receiver in time synchronization and carrier offset recovery.
- A PLCP header: it is a 31-bit coded sequence that must be inserted after the PLCP preamble to transmit information about the PHY parameters responsible for helping the receiver to decode the PSDU, essentially information about the data rate and length of the MAC frame body (data without a MAC or FCS header), and information about the following packet, whether it will be sent in BURST mode or not. These data will then be protected by a 4-bit header check sequence (HCS) (CRC-4 ITU). moreover, to improve the robustness of the PLCP header, a BCH code (31, 19,  $t = 2$ ) will be applied.
- PSDU (Physical Service Data Unit): this is the last major component of the PPDU to be transmitted, it is constructed by combining the MAC header (7 bytes), the MAC frame body with a size ranging from 0 to 255 bytes, and a 2-bit Frame Check Sequence (FCS).

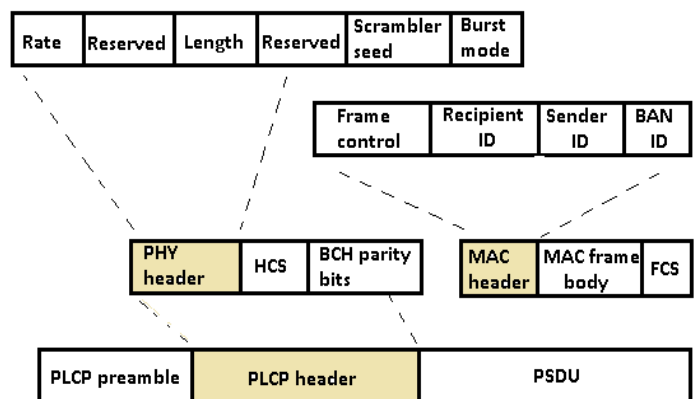


Figure1: NB PHY frame structure

### 2.3.2. NB physical layer Specifications

The IEEE 802.15.6 NB physical layer supports data transmission in a multiple frequency bands, summarized in Table 1, supporting different modulation schemes and data rates ranging from 75.9kbps to 971.4 kbps. one of the bands mentioned in the table is the 2.4 GHz band (2400-2483.5MHz) which is often the most preferred for the reasons mentioned earlier. However, several other technologies such as Wi-Fi (IEEE 802.11), Bluetooth (IEEE 802.15.1), IEEE 802.15.4 / ZigBee and others operate in this same band, and this is a real challenge due to the interference caused by the coexistence of these technologies with WBANs operating at 2.4 GHz (cross interference), and the mutual interference resulting from the coexistence of more than one WBAN in the same area [22].

### 3. Description and Settings of the Studied WBAN Model

Nodes in a WBAN can be logically arranged in different topologies, star, tree or mesh. Therefore, choosing the most suitable topology is important because of its direct impact on WBAN performances in terms of power consumption, ability to handle heterogeneity and robustness against failures [10]. As a result, most researchers assume that a one-hop star topology, where nodes send data directly to the hub without the need for relays, is the best solution [10].

The WBAN model chosen for this study operates in the NB PHY layer, at 2.45 GHz more specifically. It is composed of eleven sensor nodes, including a coordinator, which is placed at the center of the patient's belly as shown in Figure 2 [22–24]. The eleven sensors are distributed over the patient's body in a single-hop (single coordinator) star topology, where communication frames are sent directly between the coordinator and each of the other ten sensors, with no need to additional relays. Choosing these node positions is driven by the desire to monitor most of the vital signs described in the on-body medical applications of WBANS, as those proposed for IEEE 802.15.6 [25–26]. This monitoring is carried out by means of regular and programmed sampling of a number of physiological parameters such as temperature, blood pressure, pulse oximetry, ECG, EEG, etc. We have classified these applications in Table 2 into three categories: low, medium and high data rate on-body medical applications. Their technical requirements are covered and described in the references [3, 27].

#### 3.1. On-body channel model

Unlike in-body communications, where signal propagation occurs through body tissues, signals in on-body systems often propagate over body surface. Such propagation might bring up a mixture of surface, creeping, diffracted, scattered and free space waves, according to the position of the antenna [10, 28]. Furthermore, placing WBAN nodes on or around the human body, exposes them to antenna-body interaction effects, including near field coupling, distortion of the radiation pattern, and changes in the antenna impedance, which affects the functional performances of the nodes [29]. This has prompted many researchers to design on-body antennas [30–31] that take these constraints into account. In addition, patient mobility (caused by usual gestures or physiological processes such as breathing) as well as physical variations in the patient's local environment, make it difficult to model or select the appropriate on-body BAN channel, especially

since it is sensitive to temporal variations of the received signal [29]. Thus, taking into account all these issues, a good characterization of the propagation channel is required before designing any WBAN solutions.

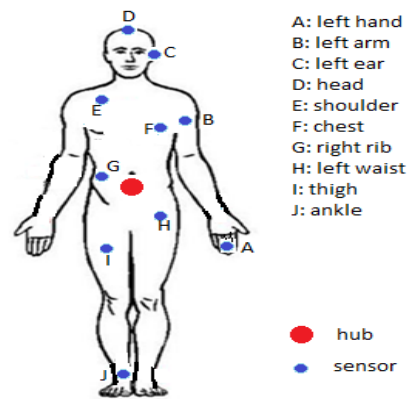


Figure 2: The proposed WBAN model

#### 3.1.1. Path loss model

To approach a realistic on-body channel modeling in our WBAN model of study, we consider that signal propagation between nodes happens on the human body surface. This brings us to distinguish two possible kinds of on-body channels:

- **The Line Of Sight channel (LOS):** characterized by the absence obstacles between sensors and the hub.
- **The Non-line Of Sight channel (NLOS)** characterized by the presence of obstacles between sensors and the hub.

Thus, for the 2.4 GHz narrow band, IEEE 802.15.6 specified two path loss models, for propagation both with and without a line-of-sight, referred to as CM3A and CM3B, respectively [32]. In this paper, the CM3B trajectory loss model is adopted, where path loss is reduced in an exponential manner around the body circumference when BAN nodes are out of line of sight. It becomes flat over longer distances due to the introduction of multipath elements from indoor conditions [32].

In this model, path loss at a distance  $d$  is defined by equation (1):

$$PL(d)[dB] = -10\log_{10} (P_0 e^{-m_0 d} + P_1) + \sigma_p n_p \quad (1)$$

Where :

- $PL(d)$  is the path loss at a distance  $d$  measured in dB
- $P_0$  represents a factor of average losses that occur close to the transmitter, and is related to the antenna design
- $P_1$  denotes the mean attenuation in indoor conditions of the elements irradiated from the body and reflected back to the receiver
- $m_0$  means the exponential mean rate of decay in dB / cm of the creeping wave diffracted around the body
- $n_p$  is a random Gaussian variable of null unit and mean
- $\sigma_p$  is the log-normal variance measured in dB

Table 1: Frequency bands and modulation parameters for the NB PHY layer

Frequency band (MHz)	Description	Modulation scheme		Information data rate (kbps)	Code rate (k/n)		Spreading factor		Symbol rate (Ksps)
		Preamble & PLCP header	PSDU		PLCP header	PSDU	PLCP header	PSDU	
402-405	Medical Implant Communication Service (MICS): Communications between implants	DBPSK	DBPSK	75.9	19/31	51/63	2	2 or 1	187.5
			DQPSK	151.8					
				303.6					
				455.4					
420-450	Wireless Medical Telemetry Services (WMTS) (Japan)	GMSK	GMSK	75.9	19/31	51/63	2	2 or 1	187.5
				151.8					
				187.5					
863-870	WMTS (Europe)	DBPSK	DBPSK	101.2	19/31	51/63	2	2 or 1	250
902-928			DQPSK	202.4					
950-958			D8PSK	404.8					
2360-2400	ISM worldwide	DBPSK	DBPSK	121.4	19/31	51/63	4	4/2/1	600
2400-2483.5			DQPSK	242.9					
				485.7					
				971.4					

Table 2: Targeted on-body medical applications

Application category	Sensor node	description	Data Rate	Duty Cycle (per device) % per time	Power Consumption	latency	Privacy
Low rate medical applications	Temperature sensor	Measures body temperature	120 bps	<1%	low	250 ms	high
	Glucose sensor (wearable)	Measures glucose level	1.6 kbps	<1%	low	250 ms	High
	Blood Pressure (BP)	The pressure of circulating blood on the walls of blood vessels.	<10 bps	< 1%	High	250 ms	Medium
Medium rate medical applications	SpO2	the oxygen saturation of hemoglobin at the blood capillaries.	32 Kbps	< 1%	Low	250 ms	High
	EEG	Electroencephalogram : an electrophysiological monitoring technique to record electrical activity of the brain.	12 leads: 43.2 kbps	< 10%	low	250 ms	high
High rate medical applications	ECG	Electrical activity of the heart.	-288 kbps (ECG-12 leads) -71 kbps (ECG-6 leads)	< 10%	Low	250 ms	High



We have measured for each node the distance separating it from the hub, and by considering that our network is under the same ambient conditions as a normal hospital room, we have calculated the values of path loss for all the links of our WBAN. We present in Table 3, the values of CM3B model parameters, under the conditions of a hospital room [32].

3.1.2. Temporal variation

Temporal variation is also another important aspect in BAN channel modeling due to the rapid changes that can occur in this type of environment. In our WBAN model of study, the mobility of body limbs is captured by a temporal variation component. To the best of our knowledge, no unique model exists in the literature describing temporal variation, but rather a general modeling of this parameter is used [33]. The description of the approach used in the majority of simulators to describe temporal variation is as follows: it is considered that the actual propagation loss may differ from the average propagation loss over time. Therefore, a probability density function (pdf) is developed to cope with this issue. The pdf is determined from the previous observed value and the time elapsed since then. This function cannot be dynamically produced from real models but must rather be derived from experimental measurements.

The actual propagation loss is calculated as the sum of the pre-calculated average propagation loss and the temporal variation.

3.2. MAC module settings

In the studied WBAN model, the beacon mode with superframe boundaries is adopted, in which a beacon is transmitted at the beginning of every active superframe. This mode allows the coordinator to divide the communication channel into multiple access phases. As a result, the MAC module configuration adopted in this work, combines both of random access based on CSMA/CA and scheduled access based on TDMA as presented in detail in our previous work [23].

3.3. Radio module settings

In the 2.4 GHz narrowband, the preamble is always modulated using  $\pi/2$ -DBPSK, and the binary information is modulated with one of the following two differential modulations:  $\pi/2$ -DBPSK or  $\pi/4$  DQPSK. For both of these, the bitstream  $b(n)$ ,  $n=0,1,\dots,N-1$  is transformed into a sequence of symbols  $s(k)$ ,  $k=0,1,\dots,(N/\text{Log}_2(M))-1$  as follows :

$$s(k) = s(k - 1)\exp(j\varphi_k) \tag{2}$$

M denotes the modulation order, and  $s(-1) = \exp(j\pi/2)$  is the reference for the first preamble symbol [20]. The transition values between the  $\varphi_k$  symbols are shown in Tables 4 and 5. for both  $\pi/2$ -DBPSK or  $\pi/4$  DQPSK. To the best of our knowledge, there is still no radio chip available in the market that has been specifically designed to meet IEEE 802.15.6 standard requirements. Therefore, our work relies on a proposed radio chip by Alan Wong et al [34], compliant with the IEEE 802.15.6 standard recommendations for the 2.4 GHz NB PHY layer. The parameters for both modulation schemes of the proposed radio are defined in Table 6.

Table 3: Values of CM3B path loss model parameters in a hospital room

Parameter	Value
Frequency (GHz)	2.45
$P_0$ [dB]	-25.8
$m_0$ [dB/cm]	2.0
$P_1$ [dB]	-71.3
$\sigma_P$ [dB]	3.6

Table 4: Symbols transition values of  $\pi/2$ -DBPSK

B(K)	$\varphi_k$
0	$+\pi/2$
1	$-\pi/2$

Table 5. Symbols transition values of  $\pi/4$  DQPSK

B(2K), B(2K+1)	$\varphi_k$
0 0	$\pi/4$
0 1	$3\pi/4$
1 0	$7\pi/4$
1 1	$5\pi/4$

4. Results and Discussion

4.1. Effects of transmission power and modulation schemes on the performance of the proposed WBAN model

To assess the impact of transmission power and modulation schemes on the studied WBAN performance, in terms of packet loss rate, throughput, and latency, we used Castalia sensor network simulator as a basis, which is based on OMNeT++ (4.6) platform and includes as standard the support of the IEEE 802.15.6 specifications, which justifies its choice in our numerical study.

In our simulations, BAN nodes send their packets at different data rates to the coordinator using a different transmission power ranging from -10 to 3 dBm, each time, and a different modulation scheme (DBPSK or DQPSK).

4.1.1. Packet Loss Rate (PLR)

Packet losses in a WBAN can be due to various possible causes such as the disconnection of a sensor from the network (no beacon reception is detected) or simply the inability to transmit data in the defined transmission tries. Therefore, The packet loss rate indicates the chance that a packet produced by a node cannot be received normally and as expected by the coordinator [35].

Nevertheless, this parameter can be impacted by different factors [36]:

- Environmental factors such as distance between sensors, activity and body posture.
- Technical factors such as transmission power, receiver

sensitivity, antenna gain or coding.

In this work, we are interested in the evaluation of two technical factors that can impact PLR in WBANs, which are modulation and transmission power. For this reason, we show in Figures 3, 4, 5, and 6, the PLR for DBPSK and DQPSK modulations as a function of different data rates when the transmission power is set to -10, -5, 0 and 3 dBm, respectively.

As expected, the analysis of the results obtained in Figures 3, 4, 5 and 6, clearly shows that packet loss rate decreases as the transmission power increases. In fact, CM3B channel model introduces relatively high path losses depending on the distance between the communicating nodes (71.3 dB for an inter-sensor distance  $d$  greater than 30cm). Therefore, to ensure a better received signal level at the coordinator (signal level higher than the minimum threshold of receiver sensitivity), corresponding to a better throughput, the transmission power of WBAN nodes must be sufficient enough. Moreover, as shown in the four figures (3, 4, 5 and 6), for low and medium data rates (below 60 kbps), DBPSK modulation shows a lower packet loss rate compared to that obtained with DQPSK. This can be explained by the difference in radio sensitivity due to the modulation scheme adopted. We can notice in table 6, that receiver sensitivity is significantly lower in DBPSK (-104 dBm) compared to DQPSK (-96.5 dBm). However, at high transmission data rates (from 60 kbps onwards), the packet reception rate appears to be better with DQPSK than with DBPSK.

We can relate the low performance of DBPSK in high data rates to the well-known problems of buffer saturation. Furthermore, DQPSK allows with 4 phases to encode two bits per symbol instead of only one in the case of DBPSK. Thus, at the same transmission data rate, DQPSK modulation allows the transmission of twice as many symbols (i.e. a multiplication of the data rate by a factor of 2) as DBPSK modulation. This explains the relatively fast saturation of the coordinator buffer at high data rate with DBPSK. This is further confirmed by the equation (3) below [10].

$$R_d = \left( \frac{R_s \cdot N}{S} * \frac{k}{n} \right) (kbps) \quad (3)$$

where  $R_d$  refers to the information data rate,  $R_s$  means the symbol rate,  $S$  is the spreading factor,  $k/n$  is the code rate BCH, and  $N$  is a variable linked to the modulation order  $M$  by  $M=2^N$ .

#### 4.1.2. Throughput

Throughput is an important metric in WBANs, it refers to the

number of packets received successfully per second (p/s). However, this parameter can be weakened due to several causes including low duty cycles, or the transmission of control packets. As demonstrated in equation (4) [35], throughput ( $S$ , expressed in bytes/s) is much higher when packet loss rate is lower.

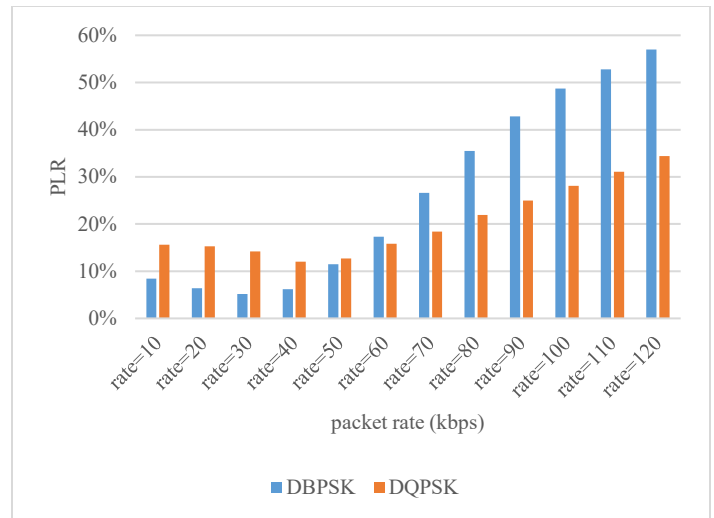


Figure 3: PLR at -10 dBm

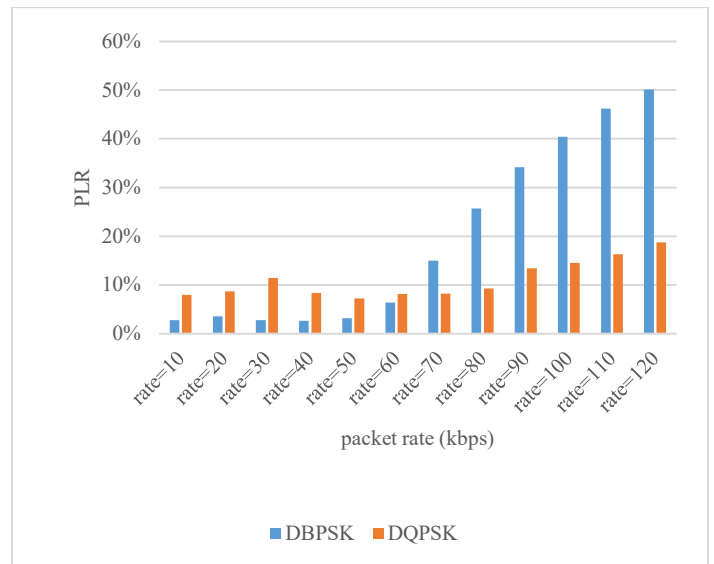


Figure 4: PLR at -5 dBm

Table 6: Settings of the modulation schemes used

Frequency band (Ghz)	Rx sensitivity (dBm)	Channel spacing (MHz)	Symbol rate (ksps)	Tx power (dBm)	Power consumption for Tx mode
2.4	-96.5(for DQPSK)	1	600	-10 to 3	• 5.9 mW (for DPSK -10 dBm).
	-104 (for DBPSK)				• 9.5 mW (for DPSK 0 dBm).
					• 12.3 mW (for DPSK 3 dBm)

At constant PLR, S is proportional to the number M of nodes in the network and to the traffic payload (P) expressed in bytes, and is inversely proportional to the length of the beacon period Bp.

$$S = \frac{M \cdot P}{B_p} (1 - PLR) \quad (4)$$

Figures 7, 8, 9 and 10 illustrate the throughput of the DBPSK and DQPSK modulations as a function of various data rates at -10, -5, 0 and 3dBm respectively.

These results show, once again, as in the earlier PLR study, that the higher the transmission power the better the throughput. Furthermore, as illustrated in the four figures, for low and medium data rates (below 60 kbps), DBPSK guarantees a better throughput compared to DQPSK modulation, as the receiver sensitivity in the case of DBPSK is lower. However, at high transmission data rates (60 kbps and more) DQPSK modulation seems to be more efficient, for the same reasons (and the same high data rate context) previously formulated with the results obtained for PLR.

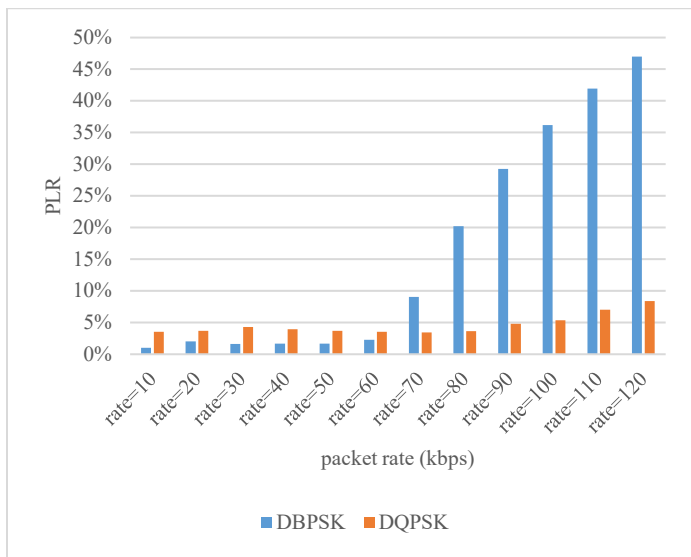


Figure 5: PLR at 0 dBm

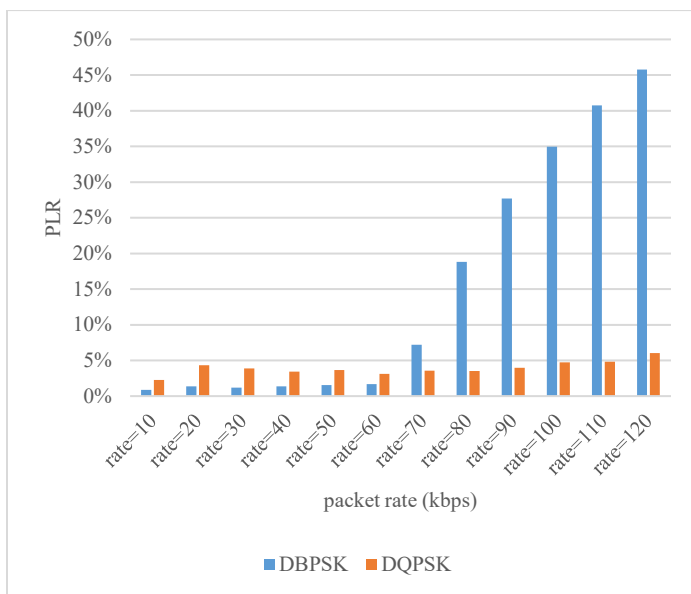


Figure 6: PLR at 3dBm

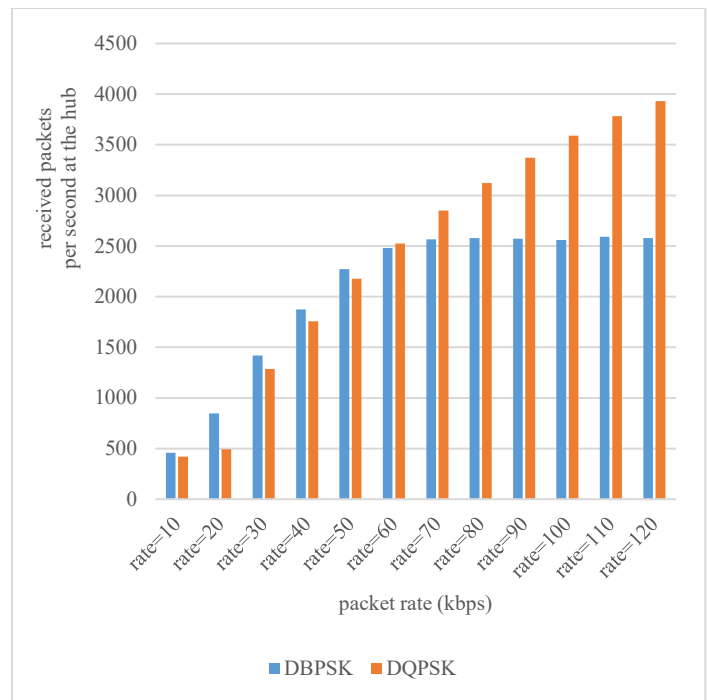


Figure 7: Throughput at -10dBm

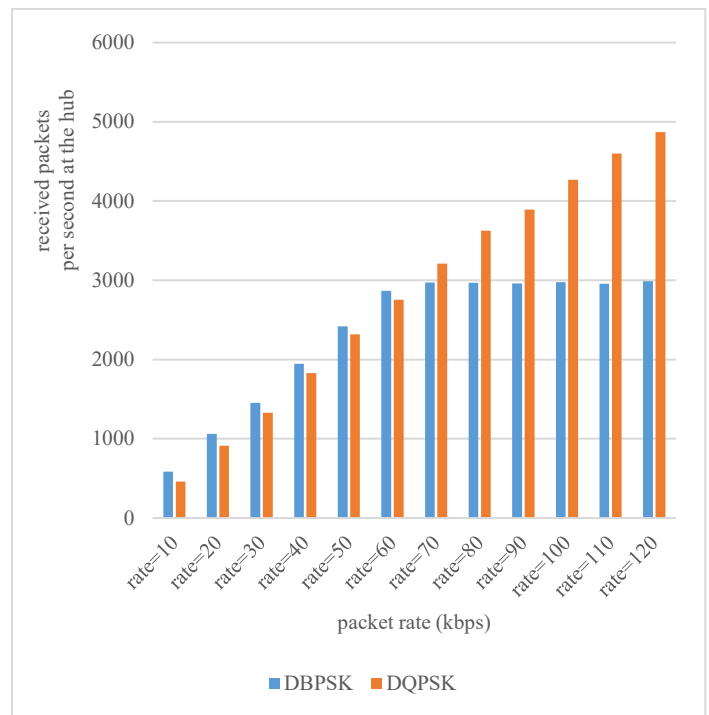


Figure 8: Throughput at -5 dBm

### 4.1.3. Latency

Latency is a metric defined, in the context of our work, as the time interval between the node receiving a beacon, and the coordinator correctly receiving the node's frame. We evaluate this metric in our WBAN model of study, at low, medium, and high data rates, for several transmission powers (-10dBm, -5dBm, 0dBm and 3dBm). For this purpose, we have considered the three following simulation scenarios:

- Transmission data rate less than 10 kbps with DBPSK modulation: to represent the case of low data rate on-body medical applications (according to the categorization in Table 2).

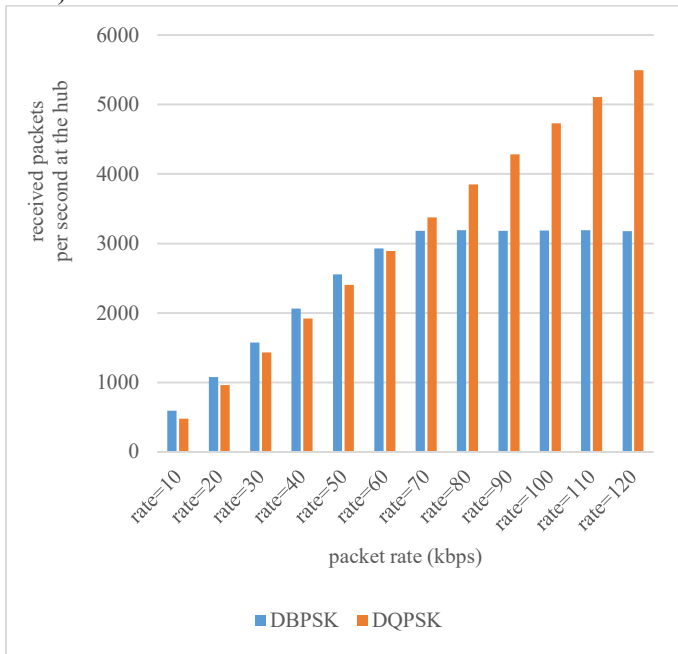


Figure 9: Throughput at 0 dBm

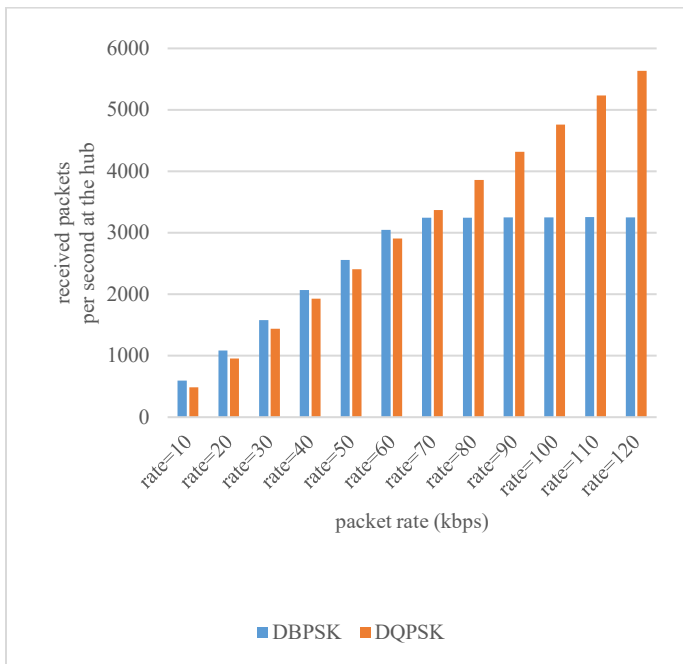


Figure 10: Throughput at 3 dBm

- Maximum transmission rate of 50 kbps with DBPSK modulation: to represent the case of medium rate on-body medical applications.
- Maximum transmission data rate of 71 kbps with DQPSK modulation: to represent the case of high data rate on-body medical applications (example of ECG monitoring with 6-electrode).

Thus, we show in Figures 11, 12, and 13 the distribution of received packets at different time intervals for the three cases respectively, when the transmission power varies from -10 to 3 dBm.

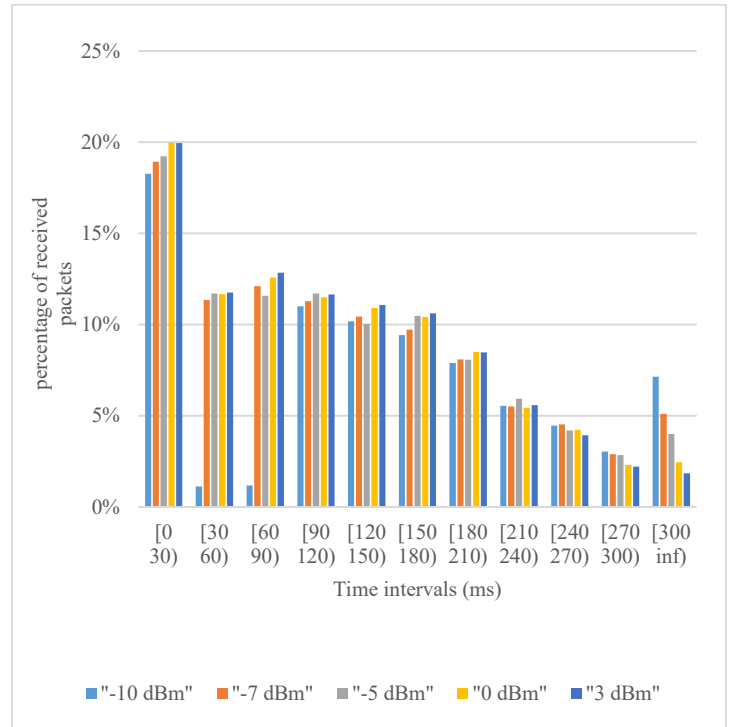


Figure 11: latency for low data rate medical applications (DBPSK)

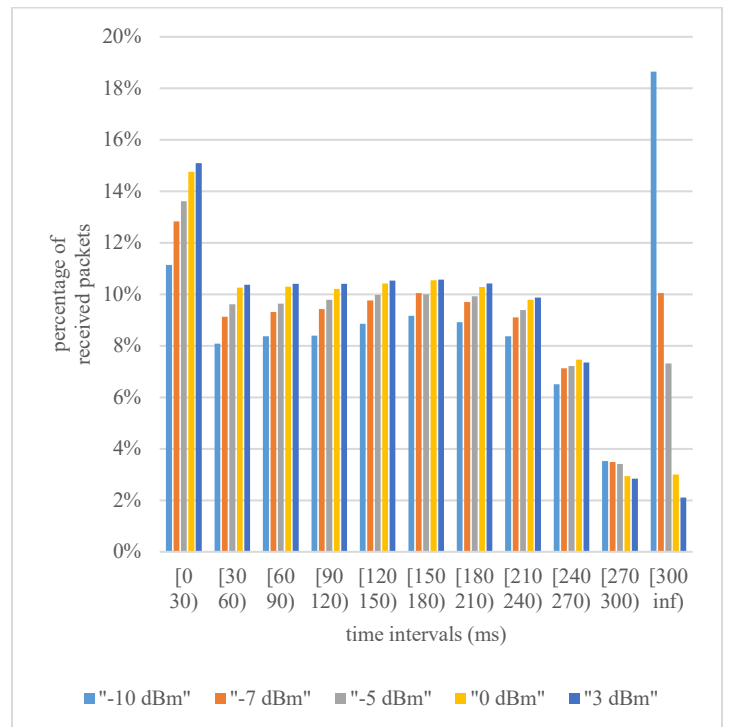


Figure 12: latency for medium data rate applications (between 32 and 43 kbps)



Table 7. Optimal modulation scheme and transmit power values for medical on-body applications

Medical applications	Optimal values of transmit power	Optimal modulation scheme
<b>Medium data rate medical applications: EEG monitoring (12 leads) and SpO2</b>	-10 dBm PLR <10 % Throughput: 2270 p/s Latency: 52 % of packets received before 240 ms	DBPSK
	-5 dBm PLR <10 % Throughput: 2418 p/s Latency: 69 % of packets received before 240 ms	
	0 dBm PLR <10 % Throughput: 2555 p/s Latency: 81 % of packets received before 240 ms	
	3 dBm PLR <1 % Throughput: 2559 p/s Latency: 84 % of packets received before 240 ms	
	-10 dBm PLR <10 % Throughput: up to 456p/s Latency: 86 % of packets received before 240 ms	
<b>Low data rate medical applications: (&lt;10 kbps)</b>	-5 dBm PLR <3 % Throughput: up to 585 p/s Latency: 90 % of packets received before 240 ms	DBPSK
	0 dBm PLR <2 % Throughput: up to 593p/s Latency: 92 % of packets received before 240 ms	
	3 dBm PLR <1 % Throughput: up to 594 p/s Latency: 93 % of packets received before 240 ms	
	-10 dBm PLR =19 % Throughput: 2850 p/s Latency: 60 % of packets received before 240 ms	
	-5 dBm PLR = 8 % Throughput: 3208 p/s Latency: 70 % of packets received before 240 ms	
<b>ECG (6 leads)</b>	0 dBm PLR <10 % Throughput: 3377 p/s Latency: 79 % of packets received before 240 ms	DQPSK
	3 dBm PLR <10 % Throughput: 2380 p/s Latency: 84 % of packets received before 240 ms	

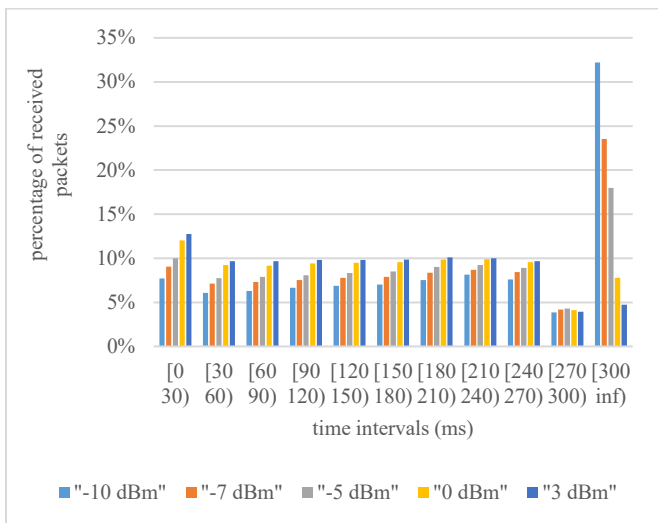


Figure 13: Latency for high data rate applications (72 kbps)

#### 4.2. Discussion and synthesis

E-health applications can be classified into real-time and near real-time applications [37–38]. Nevertheless, the QoS requirements of each of these applications may differ depending on the context in which the service is invoked. This means that in emergency cases, for example, remote medical monitoring will require real-time data transmission, with no tolerance for data loss. Whereas in non-emergency situations, vital signs transmission does not require the same level of QoS (can tolerate delay). In addition, depending on the application demands, three basic data delivery models can be distinguished [39–40] : continuous model, query model and event model. The continuous model is the most basic data delivery model, which aims to periodically transmit the data collected by sensor nodes, which can either be real-time data such as voice or image that does not tolerate delay (but packet loss can be tolerated within threshold limits), or non-real-time continuous data where latency and packet loss can be permissible. However, In the query data delivery model, the coordinator requests the transmission of packets to sensors, this model requires

reliable and timely data transmission. As for event-delivery model, data transmission only takes place when an event of interest occurs. In this case, reliable and real-time data transmission is required. Moreover, the QOS requirements also depend on the needs of the patients, for some patients it may be enough to send their signals every few minutes, while others may need to transmit the collected data every few seconds [41].

The simulation results of our WBAN model have shown that in the case of low data rate medical applications (about 10 kbps), including blood pressure, glucose and body temperature measurement, we can note that DBPSK modulation can guarantee a PLR < 10% even when nodes transmit in the lowest power range. At -10 dBm for example, a PLR of about 8% is achieved with 86% of packets received under 250 ms, and a throughput of 456 packets/s. So if the target of these applications is non-emergency continuous real-time monitoring, the choice of DBPSK modulation with the lowest power level (-10 dBm) can guarantee the intended results.

## 5. Conclusion

In this paper we have studied a realistic WBAN model based on the 2.4 GHz narrowband of the IEEE 802.15.6 standard and dedicated to on-body medical applications. With this model, we have investigated the adequacy of the modulation schemes proposed by the standard in this frequency band as well as transmission power, while considering the requirements of these applications in terms of PLR, throughput and latency. Simulation results have shown that DBPSK modulation can be a good choice for low and medium data rate applications. However, DQPSK modulation would be more suitable for high data rate medical applications. Regarding the transmission power, low data rate medical applications can accept minimum transmission powers (-10 dBm for example). However, when the required application data rate increases, higher transmission powers, e.g. 0 dBm, must necessarily be used to ensure a reduced packet loss rate, acceptable latency, and higher throughput.

The 2.4 GHz frequency band is through a band adopted by several other technologies such as Wifi, Bluetooth etc. This is likely to influence the operation of the proposed WBAN model if it is located in an environment that includes these technologies or other WBANs. This is known as mutual interference, which must be taken into account in the physical layer configuration.

## References

- [1] S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith and A. Jamalipour, "Wireless Body Area Networks: A Survey" *IEEE Commun. Surv. Tut.*, **16**(3), 1658-1686, 2014. <https://doi.org/10.1109/SURV.2013.121313.00064>.
- [2] A. Es-saqy, M. Abata, S. Mazer, M. Fattah, M. Mehdi, M. El bekkali and C. Algani, "Very Low Phase Noise Voltage Controlled Oscillator for 5G mm-wave Communication Systems" in *2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, Meknes, Morocco, 2020. <https://doi.org/10.1109/IRASET48871.2020.9092005>.
- [3] S. Ullah, "A Review of Wireless Body Area Networks for Medical Applications" *Int'l J. Commun. Netw. Syst. Sci.*, **02**(08), 797-803, 2009. <https://doi.org/10.4236/ijcns.2009.28093>.
- [4] Boulemtafes A., Badache N. "Design of Wearable Health Monitoring Systems: An Overview of Techniques and Technologies" *mHealth Ecosystems and Social Networks in Healthcare*. Springer, Cham, 2016.
- [5] R. Negra, I. Jemili, and A. Belghith, "Wireless Body Area Networks: Applications and Technologies" *Procedia Comput. Sci.*, **83**, 1274-1281, 2016. <https://doi.org/10.1016/j.procs.2016.04.266>.

- [6] B. Antonescu and S. Basagni, "Wireless body area networks: Challenges, trends and emerging technologies" in *BODYNETS 2013 - 8th Int. Conf. Body Area Networks*, 2013. <https://doi.org/10.4108/icst.bodynets.2013.253722>.
- [7] M. Boumaiz, M. El ghazi, S. Mazer, M. Fattah, A. Bouayad, M. El bekkali, and Y. Balboul, "Energy harvesting based WBANs: EH optimization methods" *Procedia Comput. Sci.*, **151**, 1040-1045, 2019. <https://doi.org/10.1016/j.procs.2019.04.147>.
- [8] Z. Jin, J. Oresko, S. Huang, and A. C. Cheng, "HeartToGo: A personalized medicine technology for cardiovascular disease prevention and detection" in *2009 IEEE/NIH Life Science Systems and Applications Workshop*, Bethesda, MD, 2009. <https://doi.org/10.1109/LISSA.2009.4906714>.
- [9] M. Hadjem, O. Salem and F. Naït-Abdesselam, "An ECG monitoring system for prediction of cardiac anomalies using WBAN" in *2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom)*, Natal, 2014. <https://doi.org/10.1109/HealthCom.2014.7001883>.
- [10] Ullah, S., Higgins, H., Braem, B. Latre, B. Blondia, C. Moerman, I. Saleem, S. Rahman, Z. and Kwak, K. S. "A Comprehensive Survey of Wireless Body Area Networks". *J Med Syst.* **36**(3), 1065-1094. <https://doi.org/10.1007/s10916-010-9571-3>.
- [11] Huzoore G., Khedo K.K., Joonas N. "Wireless Body Area Network System Architecture for Real-Time Diabetes Monitoring" In: Fleming P., Vyas N., Sanei S., Deb K. *Emerging Trends in Electrical, Electronic and Communications Engineering*. Springer, Cham, 2016.
- [12] M. Udin Harun Al Rasyid, F. A. Saputra, and A. Christian, "Implementation of blood glucose levels monitoring system based on Wireless Body Area Network" in *2016 IEEE Int. Conf. Consum. Electron*, 2016. <https://doi.org/10.1109/ICCE-TW.2016.7521005>.
- [13] Hsueh-Ting Chu, Chir-Chang Huang, Zhi-Hui Lian and J. J. P. Tsai, "A ubiquitous warning system for asthma-inducement" in *IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC'06)*, Taichung, 2006. <https://doi.org/10.1109/SUTC.2006.20>.
- [14] K. Revett and S. T. De Magalhães, *Cognitive biometrics: Challenges for the future*, Springer, Berlin, Heidelberg, 2010.
- [15] T. M. Connolly, E. A. Boyle, E. Macarthur, T. Hainey, and J. M. Boyle, "Computers & Education A systematic literature review of empirical evidence on computer games and serious games" *Comput. Educ.*, **59**(2), 661-686, 2012. <https://doi.org/10.1016/j.compedu.2012.03.004>.
- [16] K. Hasan, K. Biswas, K. Ahmed, N. S. Nafi, and M. S. Islam, "A comprehensive review of wireless body area network" *J. Netw. Comput. Appl.*, **143**, 178-198, 2019. <https://doi.org/10.1016/j.jnca.2019.06.016>.
- [17] R. Cavallari, F. Martelli, R. Rosini, C. Buratti, and R. Verdone, "A survey on wireless body area networks: Technologies and design challenges" *IEEE Commun. Surv. Tutorials*, **16**(3), 1635-1657, 2014. <https://doi.org/10.1109/SURV.2014.012214.00007>.
- [18] P. Mathew, L. Augustine, D. Kushwaha, D. Vivian and D. Selvakumar, "Hardware implementation of NB PHY baseband transceiver for IEEE 802.15.6 WBAN" in *2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom)*, Greater Noida, 2014. <https://doi.org/10.1109/MedCom.2014.7005977>.
- [19] Wang, Y., & Wang, Q. "Evaluating the IEEE 802.15.6 2.4GHz WBAN Proposal on Medical Multi-Parameter Monitoring under WiFi/Bluetooth Interference" *International Journal of E-Health and Medical Communications (IJEHMC)*, **2**(3), 48-62, 2020. <https://doi.org/10.4018/jehmc.2011070103103>.
- [20] IEEE Standard for Local and metropolitan area networks - Part 15.6: Wireless Body Area Networks," in *IEEE Std 802.15.6-2012*, 1-271, 2012. <https://doi.org/10.1109/IEEESTD.2012.6161600>.
- [21] S. AhmadSalehi, M.A Razzaque, I. Tomeo-Reyes and N. Hussain, "IEEE 802.15.6 standard in wireless body area networks from a healthcare point of view" in *2016 22nd Asia-Pacific Conference on Communications (APCC)*, 2016. <https://doi.org/10.1109/APCC.2016.7581523>.
- [22] M. Boumaiz, M. E. Bekkali, A. Bouayad and M. Fattah, "The Impact of Distance between Neighboring WBANs on IEEE 802.15.6 Performances" in *2019 7th Mediterranean Congress of Telecommunications (CMT)*, Fès, Morocco, 2019. <https://doi.org/10.1109/CMT.2019.8931341>.
- [23] M. Boumaiz, M. El Ghazi, M. Fattah, A. Bouayad, and M. El Bekkali, "Mutual coexistence in WBANs: Impact of modulation schemes of the IEEE 802.15.6 standard" *Int. J. Adv. Comput. Sci. Appl.*, **11**(5), 294-302, 2020. <https://doi.org/10.14569/IJACSA.2020.0110539>.
- [24] M. Boumaiz, M. El ghazi, A. Bouayad, M. Fattah, M. El bekkali and S. Mazer, "The impact of transmission power on the performance of a WBAN prone to mutual interference" in *2019 International Conference on Systems of Collaboration Big Data, Internet of Things & Security (SysCoIoT)*, Casablanca, Morocco, 2019. <https://doi.org/10.1109/SysCoIoT48768.2019.9028035>.
- [25] F. Di Franco, C. Tachtatzis, R. C. Atkinson, I. Timmirello and I. A. Glover, "Channel estimation and transmit power control in wireless body area

- networks" in *IET Wireless Sensor Systems*, **5**(1), 11-19, 2015. <https://doi.org/10.1049/iet-wss.2013.0070>.
- [26] K. Takizawa, A. Aoyagi, J. Takada, N. Katayama, K. Yekeh, Y. Takehiko and K. Kohno. "Channel models for wireless body area networks", in 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, 2008. <https://doi.org/10.1109/IEMBS.2008.4649465>
- [27] M. Shu, D. Yuan, C. Zhang, Y. Wang, and C. Chen, "A MAC protocol for medical monitoring applications of wireless body area networks," *Sensors (Switzerland)*, **15**(6), 12906–12931, 2015. <https://doi.org/10.3390/s150612906>.
- [28] P. S. Hall, "Antennas Challenges for Body Centric Communications" in 2007 International workshop on Antenna Technology: Small and Smart Antennas Metamaterials and Applications, Cambridge, 2007. <https://doi.org/10.1109/IWAT.2007.370076>.
- [29] S. L. Cotton, R. D'Errico, and C. Oestges, "A review of radio channel models for body centric communications". *Radio science*, **49**(6), 371–388, 2014. <https://doi.org/10.1002/2013RS005319>
- [30] B. Hazarika, B. Basu, and J. Kumar, "A Multi-Layered Dual-Band On-Body Conformal Integrated Antenna for WBAN Communication" *Aeu-int J Electron C.*, **95**, 226-235, 2018. <https://doi.org/10.1016/j.aeue.2018.08.021>.
- [31] I. Suriya and R. Anbazhagan, "Inverted-A based UWB MIMO antenna with triple-band notch and improved isolation for WBAN applications," *Aeu-int J Electron C.*, **99**, 25–33, 2019. <https://doi.org/10.1016/j.aeue.2018.11.030>.
- [32] K. Y. Yazdandoost, Channel Modeling Subcommittee Report. 2010.
- [33] M. Boumaiz, M. El ghazi, S. Mazer, M. El bekkali, A. Bouayad and M. Fattah, "Performance analysis of DQPSK and DBPSK modulation schemes for a scheduled access phase based Wireless Body Area Network" in 2019 9th International Symposium on Signal, Image, Video and Communications (ISIVC), Rabat, Morocco, 2018. <https://doi.org/10.1109/ISIVC.2018.8709232>.
- [34] A. Wong, M. Dawkins, G. Devita, N. Kasparidis, A. Katsiamis, O. King, F. Lauria, J. Schiff and A. Burdett., "A 1V 5mA multimode IEEE 802.15.6/bluetooth low-energy WBAN transceiver for biotelemetry applications" in IEEE International Solid-State Circuits Conference, San Francisco, CA, 2012. <https://doi.org/10.1109/ISSCC.2012.6177022>.
- [35] F. Martelli, C. Buratti and R. Verdone, "On the performance of an IEEE 802.15.6 Wireless Body Area Network" in 17th European Wireless 2011 - Sustainable Wireless Technologies, Vienna, Austria, 2011.
- [36] H. Zhang, F. Safaei, and L. C. Tran, "Joint transmission power control and relay cooperation for WBAN systems" *Sensors (Switzerland)*, **18**(12), 1–26, 2018. <https://doi.org/10.3390/s18124283>.
- [37] L. Skorin-Kapov and M. Matijasevic, "Analysis of QoS requirements for e-Health services and mapping to evolved packet system QoS classes" *Int. J. Telemed. Appl.*, 2010. <https://doi.org/10.1155/2010/628086>.
- [38] D. Vouyioukas, I. Maglogiannis, and D. Komnakos, "Emergency m-Health Services through High-Speed 3G Systems: Simulation and Performance Evaluation" *Simulation*, **83**(4), 329–345, 2007. <https://doi.org/10.1177/0037549707083113>.
- [39] M. A. Yigitel, O. D. Incel, and C. Ersoy, "QoS-aware MAC protocols for wireless sensor networks: A survey," *Comput. Networks*, **55**(8), 1982–2004, 2011. <https://doi.org/10.1016/j.comnet.2011.02.007>.
- [40] M. A. Razzaque, S. S. Javadi, Y. Coulibaly, and M. T. Hira, "Qos-aware error recovery in wireless body sensor networks using adaptive network coding," *Sensors (Switzerland)*, **15**(1), 440–464, 2015. <https://doi.org/10.3390/s150100440>.
- [41] A. Maroua and F. Mohammed, "Characterization of Ultra Wide Band indoor propagation". in 2019 7th Mediterranean Congress of Telecommunications (CMT), Fès, Morocco, 2019. <https://doi.org/10.1109/CMT.2019.8931367>.

## Non-Ridership Presumption Toward New Bus Rapid Transit Purwokerto-Purbalingga Execution

Fauzan Romadlon\*

*Industrial Engineering Department, Faculty of Industrial Engineering and Design, Institut Teknologi Telkom Purwokerto, Purwokerto, 53147, Indonesia*

---

### ARTICLE INFO

*Article history:*

*Received: 11 June, 2020*

*Accepted: 17 August, 2020*

*Online: 28 August, 2020*

---

*Keywords:*

*Non-ridership*

*Bus Rapid Transit*

*Presumption*

*Purwokerto-Purbalingga*

---

### ABSTRACT

*Bus Rapid Transit in Purwokerto-Purbalingga is a new mass transportation mode. Recently, the execution of the BRT has been going on for three years. In terms of service standards to ridership, the BRT has been fulfilled the requirement. However, during the execution, it shall be supported by the non-ridership (local communities) who get the impact as public engagement. The non-ridership impact is captured by observing their presumptions. This study uses quantitative method and survey technique to collect the data by spreading questionnaires to the non-ridership in Purwokerto and Purbalingga. The collected data is analyzed by Analysis of Variance (ANOVA) and Structural Equation Modelling Partial Least Square (SEM-PLS). The ANOVA results show that gender, age ratio, and residence (living area) are significant presumption factors. According to SEM-PLS model, the R-squared of non-ridership presumption variables toward the BRT execution as excellent public transportation is at 51.8% (moderate level). It is found that the economic variable affects the excellent public transportation variable is at 41.4%, and the social variable have a correlation with the excellent public transportation variable is at 36.2%, but not with the environment variable (5.6% only). Following up these findings, it is recommended that public engagement through the non-ridership presumption will lead the BRT provider to purpose some programs to improve the service and increase the occupancy. So that, the proposed program will attract the sense of awareness and public engagement of the non-ridership toward the BRT execution.*

---

### 1. Introduction

Developing worldwide Bus Rapid Transit (BRT) system are increased significantly. Especially in Indonesia, BRT has been implemented not only in the capital city but also in many big cities to support urban mobilities. One of the advantages of BRT is to provide public transportation that relies on the ease of use [1], safety, comfort, and convenience service [2]. The level of BRT service is measured based on BRT effectiveness. BRT service satisfaction is achieved by increasing service procedures, service requirements, and service distinction with conventional public transportation [3]. The quality of service is measured by customer experiences through maintaining bus stop average waiting time, evaluating the impact of operation timetable, and considering BRT capacity [4].

Bus stop facilities are compulsory in BRT execution. Bus stops have to integrate with people density which is close to job location to come up with accessibility enhancement [5] and give impacts to potential access informal job opportunities [6]. Besides, the building environment around the bus stops has to consider the density of public facilities and green areas [7]. Affordability of BRT stops shall concern spatial structure [8] and spatial heterogeneity to support BRT planning in service management and infrastructure provision [9]. Therefore, the placement of BRT stops must be considered friendly pedestrian, densely of resident, industrial park connection, and public facilities consideration [10].

In advanced development, BRT has been configured as multimodal transportation networks. It is used to promote the reduction of greenhouse gas emissions and made the diversity of land use [11]. Multimodal and intermodal transportation is enhanced to emphasize of social action while providing reliable and smooth access to the city center [12]. The multimodality index in BRT has a positive relation with BRT ridership by shaping their behavior, land use [13], and transit rate [14]. Ridership perspective

---

\*Corresponding Author: Fauzan Romadlon, Institut Teknologi Telkom Purwokerto, Email: [fauzan@ittelkom-pwt.ac.id](mailto:fauzan@ittelkom-pwt.ac.id)



and attitude (behavior) on BRT are the leading indicators to enhance BRT service. Differences between males, elder workers, workers with higher education, and income [15] make the different perspectives of the service. The perspectives are generated motivation challenge of using BRT, such as security, reducing commuting time, staff attitude, and price charges [16]. The impact of land use toward BRT affects high values of properties [17], increased urban development accesses, and optimized time reduction [18].

The success bus implementations are affected overall sustainability such as safety has increase BRT benefit, pro-poor, or lower share of daily earnings [19], and decrease road traffic crashes [20]. On the other hand, BRT implementation faces some issue such as limited government financial support, green transportation issue, and high ridership occupancy [21] and no support form vehicle industries [22]. Sometimes, the execution of BRT has a problem with existing mass transportation, so it requires an understanding of existing transportation to be reorganized [23].

Implementing BRT as public transportation shall include public engagement [24]. The public engagement is to achieve transparent decision with good input from stakeholder [25]. After the public transport is realized, the system of public transportation shall be evaluated whether has economic development impact, sustainability impact, and effect to livable communities [26].

Post-planning or operation step needs more understanding of the connection between the services and public engagement [27]. In some case, public engagement especially from marginal stakeholder who has low power and low interest [28] are less involved. The marginal stakeholder can be identified as three groups: community who has no vehicles and they captive as primary beneficiaries of BRT, community who has vehicle but reluctant, and rich community who has strong voice [29].

Marginal stakeholder engagement can be measured from their presumption. In this study, the marginal stakeholder is non-ridership that lives around BRT stops and BRT routes. Their presumption will describe public engagement through new public transportation execution. The presumption can be set as a tool to capture the existing condition of public engagement toward BRT execution especially in different perspective. The new public transportation of the study is BRT Purwokerto-Purbalingga in Indonesia. It is chosen because it merely has been executed for three years but less public awareness.

The public engagement will lead to a sense of local ownership for the BRT in Purwokerto and Purbalingga. A sense of local ownership is used to identify economic probability and community benefits especially, in the job sector, age groups of worker effect, investment time [30], and less CO2 production [31]. In the other hand, community acceptance increases awareness system to the community, maximizes in implementation, and enhances ridership capacity [32].

## 2. Methodology

### 2.1. The Covered Respondent

The covered area of the study is in Purwokerto and Purbalingga. Purwokerto and Purbalingga are placed in Central Java Province, Indonesia. The operation of BRT is conducted from Ministry of Transport in Central Java Province. The route is started from Bulupitu bus station in Purwokerto until Bukateja bus station in Purbalingga. Bulupitu bus station is central bus station in

Purwokerto that is the transit hub for ridership to travel intercity and interprovince. Bukateja is bus station that is set final destination and it will be transit hub that close to planned airport in Purbalingga. The total route of the BRT is 26.4 km with 56 bus stops [33].

The study is established by quantitative methods. The questionnaire is set to survey non-ridership who live around the BRT route and bus stops. The respondent is people who live or work around the BRT route and the BRT stop. Data collection technique is conducted by doing direct interviews, face to face, and then the result of these interview is submitted to a Google Form. The proportion of the respondent form both cities are set equal.

### 2.2. ANOVA Analysis

Demographics spread is set as proportional to both cities. The surveyed criteria are gender, residence, job, age, and routine. The presumption is measured using a Likert scale where the scale is five points; 1) Strongly disagree, 2) Disagree, 3) Moderate, 4) Agree, and 5) Strongly agree. After data collection, the presumption questionnaire result is tested for the validity and reliability. The validity test of the questionnaire uses Pearson Correlation and the reliability uses Cronbach' Alpha. Then, the multicollinearity, outliers, and normality test are performed to check whether the data is followed statistical rules or not.

The significance correlation of the demographics of non-ridership and the presumption data are approached by using ANOVA (Analysis of Variance) with Minitab 19 software (alfa is 5%). The hypothesis is set below.

**H1:** The demographics of non-ridership has significance correlation with their presumption.

### 2.3. SEM-PLS Model Evaluation

Structural Equation Modelling Partial Least Square (SEM-PLS) is used to find correlations among the presumptions. The presumption is divided into four categories; social, economic, environment, and excellent public transportation. The modelling is done by Smart-PLS 3 software. After finishing the concept, it is executed by designing a measurement model and set up the path diagram. Then, the last process is evaluated the model and tested the hypotheses.

The SEM-PLS model needs two steps to gain the fit model; model evaluation and structural model analysis. The model evaluation is measured by testing the validity and reliability of the model. The validity test needs three criteria; Convergent Validity, Average Variance Extracted (AVE), and Discriminant Validity. The criterion of reliability is Composite Reliability [34].

Convergent validity is measured based on the correlation between the indicator value and the variable value. The weight value is represented as loading factor or cross loading factor value. The loading factor is high if the value is more than 0.7. But for the initial research, the measurement scale with a loading value of 0.5 to 0.6 is considered as sufficient.

Average Variance Extracted (AVE) is the average of percentage extracted variance value from latent variable that is estimated through the loading standardize. The minimum AVE value is 0.5 (1).

$$AVE = \frac{\sum_{i=1}^n \lambda_i}{n} \quad (1)$$

Where  $\lambda$  is standardize loading factor,  $i$  means amount of indicator and  $n$  is amount of data.

Discriminant Validity can be calculated by comparing the square root of Average Variance Extracted (AVE) values. If the square root of AVE value is higher than the correlation value between latent variables, then discriminant validity can be considered as fulfilled. Then, the criterion of reliability model is Composite Reliability. The latent variable can achieve the reliability if minimum the Composite Reliability value is 0.6.

2.4. Structural Model Analysis

The structural model analysis is set up to know the relationship among constructed model, the significance value, and the model itself. The testing criteria is significance probability using R-squared ( $R^2$ ). R-Squared is a method to measure Goodness of Fit (GoF) and to appraise the influence independent latent variable with dependent latent variable. After finishing the structural model analysis, the bootstrapping process is conducted. The bootstrapping process is functioned to minimize the abnormality of the modelled study.

2.5. Model Hypothesis

The social, economic, and environmental criteria are used to set as an exogen variables and excellent public transportation as an endogen variable. Each category is built up from related presumptions that are set with significant alfa 5%. The hypothesis of the model is,

- H2:** The economic variable has positive or significance correlation with excellent public transportation system.
- H3:** The social variable has positive or significance correlation with excellent public transportation system.
- H4:** The environment variable has positive or significance correlation with excellent public transportation system.

3. Result and Discussion

3.1. Demographics of The Non-ridership BRT

In Purwokerto and Purbalingga, the non-ridership respondent of the survey is 138. The demographics of them can be seen in Table 1.

Table 1: Demographics of Non-ridership BRT

User variable		Distribution (%)
Gender	Men	60.87
	Women	39.13
Residence	Purwokerto	43.48
	Purbalingga	49.28
	Others	7.25
Job	Student	26.09
	Teacher	2.90
	Labor	7.97
	Housewife	5.80
	Others	57.25
Age Ratio	15-25	30.43
	26-35	15.22
	More than 35	54.35
Attempt	Tried	62.32
	Never	37.68

According to Table 1, non-ridership gender is dominated by men at 60.87%, and non-ridership women at 39.13%. The dominated non-ridership is caused by men to work at informal economic sector around the bus stop. Most of them rely on the crowded street to gain the customer. They work as food street sellers, fruit sellers, parking guards, and pedicab drivers. The percentage is high at 57.25% for the job sector. It is described in Table 1 as other job. Meanwhile, most non-ridership women work as students (26.09%), workers (7.97%), housewife (5.8%), and teachers (2.90%). So, their availability around the BRT stops is lower than non-ridership men. Then, most of their age ratio is more than 35. The age ratio is productive worker age who become a family mainstay. The rest is the age ratio between 15 and 25 and between 26 and 35. The age ratio between 15 and 25 is dominated by student and between 26 and 35 is dominated by labor and housewife.

Regarding the residence area, most of them are living in Purbalingga (49.28%), followed by Purwokerto (43.38%) and others (7.25%). Other places mean the non-ridership is living around both cities such as Banyumas, Banjarnegara, and Cilacap Regency. The percentage of Purbalingga is higher than Purwokerto and other places mean that, most of the non-ridership works or stay around the BRT stop. Furthermore, the BRT route in Purbalingga is longer than Purwokerto, so the covered area of the BRT is wider than Purwokerto. People in Purbalingga is more often using the BRT than Purwokerto and others because of the BRT can connect the industrial park from their living residence. The majority of respondents have tried the BRT (62.32%). They tried the BRT for many purposes, such as visiting their family, going to work or school, and just for a fun activity.

Table 2: Non-Ridership Presumptions toward BRT

Code	Presumption
Q1	BRT capacity is suitable to fulfill demands
Q2	BRT is a reliable public transportation
Q3	BRT service is better than other public transportation
Q4	BRT accommodates ridership destination
Q5	BRT routes shall be expanded
Q6	BRT can be alternative public transportation
Q7	BRT provider is followed transportation procedure service
Q8	BRT is environment friendly transportation.
Q9	BRT is on-time schedule transportation.
Q10	BRT can increase economic productivity for local people
Q11	BRT can open economic opportunity for local people
Q12	BRT can open additional job opportunity
Q13	BRT can decrease the traffic congestion
Q14	BRT improve local people mobility
Q15	BRT can reduce traffic accident
Q16	BRT can reduce fossil fuel use.
Q17	BRT facilitates to access intermodal transportation
Q18	BRT stops are able to be added and expanded.
Q19	Ticket payment shall be set e-payment

3.2. Non-Ridership Presumption

The experience of non-ridership lead to their presumption. The presumption depicts society or community knowledge of implementing BRT. The presumption is used to measure non-ridership participation and awareness even though they have never tried the BRT service.

According to Table 2, there are several criteria of non-ridership presumptions that categorized into 19 presumptions. The queries are divided into four categories. The queries are set based on the non-ridership observations toward the BRT services around their living area. First presumption is economic. The presumption is focused on punctuality of the BRT schedule (Q9), economic productivity for local (Q10), economic opportunity for local (Q11), and additional job opportunity for local (Q12).

The second presumption is social. The presumption is set to measure that the BRT can decrease traffic jams in the cities (Q13), enhance local mobility or movement (Q14) and reduce traffic accidents (Q15). Then, the environment is set for seeking the knowledge of local about green transportation (Q8) and reducing fossil fuel usability in transportation (Q16). The last is a non-ridership presumption about common presumption about excellent public transportation. The limitation of excellent public transportation is giving impact to the non-ridership living and their social activities. Excellent public transportation presumption is built up from fulfilling demands based on BRT capacity (Q1), being reliable public transportation (Q2), being excellent service bus than the existing bus (Q3), accommodating ridership destination (Q4). Moreover, BRT route shall be expanded (Q5), the BRT can be alternative transportation (Q6), the BRT provider following the service procedure (Q7), BRT facilitates to access intermodal transportation (Q17), BRT stops are able to be added and expanded (Q18), and implementing e-payment ticket (Q19).

According to Table 3, the result of validity test of the questionnaire (Q1-Q19) are valid. It is shown that the Pearson Correlation values are higher than Pearson Correlation Table (0.455) except Q19. However, Q19 is still included into measurement because the p-value is 0.01 (less than 0.05). The result of reliability test is reliable. It is shown from the table that the Cronbach' Alpha is 0.9113 where the value is higher than 0.7 as standard minimum of reliability.

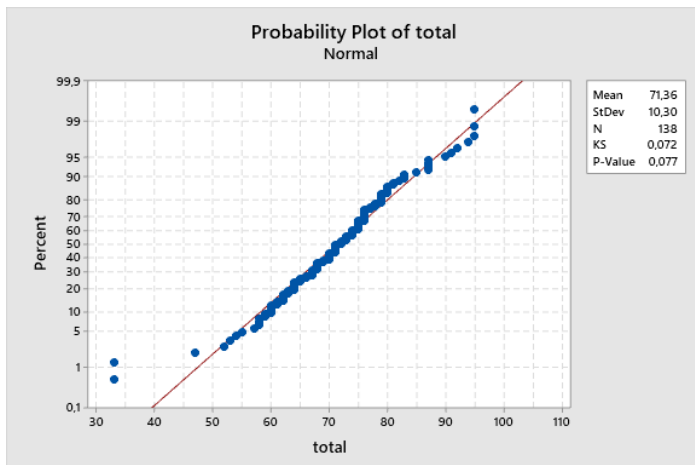


Table 3: Validity and Reliability of The Questionnaire Result

Code	Pearson Correlation	P-value	Cronbach' Alpha
Q1	0.631	0.000	0,9113
Q2	0.637	0.000	
Q3	0.619	0.000	
Q4	0.627	0.000	
Q5	0.614	0.000	
Q6	0.747	0.000	
Q7	0.682	0.000	
Q8	0.648	0.000	
Q9	0.553	0.000	
Q10	0.678	0.000	
Q11	0.695	0.000	
Q12	0.640	0.000	
Q13	0.650	0.000	
Q14	0.709	0.000	
Q15	0.559	0.000	
Q16	0.564	0.000	
Q17	0.665	0.000	
Q18	0.562	0.000	
Q19	0.286	0.001	

The multicollinearity and outliers test result can be seen in Table 4. According to the table, all of indicator has Varian Inflation Factor (VIF) value that less than 10. It means that there is multicollinearity among the indicators and it guarantees that there is no multicollinearity problem. The outliers test indicates all of p-value are more than 0.05. It means that there is no outlier data. The normality test of the data is shown in Figure 1. Regarding to the figure, the data is followed normal distribution rules because the p-values is 0.077 or greater than 0.05. So, all collected data is sufficient and qualify referring to the statistical rules.

Table 4: Multicollinearity and Outlier Test Result

Code	Multicollinearity	Outlier	
	VIF	G	P-values
Q1	2,18	2,34	1,000
Q2	2,41	2,72	0,812
Q3	1,69	2,81	0,603
Q4	1,64	4,16	0,266
Q5	1,72	3,54	0,355
Q6	2,42	4,40	0,108
Q7	2,33	4,39	0,110
Q8	2,13	3,44	0,061
Q9	1,57	2,65	1,000
Q10	2,34	2,69	0,901
Q11	2,68	2,56	1,000
Q12	1,53	2,75	0,736
Q13	1,96	2,32	1,000

Q14	1,88	3,94	0,869
Q15	1,83	1,83	1,000
Q16	1,61	3,19	0,159
Q17	1,83	2,23	1,000
Q18	1,75	3,00	0,321
Q19	1,19	1,55	1,000

### 3.3. ANOVA Result

After getting the presumptions data, ANOVA is used to test the significance between demographic factors and the non-ridership presumptions. The result is shown in Table 5.

Table 5: ANOVA Result

Code	Significance Factor	P-value
Q1	-	-
Q2	-	-
Q3	-	-
Q4	Gender	0.035
Q5	-	-
Q6	-	-
Q7	-	-
Q8	-	-
Q9	Residence	0.035
Q10	Residence	0.001
Q11	Residence	0.008
Q12	-	-
Q13	Age Ratio	0.023
Q14	Residence	0.006
Q15	Residence	0.018
Q16	Residence	0.000
Q17	-	-
Q18	-	-
Q19	Age Ratio	0.000

According Table 5, gender is significant factor for accommodation ridership destination (Q4). In this case, gender correlates with public transportation. Different gender has different travel pattern in term of using public transportation especially, for non-ridership women. Sometimes, women use more frequently use public transportation than men, so they think that the BRT can transport them to their destination.

Age ratio is a significant factor for reducing traffic congestion (Q13) and implementing e-payment for the BRT (Q19). At traffic congestion, the member of age ratio has different experience toward the BRT service. Before implementing the BRT, some of them experienced that the road condition is crowded, then they hope that after implementing the BRT, the congestion will be reduced. Furthermore, they also hope for the person who drive a private car or ride a motorcycle will switch to use the BRT as the main modes of transportation. Moreover, at e-payment system, more than 35 years old age ratio has different choice. They prefer to use cash rather than e-payment or using an e-card that is

implemented recently. Regular payment is more interesting because they do not need to install any application in mobile phones or bringing e-card for tapping. The rest is age ratio 15-25 and 26-35 are prefer to use a trended application such as e-money or e-payment because most of them have mobile phones or e-money cards.

In the context of living areas for non-ridership, it has significant factor for many queries. Residence correlates with the punctuality of the BRT schedule (Q9), increasing economic productivity for local (Q10), economic opportunity for local (Q11), improving local mobility (Q14), reducing traffic accident (Q15), and reducing fossil fuel use (Q16).

The different city has different presumption. Related to the punctuality of the BRT schedule, non-ridership who stay at Purbalingga has different presumptions with Purwokerto and the other cities. It means that the schedule is not always precise. According to the BRT service, the bus is mobilized every 15 minutes for regular hours and every 10 minutes for peak hours. The BRT does not have an exact schedule because of some limitations such as the quantity of the bus, road condition, traveling time, and break time. In some cases, the BRT will not pick up the ridership because of full capacity, so the ridership shall wait next BRT. The next presumption is enhancing local economic productivity. Purbalingga and the other cities have different presumptions with Purwokerto. The non-ridership in Purbalingga and the other cities think that BRT will increase their productivity because the places have many factories such as artificial eye flashes, wig, and motorcycle exhaust. Many of them are labor at the industries that access to industrial park. On the other hand, Purwokerto has fewer factories than Purbalingga and the other cities, yet it has more offices and banks. So, most of the people at Purwokerto work as private or public employees rather than labor. Moreover, the non-ridership hopes that there will be opened new economic access that connects government office, industrial park, bus terminal, railway station, and shopping center.

Next presumption is improving job opportunity for local people. The BRT is seen as a new job vacation even though there is a limited job opportunity in the BRT. The non-ridership hope by implementing the BRT, it will involve local participation and support. The non-ridership expects there will be additional linkage for local to access to gain a new job. Furthermore, the idea is similar to improving non-ridership mobility. By increasing local mobility, the non-ridership will have any public transportation a choice. They can take advantage of the BRT existence or not. At least, they have choice for public transportation that regular and low fare, but if their destination cannot be facilitated, they will choose other public transportation or private vehicle. In terms of opportunity for mobilizing people, among the cities has different presumptions. The non-ridership in Purbalingga and outside both cities will rely on the BRT for fulfilling their mobility Purwokerto is not too reliant on it.

The next presumption that has significant factor is reducing traffic accidents. The BRT is set for being alternative mass transportation. The logic is if many people use the BRT, the traffic is reduced because the private vehicle is few. The non-ridership in Purbalingga think that the BRT will reduce the traffic accident because many labours rely on the BRT. The BRT is more safety than using private vehicles but, the rest living area said that the BRT are not too rely on because the traffic is still lower than Purbalingga. So, in this case, private car is more suitable for going



to work or back home. The last presumption is the BRT is reduced fossil fuel consumption. This is one of a promotion for being sustainable transportation. The case is similar to reducing traffic accidents. Non-ridership thinks by reducing private vehicle the fossil fuel consumption will be low. Practically, the BRT is still not covered all destinations especially in some public area such as the central traditional market and railway station. So, it needs to develop more routes to reach the uncovered area, so that additional ridership can be facilitated. The non-ridership presumptions in Purbalingga is more agree than non-ridership who are living in the outside both cities and Purwokerto.

3.4. Convergent Validity

After setting up the hypothesis, we develop a model for analysing the correlation among environment, social, and economic criteria with excellent public transportation criteria using SEM- PLS. The model is constructed using Smart PLS 3 software (Figure 2).

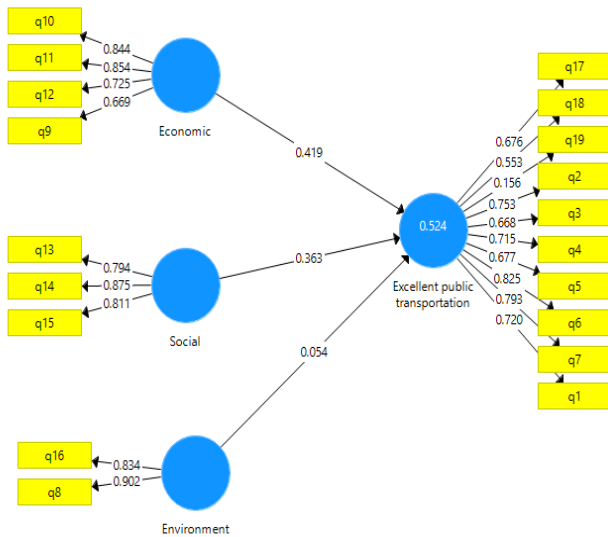


Figure 2: Initial Constructed Model

After the model is constructed, we run the model and get the convergent validity result. Convergent validity is measured based on the correlation among the estimated score component. The value is high if the weight is more than 0.7. In some case for early-stage research, the measurement scale between 0.5 to 0.6 is reliable enough [34]. The result can be seen in Table 6.

Table 6: Initial Outer Loading (Cross-Loading Factor) Result

Code	Economic	Environment	Excellent Public Transportation	Social
Q10	0.844			
Q11	0.854			
Q12	0.725			
Q13				0.794
Q14				0.875
Q15				0.811
Q16		0.834		

Q17			0.676	
Q18			0.553	
Q19			0.156	
Q2			0.753	
Q3			0.668	
Q4			0.715	
Q5			0.677	
Q6			0.825	
Q7			0.793	
Q8		0.902		
Q9	0.669			
Q1			0.720	

According to the theory, the limit of the cross-loading factor is 0.5, and we remove the factor that less than 0.5. It is found that Q19 has a cross-loading factor less than 0.5, or the value is 0.159. So, we eliminate from the constructed model and run the software again. The result is all of the loading factor values are higher than 0.5. It means that fulfilled required convergent validity. Thus, all indicators from the model are appropriate to depict each measured variable (Table 7).

Table 7: Final Outer Loading (Cross Loading Factor) Result

Code	Economic	Environment	Excellent Public Transportation	Social
Q10	0.844			
Q11	0.855			
Q12	0.724			
Q13				0.792
Q14				0.876
Q15				0.812
Q16		0.833		
Q17			0.676	
Q18			0.542	
Q2			0.757	
Q3			0.668	
Q4			0.717	
Q5			0.674	
Q6			0.827	
Q7			0.800	
Q8		0.903		
Q9	0.667			
Q1			0.724	

3.5. Average Variance Extracted (AVE)

The higher AVE value means, the better of correlation among the constructed model. The minimum required values of AVE are 0.5. According to Table 8, all of the variables have AVE values higher than 0.5. So, it can be mentioned that all of the variables are suitable to fulfil minimum requirement.

Table 8: Average Variance Extracted (AVE) Result

Variable	Cronbach's Alpha	rho A	Composite Reliability	Average Variance Extracted (AVE)	R-square
Economic	0.777	0.794	0.858	0.603	
Environment	0.679	0.708	0.860	0.754	
Excellent Public Transportation	0.877	0.884	0.902	0.510	0.518
Social	0.772	0.801	0.867	0.684	

3.6. Discriminant Validity

Discriminant validity measured how far the differences of a construct with other constructs. High value of discriminant validity can be determined as a special or unique variable and described the measured phenomenon.

According to Table 9 and Table 10, we compare the latent variable covariance with the square root of AVE. It is depicted that all of square root of AVE value are higher than latent variable covariances. So, discriminant validity criteria are fulfilled for all variables.

Table 9: Latent Variable Covariances

Variable	Economic	Environment	Excellent Public Transportation	Social
Economic	<b>1.000</b>	0.569	0.640	0.535
Environment	0.569	<b>1.000</b>	0.511	0.606
Excellent Public Transportation	0.640	0.511	<b>1.000</b>	0.618
Social	0.535	0.606	0.618	<b>1.000</b>

Table 10: Square Root of AVE

Variable	AVE	√(AVE)	Discriminant Validity
Economic	0.603	0.776531	Fulfilled
Environment	0.754	0.868332	Fulfilled
Excellent Public Transportation	0.51	0.714143	Fulfilled
Social	0.684	0.827043	Fulfilled

3.7. Composite Reliability

Reliability test is done by calculating composite reliability. The minimum value of the composite reliability is 0.6. according to Table 8, all of the construct indicator has higher value than 0.6, so it can be defined that the model has high reliability and reliable. Furthermore, Cronbach's Alpha is set an additional parameter. The minimum of Cronbach's Alpha values in the variable to be reliable is 0.7. According to Table 8, all of Cronbach's Alpha values are higher than 0.7, except the environment variable (0.679). So, the

reliability of three variables (economic, social, and excellent public transportation) is high, but the environment variable has low reliability.

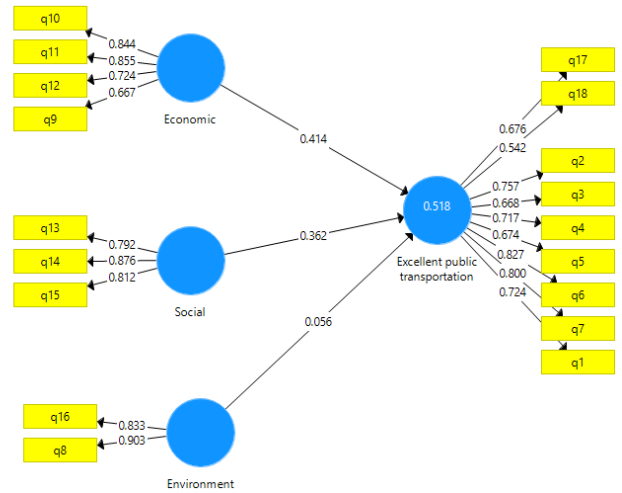


Figure 3: Final Structured Model

3.8. Structural Model Analysis

The subsequent analysis is a structural model. The analysis is conducted to measure the correlation among the constructs or variables, significance, R-squared. According to Figure 3, it is depicted that the number between the yellow rectangle and the blue circle (on the arrow) is cross loading factors value or outer loading value of the model. The description of the value can be referred to Table 7. Furthermore, the values between the blue circle of economic, social, and environment variable and the blue circle of excellent public transportation (on the arrow) is the Original Sample (O) value. Those values can be seen in Table 11. The number on the blue circle of excellent public transportation is R-squared value. The value is 0.518, it means that the expectation of contribution toward BRT execution is at moderate level. The rest is affected by other factors from the outside that are not described in this study.

Table 11: Coefficient Path Structured Model

Variable	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (O/STDEV)	P-Values
Economic -> Excellent Public Transportation	0.414	0.422	0.072	5.745	0.000
Environment -> Excellent Public Transportation	0.056	0.056	0.102	0.549	0.583
Social -> Excellent Public Transportation	0.362	0.364	0.078	4.672	0.000

The differences of Figure 2 (initial model) and Figure 3 (final model) is at removing the Q19 indicator. The removed Q19 is caused by the loading factor in convergent validity measurement is less than 0.5. The effect is the final model is more valid than the

initial model, but the R-squared of the initial model (0.524) is a little bit higher than the final model (0.518). Furthermore, the Original Sample (O) value of the initial model is not entirely different from Original Sample (O) value of the final model.

After finishing the model structural analysis, it is processed the bootstrapping. The result can be seen in Table 11. According to the table, economic variables have t-statistics at 5.745, or greater than 1.96 (alfa is 5%) or p-value is less than 0.05 (0.000). It means that the economic variable has a significant factor with excellent public transportation variable. Furthermore, the social variable has t-statistics at 4.672, and the p-value is 0.000. It means that the social variable has a significant factor toward excellent public transportation variable. Meanwhile, the environment variable has t-statistics less than 1.96 (0.549) and p-value is at 0.583 (greater than 0.05). It means that environment variable has no significant factor for excellent public transportation variable. So, only the economic and the social variables which have a significant factor for excellent public transportation variable.

Regarding the Original Sample (O) value, the correlation among those variables shows that the economic and the social variables has a significant correlation with excellent public transportation variables. The value is 0.414 and 0.362. However, for environment variable does not have a significant correlation with excellent public transportation variables and the value is 0.056. So, the economic variable affects the excellent public transportation variable is at 41.4%, and the social variable is at 36.2%. Nevertheless, the environment variable does not affect to excellent public transportation variables.

### *3.9. Implication and Recommendation*

The BRT is new and has been followed the service procedure to fulfil ridership service. The objective of the BRT is to connect ridership and local community or non-ridership. Ridership means people who use the BRT frequently. Non-ridership means the people whether they use the BRT or not, but they get the impact of the BRT execution. In this case, they are defined as marginal community. The role of them is important to support the BRT sustainability and enhance the public engagement.

The demographics result depicts that most non-ridership is non-formal occupations such as street food sellers and pedicab drivers. Furthermore, most of them are never tried the BRT. It is indicated that the BRT implementation or execution belongs to ridership in terms of the service. Indeed, the BRT is set for ridership, but non-ridership shall get the impact at least increasing the opportunity to access, and gain the new transportation mode, economic opportunity, and positive social impact.

According to the survey result of the correlation between demographics with presumptions, living areas of the non-ridership are more frequent, giving significant factors, followed age ratio and gender. The different living areas of non-ridership leads to different presumptions about the BRT schedule, economic productivity, economic opportunity, local mobility, reducing traffic accidents, and low fossil fuel use. Most of the non-ridership in Purbalingga is labour. They rely on public transportation that supported their mobility. So, they need less waiting time at the BRT stop. The BRT provider shall adjust the condition by adding more frequent bus schedules in Purbalingga at peak hours [35].

Economic productivity and opportunity factors depict that Purbalingga has lower economic status than Purwokerto. The

economic status is the ability of the non-ridership to access the industry or economic centre [30]. Actually, Purbalingga has more factories, but Purwokerto is more developed than Purbalingga, such as ease of public transportation access and strategic geographical area. So, the BRT shall open new opportunities for non-ridership to open their economic access and develop Purbalingga to be similar economic status with Purwokerto. It can be done by promoting the tourism object, traditional food, museum, and some interesting places, so that increase non-ridership economic opportunities.

Reducing traffic accidents and using fossil fuel less, the BRT can be a pioneer of public transportation that excellent for safety driving and green transportation. It will lead the use of excellent vehicles standard to support both cities to be more eco-friendly mobility and increase local or non-readership healthy.

Meanwhile, the age ratio factor influences different presumption among the non-ridership in terms of traffic congestion and e-payment. In term of traffic congestion, a different age ratio has different background. So, some of non-ridership who use private vehicle trust that the BRT can reduce traffic congestion, especially at peak hours. Furthermore, some non-ridership agree that the BRT shall provide an e-payment system. Actually, the payment system will ease the provider to record the ridership occupancy. So, it can be reliable transport and reach all society segmentation in terms of preparing a new model of the payment system.

The last is the gender that has significant presumption in term of accommodating their destination. Recently, the BRT facilities shall be improved, mainly to attract women. The non-ridership women hope that the BRT can accommodate their destination with appropriate and comfort facilities [36]. So, it will reduce non-ridership women using a private vehicle [37].

Referring to the structured model from SEM-PLS, it is indicated that economic and social variable have significant factors to excellent public transportation variable. It means that the non-ridership is strongly agree that the social and economic variables must be applied with BRT execution. The social and economic variables are perceived by non-ridership, whether direct or indirect effect such as enhance social mobility, reduce traffic accidents and congestion, and opening economic opportunity and access. According to the phenomenon, the BRT provider shall capture the opportunity by enhancing the service onboard and offboard. Onboard service is for ridership and off-board service is non-ridership. Onboard service is universal and the procedure is standardized but off-board service can be set as a sustainable society development program. Environment variable does not have a significant correlation with the excellent public transportation variables. It is indicated that the availability of the BRT is far from the green transportation campaign. Indeed, the primary purpose of the BRT is minimizing traffic congestion and promoting sustainable transportation with low CO2 emissions [38]. However, by implementing green public transportation, the BRT provider shall invest more and sometimes it will not be required recently, especially in developing countries.

According to the structural model from SEM-PLS for excellent public transportation variables, the expectation of contribution toward the BRT execution is 51.8%. It is moderate level and still far from the expectation. However, it needs more empirical study to improve the percentage, so that the phenomenon can be captured wider.

The non-ridership as marginal community is taken to offer significant improvement service by the BRT provider. The BRT provider can improve the journey time, journey time reliability, and realistic alternative mass transportation [29]. By knowing the presumption of the non-ridership, the BRT provider can purpose some program to enhance the BRT occupancy and improve the service. Meanwhile, the social, economic, and environment benefits can be gained to attract more public engagement among the programs. The program can be designed as facilitating economic opportunity program, social development program, and promoting green transportation program.

#### 4. Conclusions

BRT in Purwokerto and Purbalingga is new public transportation that is operated by the Ministry of Transport. As the new public transportation, the support from the non-ridership is a compulsory to maintain the sustainability. According to the analysis, it gives the insightful results;

- 1) The living area has dominance influenced factors, followed age ratio and gender about the BRT condition. The living area covers the presumption about the BRT schedule, economic opportunity and productivity, local mobility access, reducing traffic accident, and fossil fuel use. For the age ratio, it covers traffic congestion and e-payment system, and the gender covers the destination presumption.
- 2) The SEM-PLS model shows that social and economic criteria have correlation with excellent public transportation criteria, but not with environment criteria. So, it can be a suggestion to the BRT provider to concern more environmental issue to raise the community awareness and the public engagement.

In the BRT provider side, the suggestion will lead to purpose some programs to improve the service and increase the occupancy. In the non-ridership side, the proposed program will attract the sense of awareness and the public engagement from the BRT execution. Furthermore, the model is demonstrated at the moderate percentage, so the future research in this area can be directed to add more variables to reach excellent public transportation. The added variables such as specific involved communities, political will, regulation, gender equality, economic opportunity, and local culture to capture more public engagement.

#### Conflict of Interest

The author declares no conflict of interest.

#### Acknowledgment

This acknowledgement is given to Ministry of Education and Culture Republic of Indonesia for the funding. We thank so much to the BRT ridership who gave free time to fill our questionnaire and also surveyor team from Industrial Engineering and Information System Department, IT Telkom Purwokerto, Central Java, Indonesia.

#### References

[1] O. Afolabi, "Commuters Perception and Preferences on The Bus Rapid Transit in Lagos State," *JORIND*, **14**(2), 2016.

[2] J. Cao, X. Cao, C. Zhang, X. Huang, "The gaps in satisfaction with transit services among BRT, metro, and bus riders: Evidence from Guangzhou," *Journal of Transport and Land Use*, 97–109, 2015,

doi:10.5198/jtlu.2015.592.

[3] D. Darwis, Fatmawati, Nasrulhaq, "Hubungan Kualitas Pelayanan dengan Kepuasan Masyarakat Pengguna Bus Rapid Transit (BRT) Maminasata," *Administrasi Publik*, **3**, 2017.

[4] J.M. Bunker, "High volume bus stop upstream average waiting time for working capacity and quality of service," *Public Transport*, **10**(2), 311–333, 2018, doi:10.1007/s12469-018-0179-1.

[5] M. Islam, M. Brussel, A. Grigolon, T. Munshi, "Ridership and the Built-Form Indicators: A Study from Ahmedabad Janmarg Bus Rapid Transit System (BRTS)," *Urban Science*, **2**(4), 95, 2018, doi:10.3390/urbansci2040095.

[6] D. Oviedo, L. Scholl, M. Innao, L. Pedraza, "Do Bus Rapid Transit Systems improve accessibility to job opportunities for the poor? The case of Lima, Peru," *Sustainability (Switzerland)*, **11**(10), 2019, doi:10.3390/su11102795.

[7] D.A. Rodriguez, C.E. Vergel-Tovar, "Urban development around bus rapid transit stops in seven cities in Latin-America," *Journal of Urbanism*, **11**(2), 175–201, 2018, doi:10.1080/17549175.2017.1372507.

[8] E. Saputra, C. Widyasmara, "The effect of urban spatial structure on rapid bus transit services in Yogyakarta and Surakarta, Indonesia: A comparative study of the Trans Jogja and the Batik Solo Trans," *Geografia - Malaysian Journal of Society and Space*, **10**(2), 46–58, 2014.

[9] S. Tao, J. Corcoran, I. Mateo-Babiano, D. Rohde, "Exploring Bus Rapid Transit passenger travel behaviour using big data," *Applied Geography*, **53**, 90–104, 2014, doi:10.1016/j.apgeog.2014.06.008.

[10] L. Prayogi, "The Influence of Bus Rapid Transit System on Urban Development: An Inquiry to Boston and Seoul BRT Systems," *International Journal of Built Environment and Scientific Research*, **1**(1), 1–8, 2017.

[11] A. Koling, W.-B. Zhang, K. Zhou, H. Meng, *Bus Rapid Transit ( BRT ) Toolbox : Assessing Person Throughput to Measure Transportation Impacts for BRT Projects*, 2018.

[12] S. Tabassum, S. Tanaka, F. Nakamura, "Improving Access considering Commuters ' Perception ( A Case Study of Lahore BRT ) Keywords :," *International Journal of Innovative Research in Engineering and Management (IJIREM)*, **3**(4), 283–289, 2016.

[13] M. Woldeamanuel, C. Olwert, "The Multimodal Connectivity at Bus Rapid Transit (BRT) Stations and the Impact on Ridership," *Journal of the Transportation Research Forum*, **55**(1), 87–102, 2016, doi:10.5399/osu/jtrf.55.1.4340.

[14] B. Ferhan, X. Zhou, "Drivers , Barriers and Implementation of Social Sustainable in Supply Chain A qualitative study of SMEs," 1–83, 2016.

[15] S. Nasrin, "Acceptability of Bus Rapid Transit ( BRT ) to Commuters in Dhaka," 2015.

[16] H.I. Okagbue, M.O. Adamu, S.A. Iyase, E.A. Owoloko, "On the Motivations and Challenges Faced by Commuters Using Bus Rapid Transit in Lagos, Nigeria, *The Social Sciences*, **10**(6), 696–701, 2015, doi:10.3923/sscience.2015.696.701.

[17] J. Dubé, F. Des Rosiers, M. Thériault, P. Dib, "Economic impact of a supply change in mass transit in urban areas: A Canadian example," *Transportation Research Part A: Policy and Practice*, **45**(1), 46–62, 2011, doi:10.1016/j.tra.2010.09.002.

[18] J.B. Ingvardson, O.A. Nielsen, "Effects of new bus and rail rapid transit systems—an international review," *Transport Reviews*, **38**(1), 96–116, 2018, doi:10.1080/01441647.2017.1301594.

[19] E. Adewumi, D. Allopi, "An Appropriate Bus Rapid Transit System," *International Journal of Science and Technology Volume 3 No. 4*, **3**(4), 248–254, 2014.

[20] A.P. Ajayi, "Assessment of the Influence of Lagos Bus Rapid Transit Scheme (BRT-Lite) on Road Traffic Crashes (RTC) on Lagos Mainland-Island Corridor," *The Open Transportation Journal*, **11**(1), 102–109, 2017, doi:10.2174/1874447801711010102.

[21] A. Nikitas, M. Karlsson, "A Worldwide State-of-the-Art Analysis for Bus Rapid Transit: Looking for the Success Formula," *Journal of Public Transportation*, **18**(1), 1–33, 2015, doi:10.5038/2375-0901.18.1.3.

[22] A. Mishra, S.A.S.A. Kumar, P. Pradeep, "Study of Bus Rapid Transit system In Respect to Growing Cities of India," *International Journal of Engineering Research & Technology (IJERT)*, **2**(10), 1378–1383, 2013.

[23] M. Poku-Boansi, G. Marsden, "Bus rapid transit systems as a governance reform project," *Journal of Transport Geography*, **70**(June), 193–202, 2018, doi:10.1016/j.jtrangeo.2018.06.005.

[24] S.R. Majumdar, "The case of public involvement in transportation planning using social media," *Case Studies on Transport Policy*, **5**(1), 121–133, 2017, doi:10.1016/j.cstp.2016.11.002.

[25] E. Cascetta, F. Pagliara, "Public Engagement for Planning and Designing Transportation Systems," *Procedia - Social and Behavioral Sciences*, **87**,



- 103–116, 2013, doi:10.1016/j.sbspro.2013.10.597.
- [26] K. Rangarajan, S. Long, A. Tobias, M. Keister, “The role of stakeholder engagement in the development of sustainable rail infrastructure systems,” *Research in Transportation Business and Management*, **7**, 106–113, 2013, doi:10.1016/j.rtbm.2013.03.007.
- [27] N. Hassen, P. Kaufman, “Examining the role of urban street design in enhancing community engagement: A literature review,” *Health and Place*, **41**, 119–132, 2016, doi:10.1016/j.healthplace.2016.08.005.
- [28] J.R. Gardner, R. Rachlin, H.W.A. Sweeney, A. Sweeny, H.W. Allen, *Handbook of Strategic Planning*, Wiley, 1986.
- [29] D. Mobereola, *Lagos Bus Rapid Transit Africa ’ s first BRT scheme (The Lagos BRT-Lite System)*, 2009.
- [30] U. Dutta, “A Framework to Identify Economic and Community Benefits of the Proposed BRT System in Southeast Michigan,” *Current Urban Studies*, **05(02)**, 164–184, 2017, doi:10.4236/cus.2017.52010.
- [31] N. Duduta, C. Adriazola, D. Hidalgo, L.A. Lindau, R. Jaffe, “Understanding road safety impact of high-performance bus rapid transit and busway design features,” *Transportation Research Record*, **2317**, 8–14, 2012, doi:10.3141/2317-02.
- [32] V.J. Racehorse, G. Zhang, A. Sussman, A. Jian, T. Parker, “Bus rapid transit system deployment for high quality and cost-effective transit service: A comprehensive review and comparative analysis,” *IET Intelligent Transport Systems*, **9(2)**, 175–183, 2015, doi:10.1049/iet-its.2013.0176.
- [33] D. Prabowo, *BRT Purwokerto Terus Tambah Halte*, Radar Banyumas, 2019.
- [34] A. Rifai, “Partial Least Square-Structural Equation Modeling (PLS-SEM) untuk mengukur ekspektasi penggunaan repositori lembaga: Pilot studi di UIN Syarif Hidayatullah Jakarta,” *Al-Maktabah*, **14(1)**, 56–65, 2015.
- [35] P. Starkey, J. Hine, *Poverty and Sustainable Transport How Transport Affects Poor People with Policy Implications for Poverty Reduction: A literature review*, 2014.
- [36] D. Martinez, E. Salgado, P. Yañez-pagans, *Connecting to Economic Opportunity? The Role of Public Transport in Promoting Women ’ s Employment in Lima*, 2018.
- [37] K. Terayama, M. Odani, “Expected role of public transportation services in securing residents ’ accessibility to the city center in suburban housing development areas,” *Transportation Research Procedia*, **25**, 4258–4269, 2017, doi:10.1016/j.trpro.2017.05.237.
- [38] L. Velazquez, N.E. Munguia, M. Will, A.G. Zavala, S.P. Verdugo, “Sustainable transportation strategies for decoupling road vehicle transport and carbon dioxide emissions,” **26(3)**, 373–388, 2015, doi:10.1108/MEQ-07-2014-0120.

## Mentoring Model in an Active Learning Culture for Undergraduate Projects

Wongpanya Nuankaew<sup>1</sup>, Kanakarn Phanniphong<sup>2</sup>, Sittichai Bussaman<sup>3</sup>, Direk Teeraputon<sup>4</sup>, and Pratya Nuankaew<sup>5,\*</sup>

<sup>1</sup>Faculty of Information Technology, Rajabhat Mahasarakham University, Maha Sarakham, 44000, Thailand

<sup>2</sup>Faculty of Business Administration and Information Technology, Rajamangala University of Technology Tawan-Ok, Chonburi, 20110, Thailand

<sup>3</sup>Faculty of Science and Technology, Rajabhat Mahasarakham University, Maha Sarakham, 44000, Thailand

<sup>4</sup>School of Education, University of Phayao, Phayao, 56000, Thailand

<sup>5</sup>School of Information and Communication Technology, University of Phayao, Phayao, 56000, Thailand

### ARTICLE INFO

Article history:

Received: 17 June, 2020

Accepted: 17 August, 2020

Online: 28 August, 2020

Keywords:

Active Learning Culture  
Educational Data Mining  
Engineering Education  
Learning Analytics  
Lifelong Learning  
Mentoring Model

### ABSTRACT

Senior projects allow students to move the learning process from basic knowledge to an interdisciplinary approach. The purpose of this research is (1) to analysis attitude and perception, which is a collaboration between teachers and students to develop a model for clustering of appropriate advisors and advisee who cooperate in senior project, and (2) to develop factors that are significant to predict the right match in senior projects course. Data collection was a questionnaire consisting of 463 samples from 7 administrators, 68 lecturers, 26 staff and 362 students from two institutions: The Rajabhat Mahasarakham University, and the University of Phayao. The research methodology was designed and divided into three sections: preparation, implementation, and conclusion. The result shows that the satisfaction and the overall acceptance level were at a high level (mean = 4.04, S.D. = 0.88). Moreover, the developed model has the highest level of efficiency (accuracy = 98.06%). While the factors that are important for matching recommendations consists of 9 factors: policies of the organization, vision of the organization, mission of the organization, experience and achievements of researchers, qualifications of research team, interest in the research topics, impressions and examples in the past, technology and laboratory support, and budget support. For the future, the researchers are aimed to research on the development of students' academic achievement and aims to promote a learning culture based on the results of this research.

## 1. Introduction

In the world of learning and educational theories, developing learners to understand research problems and constructing the concepts to create research problems is most difficult and complicated in managing. Learners often have different learning styles, which are classified according to learner characteristics [1–4]. At the same time teachers have skills, abilities, experience, knowledge, expertise, and aptitude in different contexts [1]. Thus,

the response that occurs between the student and the teacher is important. It can also be used to develop a structured relationship for collaboration between students and teachers through feedback and perception of both parties [3,4].

The problem is related to the context of the curriculum that has been instructed in the field of information technology and other disciplines, which is concerned with the development of students' knowledge by using senior projects as a knowledge base learning [5]. However, a popular and applied method for developing an effective educational process is proactive learning [6] and proactive learning culture [7].

\*Corresponding Author: Pratya Nuankaew, School of Information and Communication Technology, University of Phayao, Phayao, 56000, Thailand, +66 89 961 4832 & [pratya.nu@up.ac.th](mailto:pratya.nu@up.ac.th)

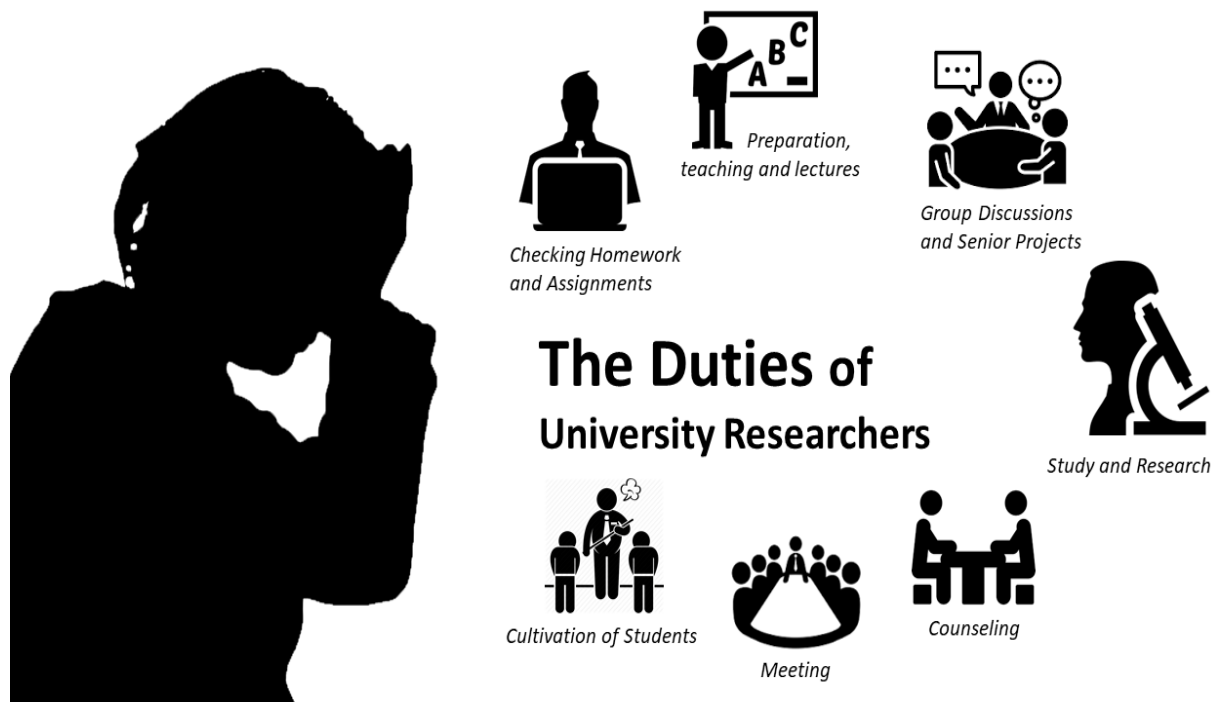


Figure 1: The Duties of University Researchers

Proactive learning or active learning is an essential feature of a person in a learning society that is aware of and responds to many events and stimuli. Having a proactive feature will allow the person to behave appropriately and produce good results. In the active learning perspective, learners play a role in acquiring knowledge and learning in an interactive way until knowledge and understanding are applied, able to analyze, synthesize, evaluate or create things and develop themselves to their full potential.

As well as organizing a learning experience to participate in discussions to practice communication skills, resulting in a high learning result [8-13]. Key aspects of learning management of active learning concepts that emphasize the role and participation of the learners. Knowledge arises from experience, building knowledge, and reviewing learners. In addition, it is a learning management aimed at developing brain potential. Learning management that allows learners to participate in learning The process of creating situations for learners and emphasizing advanced thinking skills [8-10]. It can be concluded that active learning is not only for learner development but also being used for industry development [11-12].

In addition, to creating a learning culture that promotes critical and analytical thinking [8] there needs to be a concrete order for the learner to understand the process of the problem and the solution. According to the intentions of the curriculum for modern education [1,2,9–12], it changed the role of instructors and learners to become colleagues, which is a matter of knowing and cooperating. The instructor is not responsible for educating the learner anymore. The instructor acts as a coach that guides players to play only.

At the same time, the students need to solve a specific problem in the educational problems. It seems that students will need to find [www.astesj.com](http://www.astesj.com)

solutions to educational problems in their quest to understand their own thinking processes and knowledge.

Moreover, the vast majority of researchers at the university need to perform teaching tasks, which spend most of their time for preparing in learning management and summarizing students' tasks as demonstrated in Figure 1. Teachers also need more time to evaluate the learner outcomes. This corresponds to the problem on students who wish to perform well with their senior project.

Figure 1 presents an overview of the duties and responsibilities of most researchers at universities in Thailand. It consists of teaching, conducting research, student counseling, planning classroom management, meetings, promoting students' intrinsic motivation, and so on. It seems that teachers have little time to support student projects. Whereas the collaboration of teachers and students used in senior project development consists of five phases as shown in Figure 2.

Figure 2 shows the senior project development phase which consists of five phases. Phase 1: The process of discovering senior projects based on the student's interests, relevant to teacher research and teacher interest, and the appropriate competence of the faculty member in the organization. Phase 2: The process of transferring academic results, knowledge, interests, sharing, planning, group discussion, and senior project framework development. Phase 3: Development phase, which will be in a long period of time. Phase 4: Project effectiveness testing and evaluation phase, which will test the project for improvement. Phase 5: The publishing process, which presents knowledge and new discoveries obtained from the project. All five phases need to be completed in two semesters or approximately eight months.



**Phase 1:** The process of discovering student projects based on the student's interests, relevant research and interest, and the appropriate competence of the faculty member in the organization.



**Phase 2:** The process of transferring academic results, knowledge, interests, sharing, planning, group discussion, and senior project framework development.



**Phase 3:** Development Phase



**Phase 4:** Project Effectiveness Testing and Evaluation Phase



**Phase 5:** The publishing process, which is the process of presenting knowledge and new discoveries obtained from the project.

Figure 2: Senior Project Development Process

## The Collaboration between students and teachers to develop a senior project.

Considering the abundant and heavy workload of the teacher and the limitations for the development of student projects, adopting proactive learning by creating an appropriate learning culture is important. More than that, the reason for the researchers who emphasized the model development and data analysis in this study was because the research team assembled teams from many institutions, but encountered the same problem. Our common problem is to recommend suitable teachers to students in senior projects. It is important, but difficult. Therefore, this research focused on solving problems proactively using proactive learning strategies to create opportunities for corrective action and reduce impropriety. In addition, the professional group of the researchers is in the information technology field, as researchers try to use technology to help solve the problems. The purpose of this research is (1) to analyze the attitude and perception, which is a collaboration between teachers and students to develop a model for the clustering of appropriate advisors and advisee who cooperate in the senior project, and (2) to develop factors that are significant to predict the right match in senior projects course. Moreover, the major goal of the research team is to develop a student proactive learning culture in educational institutions, which is planned for future research.

The research outline consists of five parts: a summary of the relevant research in Section 2, data collection and research methodology is presented in Section 3, research results and discussion reports are provided in Section 4. Finally, the conclusion is summarized in Section 5, respectively.

## 2. Literature Reviews

### 2.1. Cooperative Learning

Cooperative learning model is one of the learning style that aims to communicate and collaborate with both parties on academic achievements. In addition, cooperative learning is a learning activity technique that allows students to learn together in small groups, each made up of members with different abilities.

Each person is truly involved in learning [17-19]. In the group's learning and success, where the group exchanges views, shares resources, encourages one another, group members are not only responsible for their own grades but are also responsible for their own performance. Learn of all fellow members of the group, it can be called individual success, group success.

Examples of research work related to the concept of cooperative learning such as eL-PCDA model that management cooperative learning in e-Learning, based on learner's behavioral environment, scheduling, and monitoring of learning activities for cooperative learning [18]. Another example is the multi-agent system (MAS) that instead of single-agent learning to handle the performance enhancement of cooperative learning for decision making [18]. The last example is the multi-user streaming feedback system for cooperative learning application whereas teachers can monitor their students and problems in real-time, and the students have more willingness towards learning [13]. Obviously, the demonstration shows that cooperative learning can be a learning achievement, which is important in including goals that both parties emphasize.

### 2.2. Practical Learning

Practical learning is a way of learning from experience. Because work is more problematic than problem solving, it requires more than opening the textbook. It's important for learners to find a way to solve the problem appropriately be correct and reasonable, as well as being useful and applicable [20-23].

Practical learning is learning through experience, the theories used in learning management are Kolb's experiential learning theory [1-2]. Kolb defines learning as a process in which knowledge is created through the transformation of experience. Learning can be explained in four main steps: Concrete Experience (CE), Reflection Observation (RO), Abstract Conceptualization



## Research Methodology and Scope

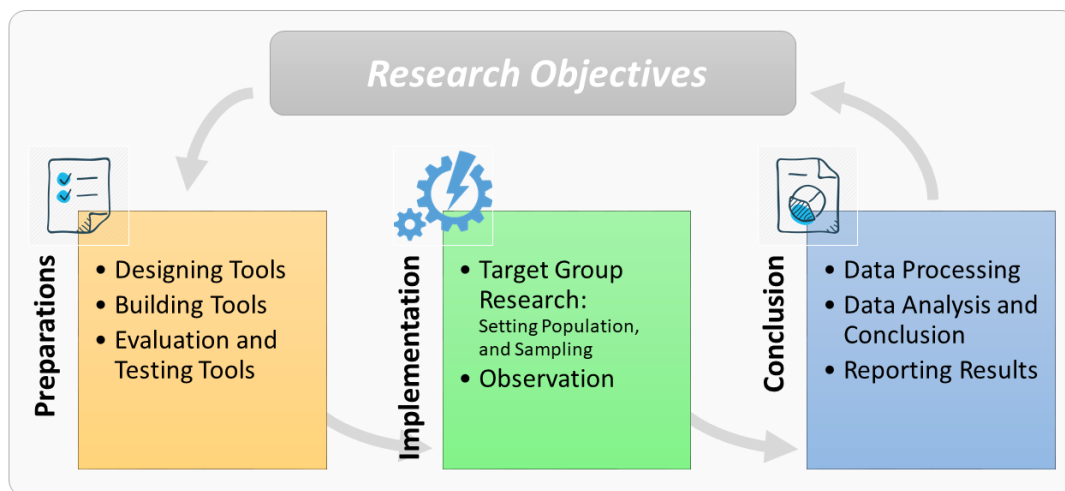


Figure 3: Research Methodology

(AC), and Active Experimentation (AE). Which is evident that learning from practice requires the use of knowledge and experience, which will help develop the learners' attitudes and behavior towards an important goal of learning.

However, practical learning is not just a matter of practice, it must Research Methodology

The research methodology is designed and shown in Figure 3.

### 2.3. Research Objectives

The purpose of the research is (1) to make an analysis on the attitude and perception, which is a collaboration between teachers and students to develop a model for the clustering of appropriate advisors and advisee who cooperate in the senior project, and (2) to develop factors that are significant to predict the right match in senior projects course.

### 2.4. Research Preparation

#### 1) Designing Tools

At the stage of designing tools, the researchers conducted a preliminary study of the research process used for finding attitudes and satisfaction, which found that the appropriate and interesting tool was a questionnaire. Researchers examined questionnaires through the researcher's organization according to procedures and standards of the research process. Questionnaires are routinely used to assess teaching and learning. It is conducted on an annual basis. In addition, the questionnaire quality assessment is described in the review stage section.

At the same time, the data used in the analysis and design of the questionnaire were used by interviewers, including teachers who had received research grants, students who presented and published senior projects, and university administrators.

#### 2) Building Tools

In the process of building tools, the questionnaire was designed into 3 parts, including, Part 1: General information of the respondents, Part 2: Satisfaction and Attitude Towards Senior [www.astesj.com](http://www.astesj.com)

Projects, and Part 3: Suggestions for Senior Projects. While the main topic of part 2 is shown in Table 1.

Table 1: Issues and Essence on Senior Projects

Main Essence	Dimensions Related to Essence: Sub-Stages
1. Organization	1.1 The policies of the organization 1.2 The vision of the organization 1.3 The mission of the organization 1.4 Management and support of the organization
2. Researcher	2.1 Experience and achievements of researchers 2.2 Aptitude and expertise of researchers 2.3 Qualifications of Research Team
3. Students	3.1 Knowledge and ability of learners 3.2 Interest in the research topics 3.3 Impressions and examples in the past
4. Project Support	4.1 Technology and laboratory support 4.2 Staff support 4.3 Budget support

The essence of Part 2 is to find out the level of opinion and attitude on the factor that encourage learners and teachers to collaborate on research, which is comprised of four main areas: organization, researcher, students and project support.

#### 3) Review Stage

After designing and constructing the tool, it was tested for confidence and accuracy by 3 experts from the Department of Information Technology, Faculty of Information Technology, Rajabhat Mahasarakham University, Thailand.

### 2.5. Evaluation and Testing Tools

The research actions were divided into two phases: 1) Target group research, which describe and define the scope of data collection, 2) Research observation, which shows the process of conducting research and gathering data.

#### 1) Target Group Research

a) Population

Data gathering is done in all dimensions, which was limited to the area from two institutions: The Rajabhat Mahasarakham University, and the University of Phayao. The population was defined by four groups: administrators, lecturer, staff and students.

b) Sampling

Sampling was randomly selected from 463 samples, which had been divided into four groups: 7 administrators, 68 lecturers, 26 staff and 362 students.

2) Research Observation

a) Respond to the objectives

After completing the data collection the data sets were already considered as information which provided covers for the purpose of the research.

b) Meet the research framework

The data set responds to the research based on the research framework and is used to test the hypothesis in its entirety.

c) Proceed with caution

In gathering information, it will only attract the attention of the respondents and collected from the reality to get the information according to the actual situation.

2.6. Research Conclusion

1) Data Processing

At this stage, the instrument used for data analysis was selected, consisting of two parts: standard deviation (S.D.) and mean. Analysis of data was divided into groups and topics has been determined and summarized in the next section.

2) Data Analysis and Conclusion

The sample set from the 463 samples was analyzed and summarized as shown in Table 2 and Table 3.

Table 2: Data collection classified by gender

Gender	Samples (n = 463)				
	Administrators (s <sub>1</sub> )	Lecturers (s <sub>2</sub> )	Staffs* (s <sub>3</sub> )	Students (s <sub>4</sub> )	Total
Male	4 (0.86%)	27 (5.83%)	13 (2.81%)	141 (30.45%)	185 (39.96%)
Female	3 (0.65%)	41 (8.86%)	13 (2.81%)	221 (47.73%)	278 (60.04%)
Total:	7 (1.51%)	68 (14.69%)	26 (5.62%)	362 (78.19%)	463 (100%)

\* staffs = Practitioners in the organization but not performing teaching duties

Table 2 shows that most of the respondents are female, with 278 people representing 60.04 percent of all data providers. The researchers found that it was consistent with the students at the University of Phayao and Rajabhat Mahasarakham University, which had mostly female students.

Table 3: Data collection classified by age

Age	Samples (n = 463)				
	Administrators (s <sub>1</sub> )	Lecturers (s <sub>2</sub> )	Staffs (s <sub>3</sub> )	Students (s <sub>4</sub> )	Total
20 - 30 years old	1 (0.22%)	24 (5.18%)	12 (2.59%)	356 (76.89%)	393 (84.88%)

Age	Samples (n = 463)				
	Administrators (s <sub>1</sub> )	Lecturers (s <sub>2</sub> )	Staffs (s <sub>3</sub> )	Students (s <sub>4</sub> )	Total
31 – 40 years old	1 (0.22%)	30 (6.48%)	9 (1.94%)	6 (1.30%)	46 (9.94%)
41 – 50 years old	1 (0.22%)	11 (2.38%)	44 (0.86%)	0 (0.00%)	16 (3.46%)
51 – 60 years old	4 (0.86%)	3 (0.65%)	1 (0.22%)	0 (0.00%)	8 (1.73%)
Total :	7 (1.51%)	68 (14.69%)	26 (5.62%)	362 (78.19%)	463 (100%)

Table 3 shows that the majority of data providers are between the ages of 20-30, with 393 people representing 84.88 percent of all data providers.

The researchers found that most of the respondents were between 20 and 30 years old, as it was the main target group the research focused on. However, other age groups are consistent with the target group, as it is the mentor group who provides consultation to the senior project.

Table 4: Data collection classified by education

Education	Samples (n = 463)				
	Administrators (s <sub>1</sub> )	Lecturers (s <sub>2</sub> )	Staffs (s <sub>3</sub> )	Students (s <sub>4</sub> )	Total
Doctorate	6 (1.30%)	9 (1.94%)	0 (0.00%)	0 (0.00%)	15 (3.24%)
Master	1 (0.22%)	55 (11.88%)	14 (3.02%)	6 (1.30%)	76 (16.41%)
Bachelor	0 (0.00%)	4 (0.86%)	12 (2.59%)	356 (76.89%)	372 (80.35%)
Total :	7 (1.51%)	68 (14.69%)	26 (5.62%)	362 (78.19%)	463 (100%)

Table 4 shows that the majority of the data providers have a bachelor's degree, with 372 people representing 80.35 percent of all data providers. The researchers found that most of the respondents had a bachelor's degree because they were students of the primary target group. In addition, respondents with a master's and doctoral degree show the proportion of mentors.

In interpreting the data according to the characterization criteria, the interpretation is based on a five-level interpretation method by comparing it with the criteria that divides the level estimation into five equal levels, as followed in Equation (1). The result of the calculation is shown in Equation (2).

$$\text{Width of the level} = \frac{\text{Maximum} - \text{Minimum}}{\text{Number of levels}} \quad (1)$$

$$\text{Width of the level} = \frac{5 - 1}{5} = 0.8 \quad (2)$$

From the calculation results in Equation (2), the interpretation results can be specified as shown in Table 5.

Table 5: Interpretation and Meaning

Width of the level	Interpretation	Meaning
1.00 – 1.80	Highly Unacceptable	The lowest level of satisfaction
1.81 – 2.60	Unacceptable	Low level of satisfaction
2.61 – 3.40	Acceptable	Satisfaction

Width of the level	Interpretation	Meaning
3.41 – 4.20	Highly Acceptable	High level of satisfaction
4.21 – 5.00	Maximum Acceptable	The highest level of satisfaction

Table 6: Satisfaction and acceptance toward the factors

Stages	Satisfaction and Acceptance (n = 463)						
	S1	S2	S3	S4	Total Mean	Total S.D.	Interpretation
<b>Stage 1 : Organization</b>							
Stage 1.1	3.71	2.87	3.04	3.69	3.53	1.12	Highly Acceptable
Stage 1.2	3.86	3.43	3.39	3.98	3.86	0.92	Highly Acceptable
Stage 1.3	3.57	3.57	3.42	3.96	3.87	0.87	Highly Acceptable
Stage 1.4	4.42	4.24	3.89	4.16	4.16	0.82	Highly Acceptable
Average	3.89	3.53	3.44	3.95	3.86	0.93	Highly Acceptable
<b>Stage 2 : Researcher</b>							
Stage 2.1	4.29	3.77	3.65	4.06	4.00	0.82	Highly Acceptable
Stage 2.2	4.00	4.10	4.08	4.17	4.15	0.78	Highly Acceptable
Stage 2.3	4.27	4.04	4.12	4.17	4.15	0.85	Highly Acceptable
Average	4.19	3.97	3.95	4.13	4.10	0.82	Highly Acceptable
<b>Stage 3 : Students</b>							
Stage 3.1	3.86	3.88	3.58	4.03	3.98	0.93	Highly Acceptable
Stage 3.2	3.71	3.50	3.39	3.91	3.82	0.89	Highly Acceptable
Stage 3.3	3.71	3.37	3.35	3.96	3.84	0.89	Highly Acceptable
Average	3.76	3.58	3.44	3.97	3.88	0.90	Highly Acceptable
<b>Stage 4 : Project Support</b>							
Stage 4.1	4.86	4.75	4.58	4.27	4.37	0.93	Highly Acceptable
Stage 4.2	4.71	4.71	4.65	4.30	4.38	0.85	Highly Acceptable
Stage 4.3	4.43	4.71	4.54	4.36	4.42	0.83	Highly Acceptable
Average	4.67	4.72	4.59	4.31	4.39	0.87	Highly Acceptable
Total Average	4.11	3.92	3.82	4.08	4.04	0.88	Highly Acceptable

From the analysis of the overall level of satisfaction and acceptance of factors, it was found at a high level of agreement (mean = 4.04, S.D. = 0.88). Moreover, the level of satisfaction with the project is the highest (mean = 4.39, S.D. = 0.87). Therefore, it can be concluded that the respondents agreed and accepted the teaching and learning as a senior project. From the data obtained from the respondents, the researchers grouped the data by clustering through the use of machine learning tools and data mining techniques, which consisted of three parts:

a) K-Means

The k-means is the most common algorithm used as an iterative refinement technique. It is also called Lloyd’s algorithm [14], especially in the computer science community. The k-mean algorithm is performed by switching between two steps: (1) the [www.astesj.com](http://www.astesj.com)

assignment step, which assigns each observation to the cluster with the closest mean. (2) updates the procedure which calculated the new means to be a centroid of observations in a cluster.

b) Decision Tree

The decision tree is one of the learning methods used in statistics, machine learning, and data mining. It works by determining the data from observations and separating data for use in data consideration and finding predictive results. The benefit is on getting important factors, which are caused by the nodes or the decision-making part of the model [15–16].

c) Model Performance

The goal of the model performance is to assess the results and review the process [15–16]. The tools are used in the research, including cross-validation methods as shown in Figure 4, and confusion matrix as mentioned in Figure 5.

Figure 4 displays the separation of data for evaluating the model. The cross-validation method divides the data into two parts. The first part is used for modeling and the remainder is applied to test the model. In addition, model evaluation requires a tool called a confusion matrix [16] to test the model’s performance, with the principles shown in Figure 6.

Figure 5 presents the composition of the confusion matrix performance, which is composed of the actual class and the predicted class. An important benefit of the performance of the confusion matrix is the ability to determine the model’s ability to predict results, such as the predictive ability or accuracy, model precision, model sensitivity, and model specificity (recall measurement). These values are used to determine the actual performance model. Moreover, Figure 5 also demonstrates the formulas and methods for calculating the various performance parameters in detail.

3) Reporting Results

From Table 6, the most attention toward the factors is the stage 4.3 budget support, which is the highest level of satisfaction (mean = 4.42, S.D. = 0.83). The second is the stage 4.2 staff support, which has a high level of satisfaction (mean = 4.38, S.D. = 0.85). The lowest is the stage 1.1 organization policies, whereas the level of satisfaction is at the middle level (mean = 3.53, S.D. = 1.12).

Interesting suggestions include that the budget is delayed and insufficient. Researchers are not very interested because they think it is difficult and not worth the money. Researchers who are truly interested in joining are few, most of whom participate because of their duties. There should be incentives to motivate participants and to develop research in the organization.

3. Research Results and Discussion

The research results classified the research report into four topics which are modelling results, model testing results, model applying results, and summarizing the significant factors.

3.1. Modelling Results

Modelling results are the reports of the various models on different criteria, including the depth of the decision tree, and types of cross-validation method tests, where the different results are shown in Table 7.

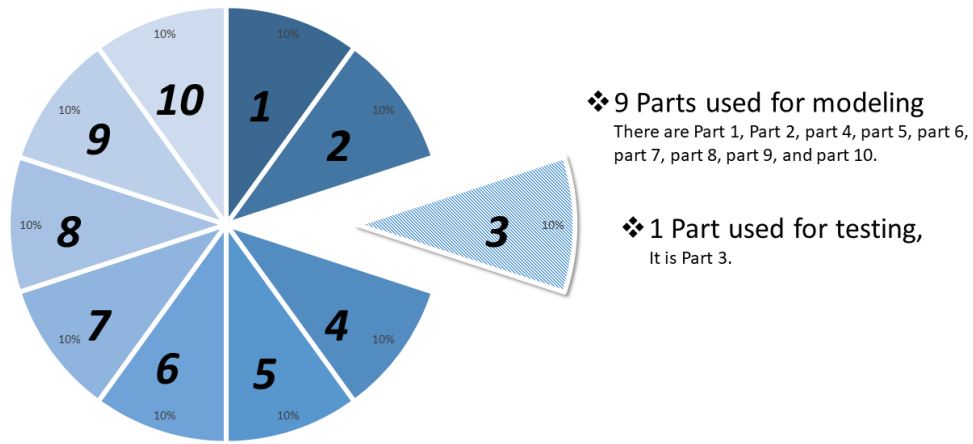


Figure 4: Cross-Validation Methods

<b>Accuracy :</b> $\frac{TP+TN}{TP+TN+FP+FN}$		<b>Actual Class</b>		<b>Precision</b> <i>Positive Predictive Value :</i> $\frac{TP}{TP+FP}$ <i>Negative Predictive Value :</i> $\frac{TN}{TN+FN}$
		<i>Positive</i>	<i>Negative</i>	
<b>Predicted Class</b>	<i>Positive</i>	True Positive : TP	False Positive: FP (Type 1 Error)	
	<i>Negative</i>	False Negative : FN (Type 2 Error)	True Negative: TN	
	<b>Recall</b>	<i>Sensitivity :</i> $\frac{TP}{TP+FN}$	<i>Specificity :</i> $\frac{TN}{TN+FP}$	

Figure 5: Confusion Matrix Performance

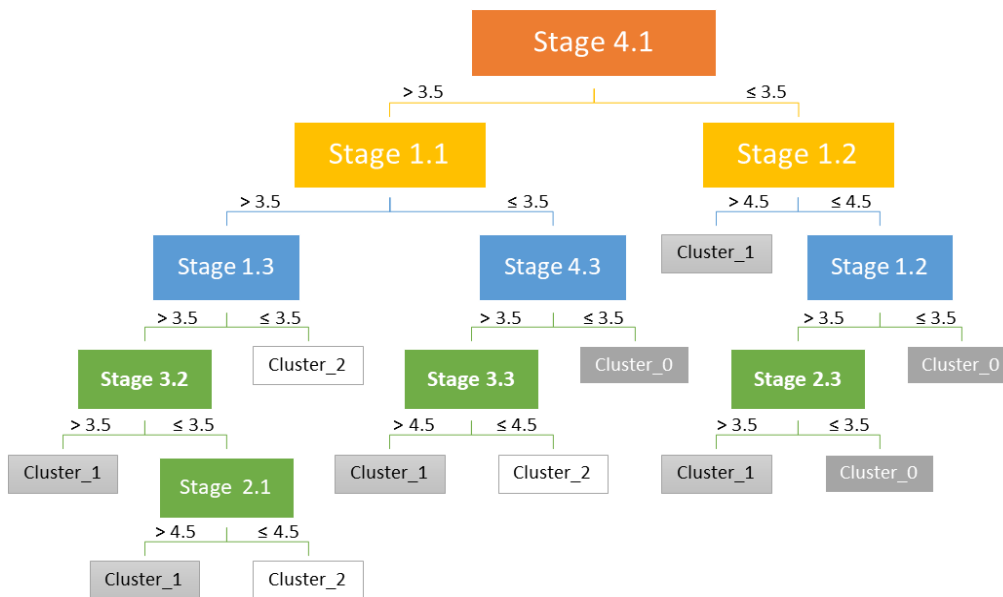


Figure 6: The decision tree model



Table 7: Model results from k-means clustering

Cluster Number	Depth of Decision Tree	Type of Cross-Validation Methods		
		5-Fold	10-Fold	Leave-one-out
3 Clusters*	Level 3	90.51%	92.21%	92.66%
	Level 4	95.91%	95.88%	96.33%
	Level 5	95.90%	96.53%	96.98%
	Level 6*	96.12%	96.75%	98.06%*
	Level 7	95.90%	96.75%	97.62%
	Level 8	95.68%	96.32%	96.76%
	Level 9	95.25%	96.32%	96.76%
4 Clusters	Level 3	71.30%	76.01%	81.21%
	Level 4	87.25%	86.63%	82.51%
	Level 5	88.33%	89.00%	84.88%
	Level 6	89.62%	89.65%	86.83%
	Level 7	89.41%	90.09%	88.12%
	Level 8*	89.85%	90.52%*	88.12%
	Level 9	90.28%	90.51%	89.20%
5 Clusters	Level 3	61.99%	61.97%	61.12%
	Level 4	77.73%	78.83%	81.64%
	Level 5	87.69%	87.25%	85.96%
	Level 6	87.04%	87.90%	85.53%
	Level 7	87.70%	88.33%	87.26%
	Level 8	88.78%	86.83%	87.26%
	Level 9*	89.21%*	89.21%*	87.69%

Table 7 shows that the k-means model with the highest accuracy is the decision tree model that is classified into 3 clusters by dividing the model testing into the leave-one-out cross-validation method with a depth of 6 levels of the decision tree model and has an accuracy of 98.06%

However, the test results classified by other clusters yield a lower accuracy. For example, the 4 clusters with the highest results are 90.52% and the 5 clusters have the highest results of 89.21%.

### 3.2. Model Testing Results

From the results of the prototype model development, it can be concluded that the model with the highest accuracy is the development of the model from k-means clustering, with the appropriate number of 3 clusters and there is a leave-one-out cross-validation result with an accuracy of 98.06%. Details of the developed model are shown in Table 8.

Table 8: The model testing results

Accuracy: 98.06% (+/- 13.82%)		Actual Class			Precision Class
		True Cluster_1	True Cluster_2	True Cluster_3	
Predicted Class	Pred. Cluster_1	220	3	1	98.21%
	Pred. Cluster_2	4	179	1	97.28%
	Pred. Cluster_3	0	0	55	100.00%
Recall Class		98.21%	98.35%	96.49%	

### 3.3. Model Applying Results

From the model that has been selected and demonstrated the performance, this section shows the decision tree model in Figure 6 as displayed in the end of paper. The decision tree rules for self-testing are shown in the test results in Table 9, and the centroid of each cluster is shown in Table 10.

Table 9: The decision tree rules for self-testing

Rule	Condition (If)	Prediction (Then)
1	If Stage 4.1 > 3.500 and Stage 1.1 > 3.500 and Stage 1.3 > 3.500 and Stage 3.2 > 3.500	Then, suitable for cluster_1 = 99.52%, and suitable for cluster_2 = 0.48%.
2	If Stage 4.1 > 3.500 and Stage 1.1 > 3.500 and Stage 1.3 > 3.500 and Stage 3.2 ≤ 3.500 and Stage 2.1 > 4.500	Then, suitable for cluster_1 = 100%.
3	If Stage 4.1 > 3.500 and Stage 1.1 > 3.500 and Stage 1.3 > 3.500 and Stage 3.2 ≤ 3.500 and Stage 2.1 ≤ 4.500	Then, suitable for cluster_1 = 12.50%, and suitable for cluster_2 = 87.50%.
4	If Stage 4.1 > 3.500 and Stage 1.1 > 3.500 and Stage 1.3 ≤ 3.500	Then, suitable for cluster_2 = 100%.
5	If Stage 4.1 > 3.500 and Stage 1.1 ≤ 3.500 and Stage 4.3 > 3.500 and Stage 3.3 > 4.500	Then, suitable for cluster_1 = 60%, and suitable for cluster_2 = 40%.
6	If Stage 4.1 > 3.500 and Stage 1.1 ≤ 3.500 and Stage 4.3 > 3.500 and Stage 3.3 ≤ 4.500	Then, suitable for cluster_1 = 1.23%, suitable for cluster_2 = 98.15%, and suitable for cluster_3 = 0.62%.
7	If Stage 4.1 > 3.500 and Stage 1.1 ≤ 3.500 and Stage 4.3 ≤ 3.500	Then, suitable for cluster_3 = 100%.
8	If Stage 4.1 ≤ 3.500 and Stage 1.2 > 4.500	Then, suitable for cluster_1 = 100%.
9	If Stage 4.1 ≤ 3.500 and Stage 1.2 ≤ 4.500 and Stage 1.2 > 3.500 and Stage 2.3 > 3.500	Then, suitable for cluster_1 = 100%.
10	If Stage 4.1 ≤ 3.500 and Stage 1.2 ≤ 4.500 and Stage 1.2 > 3.500 and Stage 2.3 ≤ 3.500	Then, suitable for cluster_3 = 100%.
11	If Stage 4.1 ≤ 3.500 and Stage 1.2 ≤ 4.500 and Stage 1.2 ≤ 3.500	Then, suitable for cluster_3 = 100%.
Correct: 456 out of 463 training examples (98.49%).		

Table 10: The average within centroid of each cluster

Stage and Cluster			
Stages	Cluster 1	Cluster 2	Cluster 3
<i>Stage 1. Organization</i>			
Stage 1.1	4.43	2.68	2.74
Stage 1.2	4.49	3.49	2.60
Stage 1.3	4.48	3.48	2.70
Stage 1.4	4.43	4.24	2.81
<i>Stage 2. Researcher</i>			
Stage 2.1	4.48	3.82	2.68
Stage 2.2	4.45	4.19	2.84
Stage 2.3	4.51	4.12	2.84
<i>Stage 3. Students</i>			
Stage 3.1	4.50	3.70	2.86
Stage 3.2	4.44	3.34	2.93
Stage 3.3	4.49	3.34	2.88

Stage 4. Project Support			
Stage 4.1	4.44	4.91	2.37
Stage 4.2	4.41	4.87	2.74
Stage 4.3	4.45	4.87	2.88
Stage 4.4	4.44	4.91	2.37
Total Number of Items: 463	224 items	182 items	57 items

### 3.4. Summarize the Significant Factors

From the data collection, there were 463 people who provided attitudes and satisfaction to the research. It can be concluded that the overall level of satisfaction is accepted (mean = 4.04, S.D. = 0.88), as shown in Table V. Therefore, it can be concluded that the respondents agreed and accepted the teaching and learning as a senior project. In addition, the prototype model has been developed for predicting the highest accuracy levels (98.06%): shown in Table VI and Table VII. Moreover, the results of the self-test data model are very accurate at the highest level, which is equal to 98.49% (Correct: 456 out of 463 training examples) as shown in Table VIII. Therefore, it can be concluded that this study was successful. It can develop highly accurate models and also have high performance model testers.

Finally, the significant factors from this research are composed of nine factors as follows: Stage 1.1 The policies of the organization, Stage 1.2 The vision of the organization, Stage 1.3 The mission of the organization, Stage 2.1 Experience and achievements of researchers, Stage 2.3 Qualifications of Research Team, Stage 3.2 Interest in the research topics, Stage 3.3 Impressions and examples in the past, Stage 4.1 Technology and laboratory support, and Stage 4.3 Budget support.

From clustering and discovering these significant factors, the researcher can use the research results to develop a program for advising advisors in a serious project course in order to effectively pair with students having a common attitude.

### 3.5. Research Discussions

The research discussions are divided into three points: (1) A proactive learning culture and learning style model development, (2) Scientific aspects of educational data mining, and (3) Utilization of research activities and research results.

- 1) A proactive learning culture and learning style model development

The reason for the researchers emphasizing on the model development and data analysis was because the research team was assembled from many institutions but encountered the same problem. Our common problem is to recommend suitable teachers to students in senior projects. Therefore, this research focused on solving problems proactively using proactive learning strategies to create opportunities for corrective action and reducing impropriety. In addition, the professional group of the researchers is in the information technology field where the effort is on using technology to help solve problems.

- 2) Scientific aspects of educational data mining

In the process of developing research instruments tool, researchers examined questionnaires through the researchers' organization according to procedures and standards of the research process. Questionnaires are routinely used to assess teaching and

learning. It is conducted on an annual basis. The past assessments are only through statistical analysis, not analyzed for improvement and problem solving. In this research, the researchers analyzed the results of stakeholder assessments in order to form a correlation group and present the stakeholders on further improving the curriculum, and teaching and learning management.

In the conclusion of scientific aspects of educational data mining, the use of available information is important and necessary.

- 3) Utilization of research activities and research results

From the research process and routine operations of the organization, the researcher found that the organizational management and its members realized the importance of conducting this research. In addition, the models in Figure 6, the decision tree model, and Table 9 decision tree rules for self-testing can be further developed into computer programs, applications, and technological applications.

One of the obvious benefits of research is engagement within the organization. It makes the people in the organization aware of the importance of applying technology. In addition, students who participate in the project can also apply the knowledge gained in future students' lives and work.

## 4. Conclusion

From this research, the researcher can summarize according to the research objectives as follows: (1) to survey attitude and perception, which is a collaboration between researchers and students, (2) to develop a model for clustering of advisors and students, and (3) to develop factors that are significant to predict the right match in senior projects course. The results of the survey attitude and perception, which is a collaboration between researchers and students is shown in Table V. The result is that respondents have a high level of overall satisfaction (mean = 4.04, S.D. = 0.88). Thus, it can be concluded that the respondents agreed and accepted the teaching and learning as a senior project.

The result of model development of the clustering model is for matching advisors and students. The result is that the model has the highest level of accuracy (98.06%) as shown in Table VI and Table VII. Moreover, the results of the self-test data model are very accurate at the highest level, which is equal to 98.49% (Correct: 456 out of 463 training examples) as shown in Table VIII.

The result with the development of the factors are significant to predict the right match in senior projects course. The result is an important factor consisting of nine factors, which are stage 1.1 the policies of the organization, stage 1.2 the vision of the organization, stage 1.3 the mission of the organization, stage 2.1 experience and achievements of researchers, stage 2.3 qualifications of research team, stage 3.2 interest in the research topics, stage 3.3 Impressions and examples in the past, stage 4.1 technology and laboratory support, and stage 4.3 budget support.

Therefore, from the research study on collecting data Rajabhat Mahasarakham University, the Maha Sarakham University, and the University of Phayao with a total of 463 samples, this work concludes that the objectives have been achieved with the four machine learning tools and data mining techniques: k-means, decision tree, cross-validation methods, and confusion matrix. For

future research projects, the researchers are committed to further studying on the development of learners' achievement and aims to promote a learning culture based on the results of this research and active learning of educational institutions.

### Author Contributions

Wongpanya Nuankaew, Kanakarn Phanniphong and Sittichai Bussaman designed the research problems and research framework; Wongpanya Nuankaew and Praty Nuankaew collected the data; Praty Nuankaew and Direk Teeraputon analyzed the data and constructed the clustering models; Praty Nuankaew wrote the paper and all authors had approved the final version; Praty Nuankaew is responsible for communicating and delivering the research to the journals.

### Conflict of Interest

The authors declare no conflict of interest.

### Acknowledgment

This research is supported by the five organizations: (1) Faculty of Information Technology, Rajabhat Mahasarakham University, Maha Sarakham, 44000, Thailand. (2) Faculty of Business Administration and Information Technology, Rajamangala University of Technology Tawan-Ok, Chonburi, 20110, Thailand. (3) Faculty of Science and Technology, Rajabhat Mahasarakham University, Maha Sarakham, 44000, Thailand. (4) School of Education, University of Phayao, Phayao, 56000, Thailand. (5) School of Information and Communication Technology, University of Phayao, Phayao, 56000, Thailand. The researchers would like to thank the advisor, lecturers, students, technicians, and all respondents for their entire support.

### References

[1] P. Nuankaew, P. Temdee, "Matching of compatible different attributes for compatibility of members and groups" *Int. J. Mob. Learn. Organ.*, **13**(1), 4–29, 2019. <https://doi.org/10.1504/IJMLO.2019.096469>.

[2] P. Nuankaew, W. Nuankaew, K. Phanniphong, S. Imwut, S. Bussaman, "Students Model in Different Learning Styles of Academic Achievement at the University of Phayao, Thailand" *Int. J. Emerg. Technol. Learn. IJET*, **14**(12), 133–157, 2019. <https://doi.org/10.3991/ijet.v14i12.10352>.

[3] F. M. van der Kleij, "Comparison of teacher and student perceptions of formative assessment feedback practices and association with individual student characteristics" *Teach. Teach. Educ.*, **85**, 175–189, Oct. 2019. <http://doi.org/10.1016/j.tate.2019.06.010>.

[4] C. Brandmiller, H. Dumont, and M. Becker, "Teacher Perceptions of Learning Motivation and Classroom Behavior: The Role of Student Characteristics" *Contemp. Educ. Psychol.*, 101893, 2020. <http://doi.org/10.1016/j.cedpsych.2020.101893>.

[5] C.-Y. Chen and P. P. Chong, "Software engineering education: A study on conducting collaborative senior project development" *J. Syst. Softw.*, **84**(3), 479–491, 2011. <http://doi.org/10.1016/j.jss.2010.10.042>.

[6] D. A. Morley, "Enhancing networking and proactive learning skills in the first year university experience through the use of wikis" *Nurse Educ. Today*, **32**(3), 261–266, 2012. <http://doi.org/10.1016/j.nedt.2011.03.007>.

[7] I. Gnizy, E. Baker William, and Grinstein Amir, "Proactive learning culture: A dynamic capability and key success factor for SMEs entering foreign markets," *International Marketing Review*, **31**(5), 477–505, 2014. <https://doi.org/10.1108/IMR-10-2013-0246>.

[8] B.-K. Joo and T. Lim, "The Effects of Organizational Learning Culture, Perceived Job Complexity, and Proactive Personality on Organizational Commitment and Intrinsic Motivation," *Journal of Leadership & Organizational Studies*, **16**(1), 48–60, Aug. 2009. <https://doi.org/10.1177/15480518093334195>.

[9] A. Ruiz-Iniesta, G. Jimenez-Diaz, M. Gomez-Albarran, "Recommendation in Repositories of Learning Objects: A Proactive Approach that Exploits Diversity and Navigation-by-Proposing" in 2009 Ninth IEEE International Conference on Advanced Learning Technologies, 543–545, 2009. <http://doi.org/10.1109/ICALT.2009.23>.

[10] D. Gallego, E. Barra, S. Aguirre, G. Huecas, "A model for generating proactive context-aware recommendations in e-Learning systems" in 2012 Frontiers in Education Conference Proceedings, 1–62012, 2012. <http://doi.org/10.1109/FIE.2012.6462246>.

[11] M. Reuter et al., "Learning Factories' Trainings as an Enabler of Proactive Workers' Participation Regarding Industrie 4.0" *Procedia Manuf.*, **9**, 354–360, 2017. <http://doi.org/10.1016/j.promfg.2017.04.020>.

[12] S. Moon, J. G. Carbonell, "Proactive learning with multiple class-sensitive labelers" in 2014 International Conference on Data Science and Advanced Analytics (DSAA), 32–38, 2014. <http://doi.org/10.1109/DSAA.2014.7058048>.

[13] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer, R. Wirth, "Promoting reading comprehension and critical-analytic thinking: A comparison of three approaches with fourth and fifth graders" *Contemp. Educ. Psychol.*, **46**, 101–115, 2016. <http://doi.org/10.1016/j.cedpsych.2016.05.002>.

[14] P. Nuankaew, P. Temdee, "Of online community: Identifying mentor and mentee with compatible different attributes and decision tree" in 2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 1–6, 2015. <http://doi.org/10.1109/ECTICon.2015.7207130>.

[15] P. Nuankaew, P. Temdee, "Determining of compatible different attributes for online mentoring model" in 2014 4th International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace Electronic Systems (VITAE), 1–5, 2014. <http://doi.org/10.1109/VITAE.2014.6934434>.

[16] P. Nuankaew, W. Nuankaew, T. Thamma, "The Recommended System for the Relationship between Educational Programs and Students' Interests" in International Conference on Digital Arts, Media and Technology, 2-3, 2016.

[17] W. Tan, A. Tang, T. Wang, W. Tan, L. Li, Z. Zhang, "An Overall Lifecycle Modeling Method for E-Learning Services and the Cooperative Mechanism Based on PCDA" in 2013 IEEE International Conference on Systems, Man, and Cybernetics, 669–674, 2013. <http://doi.org/10.1109/SMC.2013.119>.

[18] D. A. Vidhate, P. Kulkarni, "Performance enhancement of cooperative learning algorithms by improved decision making for context based application" in 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICADOT), 246–252, 2016. <http://doi.org/10.1109/ICADOT.2016.7877588>.

[19] C.-T. Chang, C.-Y. Tsai, S.-E. Peng, H.-H. Tsai, P.-T. Yu, "On the Design of Multi-User Streaming Feedback System for Application of Cooperative Learning" in 2016 International Computer Symposium (ICS), 656–659, 2016. <http://doi.org/10.1109/ICS.2016.0134>.

[20] P. Maresca, E. Gómez, J. Caja, C. Barajas, R. Ledesma, "Academic Learning Platform for Practical Classes: A Learning Model in Manufacturing Engineering" *Procedia Eng.*, **132**, 205–212, 2015. <http://doi.org/10.1016/j.proeng.2015.12.471>.

[21] A. Bouabid, P. Vidal, J. Broisin, "Integrating Learning Management Systems and Practical Learning Activities: The Case of Computer and Network Experiments" in 2009 Ninth IEEE International Conference on Advanced Learning Technologies, 398–402, 2009. <http://doi.org/10.1109/ICALT.2009.181>.

[22] A. Mathew, V. Pangracious, A. Varghese, A. Thomas, D. Mathew, J. George, "PROJECT HAWK: An innovative introduction of practical learning and entrepreneurship in engineering education" in 2012 2nd Interdisciplinary Engineering Design Education Conference (IEDEC), 34–40, 2012. <http://doi.org/10.1109/IEDEC.2012.6186919>.

[23] A. Smith, C. Bluck, "Multiuser Collaborative Practical Learning Using Packet Tracer" in 2010 Sixth International Conference on Networking and Services, 356–362, 2010. <http://doi.org/10.1109/ICNS.2010.56>.

[24] G. Hamerly J. Drake, "Accelerating Lloyd's algorithm for k-means clustering" in Partitional clustering algorithms, Springer, 41–78, 2015. [https://doi.org/10.1007/978-3-319-09259-1\\_2](https://doi.org/10.1007/978-3-319-09259-1_2).

[25] P. Nuankaew, W. Nuankaew, P. Temdee, "Institution recommendation using relationship optimisation between program and student context" *Int. J. High. Educ. Sustain.*, **2**(4), 279–302, 2019. <https://doi.org/10.1504/IJHES.2019.103365>.

[26] P. Nuankaew, W. Nuankaew, K. Phanniphong, R. Fooprateepsiri, S. Bussaman, "Analysis Dropout Situation of Business Computer Students at University of Phayao" in The Impact of the 4th Industrial Revolution on

Engineering Education, Cham, 419–432, 2020. [http://doi.org/10.1007/978-3-030-40274-7\\_42](http://doi.org/10.1007/978-3-030-40274-7_42).

- [27] P. Nuankaew, “Dropout Situation of Business Computer Students, University of Phayao” *Int. J. Emerg. Technol. Learn. IJET*, **14**(19), 2019. <https://doi.org/10.3991/ijet.v14i19.11177>.



## Performance Analysis of Grid-Connected PV Rooftop, at Sakon Nakhon Province, Thailand

Supalak Sathiracheewin<sup>1,\*</sup>, Patamaporn Sripadungtham<sup>2</sup>, Settakorn Kamuang<sup>1</sup>

<sup>1</sup>Faculty of Science and Engineering, Kasetsart University Chalermphakiat Sakon Nakhon campus, Sakon Nakhon, 47000, Thailand

<sup>2</sup>Faculty of Engineering, Kasetsart University, Bangkok, 10900, Thailand

### ARTICLE INFO

Article history:

Received: 03 July, 2020

Accepted: 23 August, 2020

Online: 28 August, 2020

Keywords:

Weather-Corrected Performance

PV Rooftop

IEC61724

### ABSTRACT

The performance ratio (PR) based on IEC 61724 standard is calculated under the influence of seasonal variations and the capability of the system. On the other hand, The National Renewable Energy Laboratory (NREL) of the U.S. Department of Energy proposed the weather-corrected performance ratio (PRcorr). This PRcorr index calculates the performance of the PV system which has taken the weather variations that influence the cell temperature into account. Both techniques were compared with the simulation program to analyze the PR index according to system conditions. The study site is the on-grid PV rooftop at Kasetsart University Chalermphakiat Sakon Nakhon (CSC). The PV panel is oriented to the southwest, 215 degrees azimuth. The angle of the inclination of the panel is 17 degrees. The result shows the trend of monthly PRcorr with little variations. At the same time, the PR value over the year has a lot of variations. The variation of PRcorr and PR is 2.39 and 5.07, respectively. PRcorr has low variability due to the correction of weather factors. The average cell temperature is an important variable. To calculate the average temperature of the panels, one year of data is needed to filter out distorted information. The system available condition is an important factor for the on-grid PV system at low voltage.

## 1. Introduction

Photovoltaic systems are currently being installed worldwide. The power received was estimated before the PV system was set up. And after installation, the energy generation potential of the PV system is measured. Due to the radiation and ambient temperature are different in each area. This paper is an extension of work originally presented in 16<sup>th</sup> ECTI-CON [1]. This paper examines the efficiency of the PV system in northeastern Thailand. The result will be a database of the energy generation potential of the PV plant in Sakon Nakhon province. According to the installation, the price per unit is much cheaper and at the same time the price of electricity is increasing. Manufacturers of equipment for the PV plant have increased, which has caused competition in the market. The amount of PV systems installation in Thailand has increased a lot, which is following the energy plan policy of the Thai government [2]. The installation of the PV system has spread throughout Thailand, both on the ground and rooftop. Therefore, many researches demonstrate the efficiency of the PV system, which influences people to install a clean energy plant.

The commonly used to assess the performance of a PV plant that is commonly used is the IEC 61724 standard. Following the IEC 61724 standard, the performance ratio (PR) is calculated by taking into account the radiation intensity, ambient temperature, wind speed, humidity, solar panels, and balance of system (BOS). The performance of the PV system depends on seasonal changes. Chong Li presented the calculation of PR with different PV types over the year in China [3]. The result shows that the performance ratio of poly-silicon PV increases from January to May and monthly array losses in summer are high. Then, the performance ratio of the PV system depends on location, PV module type, and the climate conditions. In 2019, Dhimish et.al. proposed the direction of the PV module for solar radiation that is important for energy production [4]. In addition, the azimuth angle is shown, which influences the maximum energy production. The suitable azimuth angle for the PV installation in the UK is -4 to 2 °C (the azimuth angle in the south is 0 °C). The data relation between irradiation and output power is to fitted regression. The R-square value is used to explain the variability data. High relative to R-squared, which is close to 1, means that the PV system generates the most electrical energy. Meanwhile, Singh demonstrated the

\*Sathiracheewin.S, Sakon Nakhon, +66(0)42 725 033, Email: [supalak.sat@ku.th](mailto:supalak.sat@ku.th)

performance ratio obtained by PVsyst at Naresuan University, Thailand [5]. The performance ratio in each month is quite stable, which differs from the actual operation of this variation every month. Therefore, the weather variations employ to calculate the PR index based on the IEC 61724 standard. The different climatic conditions in each month affect the PR value.

The weather-corrected performance ratio (PR<sub>corr</sub>) is presented by the national renewable energy laboratory (NREL) of the U.S. Department of Energy [6]. This report suggests that the calculation of PR values according to climate change gives a low PR value and a fluctuation value. The weather variation should calculate the only factors that affect the cell temperature. According to research by Basson, the objective of this research is to reduce the sensitivity to climate data [7]. The result shows that the temperature correction has a significant effect to reduce the seasonal changes. Compare to the PR method, its value has significant variation. Syahputra proposed the PR<sub>corr</sub> calculation in the real operation in the tropics and subtropics compare to the PR technique of IEC 61724 standard [8]. Also, the temperature coefficient for power is presented in the real operating condition as well. The value of the temperature coefficient varies in different irradiation. As a result, the range value of PR<sub>corr</sub> is greater than PR. Therefore, the PR<sub>corr</sub> index defined by NREL is also useful for benefit financing.

The design of the PV system involves an analysis of production capacity, often through a simulation program. PVsyst is another simulation program that helps the designer to define from PV sizing, the capacity of the inverter, to energy production estimation [9]. For the simulation, NASA or PVGIS weather data was always used [10]-[11]. To calculate the power ratio correctly, the weather data in the installation area should be measured. Using the estimation data in the simulation model may cause the result to deviate from the actual value [12].

The northeastern region of Thailand is in an area that receives good solar radiation. From the data of the solar map potential improvement project for Thailand 2017, it is apparent that the average monthly daily radiation intensity is 20 to 22 MJ / m<sup>2</sup> per day [13]. In particular, the Kasetsart University Chalemphakiat Sakon Nakhon (CSC) is a campus of Kasetsart University in the province of Sakon Nakhon located in the northeastern part of Thailand with 17.28661 latitude and 104.1061 degrees longitude, as shown in Figure 1. The average solar energy is 17.91 MJ / m<sup>2</sup> per day, and in April the average solar energy is up to 20.84 MJ / m<sup>2</sup> per day. Topography of Sakon Nakhon city is a flat plain with no tall buildings, and there is no large industry that causes air pollution. Therefore, the province of Sakon Nakhon is suitable for the installation of the PV system. Currently, both the public and private sectors in Sakon Nakhon Province are interested and have installed more PV systems. The production of electric power by solar cells, not only reduces the dependence on the main grid but also responds to the government policy. But many people want to know the energy efficiency of the PV system, which has different weather conditions in each area. Appropriate assessment of energy production potential to support decisions made by persons interested in the PV system.

Therefore, this research proposes an efficiency evaluation of a PV production system, based on the IEC 61724 standard and the weather-correct performance ratio of the NREL concept. The

result from both concepts were compare to the PVsyst analysis result. The measure data in the study area is input to analyze the significant factors that influence the energy production of the on-grid PV system at low voltage.



Figure 1: Location of Sakon Nakhon, Thailand on Google map.

## 2. Solar Rooftop at CSC

CSC in Sakon Nakhon covers an area of 6.4 square kilometers. The Faculty of Science and Engineering is part of the University that teaches and provides useful research results to society. One of the research sections conducts clean energy research such as on-grid PV rooftop. The schematic diagram of the grid-connected PV system is shown in Figure 2. The PV system consists of 10 modules of the series, a single-phase 3.6 kW inverter and other equipment used for the installation of the PV system. For installation details, the PV module faces southwest, 215 degrees solar azimuth (north is 0 degrees angle). The angle of inclination of the plate is 17 degrees. The energy output from the PV generation flows at low voltage into the load center of the building. The rated voltage in Thailand is 230 volts.

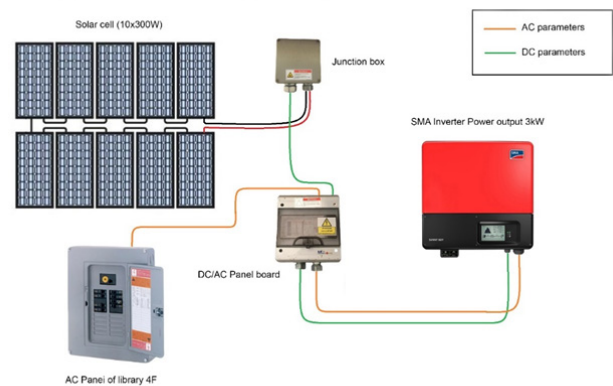


Figure 2: The schematic diagram of 3 kWp PV system

The specification for PV module and inverter is listed in Table 1. The cell type of PV module is poly-crystalline and the total area of PV panels is approximately 19.3 m<sup>2</sup>. PV panels send DC power to the junction box via a cable. The cable made of aluminum with a cross-sectional area of 4 mm<sup>2</sup>. The junction box contains a DC fuse (20 kA, 2 POLES), after which the DC power flows into the inverter and the AC power is obtained. The AC power flows back into the AC fuse in the junction box and then goes to the load

center. The inverter is a grid-tie inverter of SMA type, Sunny Boy 3600TL-21.

Table 1: PV Module and Inverter Specifications

PV module	Spec.	Inverter	Spec.
Max. Power (Wp)	300	Rated AC Output Power (kW)	36.8
Open-circuit Voltage (V)	45.14	Maximum Current (A)	16
Short-circuit Current (A)	8.74	AC Nominal Voltage (V)	230
Max. Voltage (V)	37.01	Max. Input Current	15
Max. Current (A)	8.12	Input Voltage (V)	175-500
Temp. Coefficient	-0.4574 (%/°C)	Maximum Efficiency (%)	97

According to the SMA inverter, information about electrical energy on the inverter can be found on the website, as the data is transmitted via the university's intranet. The information on this website contains only some information, this is not enough to analyze the performance of the PV system. Consequently, sensors were installed to measure relevant data such as the radiant level, ambient temperature, module temperature, AC power and DC power being recorded at the same reference time. Large amounts of data may contain data discrepancies. Therefore, both the measuring instruments as well as the data acquisition tools are required to have low tolerances [14]-[15].

### 3. Performance Index of PV System

#### 3.1. IEC 61724 Methodology

For analysis of system performance under standard IEC 61724. This standard has been used as a reference in many studies [16]. The indices specified in the standards are used to measure system performance, including assessing the losses that occur in the system.

Reference Yield ( $Y_R$ ) is referred to as the reference irradiance generate power ( $H_t$ ) under standard test condition ( $G_{STC}$ ) ( $kWh/m^2$ ), usually  $1 kW/m^2$ . It is the calculation of the energy that is generated by this light intensity. The relationship is shown in equation (1).

$$Y_R = \frac{H_t}{G_{STC}} \quad (kWh.kW_p^{-1}) \quad (1)$$

Array yield ( $Y_A$ ) is the ratio of daily energy generation from PV array ( $E_{DC}$ ) divided the kW of installed PV array ( $P_0$ ). This parameter calculates the energy output per day and uses this value to calculate the average per month. Therefore, the average power of the DC part per day of each month is obtained.  $Y_A$  value is given by equation (2).

$$Y_A = \frac{E_{DC}}{P_0} \quad (kWh.kW_p^{-1}) \quad (2)$$

Final Yield ( $Y_f$ ) is defined as the daily energy generation from PV plant ( $E_{AC}$ ) divided by the total kW of installed PV array ( $P_0$ ). This energy refers to the energy gained after passing the inverter or the AC power part.  $Y_f$  is calculated using equation (3)

$$Y_f = \frac{E_{AC}}{P_0} \quad (kWh.kW_p^{-1}) \quad (3)$$

Performance ratio (PR) indicates the actual energy production and the energy generated by the incident irradiation. Therefore, PR is the ratio of Final Yield ( $Y_f$ ) to Reference Yield ( $Y_R$ ). This parameter shows the efficiency of the system during solar conversion until the final electrical power is reached. Therefore, the PR index is influenced by all environmental factors and including the efficiency of the PV system. PR calculation is given by equation (4).

$$PR = \frac{Y_f}{Y_R} \times 100\% \quad (4)$$

According to the IEC 61724 standard, the system loss also occurring in the PV system is calculated. The loss is caused by the equipment in the system, such as the PV panel, cable wire, inverter, and other factors that cause the loss to increase.

The standard IEC 61724 also shows the loss value in the system as in equation (5). Array Capture Loss ( $L_A$ ) is the loss from PV array. It the different value between  $Y_R$  and  $Y_A$ .

$$L_A = Y_R - Y_A \quad (kWh.kW_p^{-1}) \quad (5)$$

Another loss value is system loss ( $L_S$ ) that depended on equipment effect in the system such as inverter, wire and PV panels. It is calculated in equation (6).

$$L_S = Y_A - Y_f \quad (kWh.kW_p^{-1}) \quad (6)$$

#### 3.2. Weather Corrected Performance Ratio

According to the technical report of the national renewable energy laboratory (NREL) [6], the weather-corrected performance ratio (PRcorr) was presented. PRcorr is modified from the PR concept of IEC 61724. PRcorr was created by reducing the impact of fluctuating seasonal conditions, which also causes the PR value to fluctuate. Therefore, PRcorr has improved the formula in IEC61724. PRcorr focuses on the technique to obtain the cell temperature, which affects efficient PV generation. The formula PRcorr calculates the cell temperature based on irradiation and ambient temperature. The equation is shown in (7).

$$PR_{corr} = \frac{\sum_i EN_{AC\_i}}{\sum_i \left[ P_{STC} \left( \frac{G_{POA\_i}}{G_{STC}} \right) \left( 1 - \frac{\delta}{100} (T_{cell\_typ\_avg} - T_{cell\_i}) \right) \right]} \quad (7)$$

where

PR<sub>corr</sub>: corrected performance ratio (%)

EN<sub>AC</sub>: measured AC power generation (kW)

P<sub>STC</sub>: total of installed modules' power rating at standard test condition (STC)

G<sub>POA</sub>: the radiation measurement at plane of array (POA) (kW/m<sup>2</sup>)

G<sub>STC</sub>: irradiance at STC (1,000 W/m<sup>2</sup>)

T<sub>cell</sub>: cell temperature computed from measured meteorological data (°C)

T<sub>cell\_typ\_avg</sub>: average cell temperature computed from one year of weather data using the project



weather file (°C)

$\delta$ : temperature coefficient for power (%/°C, negative in sign) that corresponds to the installed modules  
 i: a given point in time (day)

NREL shows that the use of an adapted reference cell to measure irradiance avoids the need to correct spectral variations as well. In the meantime, other weather effects such as snow losses, environmental pollution losses or the effects of variable irradiance on efficiency are not corrected. Corrections for these additional weather effects could lead to more consistent results.

### 3.3. Performance Ratio by PV<sub>system</sub> Analysis

In general, the technique and analysis of the performance of the PV system needs to be calculated before the actual installation of the system. Therefore, the use of the simulation programs to analyze a process is necessary, which leads to losses in the system, energy and production performance. The PV<sub>system</sub> simulation program can analyze these parameters and generate solar radiation and other weather data in the study site.

Therefore, when analyzing the performance of the PV system, the correct information needs to be entered compared to the actual value, including cable line, the inverter and PV module type. The installation properties are the actual value in the installation area, such as the azimuth angle and the tile angle of the module. The accuracy of the weather data entered into the simulation program must come as close as possible to the actual conditions and then the high accuracy output will be obtained. The analysis with this simulation program also analyzes the weather-corrected performance ratio [17].

## 4. Result and Discussion

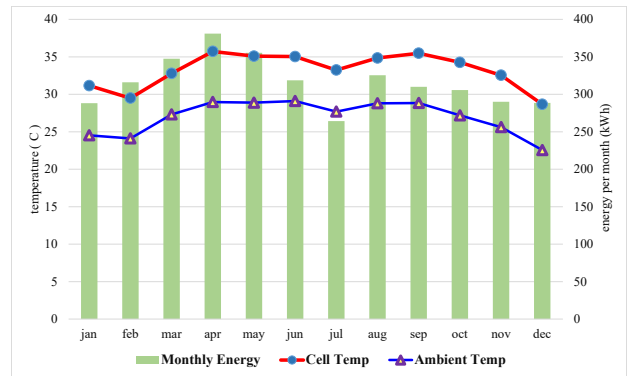
The results of this experiment are based on annual data for 2017. Figure 3 is the 3 kWp PV plant with the temperature sensor and radiation sensor. All data are collected to data logger. An energy meter is installed for the DC and AC power measurement of the PV system.



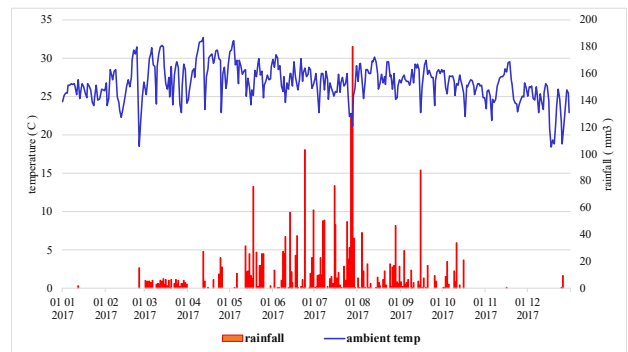
Figure 3: Plant Study at CSC, Sakol Nakhon

Figure 4 (a) shows the average cell temperature, the average ambient temperature, and the monthly energy of the PV system. The average monthly temperature data clearly shows that the panel temperature is higher than the ambient temperature. In 2017, the average ambient temperature is 26.97 °C and the average cell

temperature is 33.19 °C. Figure 4 (b) shows the amount of rainfall every day. Note that the heavy rains occur in June, July, and August. As a result, the average ambient temperature also drops. In September and October, the rainfall begins to decrease and then the average ambient temperature gradually increases. In Sakon Nakhon province, from 28 to 29 July 2017, precipitation for both days was up to 310 mm<sup>3</sup>, heavy rain caused a sudden flood. The result is a widespread power outage in Sakon Nakhon province. A major power outage meant that the inverter was also disconnected from the PV system. Therefore, the grid-connected PV system has dropped out of the main grid and does not need to generate power for a while.



(a) Cell Temperature and Ambient Temperature Compare to Monthly Energy



(b) Ambient Temperature and the Rainfall

Figure 4: The Weather Factors and Monthly Energy

Because of the PV module faces southwest, 215 degrees solar azimuth, the irradiation data are grouped in the solar azimuth or sun path throughout the year. The sun path refers to the sun that appears daily in the sky. The path of the sun has different values (degrees) every day and causes seasons. From the Sun Orbit database at CSC, each month was grouped with similar orbits. It refers to the sunrise at 6 a.m. and the sunset at 6 p.m. The summer group consists of February, March, and April with an azimuth of the sun between 76 to 97 degrees in the east and 261 to 284 degrees in the west. Rainy season groups are May, June, July, August and September. The sun orbits of this group are between 71 and 80 degrees in the east and between 277 and 292 degrees in the west. The winter group with similar approaches to the movement of the sun are January, October, November, and December with an azimuth angle between 106 and 112 in the east and between 244 and 256 in the west. Taking into account the average movement direction of the sun per month, it can be divided into three types as described above. Each group has daily routes orbit that are close to

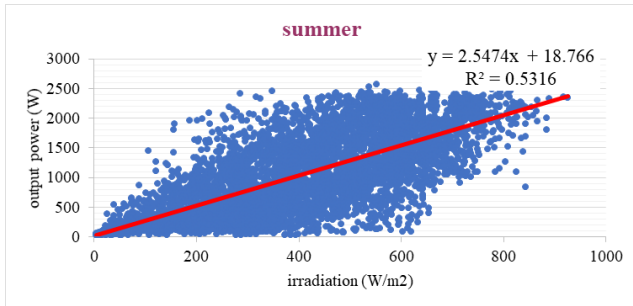


each other. It should be obtained a similar sun intensity. The three groups season relate to the irradiation, shown in Figure 5.

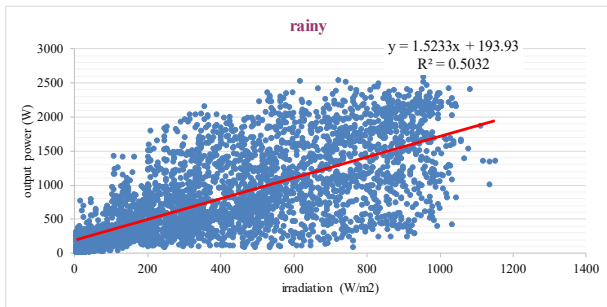
generated is 293 kWh/month. Therefore, the trend of maximum amount of energy will occur in Sakon Nakhon this month.

Table 2: Comparative Average Monthly Energy and R-square

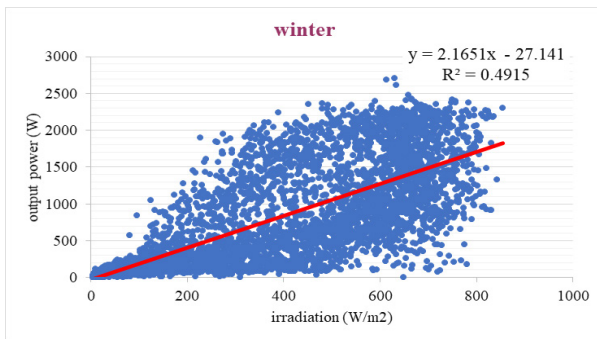
Season group	average monthly energy (kWh)	R-square
Summer	348	0.5316
Rainy	315	0.5032
Winter	293	0.4915



(a) The Summer Group



(b) The Rainy Group



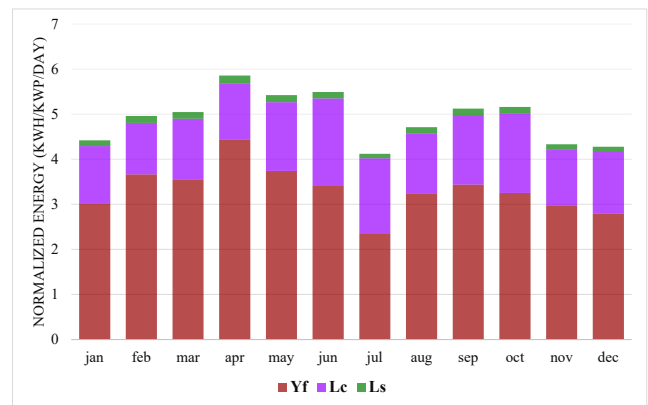
(c) The Winter Group

Figure 5: The Relationship between Solar Irradiation and Output Power

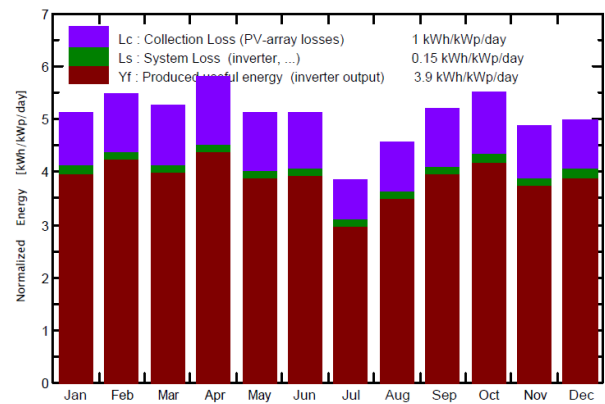
Figure 5 shows the relationship between solar irradiation ( $W/m^2$ ) and electric power produced (W) according to the monthly groupings mentioned above. Curve fitting is a statistical method used to analyze the data. The R-square indicates the degree of proximity of the data. As the R-squared value approaches 1 or 100%, it means that all data points would fall on the adjusted regression line. According to Figure 5(a), the R-squared of the summer group is higher than the others. The R-squared of Figure 5(b) and Figure 5(c) is 0.5032 and 0.4915, respectively. The winter group has a minimum of R-square value. Since the energy output is approximately linear for irradiation in the study area. This means that the electrical energy generated by the PV system at this time is most valuable.

According to the monthly energy of CSC plant, the electrical energy generated by the PV system is compared with the R-square in Table 2. The summer group has the highest average monthly production capacity of about 348 kWh, while the winter group has the smallest R-square. The average energy during the winter group

For the performance analysis according to IEC 61724 standard, Figure 6 (a) shows the average monthly value of  $Y_f$ ,  $L_c$  and  $L_s$  which is 3.33, 1.45 and 0.14 respectively. Compared to the result of PVsyst as shown in Figure 6 (b),  $Y_f$ ,  $L_c$  and  $L_s$  which is 3.9, 1.0 and 0.15 respectively. Both techniques are similar energy generation such as the low  $Y_f$  is in July. The maximum system loss occurred in April, the IEC method shows 0.19 kWh/kWp/day and PVsyst is 0.169 kWh/kWp/day. The trend of the energy value obtained from the PVsyst program is higher than in the IEC technique, where the actual value is used for the calculation. Therefore, using the prediction program, the energy estimate is inflated. Overall, the amount of energy generation from PVsyst is high. However, the PV array loss is less than with the IEC technique, possibly the dust on the PV module. The system has been installed for almost 3 years and the panel has not been cleaned. Estimating these parameters will input all data as monthly averages. The result of the analysis may be inaccurate or deviate from the actual value.



(a)  $Y_f$ ,  $L_c$  and  $L_s$  based on measure data



(b)  $Y_f$ ,  $L_c$  and  $L_s$  from PVsyst analysis

Figure 6: Result from the measure data compare to PVsyst analysis

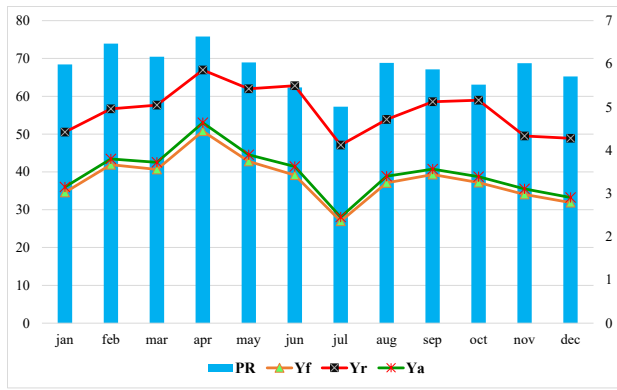


Figure 7: Performance of PV System at CSC

Figure 7 shows the calculation value of the PR index based on  $Y_F$ ,  $Y_A$ , and  $Y_R$  data. The graphic shows a comparison of the performances of the grid-connected PV system. For each month of  $Y_F$ ,  $Y_A$ , and  $Y_R$ , the line curves are similar.  $Y_F$  and  $Y_A$  are closed values, as this depends on the efficiency of the inverter and loss in cable. As a result, the average efficiency of the inverter is 96%. The maximum  $Y_R$  in April is 5.86 kWh.kW<sup>-1</sup>, which according to the monthly energy value is 381 kWh. April is considered a hot season. The average temperature is generally high and most skies are cloudless. While July which in the rainy season has a low of  $Y_R$  and similar to monthly energy as well. The average  $Y_R$  in July is 4.12 and the monthly energy is 264 kWh. On the other hand,  $Y_F$  and  $Y_R$  lines are separated from each other. It means  $Y_R$  can convert to  $Y_F$  depend on the system efficiency. It also represents the amount of loss that has occurred in the PV system. The loss consists of the loss of the PV array loss, cable loss and the system loss caused by the efficiency of the inverter.

Every month the average power ratio is calculated. The maximum PR in April, February, and March is 75.8%, 73.94%, and 70.47%, respectively. On the other side, the minimum PR in July is 57.23%. At the end of the year, which is the winter season, monthly energy production tends to decline. Note that the low PR period is during rainy and winter season. PR parameter reflects the performance of the system for converting solar energy into electricity. A high PR also indicates a small loss in the system.

Based on IEC 61724 standard, the PR calculation is caused by the effects of all environmental factors on the PV system. As a result, the PR value fluctuates quickly. Therefore, some researchers do not think about these effects. The PR index should have suitable values to assess the potential of the system. The weather-correct performance is presented by NREL, which corrects the cell temperature. The irradiation and ambient temperature data are used to recalculate the performance of a PV system that directly affects the cell temperature. At the same time, the PVsyst has been used to calculate PR values as well. By creating a simulation model with system unavailability conditions of the main grid system. It indicates the percentage of the period in which the system is unavailable, consisting of 0%, 5%, and 10%. According to the measurement data, there are outage events in a few months. For January, September, October, November, and December are defined to the month of system unavailability condition, as shown in Table 3.

Table 3 shows the PRcorr, PR index based on measured data and PR index analysis by PVsyst. The maximum PRcorr in April

is 82.05%, as PR is 75.8%. But in PVsyst calculation, the maximum PR with system unavailability 0%, 5%, and 10% are 78.9% in December, 75.7% in January and 44.7% in February and July respectively. For average PRcorr, PR and PR of PVsyst at 0% are 77.34%, 67.52%, and 77.15% respectively. The variation calculation value is based on statistical techniques. Therefore, the standard deviation is chosen. For standard deviation analysis, the PR and PRcorr are 2.39 and 5.07 respectively.

Table 3: Comparative Average Monthly Performance Ratio (%)

Month	PRcorr	PR	System Unavailability (%)		
			0	5	10
January	74.51	68.42	78.10	75.70	57.40
February	76.37	73.94	77.70	75.30	77.70
March	77.18	70.47	76.10	64.70	76.10
April	82.09	75.80	75.90	73.60	75.90
May	77.96	68.97	76.50	74.10	76.50
June	79.86	62.36	77.20	64.30	77.20
July	73.86	57.23	77.70	66.70	77.70
August	79.67	68.85	77.50	75.00	77.50
September	77.82	67.11	76.50	74.10	57.90
October	76.20	63.07	76.50	74.20	58.50
November	77.49	68.74	77.20	63.80	59.30
December	75.01	65.24	78.90	67.30	61.00
Average	77.34	67.52	77.15	70.73	69.39
Std.	2.39	5.07	0.89	4.87	9.39

PRcorr is significantly higher than PR and the line trend is quite stable, which does not change significantly during the year. While the standard deviation value of PVsyst is tended to increase with the strong value of system unavailability. The PRcorr value varies slightly because the cell temperature was influenced by radiation and ambient temperature. While other weather factors based on the NREL concept are un-correct such as pollution, snow or shading. PRcorr is therefore very valuable and encourages people to be more interested in PV production.

For the other hand, PVsyst analysis, the average PR with 0% system unavailability is similar to PRcorr, but the standard deviation value has very little variation. The standard deviation increases with high system unavailability because the energy output from the PV system cannot flow to electrical load, such as fault event or temporary power from utility. Therefore, the PR with 0% system unavailability is calculated based on the weather corrected condition.

The PR value is a performance indicator that combines the effects of various factors such as the intensity of light falling on the PV module, cell temperature, shading, or device malfunctions in the system. When analyzing these factors, some factors are not related to the weather. However, when weather factors are used to calculate the PR value, it is low and changes rapidly.

Use the PVsyst to determine the monthly energy. The monthly energy output from the PV system is compared with the result of PVsyst, shown in Table 4. The maximum energy in April is 381 kWh, similar to PVsyst, the maximum output in April is 395 kWh. When calculating deviations from the actual values, the maximum is 19% and the minimum is 13%. For average monthly with 10% system unavailability is closely the true value. Because it is stipulated that the PV system will not be able to supply electricity

according to the actual incident (fault in the utility grid). Therefore, a high energy forecast is created for electricity from a PV system that can deliver at any time. Therefore, the use of the analysis program must take these events into account. These events often occur in areas with poor electrical reliability.

Table 4: Comparative Monthly Energy (kWh) with PVsyst Prediction

Month	Monthly Energy	System Unavailability (%)		
		0	5	10
January	288	371	330	273
February	316	358	358	358
March	347	372	372	372
April	381	395	395	395
May	355	363	363	363
June	319	355	355	355
July	264	278	278	278
August	326	327	327	327
September	310	357	311	271
October	306	391	345	299
November	290	338	298	259
December	288	365	321	282
Total (kWh)	3790	4269	4054	3831
Avg. per month	316	356	338	319
Error (%)	-	19	15	13

inverter's function is to disconnect the circuit so that the energy output of the inverter is zero. Figure 9 demonstrates the real power (P) value with interval time (15 minutes). The red circle is the point of the momentary power outage. Sometimes a power failure occurs for a very short period, such as 2 or 5 minutes. The inverter cuts off the circuit and then turns it on when the main grid can be re-energized. Therefore, the average AC data, for interval time 15 minutes, is reduced. At the mentioned point, the real power (kW) decreases abnormally. These unusual weather conditions cause the inverter protection system to work. Therefore, some monthly energies are reduced even though the cell temperature is low or irradiation is high.

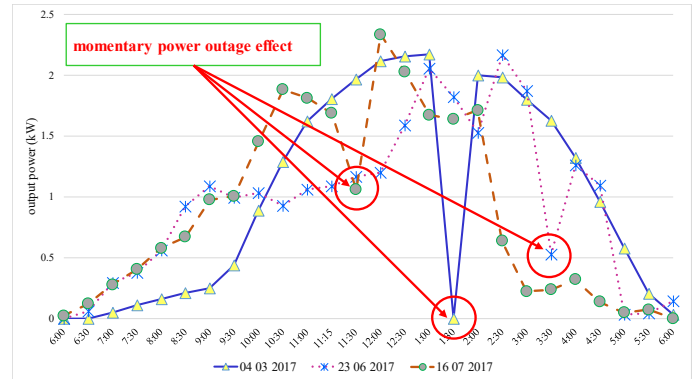


Figure 9: The momentary power outage effect in the on-grid PV system

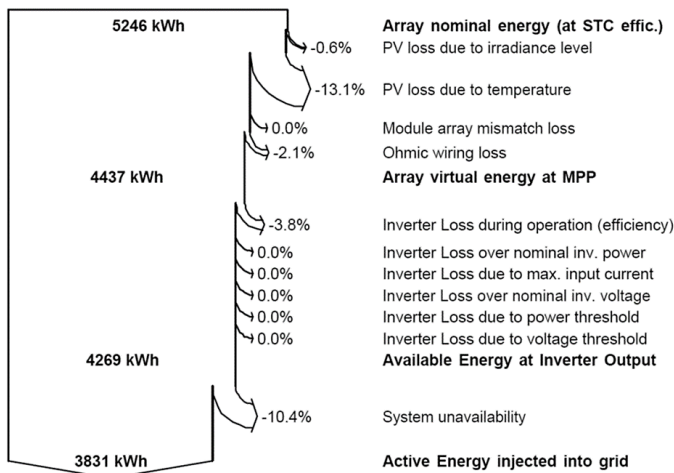


Figure 8: output from PVsyst analysis

The result from PVsyst, as shown in Figure 8, power loss calculation occurs in the PV system that has various types of loss. Nearly 13.1% power loss due to temperature. The high temperature leads to an increase in power loss and then to a reduction in the output power. The parameters used to analyze the PRcorr are derived from the device specification and the actual data in the experimental area. The important parameter is the average cell temperature (Tcell\_typ\_avg) each month which for average annual is 38.84 °C. The temperature coefficient for maximum power ( $\delta$ ) is -0.004574 which applies for over the year.

For the on-grid PV system inclement weather such as rain, strong wind will cause the main power system to always fail, such as voltage drop or momentary power outage. These unusual weather conditions cause the inverter protection to work. The

## 5. Conclusion

This article presents the technical calculation of the conversion potential of solar energy up to the final electrical power. The technique used is the IEC 61724 standard and the index proposed by the NREL. PVsyst is a simulation model to estimate the energy and the performance ratio with the real conditions. The study site is Kasetsart University Chulermphakiat Sakon Nakhon (CSC). The PV equipment consists of the total PV installation is 3 kWp and 3.6 kW single phase of grid-connected inverter. The energy generation from the PV system flow to the load center of the building at CSC.

The irradiation data show that the R-squared distribution is high, about 0.5316, in February, March, and April, which means high energy production. In April, the maximum monthly energy is 381 kWh. The minimum energy generation occurs in July, about 264 kWh, because of heavy rain and no more the irradiation. For the average of  $Y_R$ ,  $Y_A$  and  $Y_F$  is 4.91%, 3.47%, and 3.33%, respectively. The maximum performance ratio of the PV system is 75.80% in April. The average PR is 67.52%, which is influenced by weather factors. Compare to the PRcorr index, which is based on the NREL concept, the maximum PRcorr in April is like the PR index at 82.09% and the average PRcorr at 77.34%. For PVsyst calculation, the average PR with 0% system unavailability is close to the PRcorr and the average PR is decrease with more the system unavailability.

The PRcorr value tends to be less variable over the year. Therefore, PRcorr is usually higher than PR due to the weather-related effects. The effect refers to the cell temperature caused by ambient temperature and irradiation. On the other hand, the average ambient temperature is 26.97 °C and the average cell

temperature, which is normally high, is 33.19 °C. Another factor is rain. The occurrence of heavy rain, even if the temperature drops. However, this affects the amount of energy that receives from the PV production system.

Therefore, when judged by these two techniques, it was found that finding the potential of the system with PRcorr is higher than PR. Because the correct temperature of the panel results in higher production efficiency as the effects of seasonal changes are reduced. Techniques for determining the average temperature of the panel are important for the analysis of PRcorr. PVsyst is useful to estimate the performance ratio of the PV system with the measurement data in the study area. In the case of the grid-connected PV system, if the main grid has some problems due to voltage drop or outage. The PV system is disconnected from the main grid, even though the PV system capable of producing electricity. Finally, both PRcorr and PR calculations are affected.

### Conflict of Interest

The authors declare no conflict of interest.

### References

- [1] S. Sathiracheewin, K. Prommaneevat, "Short Term Analysis of PV System Performance in Different Areas of Thailand," in 2019 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 270-273, 2019, doi: 10.1109/ECTI-CON47248.2019.8955369.
- [2] Ministry of Energy Department of Alternative Energy Development and Efficiency, Thailand Alternative Energy Situation report, Bangkok, Thailand, 2017, doi: <http://www.dede.go.th>.
- [3] C. Li, "Comparative Performance Analysis of Grid-Connected PV Power Systems with Different PV Technologies in the Hot Summer and Cold Winter Zone," International Journal of Photoenergy, 2018, doi: 10.1155/2018/8307563.
- [4] M. Dhimish, S. Silvestre, "Estimating the impact of azimuth-angle variations on photovoltaic annual energy production," Clean Energy, 3, 47-58, 2019, doi: 10.1093/ce/zky022.
- [5] H. Singh, C. Sirisamphanwong, S. R. Santhi Murthy, "Effect of Tilt and Azimuth Angle on the Performance of PV Rooftop System," Applied Mechanics and Materials, 839, 159-164, 2016, doi: 10.4028/www.scientific.net/AMM.839.159.
- [6] T. Dierauf, A. Growitz, S. Kurtz, J.L.B. Cruz, E. Riley, C. Hansen, "Weather-Corrected Performance Ratio (No. NREL/TP-5200-57991)," National Renewable Energy Lab (NREL), Golden, CO, United States, 2013.
- [7] H.A. Basson, J.C. Pretorius, "Risk Mitigation of Performance Ratio Guarantees in Commercial Photovoltaic Systems," International Conference on Renewable Energy and Power Quality Journal, 14, 120-125, 2016. doi: <https://doi.org/10.24084/repqj14.244>.
- [8] E. Syahputra, R.A. Fajarin, E.A. Setiawan, "Characteristic Analysis of Photovoltaic On-Grid System in Tropical Region for Weather-Corrected Performance Ratio Calculation Method Implementation," E3S Web of Conf, 67, 1-7, 2018, doi: <https://doi.org/10.1051/e3sconf/20186701025>.
- [9] R. Srivastava, V.K. Giri, "Design of Grid Connected PV System using PVsyst," Journal on Electrical Engineering, 10, 92-96, 2016, doi: <https://doi.org/10.26634/jee.10.1.8195>.
- [10] N.M. Kumar, M.R. Kumar, P.R. Rejoice, M. Mathew, "Performance Analysis of 100 kWp Grid Connected Si-Poly Photovoltaic System Using PVsyst Simulation Tool," Energy Procedia, 117, 180-189, 2017, doi: <https://doi.org/10.1016/j.egypro.2017.05.121>.
- [11] I. Jamil, J. Zhao, L. Zhang, R. Jamil, S.F. Rafique, "Evaluation of Energy Production and Energy Yield Assessment based on Feasibility, Design, and Execution of 3× 50 MW Grid-Connected Solar PV Pilot Project in Nooriabad," International Journal of Photoenergy, 11, 2017, doi: <https://doi.org/10.1155/2017/6429581>.
- [12] L. Boughamrane, M. Boulaid, A. Tihane, A. Sdaq, K. Bouabid, A. Ihlal, "Comparative Analysis of Measured and Simulated Performance of the Moroccan First MV Grid Connected Photovoltaic Power Plant of Assa, Southern Morocco," J. Mater. Environ. Sci, 7, 4682-4691, 2016.
- [13] Ministry of Energy Department of Alternative Energy Development and Efficiency, The project to improve the solar energy potential map from satellite images for Thailand in 2017 report, Bangkok, Thailand, 2017. doi:<http://www.dede.go.th>
- [14] M.T. Boyd, "Comparative Performance and Model Agreement of Three Common Photovoltaic Array Configurations," Journal of solar energy engineering, 140(1),014503, 2018, doi: <https://doi.org/10.1115/1.4038314>.
- [15] J.S. Stein, PV Performance Modeling Methods and Practices, Sandia National Lab (SNL-NM), Albuquerque, NM, United States, 2017.
- [16] IEC 61724, Photovoltaic system performance monitoring-guidelines for measurement, data exchange and analysis, International Electrotechnical Commission, 1998.
- [17] A.Mermoud, B. Wittmer, PVSYST User's Manual, Switzerland, 2014.



# Fine Tuning the Performance of Parallel Codes

Sanaz Gheibi\*, Tania Banerjee, Sanjay Ranka, Sartaj Sahni

Department of Computer and Information Science and Engineering, University of Florida, Gainesville, 32611, USA

---

**ARTICLE INFO**
**Article history:**

Received: 03 June, 2020

Accepted: 11 August, 2020

Online: 31 August, 2020

---

**Keywords:**

Parallel speed up

Large matrices

Performance fine tuning

---

**ABSTRACT**

We propose a multilevel method to speed highly optimized parallel codes whose runtime increases faster than their workload. This method requires the ability to solve large instances by decomposing them into smaller instances. Using a simple parallel computing model, we derive a mathematical model that predicts whether or not our method can improve performance and also predicts the amount of improvement attainable. Our method is tested and shown to be effective on three highly optimized BLAS (Basic Linear Algebra Subprograms) routines from Intel's Math Kernel Library (MKL). Those routines are `cblas_dgemm`, `cblas_dtrmm` and `cblas_dsymm`. On the Intel Knights Landing (KNL) platform our method speeds `cblas_dgemm` by 33%, `cblas_dtrmm` by 50% and `cblas_dsymm` by 49% on double-precision matrices of size  $16K \times 16K$ , when the KNL's default memory-clustering configuration (cache-quadrant) is used.

---

## 1 Introduction

This paper is an extension of our previous paper on fine tuning a group of linear algebra libraries using a multi-level approach [1].

Our problems of interest are a group of parallel codes whose runtime grows faster than the problem size. The following example provides a quick demonstration of how such parallel codes behave and what motivates our proposed method.

Table 1 contains the runtime values for multiplying square matrices of different sizes on the intel KNL multi-level memory multicore computer. The memory-clustering configuration is set to its default which is cache quadrant. The multiplications are done in double precision using the highly optimized `cblas_dgemm` function from Intel MKL. The runtimes are obtained using either 32 or 64 cores. For each set of cores and each matrix size of  $2N \times 2N$ , the table also provides the runtime ratio when the matrix size increases from  $N \times N$  to  $2N \times 2N$ .

Since the matrix dimensions are increasing by a factor of 2 (thus increasing the workload by a factor of 8), we expect all the ratios to be 8. However, that is not the case and four of the ratios are greater than 8. That is where the runtime increases faster than the workload and that brings up possibilities of improvement. Here, for both 32 cores and 64 cores, we get better results if instead of performing a  $16K \times 16K$  ( $64K \times 64K$ ) multiplication, we perform 8 instances of  $8K \times 8K$  ( $32K \times 32K$ ) multiplications and combine the results (each input matrix is partitioned into 4 blocks of half the dimension; the result is obtained by doing 8 block multiplies and

4 block additions). The previous statement is valid as long as the overhead of performing 4 pairs of matrix additions is less than the runtime reduction obtained by multiplying 8 smaller instances.

We can further reduce the runtime by running the smaller instances in parallel. For example, instead of running one instance of  $32K \times 32K$  multiplication using 64 cores, we can run 4 parallel pairs of  $16K \times 16K$  multiplications using 32 cores in less total time. Again, we will have speed up if the overhead of adding 4 pairs of matrices is less than the obtained runtime reduction.

Table 1: Run times of double-precision matrix multiplications using 32 and 64 cores. Ratio is obtained by dividing the run time in the current row by the runtime in the previous row.

Matrix Size	32 Cores		64 Cores	
	Time(seconds)	Ratio	Time(seconds)	Ratio
$4K \times 4K$	0.37	-	0.32	-
$8K \times 8K$	2.32	6.27	1.51	4.71
$16K \times 16K$	20.56	8.86	12.32	8.16
$32K \times 32K$	159.01	7.73	97.52	7.92
$64K \times 64K$	1517.21	9.64	871.10	8.93

Our proposed method tries to improve the speed of the parallel codes whose runtime increases faster than the workload. This method uses multiple levels and works on applications that are of a decomposable nature. We start with a 2-level algorithm. At level 2, the problem is broken into a number of subproblems that are solved (serially or in parallel) using the 1-level algorithm and then

\*Corresponding Author: Sanaz Gheibi, Dept. of Computer and Information Science and Engineering, University of Florida, Gainesville, Florida, USA, sgheibi@ufl.edu

the results are combined to form the final solution. (By 1-level algorithm we mean the parallel and highly efficient algorithm whose runtime grows faster than its workload). Now, suppose the 2-level algorithm itself demonstrates the property that its runtime grows faster than its workload. In this case, we can gain further speed up by using a 3-level algorithm in which the problem is broken into a proper number of subproblems which are solved using the 2-level algorithm. We can continue increasing the number of levels in a similar fashion.

While our proposed method has much in common with divide and conquer, block matrix multiplication and Cannon's method for matrix multiplication [2], there are subtle differences. In a typical divide and conquer algorithm [3]- [7] the division scheme is the same across the levels (an example being the Strassen's matrix multiplication algorithm [3]); whereas in our method the number of subproblems into which the problem is broken can be different in different levels. Block matrix multiplication has been used both in serial and parallel matrix multiplication algorithms [2, 8, 9]. It can be viewed as a 1-level divide-and-conquer algorithm; while our method uses multiple levels of decomposition and combination. Our method is different from Cannon's algorithm in both the number of levels and the block's movement scheme. Cannon's movement scheme is designed for mesh-connected parallel computers with wraparound and uses row-and-column circular shifts. Our method uses a serpentine movement scheme that will be described in section 3.2.

Regardless of the above-mentioned similarities, the value of our method is that it can speed up codes whose runtimes grow faster than their workload. Here, the codes to be optimized are viewed as black boxes that can be already parallel, highly optimized and have their own complexities. As a result, this method can be applied to a range of problems broader than just linear algebra libraries.

We use a simple parallel computing model to formulate the conditions when our method results in speed up and also predict the amount of speed up. In section 5 we will demonstrate the effectiveness of our algorithm and our prediction formula on three linear algebra libraries from Intel MKL.

Our experimental results are obtained using the Knights Landing (KNL) computer. Although this architecture has been discontinued by Intel, we believe our fine-tuning method remains relevant for the following reasons. First, Knights Landing clusters are still used in many national laboratories and second, our method can be used in future architectures that have multilevel memory.

The rest of the paper is organized as follows. Section 2 summarizes some of the related work. In Section 3 we describe a solution framework for general problems and then describe its specific application to a class of linear algebra functions. Section 4 gives a brief description of KNL architecture. Section 5 presents the experimental results; and we conclude in Section 6.

## 2 Related Work

This section summarizes some of the previous methods used to speed up applications.

Tiling has been used extensively to improve data locality and cache efficiency. This method works by breaking the problem into

subproblems or tiles in such a way that the amount of memory reuse in faster levels of memory hierarchy increases. Tiling has been used in a large group of problems including sparse and dense matrix-matrix multiplication [10, 11], parallel tensor transpose [12] and LU factorization [11]. In [13] the authors use tiling along with thread batching and architecture specific optimizations to speed up Alternating Least Squares algorithm used in recommender systems. They use a fine grained tiling in order to mitigate the imbalance in threads' workload that results from the sparse nature of the problem at hand. In [14] the authors use temporal tiling alongside spatial tiling to improve cache efficiency of large scale stencil computations. In a particular class of stencil computations that is the focus of their work, each time step only depends on a limited number of previous time steps and hence temporal tiling can also be incorporated.

Data reordering has been done to change the order in which the input data are processed and increase the cache efficiency. This method performs by mapping the original data indices to the new indices and has been used to improve the efficiency of both regular applications such as dense matrix kernels [15] and irregular applications such as molecular dynamic simulation and hydrodynamic computations [16]. Other methods have applied data reordering to matrix multiplication [17] and embedded multimedia processing [18] with an extra assumption that tiling is already in place.

Reducing communication overhead has also been used to speed up applications. One way to reduce communication is through asynchrony. Asynchrony can be very effective in runtime reduction as synchronizing parallel processes results in much overhead. However, asynchrony can only be used for applications whose correctness is not hurt by lack of synchronization; an example being Stochastic Gradient Descent for sparse matrix factorization [19, 20]. Computation communication overlapping has also been used to hide the communication latency in a group of problems including three-dimensional Fast Fourier Transform [21], data-parallel training of Convolutional Neural Networks [22] and distributed Stochastic Gradient Descent [23]. Block reordering is another method used in [24] to reduce the amount of communication required in Strassen's algorithm. Reducing the communication traffic has also been used to speed up applications in large scale distributed systems such as mobile networks [25].

Load balancing has been used to improve the response time of distributed applications by evenly distributing the workload among the worker units and preventing one worker to be overloaded while the other workers are idle. Among the main areas of application are cloud computing[26] and edge computing[27]. In [28] the authors have used load balancing to speed up matrix-vector multiplication in large scale machine learning and data mining applications. The authors of [29] have developed a simulator to select the best balancing technique for scientific computing applications whose stochastic and irregular natures greatly contribute to load imbalance.

Automatic data placement can also be used to improve memory efficiency of memory bound applications. These methods work by analysing the memory access patterns of a program either in compile time [30] or during the runtime using a lightweight profiling method [31] and automatically placing data segments on optimal memory components.

All of the above-mentioned methods could be used in the 1-level algorithm that is our starting point; i.e, the algorithm for which the

runtime grows faster than the workload. What we do is essentially taking the highly optimized code as a black box and adding a layer of fine-tuning on top of that.

### 3 Proposed Approach

#### 3.1 General framework

In this section, we provide a general framework for speeding up any parallel code that is decomposable and whose runtime increases faster than its workload. In the subsequent sections, we will demonstrate how to apply our method to speed up the highly optimized codes for general, triangular and symmetric matrix multiplication.

We have already described our proposed method in the introduction section. Here we give a brief summary and present the pseudo codes. As we have already mentioned, our method works in multiple levels. A 2-level implementation divides the problem into subproblems and solves them using a 1-level code either serially, in parallel or using a hybrid of the both. Then the solutions to the subproblems are combined to form the general solution. Similarly, an  $n$ -level code divides the problem into subproblems that are solved using an  $(n - 1)$ -level algorithm and then combines the results.

Our  $n$ -level algorithm for  $n > 2$  is described in Algorithm 1. Given a problem  $A$  of size  $s$ , we decompose it into  $s/s_b$  subproblems of size  $s_b$  each. In lines 2-4, the subproblems are solved serially using the  $(n - 1)$ -level algorithm ( $GEN_{n-1}$ ). Then in line 5 the solutions to the subproblems are combined to form the final solution  $C$ . Note that multithreading can be incorporated in all the functions in this section and in subsections 3.2, 3.3 and 3.4. All the threads can be used to copy the subproblems (as we will explain in the next paragraph) and to combine the results. If subproblems are solved serially, then all the threads will be used to solve each subproblem. However, if  $b$  subproblems are solved in parallel, then  $1/b$  of the total threads will be assigned to each subproblem. Also note that combination of each partial solution with the final result can happen at the same time that the partial solution is produced. In that case, the same number of threads used to solve each subproblem are used to combine its result with the final solution.

$GEN_{n-1}$  in its turn calls  $GEN_{n-2}$  and this process continues until  $GEN_2$  is called. The last level of problem decomposition happens at  $GEN_2$  and then the subproblems are solved using  $GEN_1$  which is the highly optimized parallel code to be sped up. In order to use the fast levels of memory hierarchy, we copy the subproblems into a fast memory and that is done in level 2 which is the level at which the actual problem solving is done. Note that in many cases the fast memory is one of the levels of cache hierarchy which we can not directly access and copy blocks into it. In those cases we still make copies of the blocks in the main memory and that causes the blocks to be automatically brought into the cache. Due to the space constraints of the fast memory, the parallel solving of subproblems happens only at level 2. And that is also why we have two sets of pseudo codes one for level  $n, n > 2$  and one for level 2. Algorithm 2 describes our two-level algorithm.

Using  $GEN_2$ , a problem  $\bar{A}$  of size  $\bar{s}$  is divided into  $\bar{s}/\bar{s}_b$  subproblems of size  $\bar{s}_b$  each.  $\bar{C}$  is the output and  $b$  is the degree of concurrency. Meaning that, at each iteration, the input to  $b$  of the

subproblems are copied into the fast memory (line 3), the  $b$  subproblems are solved concurrently using  $GEN_1$  (lines 4-6) and then the results are combined to form part of the final solution (line 7).

#### 3.2 Matrix multiplication

This section illustrates our method by demonstrating how it works to speed up matrix multiplication. Software packages such as AMD Core Math Library (ACML), OpenBLAS, ATLAS and Intel Math Kernel Library (MKL) provide efficient parallel implementations of matrix multiplication as a part of the Basic Linear Algebra Subroutines (BLAS). In level 1 of our implementation, we use `cblas_dgemm` that is the double precision matrix multiplication code from MKL.

---

##### Algorithm 1 n-level Algorithm

---

```

1: function  $GEN_n(A)$   $\triangleright$  Problem  $A$  of size  $s$  is decomposed into
 $s/s_b$  subproblems of size  $s_b$  each.  $C$  is the output.  $C_p[I]$  is the
solution to the subproblem  $A[I]$ .
2:   for  $I = 1$  to  $s/s_b$  do
3:      $C_p[I] \leftarrow GEN_{n-1}(A[I])$ 
4:   end for
5:   Combine  $C_p[I], I = 1, 2, \dots, s/s_b$  with  $C$ 
6:   return  $C$ 
7: end function

```

---



---

##### Algorithm 2 Two-level Algorithm

---

```

1: function  $GEN_2(\bar{A})$   $\triangleright$  Problem  $\bar{A}$  of size  $\bar{s}$  is decomposed
into  $\bar{s}/\bar{s}_b$  subproblems of size  $\bar{s}_b$ .  $\bar{C}$  is the output and  $\bar{C}_p[I]$  is the
solution to the subproblem  $\bar{A}[I]$ .  $b$  is the degree of concurrency.
2:   while  $I \leq \bar{s}/(\bar{s}_b \times b)$  do
3:     Copy  $b$  subproblems  $\bar{A}[I]$  to  $\bar{A}[I + b - 1]$ 
4:     start  $b$  parallel function calls
5:        $\bar{C}_p[J] \leftarrow GEN_1(\bar{A}[J])$  for  $J = I, \dots, (I + b - 1)$ 
6:     wait for all functions to return
7:     Combine  $\bar{C}_p[J], J = I, \dots, (I + b - 1)$  with  $C$ 
8:      $I \leftarrow I + b$ 
9:   end while
10:  return  $\bar{C}$ 
11: end function

```

---



---

##### Algorithm 3 L-level Matrix Multiplication

---

```

1: function  $MXM_L(A, B)$   $\triangleright$  Input matrices  $A_{m \times p}$ ,  $B_{p \times n}$ 
and output matrix  $C_{m \times n}$  are divided into blocks of sizes  $s_m \times s_p$ ,
 $s_p \times s_n$  and  $s_m \times s_n$  respectively.
2:   for  $I = 1$  to  $m/s_m$  do
3:     for  $J = 1$  to  $n/s_n$  do
4:       for  $K = 1$  to  $p/s_p$  do
5:          $C[I, J] \leftarrow C[I, J] + MXM_{L-1}(A[I, K], B[K, J])$ 
6:       end for
7:     end for
8:   end for
9:   return  $C$ 
10: end function

```

---

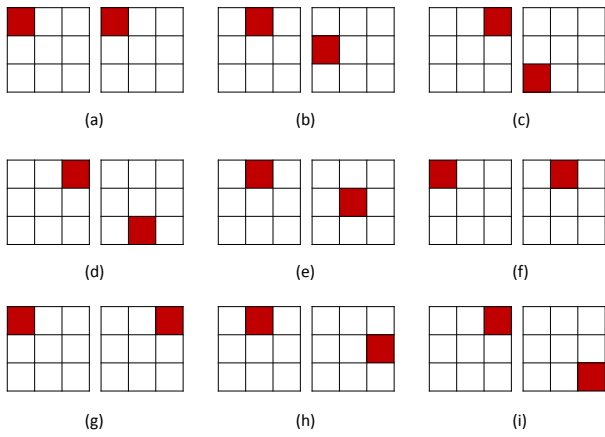


Figure 1: In parts (a)-(i) the chosen blocks in  $M1$  (the matrix on the left) and  $M2$  (the matrix on the right) are used to compute row 1 of  $M3$ . The resulting matrix  $M3$  is not shown.  $M3[1, 1]$ ,  $M3[1, 2]$  and  $M3[1, 3]$  are computed in parts (a)-(c), (d)-(f) and (g)-(i) respectively. If we follow the path of the selected blocks in each matrix, a pattern similar to snake movement will result.

Our method for the matrix multiplication problem follows the general framework presented in the previous section. Here, the input is two matrices  $A$  and  $B$  of sizes  $m \times p$  and  $p \times n$  to be multiplied and the result is stored in matrix  $C$  of size  $m \times n$ . At level  $L > 2$ , the matrices are divided into blocks of sizes  $s_m \times s_p$ ,  $s_p \times s_n$  and  $s_m \times s_n$  respectively. Product of matrix blocks  $A[I, K]$  and  $B[K, J]$  is computed using a lower level function and the result gets accumulated in the corresponding output block  $C[I, J]$ . Algorithm 3 describes our  $L$ -level algorithm ( $MXM_L$ ).

$MXM_2$  is our two-level matrix-matrix multiplication algorithm. Input matrices  $M1$  and  $M2$  are multiplied and the result is stored in matrix  $M3$ . The matrix sizes are  $\bar{m} \times \bar{p}$ ,  $\bar{p} \times \bar{n}$  and  $\bar{m} \times \bar{n}$  respectively.

As we previously mentioned, in level 2, the subproblems (matrix blocks in this case) are copied into the fast memory, so that the highly optimized level 1 code can benefit from the high bandwidth. To optimize the copying process, we use a serpentine pattern to iterate through the input sub-matrices. Figure 1 shows an example of this serpentine movement and how it reduces the amount of data transfer between the slow and the fast memories. Here, a  $3 \times 3$  grid is used to divide matrices into smaller blocks. Degree of concurrency  $b = 1$  that means one block from each matrix is selected at a time. Here, for computing  $M3[1, 1]$ , blocks  $[1, 1]$ ,  $[1, 2]$  and  $[1, 3]$  are selected from  $M1$  with their corresponding blocks from  $M2$ . Now, for computing  $M3[1, 2]$ , instead of selecting blocks of  $M1$  in the previous order, i.e. starting from  $M1[1, 1]$  and moving left-to-right, we start from  $M1[1, 3]$  and move right-to-left. As a result, we won't have to copy  $M1[1, 3]$  again as it is already in the fast memory. For  $M3[1, 3]$  we start from  $M1[1, 1]$  and move left-to-right again as then,  $M1[1, 1]$  is already present in the fast memory. If we consider the order in which we iterate the blocks of  $M1$  and their corresponding blocks in  $M2$ , each of them forms a pattern like a snake movement. Using this iteration pattern we reduce the number of copied blocks by 2 for computing the first row of  $M3$ . The effect gets more significant for larger matrices with higher degrees of concurrency.

We should note that in our implementation, we also used serpentine pattern for  $MXM_L$ ; however, as the blocks for  $L$ -level matrix

multiplication ( $L > 2$ ) are usually too large to fit in the fast levels of memory hierarchy, there is not much benefit in using this pattern. Therefore, we eliminated the details from algorithm 3 to make it more concise and readable.

Our 2-level algorithm  $MXM_2$  is described in algorithm 4. It follows the general structure of algorithm 2. Matrices  $M1$ ,  $M2$  and  $M3$  of sizes  $\bar{m} \times \bar{p}$ ,  $\bar{p} \times \bar{n}$  and  $\bar{m} \times \bar{n}$  are divided into blocks of sizes  $\bar{s}_{b1} \times \bar{s}_{b3}$ ,  $\bar{s}_{b3} \times \bar{s}_{b2}$  and  $\bar{s}_{b1} \times \bar{s}_{b2}$  respectively. In lines 8-15  $b$  blocks from each of the two input matrices are copied, then in lines 16-19  $b$  instances of our 1-level algorithm (`cblas_dgemm`) are called in parallel on each of the block pairs. Finally, the partial results are combined in lines 20, 21. The rest of the algorithm deals with the order in which the blocks are selected and whether or not they are already present in the fast memory (if a block is already present, then it should not be copied again). That is the serpentine pattern we previously talked about and is implemented by  $K_{dir}$  and  $J_{dir}$  that define the moving directions along the columns of  $M1$  and rows of  $M2$  respectively. Variables  $copy_{M1}$  and  $copy_{M2}$  define when the blocks of  $M1$  and  $M2$  should be copied.

**Algorithm 4** Two-level Matrix Multiplication

```

1: function  $MXM_2(M1, M2)$  ▷ Input
   matrices  $M1_{\bar{m} \times \bar{p}}$  and  $M2_{\bar{p} \times \bar{n}}$  and the output matrix  $M3_{\bar{m} \times \bar{n}}$  are
   divided into blocks of sizes  $\bar{s}_{b1} \times \bar{s}_{b3}$ ,  $\bar{s}_{b3} \times \bar{s}_{b2}$  and  $\bar{s}_{b1} \times \bar{s}_{b2}$ 
   respectively.  $b$  is the degree of concurrency.  $C_p[U]$  stores the
   partial result of  $M1[I, U] \times M2[U, J]$ .  $MXM_1$  is cblas_dgemm()
   in our case.
2:    $copy_{M1} \leftarrow true$ ,    $copy_{M2} \leftarrow true$ 
3:    $K \leftarrow 1$ ,    $J \leftarrow 1$ ,    $K_{dir} \leftarrow 1$ ,    $J_{dir} \leftarrow 1$ 
4:   for  $I = 1$  to  $\bar{m}/\bar{s}_{b1}$  do
5:     for  $\bar{n}/\bar{s}_{b2}$  iterations involving  $J$  do
6:       for  $\bar{p}/\bar{s}_{b3}$  iterations involving  $K$  do
7:         if  $copy_{M1}$  then
8:           Copy  $b$  blocks from  $M1$ :
9:           ( $M1[I, K]$  to  $M1[I, K + (b - 1) \times K_{dir}]$ )
10:        end if
11:       if  $copy_{M2}$  then
12:         copy  $b$  blocks from  $M2$ :
13:         ( $M2[K, J]$  to  $M[K + (b - 1) \times K_{dir}, J]$ )
14:       end if
15:       start  $b$  parallel function calls
16:          $C_p[U] \leftarrow MXM_1(M1[I, U], M2[U, J])$ 
17:         for  $U = K, \dots, K + (b - 1) \times K_{dir}$ 
18:           wait for all functions to return
19:            $M3[I, J] \leftarrow M3[I, J] + C_p[K] +$ 
20:            $\dots + C_p[K + (b - 1) \times K_{dir}]$ 
21:            $K \leftarrow K + b \times K_{dir}$ ;  $copy_{M1} \leftarrow true$ 
22:         end for
23:        $K_{dir} \leftarrow -K_{dir}$ ;  $copy_{M1} \leftarrow false$ 
24:        $J \leftarrow J + J_{dir}$ ;  $copy_{M2} \leftarrow true$ 
25:     end for
26:      $J_{dir} \leftarrow -J_{dir}$ ;  $copy_{M2} \leftarrow false$ ;  $copy_{M1} \leftarrow true$ 
27:   end for
28:   return  $M3$ 
29: end function

```

After all blocks corresponding to the  $K$  variable in the block



matrix multiplication formula

$$M3[I, J] = \sum_{K=1}^{\bar{n}/\bar{s}_{b_3}} M1[I, K] \times M2[K, J] \quad (1)$$

have been used, we change the direction in which we select those blocks for computing  $M3[I, J + 1]$ . That is when blocks from  $M1$  are reused. Similarly, after  $M3[I, 1] \dots M3[I, \bar{n}/\bar{s}_{b_2}]$  have all been computed; we compute  $M3[I + 1, J]$ s in reverse direction and  $b$  blocks from  $M2$  are reused.

### 3.3 Triangular Matrix Multiplication

This section illustrates how our methodology may be employed to multiply two matrices where one of the matrices is (upper/lower) triangular. The highly optimized function we use at the first level is `cblas_dtrmm` from Intel MKL.

#### Algorithm 5 $L$ -level Triangular Matrix Multiplication

```

1: function  $TrMXM_L(A, B)$  ▷ Input
   matrices  $A_{m \times m}$ ,  $B_{m \times n}$  and output matrix  $C_{m \times n}$  are divided into
   blocks of sizes  $s_{b_1} \times s_{b_1}$ ,  $s_{b_1} \times s_{b_2}$  and  $s_{b_1} \times s_{b_2}$  respectively.  $b$ 
   is the number of block multiplications done in parallel.
2: for  $I = 1$  to  $m/s_{b_1}$  do
3:   for  $J = 1$  to  $n/s_{b_2}$  do
4:     for  $K = 1$  to  $I$  do
5:       if  $I == K$  then
6:          $C[I, J] \leftarrow C[I, J] +$ 
7:          $TrMXM_{L-1}(A[I, K], B[K, J])$ 
8:       else
9:          $C[I, J] \leftarrow C[I, J] + MXM(A[I, K], B[K, J])$ 
10:      end if
11:    end for
12:  end for
13: end for
14: return  $C$ 
15: end function

```

#### Algorithm 6 Two-level Triangular Matrix Multiplication

```

1: function  $TrMXM_2(M1, M2)$  ▷ Input matrices  $M1_{\bar{m} \times \bar{m}}$  and
    $M2_{\bar{m} \times \bar{n}}$  and the output matrix  $M3_{\bar{m} \times \bar{n}}$  are divided into blocks
   of sizes  $\bar{s}_{b_1} \times \bar{s}_{b_1}$ ,  $\bar{s}_{b_1} \times \bar{s}_{b_2}$  and  $\bar{s}_{b_1} \times \bar{s}_{b_2}$  respectively.  $b_1$  is
   the number of triangular block multiplications done in parallel
   and  $b_2$  is the number of ordinary block multiplications done in
   parallel.  $TrMXM_1$  is cblas_dtrmm() in our case.
2: for  $I = 1$  to  $\bar{m}/\bar{s}_{b_1}$  increment  $= b_1$  do ▷ phase 1
3:   for  $J = 1$  to  $\bar{n}/\bar{s}_{b_2}$  do
4:     Copy  $b_1$  blocks of  $M1$  along the main diagonal:
5:     i.e.  $M1[I, I]$  to  $M1[I + b_1 - 1, I + b_1 - 1]$ 
6:     Copy  $b_1$  blocks of  $M2$  along column  $J$ :
7:     i.e.  $M2[I, J]$  to  $M2[I + b_1 - 1, J]$ .
8:     start  $b_1$  parallel function calls
9:        $C[U, J] \leftarrow C[U, J] +$ 
10:       $TrMXM_1(M1[U, U], M2[U, J])$ 
11:     for  $U = I, \dots, (I + b_1 - 1)$ 
12:     wait for all functions to return
13:   end for
14: end for
15:  $cnt \leftarrow 0$ 
16: for  $I = 2$  to  $\bar{m}/\bar{s}_{b_1}$  do ▷ phase 2
17:   for  $J = 1$  to  $I - 1$  do
18:     for  $K = 1$  to  $\bar{n}/\bar{s}_{b_2}$  do
19:       Copy  $M1[I, J]$ ; Copy  $M2[J, K]$ ;  $cnt \leftarrow cnt + 1$ 
20:       if  $cnt == b_2$  then
21:         start  $b_2$  parallel function calls
22:         Call  $b_2$  instances of  $\overline{MXM}$  over
23:         blocks copied from  $M1$  and  $M2$ 
24:         wait for all functions to return
25:         Sum up partial results of  $\overline{MXM}$ 
26:         to  $b_2$  corresponding blocks of  $M3$ 
27:          $cnt \leftarrow 0$ 
28:       end if
29:     end for
30:   end for
31: end for
32: return  $M3$ 
33: end function

```

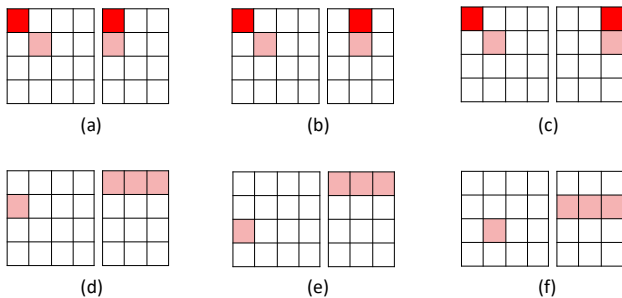


Figure 2: Phase 1 and 2 of  $TrMXM_2$ . Parts (a)-(c) show three steps from phase 1 where at each step two triangular blocks from the matrix  $M1$  are multiplied in parallel to the corresponding blocks of  $M2$ . Each pair of blocks that are multiplied at each step are shown with the same shade. The resulting matrix  $M3$  is not shown. Parts (d)-(f) show three steps from phase 2 where an ordinary block from  $M1$  is multiplied concurrently to three blocks from  $M2$ . For all the parts (a)-(f),  $M1$  is the matrix on the left,  $M2$  is the matrix on the right and  $M3$  is not shown.

Without loss of generality, we assume that the triangular matrix appears at the left side of multiplication. Therefore, in our triangular matrix multiplication problem, matrices  $A$  and  $B$  of sizes  $m \times m$  and  $m \times n$  respectively, are multiplied resulting in a matrix  $C$  of size  $m \times n$ . Notice that  $A$  being triangular means that it should be a square matrix. Again, without loss of generality, we assume  $A$  is lower triangular. The general framework is also applicable here. However, the nature of the problem implies some modifications in our implementation. Here, we divide  $A$  into blocks of size  $s_{b_1} \times s_{b_1}$ . If  $g = \frac{m}{s_{b_1}}$ , then we will have  $g \times g$  blocks,  $g$  of which are triangular,  $\frac{g \times g - g}{2}$  of which are ordinary (non-triangular) blocks and  $\frac{g \times g - g}{2}$  of which are zero blocks. The zero blocks are ignored while the ordinary blocks are involved in ordinary matrix multiplication and the triangular blocks get involved in triangular matrix multiplication.

Algorithm 5 describes the  $L$ -level algorithm ( $TrMXM_L$ ) where  $L > 2$ . Here, we break  $A$ ,  $B$  and  $C$  into blocks of sizes  $s_{b_1} \times s_{b_1}$ ,

$s_{b1} \times s_{b2}$  and  $s_{b1} \times s_{b2}$  respectively. In lines 6, 7 the triangular blocks from  $A$  are multiplied to the corresponding blocks from  $B$  whereas in line 9 the ordinary blocks from  $A$  are multiplied to the corresponding blocks from  $B$ . Here,  $MXM$  is an efficient presumably multilevel matrix multiplication algorithm.

For our two-level algorithm we make an observation: `cblas_dtrmm` that is used at the leaf level of triangular matrix multiplication is optimized to benefit from the lower amount of computations that result from the zero elements. Therefore, we can further improve the runtime by grouping the triangular blocks together and the ordinary blocks together. That way, the more efficient triangular matrix multiplications should not wait for the results of ordinary matrix multiplications. At each iteration, we operate either on  $b1$  triangular block multiplications or on  $b2$  ordinary block multiplications in parallel where  $b1$  can be different from  $b2$ .

Our algorithm uses two phases in order to correctly select and group together the matrix blocks. In the first phase,  $b1$  triangular blocks from the main diagonal of  $M1$  are copied and at each step, they are multiplied concurrently to one other block from  $M2$ . In the second phase, ordinary blocks from  $M1$  are picked in a top-to-bottom and left-to-right order. For each block  $x$  from  $M1$ , we iterate over all blocks of  $M2$  to which  $x$  should be multiplied. When the total number of block pairs reaches  $b2$ , we stop and perform  $b2$  multiplications concurrently.

The above-mentioned process is illustrated in figure 2. The top row shows three steps from phase 1 and the bottom row shows three steps from phase 2. In each part,  $M1$  is the matrix on the left and  $M2$  is the matrix on the right where a  $4 \times 4$  grid is used to block  $M1$  and a  $4 \times 3$  grid is used to block  $M2$ . Here,  $b1 = 2$  and  $b2 = 3$ . In part (a), two triangular blocks  $M1[1, 1]$  and  $M2[2, 2]$  are copied to the fast memory; and through steps (a) - (c), those blocks are concurrently multiplied to  $M2[1, J]$  and  $M2[2, J]$  respectively ( $J = 1, 2, 3$ ). In part (d), the ordinary block  $M1[2, 1]$  is copied three times in the fast memory and three parallel multiplications are done on  $M1[2, 1]$  and  $M2[1, J]$ ,  $J = 1, 2, 3$ . In parts (e) and (f) the next ordinary blocks from  $M1$  are chosen in the top-to-bottom and left-to-right order and multiplied concurrently with three blocks of  $M2$ . Note that if the degree of concurrency for the ordinary blocks ( $b2$ ) was equal to 4, then in part (d), block  $M1[3, 1]$  would also be copied along with a second copy of block  $M2[1, 1]$ . Then there would be 4 parallel multiplications, three of them involving  $M1[2, 1] \times M2[1, J]$  ( $J = 1, 2, 3$ ) and one of them  $M1[3, 1] \times M2[1, 1]$ . The cases for parts (e) and (f) would be similar.

Algorithm 6 describes our 2-level triangular matrix multiplication algorithm,  $TrMXM_2$ . We divide the input matrices  $M1_{\overline{m} \times \overline{m}}$  and  $M2_{\overline{m} \times \overline{n}}$  into blocks of sizes  $\overline{s}_{b1} \times \overline{s}_{b1}$  and  $\overline{s}_{b1} \times \overline{s}_{b2}$  respectively. The output matrix  $M3_{\overline{m} \times \overline{n}}$  will also be divided into blocks of size  $\overline{s}_{b1} \times \overline{s}_{b2}$ . Lines 2-14 implement the first phase of the algorithm where we run  $b1$  parallel `cblas_dtrmm` instances on the triangle-rectangle block multiplications. Lines 15-33 implement the second phase of our algorithm where we run  $b2$  parallel instances of  $MXM$  over the square-rectangle block multiplications where  $MXM$  is an efficient matrix multiplication algorithm. The auxiliary variable  $cnt$  is used to keep track of the square blocks being multiplied in parallel.

### 3.4 Symmetric Matrix Multiplication

In this section, we illustrate our methodology on symmetric matrix multiplication. This is when one of the matrices involved in the multiplication is symmetric. The highly optimized parallel function we use at the first level is `cblas_dsymm` from Intel MKL.

Without loss of generality, we assume that the symmetric matrix appears at the left side of multiplication. Therefore, the problem can be stated as the following: matrices  $A$  and  $B$  of sizes  $m \times m$  and  $m \times n$  respectively where  $A$  is a symmetric matrix, are multiplied and the result is stored in matrix  $C$  of size  $m \times n$ . Notice again that symmetry is only defined for square matrices; that is why  $A$  is  $m \times m$ . The same methodology is used as the general framework. Here again, similar to triangular multiplication, the nature of the problem implies some modifications to the implementation details. We divide  $A$  into blocks of size  $s_{b1} \times s_{b1}$ . If  $g = \frac{m}{s_{b1}}$ , then we will have  $g \times g$  blocks,  $g$  of which are symmetric. The ordinary blocks are involved in ordinary matrix multiplication and the symmetric blocks get involved in symmetric matrix multiplication.

Algorithm 7 describes the  $L$ -level algorithm ( $SyMXM_L$ ) where  $L > 2$ . Here, we break  $A$ ,  $B$  and  $C$  into blocks of sizes  $s_{b1} \times s_{b1}$ ,  $s_{b1} \times s_{b2}$  and  $s_{b1} \times s_{b2}$  respectively. In lines 6, 7 the symmetric blocks from  $A$  are multiplied to the corresponding blocks from  $B$  whereas in line 9 ordinary blocks from  $A$  are multiplied to the corresponding blocks from  $B$ . Here, similar to that of triangular matrix multiplication,  $MXM$  is an efficient matrix multiplication algorithm that could be multilevel.

---

#### Algorithm 7 $L$ -Level Symmetric Matrix Multiplication

---

```

1: function  $SyMXM_L(A, B)$  ▷ Input matrices
    $A_{m \times m}, B_{m \times n}$  and output matrix  $C_{m \times n}$  are divided into blocks of
   sizes  $s_{b1} \times s_{b1}, s_{b1} \times s_{b2}$  and  $s_{b1} \times s_{b2}$  respectively.
2:   for  $I = 1$  to  $m/s_{b1}$  do
3:     for  $J = 1$  to  $n/s_{b2}$  do
4:       for  $K = 1$  to  $m/s_{b1}$  do
5:         if  $I == K$  then
6:            $C[I, J] \leftarrow C[I, J] +$ 
7:              $SyMXM_{L-1}(A[I, K], B[K, J])$ 
8:         else
9:            $C[I, J] \leftarrow C[I, J] + MXM(A[I, K], B[K, J])$ 
10:        end if
11:      end for
12:    end for
13:  end for
14:  return  $C$ 
15: end function

```

---

Algorithm 8 describes our 2-level symmetric matrix-matrix multiplication algorithm,  $SyMXM_2$ . We divide  $M1_{\overline{m} \times \overline{m}}$  and  $M2_{\overline{m} \times \overline{n}}$  into blocks of sizes  $\overline{s}_{b1} \times \overline{s}_{b1}$  and  $\overline{bs}_1 \times \overline{bs}_2$  respectively. The resulting matrix  $M3_{\overline{m} \times \overline{n}}$  will also be divided into blocks of size  $\overline{s}_{b1} \times \overline{s}_{b2}$ . Similar to triangular matrix multiplication, we group the symmetric blocks together and the ordinary blocks together. The algorithm works in two phases where at each step of the first phase,  $b1$  symmetric blocks multiplications and at each step of the second phase  $b2$  ordinary block multiplications are performed in parallel.  $b1$  and  $b2$  are not necessarily equal. Lines 2-14 implement the first phase

and lines 16-33 implement the second phase of our algorithm. The auxiliary variable *cnt* is used to keep track of the ordinary blocks being multiplied in parallel. *MXM* is an efficient matrix multiplication algorithm.

**Algorithm 8** Two-level Symmetric Matrix Multiplication

```

1: function SyMXM2(M1, M2)  ▶ Input matrices M1m̄×m̄ and
   M2m̄×n̄ and the output matrix M3m̄×n̄ are divided into blocks
   of sizes sb1 × sb1, sb1 × sb2 and sb1 × sb2 respectively. b1 is
   the number of symmetric block multiplications done in parallel
   and b2 is the number of ordinary block multiplications done in
   parallel. SyMXM1 is cblas_dsymm() in our case.
2:   for I = 1 to m̄/sb1 increment = b1 do           ▶ phase 1
3:     for J = 1 to n̄/sb2 do
4:       Copy b1 blocks of M1 along the main diagonal:
5:       i.e. M1[I, I] to M1[I + b1 - 1, I + b1 - 1]
6:       Copy b1 blocks M2 along column J:
7:       i.e. M2[I, J] to M2[I + b1 - 1, J].
8:       start b1 parallel function calls
9:       C[U, J] ← C[U, J] +
10:      SyMXM1(M1[U, U], M2[U, J])
11:      for U = I, ..., (I + b1 - 1)
12:        wait for all functions to return
13:      end for
14:    end for
15:    cnt ← 0
16:    for I = 1 to m̄/sb1 do           ▶ phase 2
17:      for J = 1 to m̄/sb1, I ≠ J do
18:        for K = 1 to n̄/sb2 do
19:          Copy M1[I, J]
20:          Copy M2[J, K]
21:          cnt ← cnt + 1
22:          if cnt == b2 then
23:            start b2 parallel function calls
24:            Call b2 instances of MXM over
25:            blocks copied from M1 and M2
26:            wait for all functions to return
27:            Sum up partial results of MXM
28:            to b2 corresponding blocks of M3
29:            cnt ← 0
30:          end if
31:        end for
32:      end for
33:    end for
34:    return M3
35: end function

```

3.5 Performance Analysis

In this section, a simple shared-memory model in which processes communicate by reading/writing from/to a shared memory is used to describe the runtime of the parallel codes. Using this model, we formulate the conditions under which our method will be effective and also the speed up we can expect. Let  $T_{x,y}^i$  be the time our algorithm takes at level *i* to solve a problem of size *y* using *x* processors. Table 2 illustrates our notation for a level 1 code. Here, *b* and *c* are

integer multipliers.

Table 2: Notations used for the runtime of 1-level algorithm depend on the problem size and the number of processors.

	<i>p</i> processors	<i>bp</i> processors
problem size= <i>n</i>	$T_{p,n}^1$	$T_{bp,n}^1$
problem size= <i>cn</i>	$T_{p,cn}^1$	$T_{bp,cn}^1$

Let,

$$r = \frac{T_{bp,cn}^1}{T_{p,cn}^1} \tag{2}$$

$$k = \frac{T_{p,cn}^1}{f(c) \times T_{p,n}^1} \tag{3}$$

Where *f(c)* is defined as follows: if we decompose a problem of size *cn* into subproblems of size *n*; then *f(c)* denotes the number of those subproblems. In the case of matrix multiplication, for example, if the initial problem is multiplying two  $2n \times 2n$  matrices and we break each matrix into four matrices of size  $n \times n$ ; then we should multiply eight matrix pairs of size  $n \times n$  each. Here, we have *c* = 2 and *f(c)* = 8.

Let  $T_i = T_{bp,cn}^i$ . We wish to find the ratio  $T_i/T_1$ . We consider two cases: *i* = 2 and *i* > 2. The reason we consider two cases is that we use parallelism only at the second level.

We define *g(c, n)* to be the time it takes to combine *f(c)* subproblems of size *n* each. For a 2-level algorithm, if the subproblems run serially, then all the *bp* processors can be used for each of them. We will have:

$$T_{bp,cn}^2 = f(c)T_{bp,n}^1 + g(c, n) \tag{4}$$

However, if we run *b* subproblems in parallel at a time, then only *p* processors can be used for each of the subproblems. In that case, the runtime equation will be:

$$T_{bp,cn}^2 = \frac{f(c)}{b}T_{p,n}^1 + g(c, n) \tag{5}$$

In the rest of this subsection, we use equation 5 to describe  $T_2$ . That is

$$T_2 = \frac{f(c)}{b}T_{p,n}^1 + g(c, n)$$

Therefore,

$$\frac{T_2}{T_1} = \frac{1 + rbk\delta}{rbk} \tag{6}$$

where

$$\delta = \frac{g(c, n)}{T_1} \tag{7}$$

For level *i* > 2, we consider *c*<sub>1</sub> and *c*<sub>2</sub> to be divisors of *c*, i.e. *c* = *c*<sub>1</sub>*c*<sub>2</sub>. The time to solve a problem of size *cn* = *c*<sub>1</sub>*c*<sub>2</sub>*n* using *bp* processors and an *i*-level algorithm will be:

$$T_i = T_{bp,cn}^i = f(c_1)T_{bp,c_2n}^{i-1} + g(c_1, c_2n) \tag{8}$$

where *g(c*<sub>1</sub>, *c*<sub>2</sub>*n)* is the time to combine *f(c*<sub>1</sub>) subproblems of size *c*<sub>2</sub>*n*.

### 3.6 Speed Up Analysis

In this section, we formulate the conditions for when increasing the number of levels results in speed up. Throughout this section, we consider the problem size to be  $cn$  and the total number of available processors to be  $bp$ . Also, as mentioned before, we use different sets of conditions for  $T_2$  and  $T_i, i > 2$  and the reason is that we only run the subproblems in parallel at level 2.

#### 3.6.1 $T_2 < T_1$

The condition for a 2-level algorithm to be faster than a 1-level algorithm is as follows:

$$T_2 < T_1 \Rightarrow \frac{T_2}{T_1} < 1$$

$$\Rightarrow \frac{1 + rbk\delta}{rbk} < 1 \Rightarrow rbk(1 - \delta) > 1 \quad (9)$$

In the formulation for  $\delta$ , we assume that the combination time is negligible compared to  $T_1$  i.e.  $g(c, n) \ll T_1$ . Then  $\delta \rightarrow 0$  and the condition for  $T_2 < T_1$  will be:

$$rbk > 1 \quad (10)$$

The expected speed up is  $T_1/T_2 = rbk$  (as  $\delta \rightarrow 0$ ).

#### 3.6.2 $T_i < T_1 (i > 2)$

Using equation (8) we get the following condition for an  $i$ -level algorithm ( $i > 2$ ) to be faster than a 1-level algorithm:

$$T_i < T_1 \Rightarrow \frac{T_i}{T_1} < 1$$

$$\Rightarrow f(c_1) \frac{T_{bp,c_2n}^{i-1}}{T_1} + \frac{g(c_1, c_2n)}{T_1} < 1 \quad (11)$$

Again, assuming  $g(c_1, c_2n) \ll T_1$ , the inequality (11) will reduce to

$$f(c_1) \frac{T_{bp,c_2n}^{i-1}}{T_1} < 1 \quad (12)$$

Using equations 5 and 8 for  $i = 3$  we get:

$$T_{bp,c_2n}^{i-1} = \frac{f(c_2)}{b} T_{p,n}^{i-2} + g(c_2, n) \quad (13)$$

$$T_{i-1} = T_{bp,cn}^{i-1} = \frac{f(c_1c_2)}{b} T_{p,n}^{i-2} + g(c_1c_2, n) \quad (14)$$

and for  $i > 3$

$$T_{bp,c_2n}^{i-1} = f(c_2)T_{bp,n}^{i-2} + g(c_2, n) \quad (15)$$

$$T_{i-1} = T_{bp,cn}^{i-1} = f(c_1c_2)T_{bp,n}^{i-2} + g(c_1c_2, n) \quad (16)$$

We make the assumption that  $f(\cdot)$  is separable, meaning that  $f(c_1c_2) = f(c_1)f(c_2)$ . Assuming that  $g(c_2, n)$  and  $g(c_1c_2, n)$  are negligible compared to  $T_1$ ; for all values of  $i > 2$  we will have:

$$f(c_1) \frac{T_{bp,c_2n}^{i-1}}{T_1} = \frac{T_{i-1}}{T_1} \quad (17)$$

Combining equations (17) and (12), we get

$$\frac{T_{i-1}}{T_1} < 1 \quad (18)$$

that is (by induction)  $rbk > 1$ .

#### 3.6.3 $T_i < T_{i-1} (i > 3)$

In the previous section, the condition for  $T_i < T_1$  was established. Even if  $T_i < T_1$ , it may be better to stay at level  $i - 1$  and not to increase the level. To decide whether or not we can use another level, we can consider the  $(i - 1)$ -level code as a new 1-level code. Then recalculate  $r, k$  and  $\delta$  for this code and set  $b = 1$  (that is because we don't use parallelism for  $i > 2$ ). We call these newly computed values  $r', k', \delta'$  and  $b'$ . Then the condition for  $T_i < T_{i-1}$  will be  $r'b'k' > 1$ . As  $b' = 1$  the condition will reduce to  $r'k' > 1$ .

### 3.7 Performance Analysis: Matrix Multiplication

The general model can be applied to matrix multiplication. Without loss of generality, we assume the matrices are square  $cn \times cn$  matrices divided into blocks of size  $n \times n$ . The model details will be as follows:

- $f(c) = c^3$  therefore by plugging it in (3) we get

$$k = \frac{T_{p,cn}^1}{c^3 T_{p,n}^1} \quad (19)$$

- $g(c, n) = (c - 1)c^2a(n)$ . Here  $a(n)$  is the time it takes to add two matrix blocks of size  $n \times n$  and we should add  $(c - 1)c^2$  of them to get the final result for a  $cn \times cn$  matrix multiplication. By plugging it in (7) we will have:

$$\delta = \frac{(c - 1)c^2a(n)}{T_1} \quad (20)$$

Therefore, for matrix multiplication, equations (5) and (8) will be:

$$T_{bp,cn}^2 = \frac{c^3}{b} T_{p,n}^1 + (c - 1)c^2a(n) \quad (21)$$

$$T_i = c_1^3 T_{bp,c_2n}^{i-1} + (c_1 - 1)c_1^2 a(c_2n) \quad (22)$$

As  $f(c)$  is separable and the addition time is negligible compared to multiplication time for large enough matrices (and therefore we can safely assume that  $\delta \rightarrow 0$ ), all the speed up analysis remains valid.

### 3.8 Performance Analysis: Triangular Matrix Multiplication

Our multilevel method for triangular matrix multiplication is a compound method. Meaning that for solving the subproblems, we should call both triangular matrix multiplication and ordinary matrix multiplication functions on the smaller instances. Consider the following equations for  $T_{bp,cn}^2$  and  $T_{bp,cn}^i (i > 2)$  respectively, where - as before -  $c = c_1c_2$ :

$$T_{bp,cn}^2 = \frac{f_1(c)}{b_1} T_{p_1,n}^1 + \frac{f_2(c)}{b_2} \tau_{p_2,n} + g(c, n) \quad (23)$$

$$T_{bp,cn}^i = f_1(c_1)T_{bp,c_2n}^{i-1} + f_2(c_1)\tau_{bp,c_2n} + g(c_1, c_2n) \quad (24)$$



In both (23) and (24)  $f_1(\cdot)$  is the number of triangular matrix multiplication subproblems and  $f_2(\cdot)$  is the number of ordinary matrix multiplication subproblems.  $b_1$  is the degree of concurrency for triangular block multiplication and  $b_2$  is the degree of concurrency for ordinary block multiplication.  $p_1 = (bp)/b_1$  and  $p_2 = (bp)/b_2$ . We use only one  $g(\cdot, \cdot)$  function because the result of multiplying a triangular block with an ordinary block will not necessarily be triangular and therefore there is no need to distinguish between the combination times. Also,  $\tau_{x,y}$  is the runtime of an ordinary matrix multiplication function on blocks of size  $y$  using  $x$  processors.

Let,

$$\theta = \frac{b_2}{b_1} \tag{25}$$

$$\lambda = \frac{\tau_{p_2,n}}{T_{p_1,n}^1 \times \theta} \tag{26}$$

Now, if we define:

$$b = b_1 \tag{27}$$

$$f(x) = f_1(x) + \lambda f_2(x) \tag{28}$$

equations (23) and (24) will turn to

$$T_{bp,cn}^2 = \frac{f(c)}{b} T_{p,n}^1 + g(c, n) \tag{29}$$

$$T_{bp,cn}^i = f(c_1) T_{bp,c_2n}^{i-1} + g(c_1, c_2n) \tag{30}$$

that are the same as equations 5 and 8. Again, without loss of generality, we assume that the matrices are square, their sizes are  $cn \times cn$  and their block sizes are  $n \times n$ . As a result,

- $f_1(c) = c^2$  and  $f_2(c) = \frac{c^3 - c^2}{2}$ . By using (28) we get:

$$f(c) = \frac{\lambda c^3 + (2 - \lambda)c^2}{2} \tag{31}$$

- $g(c, n) = \frac{c^3 - c^2}{2} a_n$  where  $a_n$  is the time to add two  $n \times n$  blocks.

We can safely assume that  $g(c, n) \ll T_1$  for large enough matrices and therefore the condition for  $T_2 < T_1$  still holds. However,  $f(c)$  is not separable, therefore the conditions for  $T_i < T_1$  for  $i > 2$  will not necessarily work.

### 3.9 Performance Analysis: Symmetric Matrix Multiplication

The performance analysis for symmetric matrix multiplication has a lot in common with the performance analysis for triangular matrix multiplication. Equations 23 - 30 are valid for symmetric matrix multiplication. The definitions for  $f_1(\cdot)$ ,  $f_2(\cdot)$ ,  $b_1$ ,  $b_2$ ,  $\theta$  and  $\lambda$  are the same as those in section 3.8 and the same equations  $b = b_1$  and  $f(x) = f_1(x) + \lambda f_2(x)$  work for symmetric matrices. The differences are in details as described in the following (again, without loss of generality, we assume that the matrices to be multiplied are both square matrices of sizes  $cn \times cn$  which are divided into blocks of sizes  $n \times n$ ):

- $f_1(c) = c^2$  and  $f_2(c) = c^3 - c^2$ . By using (28) we get:

$$f(c) = \lambda c^3 + (1 - \lambda)c^2 \tag{32}$$

- $g(c, n) = c^2(c - 1)a_n$  with  $a_n$  being the time to add two  $n \times n$  blocks.

Once again, we can assume that  $g(c, n) \ll T_1$  for large enough matrices and therefore the condition for  $T_2 < T_1$  holds. However, the function  $f(c)$  is only separable if  $\lambda = 1$  and that is the only case where the conditions  $T_i < T_1 (i > 2)$  will work.

## 4 KNL Architecture

### 4.1 Architecture Overview

Knights Landing (KNL) [32] is the codename for the second generation Intel Xeon Phi product family. It is a many-core architecture enabling highly parallel workloads. KNL CPU includes up to 36 active tiles. Each tile includes 2 cores, 2 vector processing units per core and one  $L_2$  cache shared between the cores. A 2D mesh interconnect provides connection between the tiles, memory controllers and other on-board elements. The mesh supports the MESIF (modified, exclusive, shared, invalid, forward) cache-coherence protocol and uses a distributed tag directory to keep all the  $L_2$  caches coherent in all the tiles.

KNL includes 2 types of memory: 16GB of multichannel DRAM (MCDRAM) and up to 384 GB of double data rate (DDR) memory. MCDRAM is a high bandwidth memory (HBM) that provides an aggregate bandwidth of more than 450 GBps. The aggregate bandwidth of DDR is more than 90 GBps.

### 4.2 Memory Modes

The two memory modules can be configured in three modes explained below:

#### 4.2.1 Flat Mode

In this mode, MCDRAM is treated as an addressable memory alongside DDR. Flat mode gives the user the ability to allocate data either from DDR or MCDRAM. The downside is that it requires software modification.

#### 4.2.2 Cache Mode

In cache mode, MCDRAM is configured as a memory-side cache for the whole DDR. Here, the user has no control over MCDRAM usage, but no software modification is required.

#### 4.2.3 Hybrid Mode

This mode is a combination of flat mode and cache mode. In hybrid mode, a portion of MCDRAM (either 0.25 or 0.5) is used in cache mode and the remaining portion is used in flat mode.

### 4.3 KNL Clustering Modes

The mesh interconnect provides five different clustering modes. Each of these clustering modes defines the affinity properties of tiles, tag directories and memory controllers. These modes are explained in the following:

### 4.3.1 All-to-All

This is the most general mode that lacks any affinity between tiles, directory and memory.

### 4.3.2 Quadrant

This mode divides the KNL chip into four virtual quadrants. A request from any tile can go to any directory, but the directory can only access the memory in its own quadrant. Meaning that this mode provides affinity between directory and memory. However, there is no affinity between a tile and the directory or memory.

### 4.3.3 Hemisphere

This mode is similar to the quadrant mode, with the only difference that the chip is divided into 2 hemispheres instead of 4 quadrants.

### 4.3.4 SNC4

SNC stands for Sub-NUMA Clustering. SNC4 is a more restricted version of the quadrant mode where the quadrants are viewed by the OS as nonuniform memory access (NUMA) nodes (or clusters). Here, a request from a tile accesses the directory in the same cluster and the directory will also access a memory controller in that same cluster. Meaning that this mode provides affinity between tiles, memory directories and memory controllers.

### 4.3.5 SNC2

This mode is similar to SNC4 the only difference being the number of clusters. Here the number of clusters is two instead of four.

## 5 Experimental Results

We implemented our algorithms using C++ and OpenMP. The total number of threads was set to 64. Using a larger number of threads led to performance degradation. We used the Intel icpc compiler with the flags setting "-O3 -xMIC-AVX512 -mkl -lmemkind -qopenmp". In "flat" memory mode, we have to manually allocate memory from MCDRAM. This was done using "hbw\_posix\_memalign" command. To allocate memory from DDR, the command "posix\_memalign" was used.

For all three of our algorithms, namely matrix-matrix multiplication, triangular matrix-matrix multiplication and symmetric matrix-matrix multiplication, we used double precision square matrices as test data for both the operands. For matrix pairs of sizes up to  $16K \times 16K$ , the reported times are the average over 10 run times. For matrix pairs of sizes  $32K \times 32K$  and  $64K \times 64K$ , we used only 5 runs. That is because those runs were very time consuming and also because at those sizes, the runtime values were quite stable and 5 runs seemed to be sufficient. The error bound for our reported times is at most 4%.

With 3 memory modes and 5 clustering modes, there are 15 possible architectural configurations in the KNL. Of these, the *cache, all-to-all* mode is not supported. So, in reality, the KNL has 14 architectural configurations. In our experiments, we do not consider the hybrid memory mode. This leaves us with 9 configurations to

consider. There are at least three different scenarios for configuration selection; the selection of scenario being limited by system and application constraints.

1. S1: Each algorithm can select the configuration to run on based on the size of the matrices to be multiplied.
2. S2: The configuration is determined by the application and cannot be changed by the algorithm.
3. S3: Each algorithm must use the same configuration for all instances; different algorithms can use different configurations.

In Sections 5.1-5.3, we give the measured average run times for the three versions of matrix multiplication considered in this paper for all 9 of the architectural configurations considered. A comparison of these run times for each of the three scenarios (S1-S3) of configuration selection is done in Section 5.4. The speed up values reported in Section 5.4 are the percentage of runtime reduction  $(T_{old} - T_{new}) / T_{old}$ .

### 5.1 Matrix Multiplication Times

Table 3 shows the runtime values for the code we use as the 1-level matrix multiplication algorithm (cblas\_dgemm) for different matrix sizes and different KNL memory/clustering configurations.

Table 4 contains the runtimes for our 2-level algorithm. For all the matrix sizes, the input and output matrices are divided into sub-matrices using a  $4 \times 4$  grid. The degree of concurrency used in our 2-level implementation is ( $b = 4$ ), meaning that at each step 4 pairs of matrix blocks are multiplied in parallel. We have also experimented with other block arrangements such as  $2 \times 2$  and other degrees of concurrency such as 2 and 8; however, the current configuration resulted in the best runtimes.

Tables 5 and 6 show the runtimes for our 3 and 4-level algorithms respectively. For both the levels, we used  $2 \times 2$  grids to divide the input and output matrices into sub-matrices. The block pairs were multiplied sequentially.

### 5.2 Triangular Matrix Multiplication Times

Table 7 shows the run times for the one-level code (cblas\_dtrmm). This table together with Table 3 show the speed up potential using our method. Consider  $64K \times 64K$  multiplications as an example. Instead of performing one  $64K \times 64K$  triangular matrix multiplication; we can multiply six  $32K \times 32K$  blocks and combine the results. Out of those block multiplications, four of them are triangular matrix multiplication and two of them are ordinary matrix multiplication. If we consider the combination time to be small enough, then on all the memory-clustering configurations, the time to sequentially perform two ordinary multiplications of  $32K \times 32K$  matrices plus the time to sequentially perform four triangular multiplications of  $32K \times 32K$  matrices is less than the time to perform one  $64K \times 64K$  triangular matrix multiplication. That notifies us about the potential of speed up. Note that this is just an initial evaluation and in practice we can also use parallelism and different settings for block sizes.

Table 8 shows the run time values for our two-level algorithm. Here, matrix  $A$  is divided using a  $2 \times 2$  grid and matrices  $B$  and  $C$  are divided using a  $2 \times 1$  block arrangement. Degree of concurrency for

Table 3: Run time (in seconds) of the 1-level matrix multiplication using cblas\_dgemm

Matrix Dimensions	Flat-SNC4	Flat-SNC2	Flat-All2All	Flat-Quadrant	Flat-Hemisphere
4K × 4K	0.65	0.39	0.30	0.33	0.30
8K × 8K	3.41	1.99	1.64	1.49	<b>1.39</b>
16K × 16K	23.21	15.42	12.74	11.92	<b>10.25</b>
32K × 32K	119.81	121.47	102.53	101.58	102.63
64K × 64K	1002.30	843.25	843.04	826.23	824.15
Matrix Dimensions	Cache-SNC4	Cache-SNC2	Cache-All2All (not supported)	Cache-Quadrant	Cache-Hemisphere
4K × 4K	0.48	0.42	—	0.32	<b>0.29</b>
8K × 8K	1.94	1.71	—	1.51	<b>1.39</b>
16K × 16K	10.89	11.32	—	12.32	10.36
32K × 32K	88.90	87.39	—	97.52	<b>80.18</b>
64K × 64K	<b>618.26</b>	781.83	—	871.10	652.60

Table 4: Run time (in seconds) of the 2-level matrix multiplication algorithm

Matrix Dimensions	Flat-SNC4	Flat-SNC2	Flat-All2All	Flat-Quadrant	Flat-Hemisphere
4K × 4K	0.36	0.32	0.30	0.30	<b>0.28</b>
8K × 8K	2.42	1.52	1.27	1.21	1.25
16K × 16K	113.51	14.82	10.12	8.87	8.85
Matrix Dimensions	Cache-SNC4	Cache-SNC2	Cache-All2All (not supported)	Cache-Quadrant	Cache-Hemisphere
4K × 4K	0.31	0.29	—	0.30	0.32
8K × 8K	1.62	1.36	—	<b>1.17</b>	1.21
16K × 16K	11.71	11.54	—	<b>8.15</b>	8.61

Table 5: Run time (in seconds) of the 3-level matrix multiplication algorithm

Matrix Dimensions	Flat-SNC4	Flat-SNC2	Flat-All2All	Flat-Quadrant	Flat-Hemisphere
32K × 32K	151.12	119.23	80.62	70.50	70.28
64K × 64K	10540.52	884.25	768.97	720.37	653.97
Matrix Dimensions	Cache-SNC4	Cache-SNC2	Cache-All2All (not supported)	Cache-Quadrant	Cache-Hemisphere
32K × 32K	93.53	94.26	—	<b>64.98</b>	68.35
64K × 64K	729.14	805.44	—	686.36	<b>651.94</b>

Table 6: Run time (in seconds) of the 4-level matrix multiplication algorithm

Matrix Dimensions	Flat-SNC4	Flat-SNC2	Flat-All2All	Flat-Quadrant	Flat-Hemisphere
64K × 64K	1292.55	945.01	640.35	560.46	557.92
Matrix Dimensions	Cache-SNC4	Cache-SNC2	Cache-All2All (not supported)	Cache-Quadrant	Cache-Hemisphere
64K × 64K	744.34	745.16	—	<b>517.99</b>	544.31

Table 7: Run time (in seconds) of the 1-level triangular matrix multiplication using cblas\_dtrmm

Matrix Dimensions	Flat-SNC4	Flat-SNC2	Flat-All2All	Flat-Quadrant	Flat-Hemisphere
4K × 4K	0.46	0.40	<b>0.34</b>	0.36	0.41
8K × 8K	2.43	1.50	1.31	1.41	<b>1.12</b>
16K × 16K	10.97	10.38	7.83	9.12	6.21
32K × 32K	73.78	79.32	62.85	71.24	62.82
64K × 64K	556.95	588.33	503.83	556.77	501.58
Matrix Dimensions	Cache-SNC4	Cache-SNC2	Cache-All2All (not supported)	Cache-Quadrant	Cache-Hemisphere
4K × 4K	0.49	0.38	—	0.37	0.38
8K × 8K	2.04	1.43	—	1.61	<b>1.12</b>
16K × 16K	7.99	7.93	—	15.40	<b>6.20</b>
32K × 32K	54.03	59.04	—	77.93	<b>50.15</b>
64K × 64K	450.17	445.05	—	596.00	<b>389.26</b>

Table 8: Run time (in seconds) of the 2-level triangular matrix multiplication algorithm

Matrix Dimensions	Flat-SNC4	Flat-SNC2	Flat-All2All	Flat-Quadrant	Flat-Hemisphere
$4K \times 4K$	0.44	0.40	0.35	<b>0.33</b>	0.35
$8K \times 8K$	1.65	1.32	1.16	1.33	<b>1.04</b>
$16K \times 16K$	10.02	9.52	6.91	7.72	6.02
$32K \times 32K$	69.81	59.07	52.05	51.58	51.64
$64K \times 64K$	868.32	566.61	382.54	380.42	380.55
Matrix Dimensions	Cache-SNC4	Cache-SNC2	Cache-All2All (not supported)	Cache-Quadrant	Cache-Hemisphere
$4K \times 4K$	0.35	0.35	—	<b>0.33</b>	0.32
$8K \times 8K$	1.11	1.13	—	1.26	<b>1.04</b>
$16K \times 16K$	6.87	6.60	—	7.64	<b>5.97</b>
$32K \times 32K$	50.60	50.57	—	49.77	<b>42.98</b>
$64K \times 64K$	402.92	414.56	—	422.38	<b>356.40</b>

triangular blocks is  $b_1 = 2$  and for ordinary blocks is  $b_2 = 1$ . Moreover, for square-rectangle block multiplication, we called  $MXM_2$  instead of `cblas_dgemm`. In the  $MXM_2$  algorithm we used, the grids imposed on the matrices were  $4 \times 4$ ,  $4 \times 8$  and  $4 \times 8$  for  $A$ ,  $B$  and  $C$  respectively.

### 5.3 Symmetric Matrix Multiplication Times

Table 9 shows the run time values for the one-level symmetric matrix-matrix multiplication code (`cblas_dsymm`). Similar to what we did for triangular matrix multiplication, here again we can use this table along with Table 3 to make sure that there is the potential for performance improvement. Consider the multiplication time for  $64K \times 64K$  matrices as an example. We can break matrices into  $32K \times 32K$  blocks and perform 8 block multiplications 4 of which are symmetric matrix multiplication and the remaining 4 are ordinary matrix multiplication. For all memory-clustering configurations (except for flat-snc2) the time to sequentially perform four ordinary  $32K \times 32K$  multiplication plus the time to sequentially perform four symmetric  $32K \times 32K$  multiplications is less than the time to perform one symmetric  $64K \times 64K$  multiplication. As a result, the initial test has passed and the potential for speed up exists for a simple sequential model using  $2 \times 2$  grids to block the matrices. We can also use parallelism and different settings for block dimensions.

Table 10 shows the run time values for our two-level algorithm. In this implementation, matrix  $A$  is divided using a  $2 \times 2$  grid and matrices  $B$  and  $C$  are divided using a  $2 \times 1$  grid each. The degree of concurrency for both symmetric and ordinary blocks is  $b_1 = b_2 = 2$ .

### 5.4 Run Time Comparison For Different Configuration Selection Scenarios

Throughout this section, we use the layout convention given in Table 11 when we refer to “our algorithm”. For example, the 3-level algorithm for matrix multiplication is the same one for which we reported runtimes in Section 5.1 and has the same parameters.

Figures 3, 4 and 5 compare the algorithms under scenario S1 (each algorithm uses the best configuration for each matrix size), for matrix multiplication, triangular matrix multiplication and symmetric matrix multiplication respectively. For matrix multiplication, our algorithm results in 20.5% speed up for  $16K \times 16K$  matrices,

19.0% for  $32K \times 32K$  matrices and 16.2% for  $64K \times 64K$  matrices. The corresponding speed up values are 3.7%, 14.3% and 8.4% for triangular matrix multiplication and -0.5%, 11.5% and 10.5% for symmetric matrix multiplication.

Tables 12, 13 and 14 show the comparison results using scenario S2 (all algorithms use the same configuration) for matrix multiplication, triangular matrix multiplication and symmetric matrix multiplication respectively. The comparisons are expressed as the percentage runtime reduction obtained by our algorithm over the 1-level code for each of the above-mentioned applications. In all the three tables, the negative percentages (where our algorithm has performed worse than the 1-level code) appear at the configurations involving snc2 or snc4 (Table 14 is exception, but the negative speed up values we get for other configurations have very small absolute values). We will briefly explain in Section 5.5 why our method performs poorly in the Sub-NUMA Clustering modes. If we ignore the modes involving snc2 and snc4, then for matrix multiplication on  $16K \times 16K$  matrices, the speed up we get is in range [13.7%, 33.8%]; for  $32K \times 32K$  matrices it is in range [14.8%, 33.4%] and for  $64K \times 64K$  matrices it is in range [16.6%, 40.5%]. The corresponding ranges for triangular matrix multiplication are [3.1%, 50.4%], [14.3%, 36.1%] and [8.4%, 31.7%] and for symmetric matrix multiplication are [-0.9%, 50.9%], [10.3%, 28.2%] and [11.6%, 30.1%].

For the third scenario S3, we make the following observations ( $conf_1$  is the best configuration for the 1-level algorithm and  $conf_m$  is the best one for our multilevel algorithm):

- For matrix multiplication,  $conf_1$  : cache-hemisphere and  $conf_m$  : cache-quadrant (for  $conf_m$ , results of 2-level, 3-level and 4-level algorithms have been considered collectively).
- For triangular matrix multiplication,  $conf_1$  : cache-hemisphere and  $conf_m$  : cache-hemisphere.
- For symmetric matrix multiplication,  $conf_1$  : cache-hemisphere and  $conf_m$  : cache-hemisphere.

Figures 6, 7 and 8 compare our algorithms vs. the 1-level code for scenario S3. Using our speed up method on matrix multiplication, we get 21.3% speed up for  $16K \times 16K$  matrices, 19.0% for  $32K \times 32K$  matrices and 20.6% for  $64K \times 64K$  matrices. The corresponding speed up values are 3.7%, 14.3% and 8.4% for triangular matrix multiplication and -0.9%, 10.3% and 11.6% for symmetric



Table 9: Run time (in seconds) of the 1-level symmetric matrix multiplication algorithm using cblas\_dsymm

Matrix Dimensions	Flat-SNC4	Flat-SNC2	Flat-All2All	Flat-Quadrant	Flat-Hemisphere
4K × 4K	0.67	0.57	<b>0.47</b>	0.54	0.56
8K × 8K	4.47	2.63	2.33	2.97	1.99
16K × 16K	24.78	20.75	15.12	32.13	<b>11.96</b>
32K × 32K	130.27	152.23	122.94	120.16	123.49
64K × 64K	1006.21	957.23	994.72	1002.23	1005.31
Matrix Dimensions	Cache-SNC4	Cache-SNC2	Cache-All2All (not supported)	Cache-Quadrant	Cache-Hemisphere
4K × 4K	0.67	0.55	—	0.53	0.55
8K × 8K	2.98	2.50	—	2.87	<b>1.97</b>
16K × 16K	13.43	13.98	—	31.66	12.00
32K × 32K	99.14	109.71	—	153.31	<b>98.17</b>
64K × 64K	791.85	<b>778.54</b>	—	1346.62	797.71

Table 10: Run time (in seconds) of 2-level symmetric matrix multiplication algorithm

Matrix Dimensions	Flat-SNC4	Flat-SNC2	Flat-All2All	Flat-Quadrant	Flat-Hemisphere
4K × 4K	0.75	0.51	0.41	0.39	0.39
8K × 8K	3.59	2.59	2.15	2.27	1.80
16K × 16K	24.09	19.95	14.78	15.78	<b>12.02</b>
32K × 32K	93.36	164.51	89.49	89.71	90.72
64K × 64K	1631.96	1312.09	711.02	711.50	709.60
Matrix Dimensions	Cache-SNC4	Cache-SNC2	Cache-All2All (not supported)	Cache-Quadrant	Cache-Hemisphere
4K × 4K	0.45	0.45	—	<b>0.36</b>	0.38
8K × 8K	2.08	2.16	—	2.19	<b>1.77</b>
16K × 16K	12.85	13.41	—	16.05	12.11
32K × 32K	94.60	100.01	—	110.11	<b>88.03</b>
64K × 64K	826.87	765.08	—	941.59	<b>704.79</b>

matrix multiplication. Note that for the latter two applications as  $conf_1 = conf_m$ , the speed up values are the same as the ones in the “cache-hemisphere” column of Tables 13 and 14 respectively.

Table 11: The algorithm we use (number of levels) depends on the target application and the size of the input matrices.

	16K × 16K matrices	32K × 32K matrices	64K × 64K matrices
Matrix Multiplication (MXM)	2-level algorithm	3-level algorithm	4-level algorithm
Triangular MXM	2-level algorithm		
Symmetric MXM	2-level algorithm		

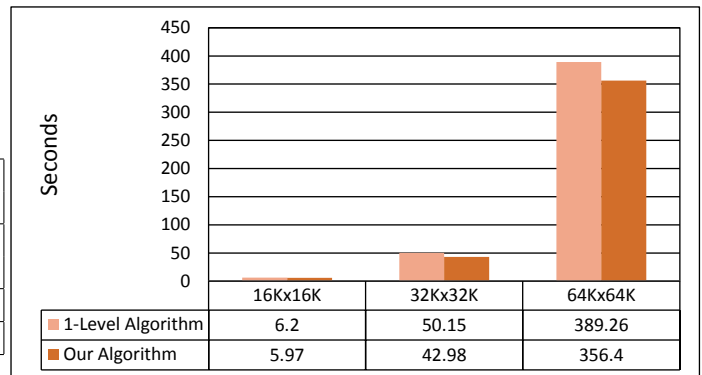


Figure 4: Comparison of our method against the 1-level triangular matrix multiplication algorithm (cblas\_dtrmm) using scenario S1.

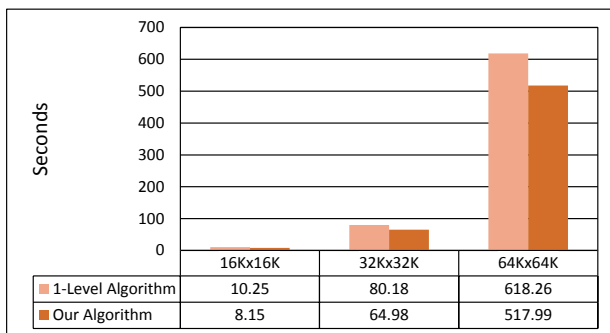


Figure 3: Comparison of our method against the 1-level matrix multiplication algorithm (cblas\_dgemm) using scenario S1.

### 5.5 Model Validation

In this section, we validate the mathematical model we developed in Section 3.5 to predict performance improvement. The validation is done on matrix multiplication. It is similar for triangular and symmetric matrix multiplication problems. We show that the  $rbk > 1$  condition correctly predicts the potential for speed up. Also, we compare the actual  $\frac{T_i}{T_1}$  ratios with those predicted using our formulation. For predicting the ratios, we use equations (5) and (8) ignoring the  $g(.,.)$  terms. We ignore the  $g(.,.)$  terms by assuming

that  $g(c, n) \ll T_1$  (the same assumption as the one we made several times in Section 3.5). Meaning that we use (33) and (34) to predict the runtime ratios.

$$\frac{T_2}{T_1} = \frac{f(c)}{b} \times \frac{T_{p,n}^1}{T_1} \tag{33}$$

$$\frac{T_i}{T_1} = \frac{f(c_1)T_{bp,c_2n}^{i-1}}{T_1} \tag{34}$$

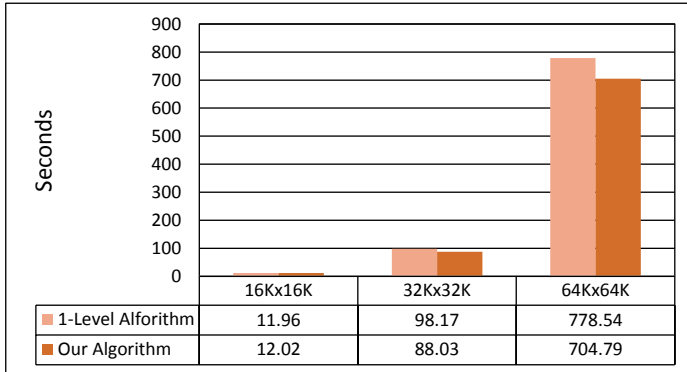


Figure 5: Comparison of our method against the 1-level symmetric matrix multiplication algorithm (cblas\_dsymm) using scenario S1.

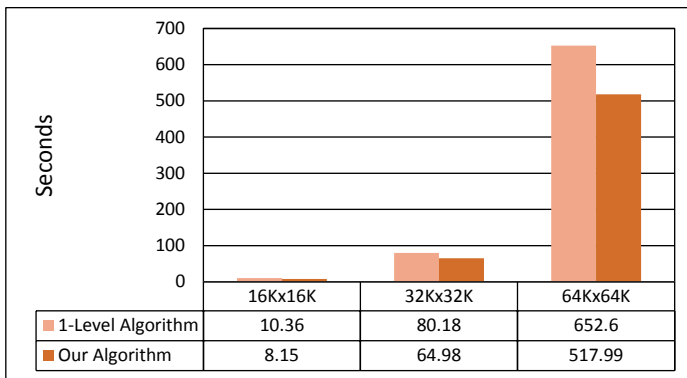


Figure 6: Comparison of our method against the 1-level matrix multiplication algorithm (cblas\_dgemm) for scenario S3.

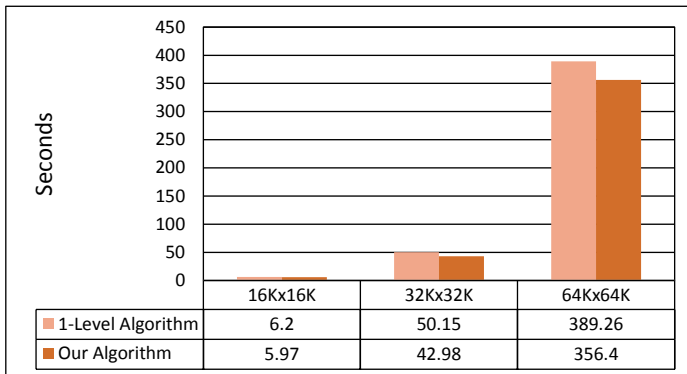


Figure 7: Comparison of our method against the 1-level triangular matrix multiplication algorithm (cblas\_dtrmm) for scenario S3.

In our analysis we ignore matrix addition time as well as the time spent on copying the matrix blocks into the fast memory. The reason is that theoretically those are  $O(n^2)$  operations while matrix multiplication is an  $O(n^3)$  operation. In practice, for our two-level matrix multiplication code in the default KNL mode (cache-quadrant) for instance, the ratio (multiplication time)/(data transfer

time) is 3 for matrices of size  $4K \times 4K$  and 42 for matrices of size  $16K \times 16K$ . The ratio (multiplication time)/(addition time) is 21 and 199, for matrix sizes of  $4K \times 4K$  and  $16K \times 16K$  respectively. Similar ratios can be obtained for other matrix sizes and using other configurations.

Table 15 shows the validation results for our  $T_2$  formula in Equation 5. Here,  $cn = 16K$ , number of blocks  $c = 4$  and therefore  $n = 4K$ . The degree of concurrency  $b$  is 4. Here and also in the following validation experiments, the total number of processors  $bp = 64$ . The error between the computed value of  $T_2/T_1$  and the measured value ranges from 1.27% (Flat-all2all) to 19.44% (Cache-snc4) in all the modes except for Flat-snc4 (86.09%).

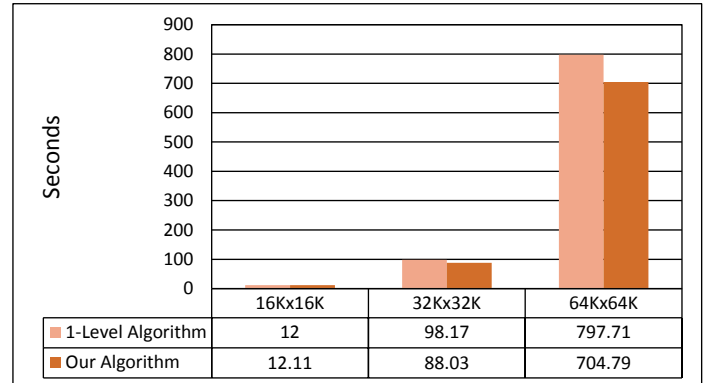


Figure 8: Comparison of our method against the 1-level symmetric matrix multiplication algorithm (cblas\_dsymm) for scenario S3.

Table 16 shows the validation results for our  $T_3$  formula. Here, matrix size  $cn = 32K$ , number of blocks in the first (outermost) level  $c_1 = 2$  (arranged as a  $2 \times 2$  grid) and therefore the outermost blocks are of size  $16K \times 16K$ . The error between computed  $T_3/T_1$  and the observed value ranges from 0% (Flat-snc2, Flat-all2all) to 1.85% (Cache-snc2) for all the modes except for Flat-snc4 where it is 502%.

Table 17 shows the validation results for our  $T_4$  formula. Here, matrix size  $cn = 64K$ , number of blocks in the first (outermost) level  $c_1 = 2$  (arranged as a  $2 \times 2$  grid) and therefore the outermost blocks are  $32K \times 32K$ . Note that in tables 15, 16 and 17 the computed and the actual runtime ratios have been rounded to the nearest hundredth and that the error values are computed based on those rounded ratios. Here, the error between the computed and the observed values ranges from 0% (Flat-quadrant) to 1.69% (Cache-quadrant) in all the modes except for the Flat-snc4 mode where it is 6.2%.

Figures 9, 10 and 11 visually show the actual values versus the computed values for  $T_3/T_1$ ,  $T_2/T_1$  and  $T_4/T_1$  respectively.

Our validation results show that our model works fairly well for all the KNL configurations except for those involving snc2 or snc4. The reason is that in the Sub-NUMA Clustering modes, each half (or quadrant) of the chip is exposed as a separate domain to the OS, essentially making the chip analogous to a 2 or 4 socket Xeon. This restricts data sharing and memory access time will increase as a result.

The condition for speed up, namely  $rbk > 1$  works correctly for 8 out of 9 KNL configurations (the only exception being Flat-snc4). Note that the  $rbk$  column is the same in all the three tables. That is because its value only depends on the parameters of levels 1 and 2.

Table 12: The speed up obtained by our algorithm over the 1-level matrix multiplication code (cblas\_dgemm) for scenario S2,

Matrix Dimensions	Flat-SNC4	Flat-SNC2	Flat-All2All	Flat-Quadrant	Flat-Hemisphere
16K × 16K	-389.0%	3.9%	20.6%	25.6%	13.7%
32K × 32K	-26.1%	1.8%	21.4%	30.6%	31.5%
64K × 64K	-29.0%	-12.1%	24.0%	32.2%	32.3%
Matrix Dimensions	Cache-SNC4	Cache-SNC2	Cache-All2All (not supported)	Cache-Quadrant	Cache-Hemisphere
16K × 16K	-7.5%	-1.9%	—	<b>33.8%</b>	16.9%
32K × 32K	-5.2%	-7.9%	—	<b>33.4%</b>	14.8%
64K × 64K	-20.4%	4.7%	—	<b>40.5%</b>	16.6%

Table 13: The speed up obtained by our algorithm over the 1-level triangular matrix multiplication code (cblas\_dtrmm) for Scenario S2.

Matrix Dimensions	Flat-SNC4	Flat-SNC2	Flat-All2All	Flat-Quadrant	Flat-Hemisphere
16K × 16K	8.7%	8.3%	11.7%	15.4%	3.1%
32K × 32K	5.4%	25.5%	17.2%	27.6%	17.8%
64K × 64K	-55.9%	3.7%	24.1%	<b>31.7%</b>	24.1%
Matrix Dimensions	Cache-SNC4	Cache-SNC2	Cache-All2All (not supported)	Cache-Quadrant	Cache-Hemisphere
16K × 16K	14.0%	16.8%	—	<b>50.4%</b>	3.7%
32K × 32K	6.3%	14.3%	—	<b>36.1%</b>	14.3%
64K × 64K	10.5%	6.9%	—	29.1%	8.4%

Table 14: The speed up obtained by our algorithm over the 1-level symmetric matrix multiplication code (cblas\_dsymm) for Scenario S2.

Matrix Dimensions	Flat-SNC4	Flat-SNC2	Flat-All2All	Flat-Quadrant	Flat-Hemisphere
16K × 16K	2.8%	3.9%	2.2%	<b>50.9%</b>	-0.5%
32K × 32K	<b>28.3%</b>	-8.1%	27.2%	25.3%	26.5%
64K × 64K	-62.2%	-37.1%	28.5%	29.0%	29.4%
Matrix Dimensions	Cache-SNC4	Cache-SNC2	Cache-All2All (not supported)	Cache-Quadrant	Cache-Hemisphere
16K × 16K	4.3%	4.1%	—	49.3%	-0.9%
32K × 32K	4.6%	8.8%	—	<b>28.2%</b>	10.3%
64K × 64K	-4.4%	1.7%	—	<b>30.1%</b>	11.6%

Table 15: Evaluation of  $T_2$  formula against different KNL configurations - All time values are in seconds

	$T_1$	$T_2$	$T_{p,n}^1$	$T_{bp,n}^1$	$T_{p,cn}^1$	Computed $T_2/T_1$	Actual $T_2/T_1$	Error	$rbk$
Flat - SNC4	23.21	113.51	0.98	0.65	46.04	0.68	4.89	86.09%	1.48
Flat - SNC2	15.42	14.82	0.84	0.39	51.67	0.87	0.96	9.37%	1.15
Flat - All2All	12.74	10.12	0.64	0.3	52.88	0.80	0.79	1.27%	1.24
Flat - Quadrant	11.92	8.87	0.57	0.33	35.62	0.77	0.74	4.05%	1.31
Flat - Hemisphere	10.25	8.85	0.59	0.3	38.02	0.92	0.86	6.98%	1.09
Cache - SNC4	10.89	11.71	0.88	0.48	38.03	1.29	1.07	19.44%	0.77
Cache - SNC2	11.32	11.54	0.81	0.42	42.56	1.14	1.02	11.76%	0.87
Cache - Quadrant	12.32	8.15	0.55	0.32	35.79	0.71	0.66	7.58%	1.40
Cache - Hemisphere	10.36	8.61	0.56	0.29	38.33	0.86	0.83	3.61%	1.16

Table 16: Evaluation of  $T_3$  formula against different configurations - All time values are in seconds

	$T_1$	$T_{bp,c2n}^2$	$T_3$	Computed $T_3/T_1$	Actual $T_3/T_1$	Error	$rbk$
Flat - SNC4	119.81	113.51	151.12	7.58	1.26	501.59%	1.48
Flat - SNC2	121.47	14.82	119.23	0.98	0.98	0.00%	1.15
Flat - All2All	102.53	10.12	80.62	0.79	0.79	0.00%	1.24
Flat - Quadrant	101.58	8.87	70.5	0.70	0.69	1.45%	1.31
Flat - Hemisphere	102.63	8.85	70.28	0.69	0.68	1.47%	1.09
Cache - SNC4	88.9	11.71	93.53	1.05	1.05	0.00%	0.77
Cache - SNC2	87.39	11.54	94.26	1.06	1.08	1.85%	0.87
Cache - Quadrant	97.52	8.15	64.98	0.67	0.67	0.00%	1.40
Cache - Hemisphere	80.18	8.61	68.35	0.86	0.85	1.18%	1.16

Table 17: Evaluation of  $T_4$  formula against different KNL configurations - All time values are in seconds

	$T_1$	$T_{bp,c,n}^3$	$T_4$	Computed $T_4/T_1$	Actual $T_4/T_1$	Error	$rbk$
Flat - SNC4	1002.3	151.12	1292.55	1.21	1.29	6.20%	1.48
Flat - SNC2	843.25	119.23	945.01	1.13	1.12	0.89%	1.15
Flat - All2All	843.04	80.62	640.35	0.77	0.76	1.32%	1.24
Flat - Quadrant	826.23	70.5	560.46	0.68	0.68	0.00%	1.31
Flat - Hemisphere	824.15	70.28	557.92	0.68	0.68	0.00%	1.09
Cache - SNC4	618.26	93.53	744.34	1.21	1.20	0.83%	0.77
Cache - SNC2	781.83	94.26	745.16	0.96	0.95	1.05%	0.87
Cache - Quadrant	871.1	64.98	517.99	0.60	0.59	1.69%	1.40
Cache - Hemisphere	652.6	68.35	544.31	0.84	0.83	1.20%	1.16

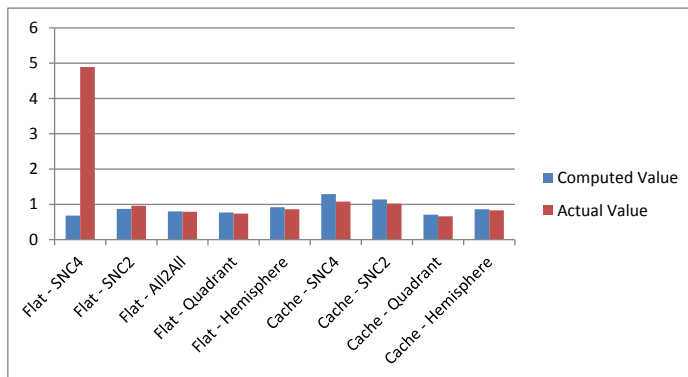


Figure 9:  $T_2/T_1$ : Computed Values vs. Actual Values. Demonstrated over KNL memory-clustering configurations.

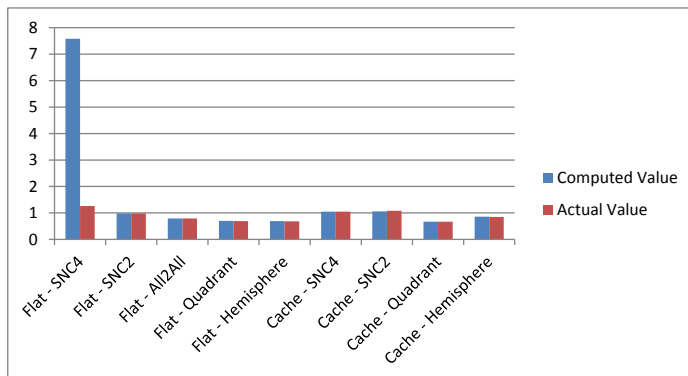


Figure 10:  $T_3/T_1$ : Computed Values vs. Actual Values. Demonstrated over KNL memory-clustering configurations.

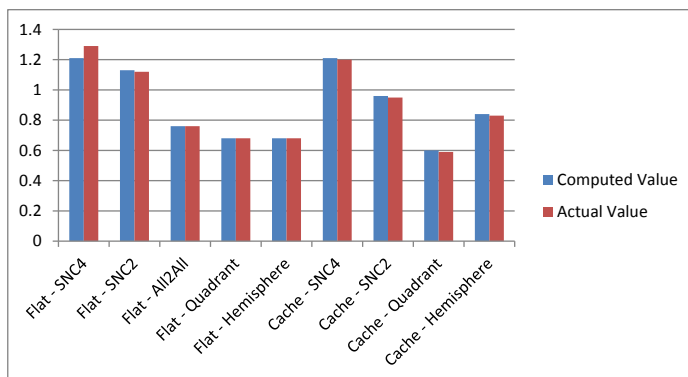


Figure 11:  $T_4/T_1$ : Computed Values vs. Actual Values. Demonstrated over KNL memory-clustering configurations.

If we leave out the Flat-snc4 configuration, then the speed up predictions of our model (either negative or positive) are within 2%, 2% and 20% of the actual values for matrix sizes  $64K \times 64K$ ,  $32K \times 32K$  and  $16K \times 16K$ , respectively.

## 6 Conclusions

We have proposed a multilevel method to speed parallel codes whose runtime grows faster than their workload. The target problems are the ones that can be decomposed into smaller subproblems. The original highly optimized parallel code is treated as a black box and no change is made to it. Using a simple parallel computing model, we derived formulas to predict whether or not any speed up is possible and also the amount of attainable speed up.

We demonstrated the effectiveness of our multilevel method on the highly optimized parallel BLAS routines `cblas_dgemm`, `cblas_dtrmm` and `cblas_dsymm` that are in the MKL library using the Intel KNL platform. Our proposed method is, however, general and can be potentially applied to other optimized parallel codes and on other platforms.

In our experiments, we considered 3 possible application scenarios. In the first of these (each algorithm is run using the best memory/clustering configuration for each problem size), we obtained speed up values of 16.2%, 8.4%, and 10.5%, respectively, for matrix multiplication, triangular matrix multiplication, and symmetric matrix multiplication for  $64K \times 64K$  matrices. In the second scenario (all algorithms use the same configuration), these percentages are up to 40.5%, 31.7%, and 30.1%. The default mode in the KNL is cache - quadrant configuration. For this mode, the speed up percentages are 40.5%, 29.1%, and 30.1%. In the third scenario (each algorithm uses the same configuration regardless of problem size; different algorithms may use different configurations), the speed up values were 20.6%, 8.4%, and 11.6%.

For the matrix multiplication problem, the simple predictive mathematical model developed by us is able to correctly predict whether or not speed up is possible for 8 out of the 9 memory/clustering configurations tested. Also, for 8 out of the 9 configurations and for the matrix sizes of  $64K \times 64K$ , the amount of runtime increase or decrease our formula predicted for matrix multiplication was within 2% of the actual value. For both predicting the potential of speed up and predicting the amount of attainable speed up, Flat-snc4 was the configuration for which our model failed.

**Conflict of Interest** The authors declare no conflict of interest.



**Acknowledgment** This work was funded, in part, by the National Science Foundation, under Contract No. 1748652.

## References

- [1] S. Gheibi, T. Banerjee, S. Ranka, S. Sahni, "Multilevel Approaches to Fine Tune Performance of Linear Algebra Libraries," in 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 1–6, IEEE, 2019.
- [2] L. E. Cannon, A cellular computer to implement the Kalman filter algorithm, Ph.D. thesis, Montana State University-Bozeman, College of Engineering, 1969.
- [3] V. Strassen, "Gaussian elimination is not optimal," *Numerische mathematik*, **13**(4), 354–356, 1969.
- [4] H. Prokop, Cache-oblivious algorithms, Ph.D. thesis, Massachusetts Institute of Technology, 1999.
- [5] R. D. Blumofe, M. Frigo, C. F. Joerg, C. E. Leiserson, K. H. Randall, "An analysis of dag-consistent distributed shared-memory algorithms," in SPAA, volume 96, 297–308, 1996.
- [6] J. Demmel, D. Eliahu, A. Fox, S. Kamil, B. Lipshitz, O. Schwartz, O. Spillinger, "Communication-optimal parallel recursive rectangular matrix multiplication," in 2013 IEEE 27th International Symposium on Parallel and Distributed Processing, 261–272, IEEE, 2013.
- [7] B. Lipshitz, G. Ballard, J. Demmel, O. Schwartz, "Communication-avoiding parallel strassen: Implementation and performance," in SC'12: Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, 1–11, IEEE, 2012.
- [8] G. H. Golub, C. F. Van Loan, *Matrix computations*, volume 3, JHU press, 2012.
- [9] R. A. Van De Geijn, J. Watts, "SUMMA: Scalable universal matrix multiplication algorithm," *Concurrency: Practice and Experience*, **9**(4), 255–274, 1997.
- [10] M. M. A. Patwary, N. R. Satish, N. Sundaram, J. Park, M. J. Anderson, S. G. Vadlamudi, D. Das, S. G. Pudov, V. O. Pirogov, P. Dubey, "Parallel efficient sparse matrix-matrix multiplication on multicore platforms," in International Conference on High Performance Computing, 48–57, Springer, 2015.
- [11] Q. Xiangzhen, "Cache performance and algorithm optimization," in High Performance Computing on the Information Superhighway, 1997. HPC Asia'97, 12–17, IEEE, 1997.
- [12] D. I. Lyakh, "An efficient tensor transpose algorithm for multicore CPU, Intel Xeon Phi, and NVidia Tesla GPU," *Computer Physics Communications*, **189**, 84–91, 2015.
- [13] J. Chen, J. Fang, W. Liu, T. Tang, C. Yang, "clmf: A fine-grained and portable alternating least squares algorithm for parallel matrix factorization," *Future Generation Computer Systems*, **108**, 1192–1205, 2020.
- [14] C. Yount, A. Duran, "Effective use of large high-bandwidth memory caches in HPC stencil computation via temporal wave-front tiling," in Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems (PMBS), International Workshop on, 65–75, IEEE, 2016.
- [15] S. Chatterjee, V. V. Jain, A. R. Lebeck, S. Mundhra, M. Thottethodi, "Nonlinear array layouts for hierarchical memory systems," in Proceedings of the 13th international conference on Supercomputing, 444–453, ACM, 1999.
- [16] J. Mellor-Crummey, D. Whalley, K. Kennedy, "Improving memory hierarchy performance for irregular applications using data and computation reorderings," *International Journal of Parallel Programming*, **29**(3), 217–247, 2001.
- [17] E. Athanasaki, N. Koziris, "Fast indexing for blocked array layouts to improve multi-level cache locality," in Interaction between Compilers and Computer Architectures, 2004. INTERACT-8 2004. Eighth Workshop on, 107–119, IEEE, 2004.
- [18] C. Kulkarni, C. Ghez, M. Miranda, F. Catthoor, H. De Man, "Cache conscious data layout organization for embedded multimedia applications," in Proceedings of the conference on Design, automation and test in Europe, 686–693, IEEE Press, 2001.
- [19] B. Recht, C. Re, S. Wright, F. Niu, "Hogwild: A lock-free approach to parallelizing stochastic gradient descent," in Advances in neural information processing systems, 693–701, 2011.
- [20] W.-S. Chin, Y. Zhuang, Y.-C. Juan, C.-J. Lin, "A fast parallel stochastic gradient method for matrix factorization in shared memory systems," *ACM Transactions on Intelligent Systems and Technology (TIST)*, **6**(1), 2, 2015.
- [21] S. Song, J. K. Hollingsworth, "Designing and auto-tuning parallel 3-D FFT for computation-communication overlap," in Proceedings of the 19th ACM SIGPLAN symposium on Principles and practice of parallel programming, 181–192, 2014.
- [22] S. Lee, D. Jha, A. Agrawal, A. Choudhary, W.-k. Liao, "Parallel deep convolutional neural network training by exploiting the overlapping of computation and communication," in 2017 IEEE 24th International Conference on High Performance Computing (HiPC), 183–192, IEEE, 2017.
- [23] H. Wang, S. Guo, R. Li, "Osp: Overlapping computation and communication in parameter server for fast machine learning," in Proceedings of the 48th International Conference on Parallel Processing, 1–10, 2019.
- [24] J. Huang, T. M. Smith, G. M. Henry, R. A. van de Geijn, "Strassen's algorithm reloaded," in Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 59, IEEE Press, 2016.
- [25] R.-I. Ciobanu, C. Dobre, M. Bălănescu, G. Suciu, "Data and task offloading in collaborative mobile fog-based networks," *IEEE Access*, **7**, 104405–104422, 2019.
- [26] V. Priya, C. S. Kumar, R. Kannan, "Resource scheduling algorithm with load balancing for cloud service provisioning," *Applied Soft Computing*, **76**, 416–424, 2019.
- [27] D. Puthal, R. Ranjan, A. Nanda, P. Nanda, P. P. Jayaraman, A. Y. Zomaya, "Secure authentication and load balancing of distributed edge datacenters," *Journal of Parallel and Distributed Computing*, **124**, 60–69, 2019.
- [28] A. Mallick, M. Chaudhari, U. Sheth, G. Palanikumar, G. Joshi, "Rateless codes for near-perfect load balancing in distributed matrix-vector multiplication," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, **3**(3), 1–40, 2019.
- [29] A. Mohammed, A. Eleliemy, F. M. Ciorba, F. Kasielke, I. Banicescu, "An approach for realistically simulating the performance of scientific applications on high performance computing systems," *Future Generation Computer Systems*, **111**, 617–633, 2020.
- [30] I.-J. Sung, J. A. Stratton, W.-M. W. Hwu, "Data layout transformation exploiting memory-level parallelism in structured grid many-core applications," in Proceedings of the 19th international conference on Parallel architectures and compilation techniques, 513–522, ACM, 2010.
- [31] G. Chen, B. Wu, D. Li, X. Shen, "PORPLE: An extensible optimizer for portable data placement on GPU," in Proceedings of the 47th Annual IEEE/ACM International Symposium on Microarchitecture, 88–100, IEEE Computer Society, 2014.
- [32] A. Sodani, R. Gramunt, J. Corbal, H.-S. Kim, K. Vinod, S. Chinthamani, S. Hutsell, R. Agarwal, Y.-C. Liu, "Knights landing: Second-generation intel xeon phi product," *Ieee micro*, **36**(2), 34–46, 2016.

# Generic Decision Support Model for Determining the Best Marketer

Ditdit Nugeraha Utama\*, Sherly Oktafiani

Computer Science Department, BINUS Graduate Program, Bina Nusantara University, Jakarta, 11480, Indonesia

## ARTICLE INFO

Article history:

Received: 03 June, 2020

Accepted: 05 August, 2020

Online: 31 August, 2020

Keywords:

Marketer

Decision Support Model

Fuzzy Logic

Profile Matching

## ABSTRACT

For marketing company, determining the best marketing is a strategic decision. It is able to affect the company's future purposefully. The study conducted to propose a novel decision support model (DSM) for determining the best marketer. Seventeen parameters were considered into model and two methods fuzzy logic and profile-matching (PM) were combined as main method in constructing the model theoretically. The model was applied in medical appliance marketing company. It is finally also able to measure and present the best ranking of marketers that is able to make company for easily making a crucial decision.

## 1 Introduction

Today discussing department of marketing in a company cannot be ignored practically. Where the department is able to be compared to spearhead of the company, marketing itself has a very imperative function that is able to support the company's prosperity and profits [1]. Marketing allows interactions involving factors that provide attractiveness that are tailored to the company's strategy [2], as well as making a positive contribution through the process of developing long-term relationships through marketing interaction activities [3].

Thus, being able to give a stronger priority to marketing strategies and have positive implications for business [3], which are going to greatly affect the growth of a company. If the marketing performance is good, the company's growth will also be much better. When marketing increases business performance, a creative marketing strategy will be produced, to achieve the effectiveness of the implementation of corporate strategy [4]. Basically, a successful marketing department can be obtained by being success in supporting customers to buy back the products [5]. It can make customers become loyal customers [6].

Furthermore, the high level of competition in product marketing makes marketing department responsible in increasing sales value [7]. Marketing is a media that is very close to building business relationships to support companies. Helping companies in providing support and motivation to the marketing team to improve their ability to work is good effort to increase the company's income.

Based on technological developments, many ways have been

taken by several companies to help or optimize marketing activities. Among other things, by implementing a decision support system (DSS) in e-tourism applications to help decide on the destination that best suits the interests of customers [8], the application of the Mehrabian-Rusel (MR) method to marketing DSS (MDSS) to examine weather factors in influencing consumer buying interest [7]. In addition, DSS is also combined with geographical information system (GIS) in marketing information systems as a tool for the decision making process to determine strategic locations to increase attractiveness for customers [9], and is also applied based on the concept of customer decision making (CDM) by using Model P6 for customers to determine web services that fit their needs [10].

In studies on the spread of advertising through social media, DSS has been successfully implemented too [11],[12]. But from a number of cases that help improve the quality of the company's business, there hasn't been any specific discussion in the marketing department itself. Because as we all know, marketing is an important point in improving the quality of a company's business [13]; then research on how to market to find out who is the best marketing person in a company through DSS is academically feasible to do.

Specifically for companies that distribute medical appliances in West Jakarta (Indonesia), which are the object of research, the company has difficulty in determining the ranking of marketers. This condition has an impact on the objectivity in determining the commission value of its marketers. This condition also, directly or indirectly, has an impact on the motivation of marketers' performance. Therefore, this research is carried out with the main

\*Corresponding Author: Ditdit Nugeraha Utama, Computer Science Department, BINUS Graduate Program - Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480, +81289614291 & ditdit.utama@binus.edu

objective to build a decision support model (DSM) that can determine objectively top-marketers.

Furthermore, DSM to determine the best marketers (top-marketers) is able to be used to provide incentives (bonuses) made already. This is a novel model to select the best marketer in specific pharmaceutical marketing company. Thus, the company is able to operate the best marketer suggested to next action, such as annual bonus giving. By considering seventeen selected-verified parameters and also applying a combination of fuzzy logic (FL) and profile matching (PM) methods, a constructed DSM is new scientific-proposed model. It is able to enrich a knowledge of computer science domain. Practically, it is able to become a new procedure to determine the best marketer specifically or employee commonly. The procedure is functioned to make the decision more objectively.

Moreover, for designing the model, object-oriented (OO) development method was functioned. This OO combined the data structure and behavior of all entities involved in the model through the abstraction of real world objects. OO was able to prove the efficiency for design modeling, and the implementation of a small scale object-based component information system involves segmentation steps [14]. Where the data analysis itself was carried out using data marketers at one of the medical distributor companies.

## 2 Theoretical Framework

### 2.1 Fuzzy Logic (FL)

FL was discovered by Prof. Lotfi Zadeh from the University of California at Berkeley in the mid-1960s, that the laws of right and wrong from Boolean logic did not take into account real conditions [15], Boolean reasoning itself conceptually related to unclear truth [16]. The main contribution of fuzzy logic is to provide a basis for the development from binarism to pluralism [17].

FL is part of artificial intelligence (AI) that is able to capture uncertainty or ambiguity into truth values. FL is used in assessing opinions based on human assumptions [18]. Fuzzy method itself is a classification and algorithm of ambiguity in the grouping of data [19]. Likewise what was said [20] that fuzzy can be used as a method in developing DSM based on gray and unexplained human language. The application of fuzzy methods is done because of its ability to capture uncertainty [21]. FL describes the variables that are the most unclear so that it can help in estimating reasons [22].

Compared to general methods that still provide subjective judgments in technical justification, through FL feelings of subjectivity can be avoided scientifically [23], where fuzzy-based models can provide realistic estimates [24]. Such as the application of fuzzy logic in predicting the location of mineralization that is known in the study area through complete and incomplete data [25]. The application of FL in considering different noise levels in a number of alpha planes and four types of membership functions in a comparison simulation to analyze the type-2 fuzzy approach when applied in fuzzy control [26]. Application of fuzzy sets in chaos control and chaos synchronization such as turbulence control, laser dynamics, plasma, chemical reactions, ecological systems and arrhythmias [27].

Then there is also the design of FL controller complexity with some rules embedded as energy management for microgrids that are connected to residential networks including renewable energy sources and storage capabilities [28]. In the proliferation of cellular devices and heterogeneous wireless communication environments in increasing the need for additional spectrum for data transmission is done by fuzzy-based decision making (FLB-DSS) which together deal with channel selection and channel switching to increase overall device throughput [29]. The selection of the right cluster head in reducing energy consumption dramatically through the fuzzy inference engine in choosing the opportunity that is most likely to overcome this [30]. The design of an index based on FL is aimed specifically at assessing the dynamic quality of agricultural land - Dynamic Quality Index (S-DQI) [31]. Overcoming the issue of merger and acquisition (M&A) assessments and investors in terms of configuration of factors that cause positive and negative reactions [32]. Fuzzy convergence as a control of diversity in Ant Colony Optimization, a kind of population-based meta-heuristic that exploits a form of memory of past performance inspired by the real behavior of ants in foraging, with the aim of maintaining diversity at a certain level in order to avoid premature convergence created [33]. The following is explained in equation (1); where  $Mt(x)$  is a fuzzy value,  $x$  is inputted as a fuzzy value,  $x_i$  is the lower bound and  $N$  is the upper limit.

$$Mt(x) = \begin{cases} x - 1, & 1 \leq x < N \\ 1 & \\ x_i, & A < x \leq N \end{cases} \quad (1)$$

### 2.2 Profile Matching

Profile Matching (PM) is a decision making method assuming that decision making process considers a fulfilling of ideal value for predictor variable [34]. The decision maker should have an ideal value for each parameter operated to check all value of alternative decisions' parameter. The comparison between ideal and current value is going to produce gap value [35]-[36], like a distance of two parameters' value. Here, the marketing company has a standard value compared broadly with all aspects and then identify such a gap. The gap ( $\delta$ ) itself is measured by subtracting the profile of the marketer ( $\mu$ ) by the target value ( $\tau$ ).

After getting a gap value for each marketer profile, each profile is given a weight value by measuring the difference in weight values. The grading of weights can be determined as follows; if the difference is 0 then the weight value is 5, if the difference is 1 then the weight value is 4.5, if the difference is -1 then the weight value is 4, if the difference is 2 then the weight value is 3.5, if the difference is -2 then the weight value is 3, if the difference is 3 then the weighted value is 2.5, if the difference is -3 then the weighted value is 2, if the difference is 4 then the weighted value is 1.5 and furthermore if the difference is -4 then the weighted value is 1 [35].

After determining the gap value in the difference between marketing profile and target value, then dividing aspects into two groups, namely core factor ( $CF$ ) and secondary factor ( $SF$ ), is done.  $CF$  is a major factor in determining the criteria factors in the assessment process that can produce optimal performance as in equation (2); with  $NCF$  as the average value of  $CF$ ,  $NC(a, n)$  as the total number

of *CF* and *IC* values as the number of *CF* items. Whereas *SF* itself is a secondary factor apart from factors in the main criteria [37], can be calculated through equation (3); with *NSF* as the average value of *SF*, *NS(a, n)* as the total number of *SF* values and *IS* as the number of *IS* items.

$$NCF = \frac{\sum NC(a, n)}{IC} \tag{2}$$

$$NSF = \frac{\sum NS(a, n)}{IS} \tag{3}$$

Calculation of the total value is obtained based on the calculation of *CF* and *SF* value calculations in determining the ranking of each marketing. The total value is the ranking of each marketing in a particular position [37]. Examples of calculations can be seen in the formula equation (4); where  $(a, n)$  is the marketing profile and target value,  $N(a, n)$  is the total value of the aspect,  $NCF(a, n)$  is the average value of *CF*,  $NSF(a, n)$  is the average value of *SF* and  $(x)\%$  is the percentage value inputted.

$$N(a, n) = (x)\%NCF(a, n) + (x)\%NSF(a, n) \tag{4}$$

### 3 Research Methodology

Based on the existing research framework, it is necessary to explain further about the stages of the research conducted. Where this stage starts from determining the objectives of the study, then finding the source of the existing problems, which is followed by conducting a literature study so that related parameters are obtained in dealing with the problems that occur. Then after that, the related parameters are analyzed further to further determine the appropriate method in handling the existing problems, until then carried out implementation and testing.

#### 3.1 Identification of Research Problems

In medical appliance distributor companies, the problem that has been identified is related to the company's marketers. Where the motivation of the marketer itself is used as one of the sources in increasing company profits, one of which is by giving a commission. But technically the commission giving that is happening at the moment is still considered unsatisfactory because it tends to be an assessment carried out by the top-level subjective nature with the assumption that feelings tend to be more dominant so that social factors of jealousy often occur among fellow marketers. This has an impact on the decrease in motivation for some marketers who feel unacceptable to the existing assessment. For this reason, this research is conducted so that the research that takes place prioritizes the reality of the facts compared to the value of assumptions that have not been proven to be true.

#### 3.2 Literature Study

Literature study is carried out to collect variables and problems that arise and the needs that are expected in determining top-marketers in the company. As this stage is carried out through a systematical literature review (SLR) to obtain relevant information in accordance

with the research topic. SLR consists of identifying, evaluating and interpreting findings in research topics in order to answer pre-determined research statements [38]. SLR was done by collecting related data through Scencedirect literature library, with searches limited to the last 10 years. Where the title, abstract or keyword is also limited to the "decision support system" and marketing. Not only that, the data was also collected through Googlescholar library related to the data of the last five years with searches limited to only about multi-criteria DSM, so a number of data were obtained for analysis.

#### 3.3 Parameter Determination and Data Collection

Henceforth the data obtained through the literature review stage are grouped again based on the objectives and research themes that have been determined. Through this, the influencing factors and related parameters are able to provide an initial picture of why this research is important to do. The data that has been collected correlatively, then also analyzes the factors that affect the performance or performance of marketers, until some parameters or criteria are identified that might affect the performance of each existing marketer.

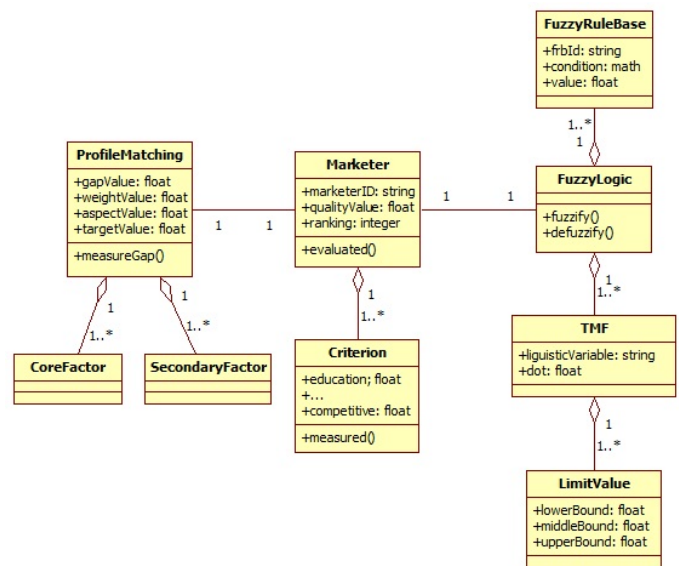


Figure 1: Class Diagram of Constructed Model

#### 3.4 Parameter Priority

Based on the identification and collection of data that has been presented previously, it is necessary to analyze the data first. It is intended that problem solving can produce a solution, not a new problem. The decision support system using the PM and FL method is the right solution after an analysis of the problems often faced by one of the medical equipment distributor companies, so that the right decision can be generated easily based on available data.

Henceforth, the FL method with PM is a method that will be used as a solution to the existing problem, with FL functioning as the initial process in defining large amounts of data so that a mathematical definition is obtained which will be used as input to the PM method. Through the FL method linguistic terms defined



membership function, where the membership function has a membership interval value from 0 to 1. In order to solve the problem is able to produce the expected solution, the PM method is used as a right solution after FL analysis is done to obtain a ranking value so that it is easily determined top-marketer in medical equipment distributors.

## 4 Result and Discussion

### 4.1 Proposed Model

#### 4.1.1 High Level Configuration of Constructed Model

The constructed model clearly depicted via class diagram in Figure 1. Mainly, it consists of three type of central classes: Marketer, Fuzzy-Logic, and ProfileMatching. Class criterion describes parameters considered in the model. Here, interconnections among parameters are described obviously. They operated to measure the quality and ranking of each marketer. All seventeen selected parameters are input of fuzzy rules and ranking is an output of fuzzy rules. The membership function used in the model is triangular membership function (TMF) with specific linguistic variables and limit value.

#### 4.1.2 Selected Parameters/Criterion

Fuzzy values are formulated to help support the decision value by declaring right or wrong as an absolute input in the PM method to eliminate the value of ambiguity. FL is implemented on the marketing criteria attribute as shown in Table 1, so that the relevant values obtained as input to the PM method are obtained.

For the fuzzy membership value, the Education criteria are configured in Figure 2, where the level of member ( $X$ ) that intersects the axis ( $Y$ ) is divided into several level categories namely; Diploma, Bachelor, and Master with each fuzzy value (1, 1, 2), (1, 2, 3), and (2, 3, 3). For the pair of measurement criteria 1, there are three levels of Goodness as shown in Figure 3. Where each level of membership ( $X$ ) which intersects the axis ( $Y$ ) is divided as follows; Not Good ( $NG$ ), Good ( $G$ ), and Very Good ( $VG$ ) with each fuzzy value; i.e. (1, 1, 2), (1, 2, 3), and (2, 3, 3). The fuzzy rules for Education VS Goodness criteria are then made as in Algorithm 1.

---

#### Algorithm 1: Fuzzy Rules for Education VS Goodness

---

```

Result: Goodness value
if (Edu is Diploma) then
  | Goodness is NG;
else
  | if (Edu is Bachelor) then
  | | Goodness is G;
  | else
  | | Goodness is VG;
  | end
end

```

---

#### 4.1.3 Process of Fuzzy Logic - Profile Matching

Calculation of the value of the selected object gap has different calculation values [44]. Where the processing of raw data is carried

out as follows; for marketers with parameter P01, the value of the raw data entered refers as in Figure 1 with diploma education level having a value of 1, bachelor education level having a value of 2 and master education level having a value of 3.

For parameter P02, value input refers to the number of customers handled by each marketer. For example, Marketer M01 handles two customers who are in the presence, the raw data input value is 2. As for the parameter P03, the input value is based on the current billing amount. Where if there is 1 smooth customer bill, the value is 1 and for 1 non-current, worth -1. Then the next value is added up as raw data input.

For the parameter P04, the input value of raw data is obtained based on the accumulation of the number of machines that can be installed on each customer. For example, Marketer M01 the number of machines installed on the customer he has is 4, the raw data input value is 4. Then for the parameter P05 input value, the value is inputted based on the service class owned by the marketer. For example, Marketer M01 service class service is 1, then the value of raw data input is also worth 1. Furthermore, for the value of parameter P06 input, inputted based on the amount of reduction in the machine to the customer owned by each marketer. Missal marketer M01 has 1 engine reduction in the customer he has, so the value of raw data input is 1.

Followed by the parameter P07, the input of raw data is accumulated based on the year of joining the company. For example, marketer M01 started to join the company in 2018, so based on the current year (2019), the marketer has only joined the company for 1 year. So the raw data input value is 1. Furthermore, for parameter P08, raw data input is obtained based on the average value of customer actions owned by each marketer. If the marketer M01 has an average value of customer actions of 15, the raw data input value is also 15. Then the parameter P09 input value is based on the consistency value of the promise fill. If the Marketer M01 has the consistency value of promise filling is 0, then the raw data input is also worth 0. And for the parameter P10, if the percentage value of achieving the Marketer M01 target is 60%, the raw data input value is also worth 60.

Then for the parameter P11, input is based on the number of additional machines that can be done by marketers. If the addition of machines installed by Marketer M01 amounts to 0, then the value of inputting raw data is also worth 0. For the input of the P12 criterion is based on the average achievement of visits to the number of customers handled. If the marketer M01, the visit is 0, the raw data input value is also 0. And for the parameter P13, based on the level of productivity. If the Marketer M01 has a productivity value of 0, the raw data input is also 0.

For parameter P14 itself, raw value input is based on the number of HD products that can be sold in each month. For example, a Marketer M01 is able to sell 120 HD products, so the raw data input value is also 120. As for the Marketer M01's P15, the value of raw data input is based on the amount of cooperation that can be built by the relevant marketer. If Marketer M01 does not get a cooperative relationship with any customer, then the raw data input value is 0.

Table 1: Selected Parameters

Code	Parameter	Reference
P01	Education	[39]
P02	Verbalization idea	[39][40][41]
P03	Reasoning and real solution	[39]
P04	Concentration	[39][40][41][42]
P05	Creativity of imagination	[39]
P06	Anticipation	[39]
P07	Years of service	[39][43]
P08	Accuracy and responsibility	[39]
P09	Behaviour	[39]
P10	Work achievement	[39][40][43]
P11	Vitality and planer	[39][40][41]
P12	Discipline	[39]
P13	Attitude toward company	[40]
P14	Quantity of work	[40]
P15	Work setting	[40][41]
P16	Reliable	[40]
P17	Competitive	[41]

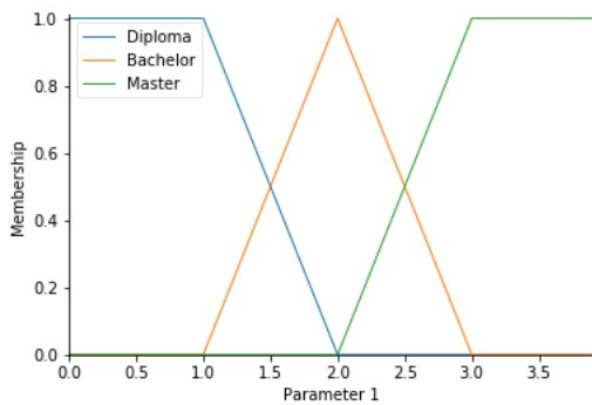


Figure 2: Membership Functions for Parameter 1

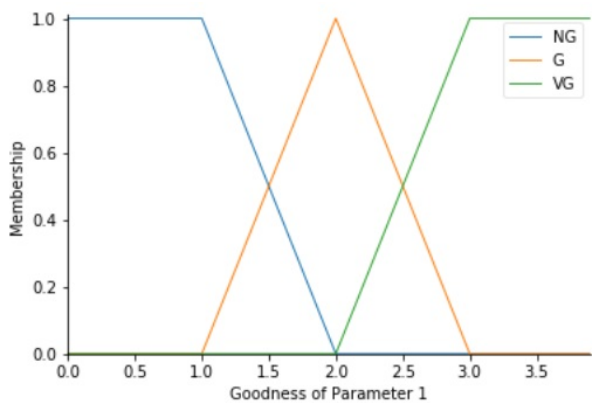


Figure 3: Membership Function for Goodness of Parameter 1

Then for parameter P16, raw data input is based on the number of new HD Center additions. If the Marketer M01 cannot get the addition of a new HD Center, then the raw data input value is 0.

And finally for the parameter P17, the raw data input value is based on the marketer’s ability to reduce the number of machines installed at the customers they have. For example, Marketer M01 cannot reduce the number of competitor machines installed on the customer he is handling, so the raw data input value is 0. Likewise for raw data input on other marketers, the data input process is the same as inputting the Marketer M01’s raw data.

Next we can see in Table 2, an example of the data of 21 marketers that we will use as a sample calculation of FL and PM values. The processing of raw data from each marketer is presented hereinafter to be inputted to FL which then produces output which will be inputted to the PM method.

Table 2: Sample Data of Marketers

Marketer	P01	P02	...	P16	P17
M01	2	2	...	0	0
M02	2	4	...	0	0
...	...	...	...	...	...
M20	2	2	...	0	0
M21	2	10	...	0	0

The first step is to determine the polar interpolation of this value, then proceed to determine the center of gravity with the following formula. Equation (5) is a formula for determining linear interpolation of marketer criteria. Also equation (6) is a formula for determining Central of Gravity (CoG) on marketer’s criteria.

$$Y = Y0 + (X - X0) \frac{Y1 - Y0}{X1 - X0} \tag{5}$$

$$CoG = \frac{\sum(Fot)(FGrade)}{\sum(Fot)} \tag{6}$$

We see fuzzy calculations for the Education criteria for Marketers with the M01 code. Numeric values in the basic rules will be processed to get the value of  $Y(S1)$ . Through the description of Figure 1 we get the input for  $Y0$  is 1,  $Y1$  is 0,  $X$  is 2,  $X0$  is 2 and  $X1$  is 3, so after being entered into equation (5) following  $Y(S1) = 1 + (2 - 2)(0 - 1)/(3 - 2)$  the result is 1. After the Education criteria interpolation results are obtained, then adjusted to rules 1 and adjust fuzzy interpolation input with Table 3, *EIS EIF Edu is 1 THEN Goodness is 1*. Then the CoG value is determined according to equation (6),  $CoG = ((1 * 3) + (1 * 3))/(1 + 1)$  so that the value obtained is 3.

Next we look at fuzzy calculations for the Verbalization Idea criteria on Marketers with the M01 code. Numeric values in the basic rules will be processed to get the value  $Y(K)$ . Through the description of Figure 3 we get the input for  $Y0$  is 1,  $Y1$  is worth 1,  $X$  is worth 2,  $X0$  is worth 5, and  $X1$  is worth 10, so after it is entered into equation (5) following  $Y(K) = 1 + (2 - 5)(1 - 1)/(10 - 5)$  the result is 1.

After interpolation results obtained by the Verbalization Idea criteria, then adjusted to AF 02 rules and adjusting fuzzy interpolation input with Table 3, *IF (VI is 1) THEN Goodness is 1*. Then the CoG value is determined according to equation (6),  $CoG = ((1 * 1) + (1 * 1))/(1 + 1)$  so that the value obtained is 1. Henceforth we do the same steps in the sample data Table 2. by following the calculation method of P1 and P2 on Marketer M01.

After the calculation is done for all criteria and marketers, the overall results can be seen in Table 3.

Table 3: Result of FL Process

Marketer	P01	P02	...	P16	P17
M01	3.0	1.0	...	1.0	1.0
M02	3.0	1.0	...	1.0	1.0
...	...	...	...	...	...
M20	3.0	1.0	...	1.0	1.0
M21	3.0	3.0	...	1.0	1.0

### 4.2 The Result Applied to Medical Company

After the fuzzy value is determined, the next aspect is divided into two groups, Core Factor (CF) and Secondary Factor (SF) as in Table 4. CF is the main factor in determining the assessment process that can produce optimal performance, while SF is a secondary or supporting factor in addition to factors in the main criteria [36]-[37].

Table 4: Assessment Aspects

Core Factor	Secondary Factor
Accuracy and responsibility	Work achievement
	Education
	Concentration
	Quantity of work
	Reliable
Verbalization idea	Reasoning and real solution
	Creativity of imagination
	Anticipation
	Vitality and planner
	Competitive
Attitude toward company	Years of service
	Behaviour
	Discipline
	Work setting

Table 5: Calculation of GAP Value for Each Parameter

Marketer	P01	P02	...	P16	P17
M01	3.0	1.0	...	1.0	1.0
M02	3.0	1.0	...	1.0	1.0
...	...	...	...	...	...
M20	3.0	1.0	...	1.0	1.0
M21	3.0	3.0	...	1.0	1.0
GAP	3.0	3.0	...	3.0	3.0
M01	0.0	-2.0	...	-2.0	-2.0
M02	0.0	-2.0	...	-2.0	-2.0
...	...	...	...	...	...
M20	0.0	-2.0	...	-2.0	-2.0
M21	0.0	0.0	...	-2.0	-2.0

gap value for each criterion based on equation (2). Where we can see an example of determining the Gap value for M01 marketers in the Education criteria column, 3 minus the gap value 3 (the standard ideal value in the Education criteria) the result is 0. With the overall marketer's results can be seen in Table 5. After the gap value is obtained, the weight value is determined based on the provisions in Table 1. With each employee there is a weight value table [44], which looks like in Table 6. After the difference results are obtained, each aspect is grouped into 2 groups, namely by determining the value of CF and SF [44]. How to calculate the average value according to equation (3) and equation (4) so that the results obtained are as in Table 7 and Table 8.

Table 6: Weight for each Parameter

Marketer	P01	P02	...	P16	P17
M01	5.0	3.5	...	3.5	3.5
M02	5.0	3.5	...	3.5	3.5
...	...	...	...	...	...
M20	5.0	3.5	...	3.5	3.5
M21	5.0	0.0	...	3.5	3.5

Table 7: CF Value

Marketer	VI	AR	Atc	CF
M01	3.50	2.00	3.50	3.00
M02	3.50	4.00	3.50	3.67
...	...	...	...	...
M20	3.50	1.00	3.50	2.67
M21	0.00	3.50	3.50	1.75

Table 8: SF Value

Marketer	P01	P02	...	P16	P17	SF
M01	5.0	3.5	...	3.5	3.5	3.25
M02	5.0	3.5	...	3.5	3.5	2.64
...	...	...	...	...	...	...
M20	5.0	3.5	...	3.5	3.5	2.79
M21	5.0	0.0	...	3.5	3.5	3.29

Table 9: Ranking for each Marketer

Marketer	CF	SF	N (a, n)	Ranking
M01	3.00	3.25	3.10	5
M02	3.67	2.64	3.26	4
M03	0.17	3.15	1.36	21
...	...	...	...	...
M09	3.67	3.61	3.65	1
...	...	...	...	...
M20	2.67	2.79	2.72	8
M21	1.75	3.29	2.37	14

Calculation of the total value of each marketer is obtained based on equation (5) with the percentage of CF is 65% and the percentage of SF is 45%. The total value itself is a ranking of each marketer in

After classifying CF and SF, the next step is to determine the

a certain position [37], as shown in Table 9. In addition, for more details, we can see the value of marketer ranking through Figure 4. It can be seen on the graph that the highest value achieved by Marketer M09 is ranked first with a ranking value of 3.65, while the lowest value can be seen obtained by Marketer M03 with a rating of 1.36. The result of marketer ranking is new idea to see marketer objectively. Then, the company can take a proper decision to give a reward.

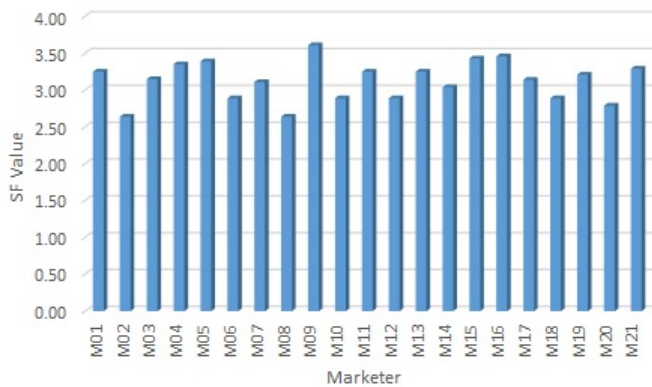


Figure 4: Marketers' SF Value Charts

## 5 Conclusion and Furtherworks

A generic DSM for determining the best marketer of particular medical equipment distributor company based on main methods FL and PM was methodically developed. Seventeen verified-selected parameters considered in the model; where they were scientifically adopted from five scientific manuscripts. The model is a novel model that is able to enrich the knowledge domain of computer science. Also, the study result can be functioned as a formal procedure in evaluating the employees' performance generally. Thus it contributes to the real world practically.

In addition, the application of FL and PM methods is very useful in determining the value of marketer ranking. Because the level of accuracy of the FL method is able to minimize the value of ambiguity, so the expected final results will be more accurate and rational. The model is very useful in helping top-level decision making. Because here the top-level is required to be more professional by putting aside feelings and assumptions in deciding something. Through the results of research that have been obtained, top-level companies where research is conducted can make decisions more quickly and effectively when determining who has the right to become a Top-Marketer of the company. So that with this also doubts can be minimized and can reduce mental burdens or mistakes that may be more dominated by the feeling element.

Finally, new parameters should be taken in to account for next study (e.g. communication skill, informal education, etc.). To do so, the deep literature study is possibly to conduct. And also, for implementing other methods is still open to conduct (e.g. optimization method, simple additive weighting, etc.). The proper and more detail methods implementation is able to make the model more reliable and fruitful scientifically.

## References

- [1] V. Chouliaras, A. Gazepis, T. Kargidis, "Marketing's Contribution to the Profitability of Greek Enterprises During the Economic Crisis" *Procedia Economics and Finance*, **19**, 217–225, 2015. doi: 10.1016/S2212-5671(15)00023-4
- [2] H. Lim, J. G. Choi, A. Akhmedov, J. Chung, "Predicting future trends of media elements in hotel marketing by using Change Propensity Analysis" *International Journal of Hospitality Management*, 278–431, 2018. doi: 10.1016/j.ijhm.2018.10.001
- [3] F. Y. Lo, N. Campos, "Blending internet-of-things (IoT) solutions into relationship marketing strategies" *Technological Forecasting and Social Change*, **137**, 10–18, 2018. doi: 10.1016/j.techfore.2018.09.029
- [4] S. F. Slater, G. T. M. Hult, E. M. Olson, "Factors influencing the relative importance of marketing strategy creativity and marketing strategy implementation effectiveness" *Industrial Marketing Management*, **39**(4), 551–559, 2010. doi: 10.1016/j.indmarman.2008.03.007
- [5] J. Y. Ho, M. Dempsey, M. "Viral marketing: Motivations to forward online content" *Journal of Business research*, **63**(9-10), 1000–1006, 2010. doi: 10.1016/j.jbusres.2008.08.010
- [6] A. Sela, D. Goldenberg, I. Ben-Gal, E. Shmueli, "Active viral marketing: Incorporating continuous active seeding efforts into the diffusion model" *Expert Systems with Applications*, **107**, 45–60, 2018. doi: 10.1016/j.eswa.2018.04.016
- [7] J. Tian, Y. Zhang, C. Zhang, "Predicting consumer variety-seeking through weather data analytics" *Electronic Commerce Research and Applications*, **28**, 194–207, 2018. doi: 10.1016/j.elerap.2018.02.001
- [8] G. Bykzkan, B. Ergn, "Intelligent system applications in electronic tourism" *Expert systems with applications*, **38**(6), 6586–6598, 2011. doi: 10.1016/j.eswa.2010.11.080
- [9] T. Turk, O. Kitapci, I. T. Dorytol, "The usage of geographical information systems (GIS) in the marketing decision making process: a case study for determining supermarket locations" *Procedia-Social and Behavioral Sciences*, **148**, 227–235, 2014. doi: 10.1016/j.sbspro.2014.07.038
- [10] M. Moradi, A. Aghaie, M. Hosseini, "Knowledge-collector agents: Applying intelligent agents in marketing decisions with knowledge management approach" *Knowledge-Based Systems*, **52**, 181–193, 2013. doi: 10.1016/j.knosys.2013.08.014
- [11] C. Kiss, M. Bichler, "Identification of influencers measuring influence in customer networks" *Decision Support Systems*, **46**(1), 233–253, 2008. doi: 10.1016/j.dss.2008.06.007
- [12] S. L. France, M. S., Vaghefi, H. Zhao, "Characterizing viral videos: Methodology and applications" *Electronic Commerce Research and Applications*, **19**, 19–32, 2016. doi: 10.1016/j.elerap.2016.07.002
- [13] M. Ripolls, A. Blesa, "International new ventures as small multinationals: The importance of marketing capabilities" *Journal of World Business*, **47**(2), 277–287, 2012. doi: 10.1016/j.jwb.2011.04.015
- [14] L. Khiali, D. Ienco, M. Teisseire, "Object-oriented satellite image time series analysis using a graph-based representation" *Ecological Informatics*, **43**, 52–64, 2018. doi: 10.1016/j.ecoinf.2017.11.003
- [15] D. N. Utama, *Sistem Penunjang Keputusan: Filosofi, Teori, dan Implementasi*. Garudhawaca, 2017.
- [16] F. Gunawan, G. Wang, D. N. Utama, S. Komsiyah, "Decision Support Model for Supplier Selection Using Fuzzy Logic Concept" in *International Conference on Information Management and Technology*, 394–399, 2018. doi: 10.1109/ICIMTech.2018.8528103
- [17] L. A. Zadeh, "Fuzzy logica personal perspective" *Fuzzy sets and systems*, **281**, 4–20, 2015. doi: 10.1016/j.fss.2015.05.009
- [18] D. N. Utama, E. Rustamaji, "Fuzzy Decision Support Model for Human Resources Performance Appraisal" *Proceedings of 2018 International Conference on Information Management and Technology*, 272–276, 2018. doi: 10.1109/ICIMTech.2018.8528145



- [19] J. W. Liu, "Using big data database to construct new GFuzzy text mining and decision algorithm for targeting and classifying customers" *Computers & Industrial Engineering*, 2018. doi: 10.1016/j.cie.2018.04.003
- [20] Q. Zhou, Z. Xu, N. Y. Yen, "User sentiment analysis based on social network information and its application in consumer reconstruction intention" *Computers in Human Behavior*, **100**, 177–183, 2018. doi: 10.1016/j.chb.2018.07.006
- [21] A. Vafadarnikjoo, N. Mishra, K. Govindan, K. Chalvatzis, "Assessment of Consumers' Motivations to Purchase a Remanufactured Product by Applying Fuzzy Delphi Method and Single Valued Neutrosophic Sets" *Journal of Cleaner Production*, **196**, 230–244, 2018. doi: 10.1016/j.jclepro.2018.06.037
- [22] K. Howells, A. Ertugan, "Applying fuzzy logic for sentiment analysis of social media network data in marketing" *Procedia computer science*, **120**, 664–670, 2017. doi: 10.1016/j.procs.2017.11.293
- [23] D. N. Utama, U. Taryana, "Fuzzy Logic For Simply Prioritizing Information In Academic Information System" *International Journal of Mechanical Engineering and Technology (IJMET)*, **10**(2), 1594-1602 2019.
- [24] L. Suganthi, S. Iniyar, A. A. Samuel, "Applications of fuzzy logic in renewable energy systemsa review" *Renewable and sustainable energy reviews*, **48**, 585–607, 2015. doi: 10.1016/j.rser.2015.04.037
- [25] A. Ford, J. M. Miller, A. G. Mol, "A comparative analysis of weights of evidence, evidential belief functions, and fuzzy logic for mineral potential mapping using incomplete data at the scale of investigation" *Natural Resources Research*, **25**(1), 19–33, 2016.
- [26] O. Castillo, L. Amador-Angulo, J. R. Castro, M. Garcia-Valdez, "A comparative study of type-1 fuzzy logic systems, interval type-2 fuzzy logic systems and generalized type-2 fuzzy logic systems in control problems" *Information Sciences*, **354**, 257–274, 2016. doi: 10.1016/j.ins.2016.03.026
- [27] S. Vaidyanathan, A. T. Azar, "Takagi-Sugeno fuzzy logic controller for Liu-Chen four-scroll chaotic system" *International Journal of Intelligent Engineering Informatics*, **4**(2), 135–150, 2016. doi: 10.1504/IJIEI.2016.076699
- [28] D. Arcos-Aviles, J. Pascual, L. Marroyo, P. Sanchis, F. Guinjoan, "Fuzzy logic-based energy management system design for residential grid-connected microgrids" *IEEE Transactions on Smart Grid*, **9**(2), 530–543, 2016. doi: 10.1109/TSG.2016.2555245
- [29] A. Ali, L. Abbas, M. Shafiq, A. K. Bashir, M. K. Afzal, H. B. Liaqat, ..., K. S. Kwak, "Hybrid fuzzy logic scheme for efficient channel utilization in cognitive radio networks" *IEEE Access*, **7**, 24463–24476, 2019. doi: 10.1109/ACCESS.2019.2900233
- [30] R. Logambigai, A. Kannan, "Fuzzy logic based unequal clustering for wireless sensor networks" *Wireless Networks*, **22**(3), 945–957, 2016.
- [31] E. Rodriguez, R. Peche, C. Garbisu, I. Gorostiza, L. Epelde, U. Artetxe, ..., J. Etxebarria, "Dynamic Quality Index for agricultural soils based on fuzzy logic" *Ecological indicators*, **60**, 678–692, 2016. doi: 10.1016/j.ecolind.2015.08.016
- [32] J. T. Campbell, D. G. Sirmon, M. Schijven, "Fuzzy logic and the market: A configurational approach to investor perceptions of acquisition announcements" *Academy of Management Journal*, **59**(1), 163–187, 2016. doi: 10.5465/amj.2013.0663
- [33] O. Castillo, H. Neyoy, J. Soria, P. Melin, F. Valdez, "A new approach for dynamic fuzzy logic parameter tuning in ant colony optimization and its application in fuzzy control of a mobile robot" *Applied soft computing*, **28**, 150–159, 2015. doi: 10.1016/j.asoc.2014.12.002
- [34] J. Fitriana, E. F. Ripanti, Tursina, "Sistem Pendukung Keputusan Pemilihan Mahasiswa Berprestasi dengan Metode Profile Matching (Studi Kasus: Fakultas Teknik UNTAN)" *Jurnal Sistem dan Teknologi Informasi*, **6**(4), 153–160, 2018.
- [35] S. Oktafiani, D. N. Utama, "Generic Model of Fuzzy Profile Matching for Determining the Best Marketer" *International Journal of Advanced Trends in Computer Science and Engineering*, **9**(1), 859–869, 2020. doi: 10.30534/ijatcse/2020/123912020
- [36] A. Verdian, A. Wantoro, "Komparasi Metode Profile Matching Dengan Fuzzy Profile Matching Pada Pemilihan Wakil Kepala Sekolah" *Jurnal Ilmiah Media Sisfo*, **13**(2), 97–105, 2019.
- [37] L. Sopianti, N. Bahtiar, "Student Major Determination Decision Support Systems using Profile Matching Method with SMS Gateway Implementation" *Jurnal Sains dan Matematika*, **23**(1), 14–24, 2015.
- [38] R. B. Gandara, G. Wang, D. N. Utama, "Hybrid Cryptography on Wireless Sensor Network: A Systematic Literature Review" in 2018 International Conference on Information Management and Technology, 2018. doi: 10.1109/ICIMTech.2018.8528147
- [39] A. A. T. Susilo, "Penerapan Metode Profile Matching pada Sistem Pendukung Keputusan Pemilihan Ketua Program Studi (STUDI Kasus: Program Studi Teknik Informatika STMIK Musi Rawas)" *JUITA: Jurnal Informatika*, **5**(2), 87–93, 2018.
- [40] R. Nasriyah, Z. Arham, Q. Aini, "Profile matching and competency based human resources management approaches for employee placement decision support system (case study)", *Asian J. Appl. Sci*, **9**(2), 75–86, 2016. doi: 10.3923/ajaps.2016.75.86
- [41] T. Susilo, A. Anto, "Penerapan Metode Composite Performance Index (CPI) pada Pemilihan Hotel di Kota Lubuklinggau" *Jurnal Rekayasa Sistem dan Teknologi Informasi*, **1**(3), 204–210, 2017.
- [42] B. W. Sari, "Perbandingan Metode Profile Matching Dan Simple Additive Weighting Pada Penentuan Jurusan Siswa Kelas X SMA N 2 Ngaglik" *Data Manajemen dan Teknologi Informasi (DASI)*, **16**(1), 2015.
- [43] A. Nurdianto, S. J. Hartati., Y. M. Maulana, "Rancang Bangun Sistem Pendukung Keputusan Penempatan Tenaga Pendidik Menggunakan Metode Profile Matching" *Jurnal Sistem informasi dan Komputer Akuntansi*, **5**(2), 1–8, 2016.
- [44] T. Susilowati, E. Y. Anggraeni, W. Andewi, Y. Handayani, A. Maseleno, "Using Profile Matching Method to Employee Position Movement" *Int. J. Pure Appl. Math*, **118**(7), 415–423, 2018.

## Design of Aerial Panoramic Photography: Contrast between Industrialized and Natural Zones

Wei Zhong Feng<sup>1</sup>, Yu-Che Huang<sup>2,\*</sup>, Fang-Lin Chao<sup>2</sup>

<sup>1</sup>Department of Visual Communication, Chaoyang University of Technology, Taichung, 436, Taiwan, R.O.C.

<sup>2</sup>Department of Industrial Design, Chaoyang University of Technology, Taichung, 436, Taiwan, R.O.C.

### ARTICLE INFO

Article history:

Received: 06 July, 2020

Accepted: 25 July, 2020

Online: 31 August, 2020

Keywords:

Panoramic

Aerial film

Creation

Environmental awareness

Planet image

### ABSTRACT

*This study utilized aerial panoramic photography and 3D virtual reality to demonstrate the environmental changing issues. Environmental problems have occurred in the surroundings for a long time, but the public has gradually become numb. The wide view through aerial photography helps to show the difference. The proposed self-assembled drone equipped with an antishock bracket containing a six-axis aerial camera was used to capture videos with the critical environmental problems in Taiwan. The captured panoramic video was processed using post-production software to create a unique documentary revealing the adverse effects of pollution. The video portrayed a sharp contrast between the eastern and western coastline of Taiwan by presenting planet images. Most participants of a survey reported increased awareness of environmental protection after watching the aerial panoramic films.*

## 1. Introduction

Taiwan's limited land resources are often annexed by industries to satisfy economic needs. The industrialization has considerably enriched people's lifestyles but also permanently changed the environment. Because changes occur gradually over an area, people habitually endure and ignore them.

Farmlands are gradually being converted into industrial and residential zones. This unbalanced utilization of environmental resources poses a threat to the sustainability of society. Although laws related to environmental protection have been enacted, the conversion of farmland into industrial zones has not abated. Although institutions have attempted to preserve cultivated land [1], recent manifestation of environmental problems indicates that development policies require reform. The education of citizens regarding environmental protection is crucial for the development of sustainability-related social consciousness. Agriculture and health sustainability need public policy support, especially in the context of investment and regional technology development [2].

### 1.1. Technology Background

Aerial panoramic photography and 3D virtual reality technologies are used in artistic creation and movies to demonstrate the environmental changing issues. These technologies contrasted landscape transformed by industrialization

to evoke reflection on the sustainable utilization of resources. Panoramic art has been employed to present ecological changes to the public to induce such reflection. Dramatic representation using artistic techniques causes people to think about the relationship between humans and the environment in the process of modernization. People's perception helps to identify and weigh social priorities and maintain desired sustainable relationships.

New technology allows us to experience situations from a first-hand perspective. To achieve this goal, the following techniques were used to present design concepts:

- Aerial photography: Aerial photography provides a broad view of the changes in a landscape. However, aerial photography cannot offer a microscopic perspective of the changes.
- Digital image processing: This processing technology enables us to show the overall relationship. Post-production can considerably transform captured scenes; for example, a remote 360° perspective can be employed to convey an idea effectively.
- Multiple projectors connected in series: Images can be divided to depict various perspectives to prompt the viewer's emotional investment.
- Virtual reality (VR) enabled by mobile phones: Mobile phones have built-in gyroscopes and orientation sensors. The

\*Corresponding Author: Y.C. Huang, 79 Futong St., Taichung 436, Taiwan R.O.C., 886-910673898, hychc@cyut.edu.tw

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj050499>

orientation of mobile phone changes the viewing angle, and the image displayed on the mobile phone screen changes. VR can be used to take advantage of these devices and present extraordinary experiences without the need for additional helmets and induction devices.

VR video technology has emerged as a new medium of documentary filmmaking. More news editing teams are using immersive technologies to relate news events effectively. These teams are usually composed of technical and journalism teams [3]. Although journalism is a time-critical field, limited research has been conducted on immersive technology. VR provides a human-computer interactive experience and has, therefore, become a new avenue for documentary filmmaking. The appropriate application requires an understanding of advanced interdisciplinary methods and proper design and techniques [4].

The configured aerial imaging platform provides a new perspective and enables the user to investigate high-altitude locations [5]. When a user moves along a high-altitude path, the device receives location data from the camera sensor. Unidentified aerial system (UAS) provides instant processing under controlled conditions. The cameras installed on the platform have the following specifications[6]: GoPro Hero and DJI [7]. UAS images can clearly show areas of concrete and steel bridges that are generally inaccessible. Furthermore, non-destructive damage can be detected, which is suitable for conveying environmental problems. Observing the structure and dynamics of social ecosystems in specific geographic areas [8] can help policymakers. However, drones pose security risks and privacy concerns [9]. Therefore, the government strictly regulates their use.

### 1.2. Related Studies

Considerable progress has been achieved in the creation of static spherical panoramic images and full-area guided VR tours. These technologies are now user friendly and widely employed [10]. Digital camera technology has evolved considerably over the years, resulting in improved image resolution. Image pixels are no longer the main distinguishing factor of mobile visual communication. User-friendly interconnectivity is essential for creating immersive user experiences and visual interfaces [11]. For example, Google Street View, which provides location photos accumulated through user participation, is a popular platform. Street View is also used in virtual tours, even allowing advanced navigation functions in designated regions.

## 2. Concept and implementation

### 2.1. Concept

Because Taiwan relied on thermal power generation, coupled with industrial intermediate byproducts and vehicle exhaust, fine particles are currently the country's first environmental problem. The high number of suspended particles in the vicinity of thermal power plants is a concern for public health. In Taiwan, thermal power plants are located along the central coast of Taiwan (Figure 1a). However, the considerable suspended matter has been observed in Puli, a mountainous area. A cumulative effect has thus occurred, caused by the accumulation of fine floating particles in the basin area with a range of more than 100 km.

Taiwan's main petrochemical and medium-sized industries are located in Kaohsiung City in southern Taiwan (Figure 1b). The industrial zone and Petro-China Kaohsiung Refinery constitute a major petrochemical town of south Taiwan. This industrial area borders a river in the east, and the Taiwan Strait is located to the

south approximately 25 km from city. The total number of manufacturers in the area is 33, including 19 chemical material and four chemical product manufacturers. The coast was originally an ecological zone full of mangroves but is being slowly converted into an industrial area.

The staggered existence of agricultural and industrial areas is a controversial issue. Agricultural land becomes an option for industrialization when land acquisition is difficult. The sewage discharge and waste treatment of factories in agricultural areas affect the lives of nearby families.

Because of the three aforementioned observations, we selected the following as the focus of our documentary: (1) thermal power plants (Figure 1a), and (2) major petrochemical industrial areas (Figure 1b). Factories scattered in rice fields in suburbs of the city (Figure 1c). Pollution images were captured through aerial photography and used to compare pollution levels in the industrial and agricultural sectors, which could then be presented to the public to understand the situation.



Figure 1: (a) Thermal power plant, (b) Refinery Industrial Zone, (c) Factories scattered in rice fields



## 2.2. Visual Design Setting

Figure 2 depicts the design process of image production. An arranged screen spliced by projectors. A shockproof photography bracket was mounted on a drone. GoPro HERO3+ motion recorders (six underneath) were installed. The obtained 4k panoramic video was then edited using Kolor-Autopano software, with the landscape images transformed into a planet-shaped image displaying a sharp contrast between industrial and agricultural areas.

The video-capturing method of a 360° panoramic camera is different from that of an ordinary camera. Irrespective of lens rotation and tilting, the picture remains horizontal and faces the same direction. Panoramic shooting mainly provides high and low or left and right viewing angles. The 720° VR display is a 360° horizontal and vertical guide, allowing the user to decide the perspective of the scene or view of the image.

Kolor-Autopano software is a video-editing software program that can be used to stitch together two or more panoramic videos of overlapping areas to correct the horizon control points. A parallax compensation algorithm for stitching videos enables the user to reduce the time required to fix stitching issues in post-production. A planet-shaped image, called a small planet with a parallel display, is the desired viewing style in power plant projects. These tiny planet photos are a novel means of displaying panoramas and landscape photos.

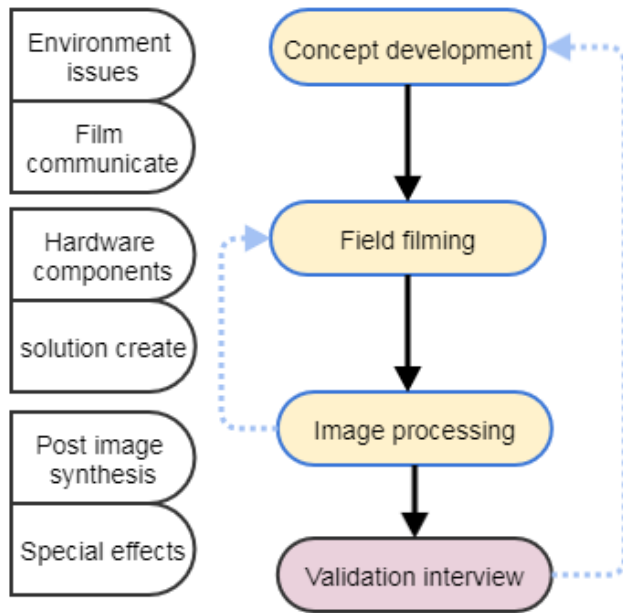


Figure 2: Iterative design and study process.

Hardware implementation (Figure 3) utilized several facilities with the flying drone to ensure the image quality. From the top to the bottom of the fuselage, there are Kodak Pixpro SP360 Camera, three-axis gimbal stabilizer for VR Camera, three-axis gimbal stabilizer for VR Camera, and 360° Spherical VR bracket for six GoPro HERO4. The GoPro is small and weighs only 74 g. This single-leg fixed accessory can easily fit on an unmanned aerial vehicle. The camera has a wide-angle lens that resembles a fisheye lens, obtaining an extended depth of field. The largest aperture is f/2.8, and the camera has a full 4k cinema resolution at 12 fps.

The aerial camera captured videos from different heights to determine the optimum height at which to shoot the video; the results of the experiment indicated that the optimum camera height was 500 m, slightly higher than the clouds. Next, the panoramic view experiments were performed. An appropriate number of spatial samples were selected during photography to obtain high-quality images. In the selected shooting area, the researchers avoided restricted flight areas and designed the image acquisition method according to the selected flight path. The flight path was controlled to record multiple photos at a selected angle. Typically, more than 40 images were captured, which were then combined in the post-processing step.



Figure 3: Hardware implementation and utilization.

## 3. Petrochemical industrial areas

### 3.1. Spherical Transformation

A spherical matrix operator was developed for panoramic image processing based on spherical panoramas. The photogrammetric equations were reviewed and matched with the geometry of spherical panoramic images. The shape of the line was a two-dimensional mapping of a three-dimensional space. The rotation matrix of spherical photogrammetry is an analytical approach that is used with spherical images [12].

The Laplacian operator is employed to obtain the relationship between omnidirectional panoramic and spherical images (tiny planet images).

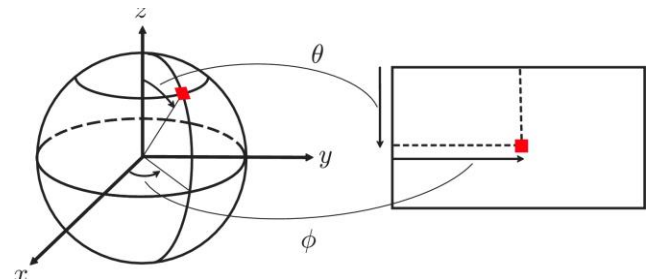


Figure 4: Coordinate the relationship between omnidirectional panoramic and spherical images [13].

An eight-neighbor Laplacian operator matrix has been reported for omnidirectional panoramic images. An extension of the Laplacian operator for omnidirectional panoramic images covering 360° views solves the problem of nonuniformity of spatial resolution [13]. The spherical images were mapped by equirectangular projection in which the meridians and parallels are



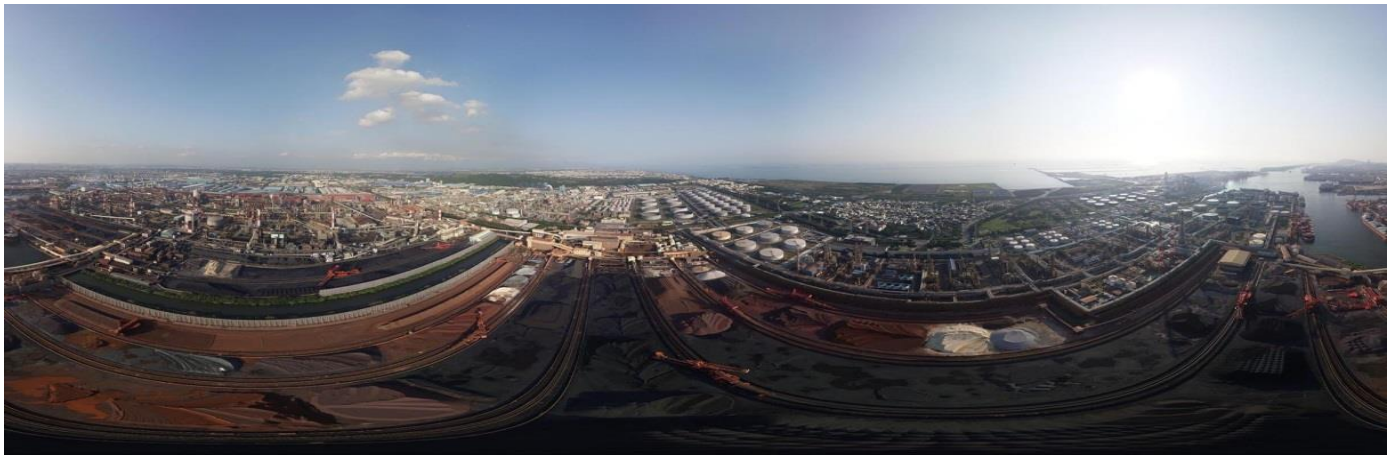


Figure 5: Panoramic aerial photo of Kaohsiung's Dalinpu village surrounded by the industrial zone.

represented by vertical and horizontal lines. The position of an image pixel related to a corresponding point on the spherical image. The computational time required (decomposition and integration processes) was considerably reduced by preparing a correspondence table of grid images in advance. This theoretical analysis considerably reduces the mathematical difficulties encountered by art producers.

We captured a top view of the industrial zone to trigger the viewer's thinking. The industrial area was shown in the images to appear calm in the morning (Figure 5). With spherical appearance, a tiny planet image is a metaphor and can prompt people to feel the pressure placed on the environment by the extensive industry comprising the tiny planet. The spherical projection viewer can then reflect on various appearances at different times (Figure 6).

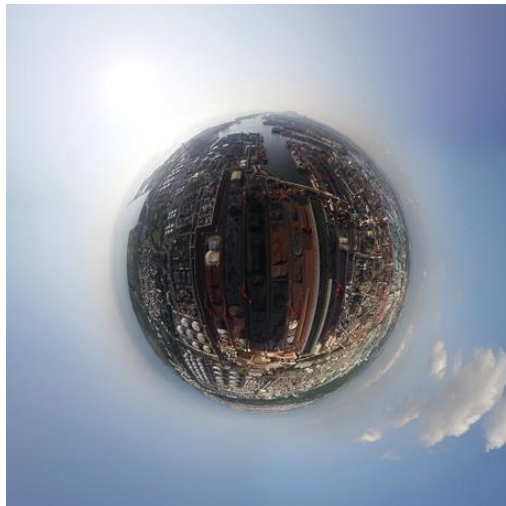


Figure 6: Spherical projection showing a tiny planet image.

### 3.2. Dual-image Contrast

The "double-sided" contrast image shows a sharp contrast between the industrialized world and untouched natural environment on the eastern and western coasts of Taiwan, respectively. The eastern coast is mountainous. From north of Hualien Creek to the south to Xiaoyeliu in Taitung County, facing the Pacific Ocean, the mountains, deep valleys, and coastal terrain are populated with indigenous people. Long-term differential erosion on the northern coast has created terraces and reef banks. Stones are scattered randomly over the coastline, showing pure and

primitive beauty. Unusual geological patterns have been formed by geological movement, long-term weathering, and seawater erosion (Figure 7). The strata consist of interlayers of sandstone and shale from top to bottom.

The "two-sided" contrast image is a merging of the two hemispheres and shows the polarization between the industrial world and the untouched natural environment. Figure 8 presents a dual half-sphere image of the eastern region versus the western industrial zone. In a matrix developed for the region surrounding the factory, suspended particles and black smoke cover the original blue sky and white clouds. The blue sky and white clouds symbolize the need to breathe freely and deep inside desire. The matrix arrangement of industrial plants symbolizes regularity and suffocation, while the gray sky makes people feel oppressed.

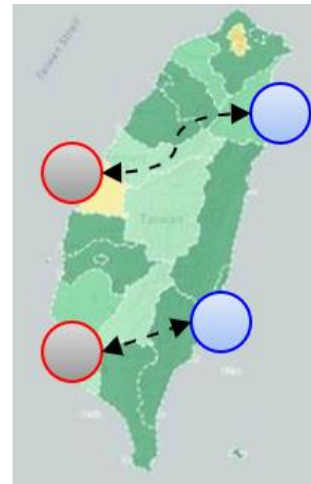


Figure 7: Contrasting locations on the eastern and western coasts of Taiwan

Figure 9 shows a dual image of rocky topography in the northern region versus a dense industrial area on the western coast. When the two regions are merged in a single image, unevenness exists along the seam. In post-production, elastic zoom editing was employed to adjust the natural edges of the two images so that the seam is not apparent. The untouched sandstone coast intrigues the viewer and reminds them that it is an indispensable part of a sustainable life. The contrast between the natural etched space and the industrial area evokes strong reactions from the viewer.



Figure 8: Green space in the eastern region versus an industrial zone in the west.



Figure 9: Dual images of rocky topography in the north versus a dense industrial area on the western coast.

## 4. Thermal power plants

### 4.1. Smoky Tower

The theme of the project was to convey the following: “God, please look at how the people have burned the planet you have created to satisfy their desires. Heaven or hell? Who can live in this place?” The film starts with the narration of a fictional ancient legend, reminding people that they should be in awe of nature, and

shows a bird’s-eye view of a soul floating in the sky, presenting a scene of overdevelopment of Taiwan’s western coast industry. The monologue is written in a narrative style of mystery, with air pollution issues illustrated through aerial photography.

Incense burning is a traditional eastern custom but also associated with thermal power plants, towering columnar chimneys, and smoke. As the drone gradually approached the power plant, the first striking feature was numerous trenches containing burning coal. Symbolically, these trenches depict ashes formed by incense in exchange for satisfying human avarice. Figure 10 presents an image captured when the drone was at half the height of the chimney and depicted the chimney and other combustion units from a distance. The chimney is shown to be coated with images of animals (black-faced spoonbill birds) to confuse the public cognition. Finally, the drone reached the top of the chimney and surveyed the deep darkness, a strong symbolic representation of the endless desire for burning.



Figure 10: Chimney painted with images of animals [14].

### 4.2. Panoramic Layout

The natural coast of the east and industrial area of the west were combined to form a sharply contrasting image. The corresponding hemispheres on the left and right were inserted between the boundary; by overlapping the picture of the plane and 3D hemispheres. The combination of multiple post-production images increases the variability and also forms a sharp contrast between the blocks. This method of flipping and superimposing the images created the effect of alternation between heaven and earth (Figure 11).

In post-production, the bulge of the landscape was enhanced using a four-axis aerial camera passing through the clouds, creating a bird’s-eye view (Figure 12). The images show the obstruction of the air in the sky, resulting in a heavy feeling. The cross-domain surreal composition encourages the viewer to engage in dialogue. The themes resonate with the audience and transform the interaction from a digital framework directly into the discussion on air pollution from a perspective of environmental protection. Because no specific shooting space was used, the creator had to relinquish control of the visual frame, and the viewer is encouraged to discover for themselves. The design can be used to evoke empathy and ensure that the viewer has understood the message hidden in the image.





Figure 11: Comparison of an industrial area with a natural coastal beach [15].



Figure 12: Bird's-eye view of severe air pollution, which blurs the view.

## 5. Survey and discussion

### 5.1. Questionnaire and Analysis

Thirty-seven college students were recruited for a survey. They were requested to watch the video on a VR headset connected to their smartphone. After watching the video, they completed a questionnaire, rating the video presentation, their level of feeling, and environmental concerns. We used SPSS software to analyze the participants' reported feelings. The responses to the unstructured questionnaire were then summarized. The content of the questionnaire is shown in below.

#### Part A: Feeling (Likert scale)

- q1. The ground view is different from the ground.
- q2. You feel yourself flying when watching the 360 film.
- q3. I like to explore the environment in a 360-degree scene.
- q4. Exploring the environment in 360 degrees is time-consuming.
- q5. I will rotate my body to watch the rear during 360-degree exploration.
- q6. I feel dizzy while watching.
- q7. The way of viewing the left and right split screen is very creative.
- q8. The spherical image that appears in the film helps to understand the theme of the film.
- q9. The smog in the air is scary.
- q10. I'm worried about air pollution in central Taiwan.
- q11. I am willing to participate in PM2.5 environmental protection actions
- q12. The visual communication technique is updated quickly, I can accept VR movies in the future.

#### Part B. Opinions on film creation:

- o1. After watching the 360-degree image in the sky, my feeling is...
- o2. The three chapters in the movie, which one do you like and why?

### 5.2. Results

The results of the 12-item narrative statistics are shown in table 1. The results with the largest mean are questions 1 (4.5946), 12 (4.5405), and 5 (4.4054), respectively. This represents a strong agreement: the difference between watching in the air, movies that can already accept VR, will rotate the body to watch when watching with a mobile phone. Show that younger generations can accept VR skills and are willing to try them. The results with the smallest mean are questions 4 (2.8378), 6 (3.000), respectively. This represents disagreement: observation with a mobile phone is not time-consuming and no dizziness. This value shows the convenience of mobile phones to observe VR.

Table 1: Descriptive statistics of the items

Items	Average	Std. deviation
q1	4.5946	0.4977
q2	3.5676	0.8673
q3	4.0811	0.7218
q4	2.8378	1.1184
q5	4.4054	0.8963
q6	3.0000	1.2693
q7	3.8649	0.7875
q8	3.4865	0.9013
q9	3.6757	1.0015
q10	4.2162	0.8862
q11	3.9459	0.9703
q12	4.5405	0.8365

We merged three questions with similar meanings to understand the subjects' attitudes through the Bi-variable correlation analysis. Among them, t1 is the addition of three questions 1, 3, and 5, which means "like to experience 360-degree videos"; t2 is the question 9, 10, and 11, and express willingness to do environmental protection. The participants all appreciated

and wanted to watch the VR audiovisual presentation (Figure 13). The sounds in the film, such as the buzzing of the factory and sounds of the wind and heartbeat, helped them explore the content of the film. The audience was deeply affected by the film and reported that they were more likely to participate in environmental protection activities (Figure 14). We used an independent sample T-test to test whether male and female participants felt differently about the above issues; there were no significant difference.

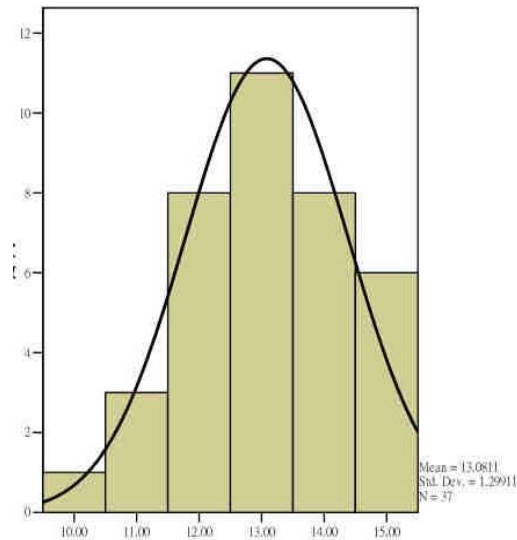


Figure 13: Histogram of participants' reactions (t1) to the film.

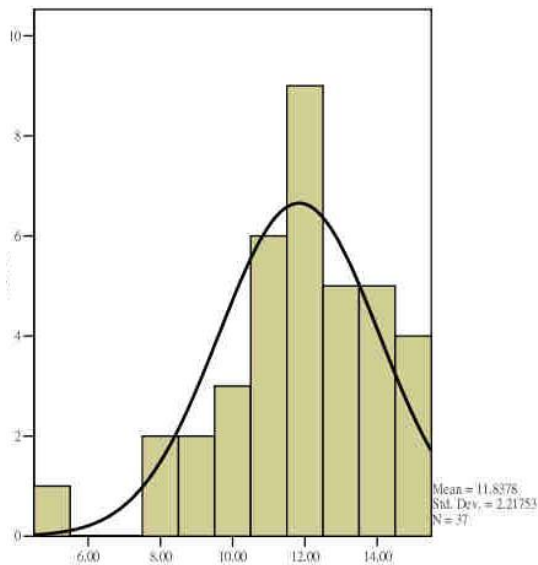


Figure 14: Participants' willingness to participate (t2) in the conservation event.

The participants reported that the images of a peaceful offshore fishing port surrounded by dense factories were depressing; the balance between the environment and factories was crucial. The smoke from the factory-made them feel as if they had entered the chimney and caused anxiety about the industrial areas scattered along the coast of Taiwan. The participants also reported that the view from the sky was very different from the view in their daily experience. The left versus right image segmentation produced a sharp contrast; the spherical display was creative, easy to understand, and added tension.

We use text analysis to count and observe their individual independent opinions. Among them, the count number of "willing to try to use VR" is 5, but "I don't know where to watch in the movie" and "feel tired or exhausted," still have eight counts. It shows that VR does not have a special visual focus so that people often enter a fantasy state without knowing where to look during exploration. This requires strengthening guidance in the design. The novelty feeling of flying in high altitude (skip from the existing state) and enter a broader experience counted 14 times, which shows the strong willingness to use this media.

## 6. Conclusions

Young students had a strong willingness to use VR media. Mobile phones perform well in changing the viewing angle for VR experiences. The "double-sided" contrast image shows a sharp contrast between the industrialized world and untouched natural environment on the eastern and western coastline of Taiwan. The 3D view of the digitally processed panoramic image enriches the spatial range of perception.

Post-production transformed the captured scenes and highlighted the contrast. The combined image also appeared on the east's natural coastline with the industrial landscape. Overlapping the planes in a 3D space evokes positive reactions from the viewer. Comparing natural space and an industrial area encourages the viewer to explore the differences between the two environments. The observation with a mobile phone is not time-consuming. The participants reported skipping from the present status and entering a broader experience.

## Conflict of Interest

The authors declare no conflict of interest.

## References

- [1] T. Zhong, B. Mitchell, S. Scott, X. Huang, X., Y. Li, X. Lu., "Growing centralization in China's farmland protection policy in response to policy failure and related upward-extending unwillingness to protect farmland since 1978." *Environment and Planning C: Politics and Space*, **35**(6), 1075-1097. <https://doi.org/10.1177/0263774X16682958>
- [2] J. Y. Hsu, D. W. Gimm, J. Glassman, "A tale of two industrial zones: A geopolitical economy of differential development in Ulsan, South Korea, and Kaohsiung, Taiwan." *Environment and Planning A: Economy and Space*, **50**(2), 457-473, 2018. <https://doi.org/10.1177/0308518X16680212>
- [3] R. Mabrook, and J. B. Singer, "Virtual Reality, 360° video, and journalism studies: conceptual approaches to immersive technologies," *Journalism Studies*, 1-17, 2019. <https://doi.org/10.1080/1461670X.2019.1568203>
- [4] E. Sirkkunen, and P. P. Rezaei, "Journalism in virtual reality: opportunities and future research challenges," In *Proceedings of the 20th International Academic Mindtrek Conference*, 297-303, ACM, 2016. <https://doi.org/10.1145/2994310.2994353>
- [5] R. Chandra, Manohar Swaminathan, "Aerial imaging of a region using above ground aerial camera platform," US Patent Application 15/414,949, filed July 26, 2018. <https://patents.google.com/patent/US20180213187A1/en>
- [6] S. Dorafshan, M. Maguire, V. Nathan, C. Coopmans, "Challenges in bridge inspection using small unmanned aerial systems: Results and lessons learned," In *2017 International Conference on Unmanned Aircraft Systems (ICUAS)*, 1722-1730, IEEE, 2017. <https://doi.org/10.1109/ICUAS.2017.7991459>
- [7] DJI, <https://www.dji.com>, access date: 25 July, 2020, 12:20.
- [8] Zuppinger-Dingley, Debra, Cornelia B. Krug, Owen Petchey, Bernhard Schmid, Norman Backhaus, and Michael E. Schaepman, "Editorial overview: environmental change issues: integrated global change and biodiversity research for a sustainable future," **vii-xi**, 2017. <https://doi.org/10.1016/j.cosust.2018.05.019>
- [9] I. Budinska, "On Ethical and Legal Issues of Using Drones" In: Aspragathos N., Koustoumpardis P., Moulianitis V. (eds) *Advances in Service and Industrial Robotics. RAAD 2018. Mechanisms and Machine Science*, **67**. Springer, Cham. 2018. [https://doi.org/10.1007/978-3-030-00232-9\\_74](https://doi.org/10.1007/978-3-030-00232-9_74)



- [10] Ippoliti, Elena, and M. Calvano, "Enhancing the cultural heritage between visual technologies and virtual restoration: case studies to models for visual communication," In *Digital Curation: Breakthroughs in Research and Practice*, 309-348, IGI. Global, 2019. <https://doi.org/10.4018/978-1-5225-6921-3.ch015>
- [11] L. Brun, "Enabling 360 visual communications: next-level applications and connections," *IEEE Consumer Electronics Magazine*, **5**(2), 38-43, 2016. doi: 10.1109/MCE.2016.2519050
- [12] K. Hara, K. Inoue, and K. Urahama, "Gradient operators for feature extraction from omnidirectional panoramic images," *Pattern Recognition Letters*, **54**, 89-96, 2015. <https://doi.org/10.1016/j.patrec.2014.12.010>
- [13] G. Fangi and C. Nardinocchi, "Photogrammetric processing of spherical panoramas," *The Photogrammetric Record*, **28**(143), 293-311, 2013. <https://doi.org/10.1111/phor.12031>
- [14] <https://www.youtube.com/watch?v=eFlfVw5A6XY>, access date: 25 July, 2020, 12:25.
- [15] <https://youtu.be/z3IMRkr6iHw>, access date: 25 July, 2020, 12:30.