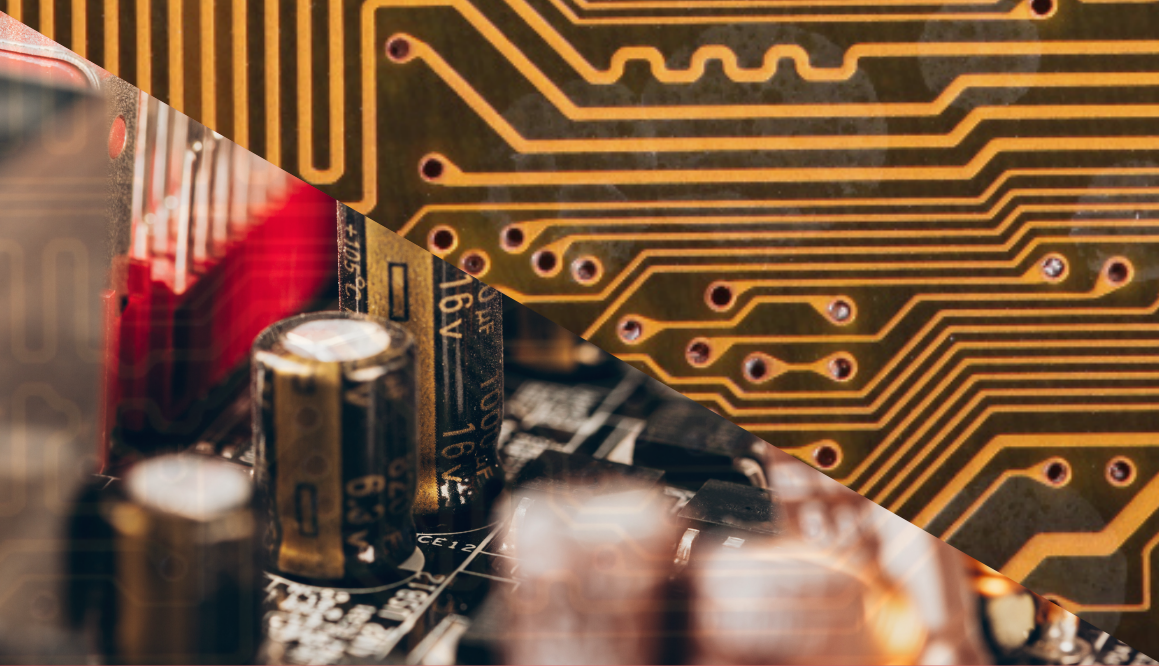


Advances in Science, Technology & Engineering Systems Journal



VOLUME 5-ISSUE 3 | MAY-JUNE 2020

www.astesj.com

ISSN: 2415-6698

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

Advances 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 73 accepted papers in Computer Science domain.

Editor-in-chief

Prof. Passerini Kazmersk

ADVANCES IN SCIENCE, TECHNOLOGY AND ENGINEERING SYSTEMS JOURNAL

Volume 5 Issue 3

May-June 2020

CONTENTS

<i>Evaluation of Uncertainty Measurement Calculation for Vector Network Analyzer From 300 kHz to 8.5 GHz</i>	01
Tan Ming Hui, Ahmad Yusairi Bani Hashim	
<i>Parameter Estimation on Two-Dimensional Advection-Dispersion Model of Biological Oxygen Demand in Facultative Waste Water Stabilization Pond: Case Study at Sewon Wastewater Treatment Facility</i>	11
Sunarsih Sunarsih, Dwi Purwanto Sasongko, Sutrisno Sutrisno	
<i>A Novel Hybrid Method for Segmentation and Analysis of Brain MRI for Tumor Diagnosis</i>	16
Kapil Kumar Gupta, Namrata Dhanda, Upendra Kumar	
<i>Centralized System of Universities Learning Materials</i>	28
Ruslan Vynokurov, Volodymyr Tigariev, Oleksii Lopakov, Kateryna Kirkopulo, Olena Pavlyshko	
<i>Dynamic Objects Parameter Estimation Program for ARM Processors Based Adaptive Controllers</i>	34
Vasiliy Olonichev, Boris Staroverov, Maxim Smirnov	
<i>Design, Implementation and Performance Analysis of a Dual Axis Solar Tracking System</i>	41
Ba Thanh Nguyen, Hong-Xuyen Thi Ho	
<i>Using Leader Election and Blockchain in E-Health</i>	46
Basem Assiri	
<i>Organizational, Social and Individual Aspect on Acceptance of Computerized Audit in Financial Audit Work</i>	55
Bambang Leo Handoko, Nada Ayuanda, Ari Tihar Marpaung	
<i>Performance Effects of Algorithmic Elements in Selected MANETs Routing Protocols</i>	62
Mutuma Ichaba, Felix Musau, Simon Nyaga Mwendia	
<i>Intrusion Detection in Cyber Security: Role of Machine Learning and Data Mining in Cyber Security</i>	72
Gillala Rekha, Shaveta Malik, Amit Kumar Tyagi, Meghna Manoj Nair	
<i>Evolution of Privacy Preservation Models in Location-Based Services</i>	82
A B Manju, Sumathy Subramanian	

<i>A Harmonized European Drone Market? – New EU Rules on Unmanned Aircraft Systems</i>	93
Anna Konert, Tadeusz Dunin	
<i>Personality Measurement Design for Ontology Based Platform using Social Media Text</i>	100
Andry Alamsyah, Sri Widiyanesti, Rizqy Dwi Putra, Puspita Kencana Sari	
<i>Experimental Studies of the Silicon Photomultiplier Readout Electronics Based on the Array Chip MH2XA030</i>	108
Oleg Dvornikov, Vladimir Tchekhovski, Yaroslav Galkin, Alexei Kunz, Nikolay Prokopenko	
<i>University Students Result Analysis and Prediction System by Decision Tree Algorithm</i>	115
Md. Imdadul Hoque, Abul kalam Azad, Mohammad Abu Hurayra Tuhin, Zayed Us Salehin	
<i>Alternative Real-time Image-Based Smoke Detection Algorithm</i>	123
Sally Almanasra, Ali Alshahrani	
<i>Improvement of Desirable Thermophysical Properties of Soybean Oil for Metal Cutting Applications as a Cutting Fluid</i>	129
Putta Nageswara Rao, Suresh Babu Valer, Koka Naga Sai Suman	
<i>Socioeconomic and Productive Disparity in Child Stunting in the Central Andes of Peru, Taking as a Model the Community of Tunanmarca, Jauja</i>	135
Jorge Castro-Bedriñana, Doris Chirinos-Peinado, Elva Ríos Ríos	
<i>Performance Analysis of Joint Precoding and Equalization Design with Shared Redundancy for Imperfect CSI MIMO Systems</i>	142
Bui Quoc Doanh, Ta Chi Hieu, Truong Sy Nam, Pham Thi Phuong Anh, Pham Thanh Hiep	
<i>Non Parallelism and Cayley-Menger Determinant in Submerged Localization</i>	150
Anisur Rahman	
<i>Analysis and Improvement of an Innovative Solution Through Risk Reduction: Application to Home Care for the Elderly</i>	158
Linda Acosta-Salgado, Auguste Rakotondranaivo, Eric Bonjour	
<i>A Perturbation Finite Element Approach for Correcting Inaccuracies on Thin Shell Models with the Magnetic Field Formulation</i>	166
Vuong Dang Quoc, Quang Nguyen Duc	

<i>Digestibility, Digestible and Metabolizable Energy of Earthworm Meal (Eisenia Foetida) Included in Two Levels in Guinea Pigs (Cavia Porcellus)</i>	171
Jorge Castro-Bedriñana, Doris Chirinos-Peinado, Hanz Sosa-Blas	
<i>Piezoelectric Teeth Aligners to Accelerate Orthodontics Treatment</i>	178
Muath Bani-Hani, M. Amin Karami	
<i>Generating a Blockchain Smart Contract Application Framework</i>	191
Arif Furkan Mendi, Tolga Erol, Emre Safak	
<i>Factors Influencing Social Knowledge Management in Social Society: A Systematic Literature Review</i>	198
Erick Fernando, Meyliana, Achmad Nizar Hidayanto, Harjanto Prabowo	
<i>Automated Abaca Fiber Grade Classification Using Convolution Neural Network (CNN)</i>	207
Neptali Montañez, Jomari Joseph Barrera	
<i>Machine Learning Model to Identify the Optimum Database Query Execution Platform on GPU Assisted Databasev</i>	214
Dennis Luqman, Sani Muhamad Isa	
<i>valuation of Type A Uncertainty in a Network Analyzer From 300 kHz to 8.5 GHz</i>	226
Tan Ming Hui, Ahmad Yusairi Bani Hashim, Mohd Rizal Salleh	
<i>A Review on Autonomous Mobile Robot Path Planning Algorithms</i>	236
Noraziah Adzhar, Yuhani Yusof, Muhammad Azrin Ahmad	
<i>Study of Wrinkling and Thinning Behavior in the Stamping Process of Top Outer Hatchback Part on the SCGA and SPCC Materials</i>	241
Sri Wahyanti, Agus Dwi Anggono, Waluyo Adi Siswanto	
<i>Improved Nonlinear Fuzzy Robust PCA for Anomaly-based Intrusion Detection</i>	249
Amal Hadri, Khalid Choug dali, Raja Touahni	
<i>Enhanced Collaborative Constellation for Visible Light Communication System</i>	259
Manh Le Tran, Sunghwan Kim	
<i>Trajectory Tracking Control of a DC Motor Exposed to a Replay-Attack</i>	264
Reda El Abbadi, Hicham Jamouli	
<i>Digital Sovereignty Between "Accountability" and the Value of Personal Data</i>	270
Nicola Fabiano	
<i>Angular Orientation of Steering Wheel for Differential Drive</i>	275
Rajesh Kannan Megalingam, Deepak Nagalla, Ravi Kiran Pasumarthi, Vamsi Gontu, Phanindra Kumar Allada	

<i>Controller Design Using Backstepping Algorithm for Fixed-Wing UAV with Thrust Vectoring System</i>	284
Shogo Hirano, Kenji Uchiyama, Kai Masuda	
<i>Based on Reconfiguring the Supercomputers Runtime Environment New Security Methods</i>	291
Andrey Molyakov	
<i>Analysis of Local Rainfall Characteristics as a Mitigation Strategy for Hydrometeorology Disaster in Rain-fed Reservoirs Area</i>	299
Kartono Kartono, Purwanto Purwanto, Suripin Suripin	
<i>Multimode Control and Simulation of 6-DOF Robotic Arm in ROS</i>	306
Rajesh Kannan Megalingam, Raviteja Geesala, Ruthvik Rangaiah Chanda, Nigam Katta	
<i>A Hybrid Approach for Intrusion Detection using Integrated K-Means based ANN with PSO Optimization</i>	317
Jesuretnam Josemila Baby, James Rose Jeba	
<i>ANN Based MRAC-PID Controller Implementation for a Furuta Pendulum System Stabilization</i>	324
Efrain Mendez, German Baltazar-Reyes, Israel Macias, Adriana Vargas-Martinez, Jorge de Jesus Lozoya-Santos, Ricardo Ramirez-Mendoza, Ruben Morales-Menendez and Arturo Molina	
<i>Balance as One of the Attributes in the Customer Segmentation Analysis Method: Systematic Literature Review</i>	334
Uus Firdaus, Ditdit Nugeraha Utama	
<i>Estimation of Influential Parameter Using Gravitational Search Optimization Algorithm for Soccer</i>	340
J. Vijay Fidelis, E. Karthikeyan	
<i>Sentence Retrieval using Stemming and Lemmatization with Different Length of the Queries</i>	349
Ivan Boban, Alen Doko, Sven Gotovac	
<i>Design and Optimization of Dual-Band Branch-Line Coupler with Stepped-Impedance-Stub for 5G Applications</i>	355
Ayyoub El Berbri, Adil Saadi, Seddik Bri	
<i>A Survey on Image Forgery Detection Using Different Forensic Approaches</i>	361
Akram Hatem Saber, Mohd Ayyub Khan, Basim Galeb Mejbil	
<i>A Model for Operationalizing the Information Technology Strategy Based on Structuration View</i>	371
Thami Batyashe, Tiko Iyamu	

<i>Applications of Causal Modeling in Cybersecurity: An Exploratory Approach</i>	380
Suchitra Abel, Yenchih Tang, Jake Singh, Ethan Paek	
<i>Racial Categorization Methods: A Survey</i>	388
Krina B. Gabani, Mayuri A. Mehta, Stephanie Noronha	
<i>Enhancing Decision Making Capabilities in Humanitarian Logistics by Integrating Serious Gaming and Computer Modelling</i>	402
Za'aba Bin Abdul Rahim, Giuseppe Timperio, Robert de Souza, Linda William	
<i>A Framework for Measuring Workforce Agility: Fuzzy Logic Approach Applied in a Moroccan Manufacturing Company</i>	411
Fadoua Tamtam, Amina Tourabi	
<i>The Design of an Experimental Model for Deploying Home Area Network in Smart Grid</i>	419
Fatima Lakrami, Najib El Kamoun, Hind Sounni, Ouidad Albouidya, Khalid Zine-Dine	
<i>Prognosis of Failure Events Based on Labeled Temporal Petri Nets</i>	432
Redouane Kanazy, Samir Chafik, Eric Niel	
<i>Performance of Robust Confidence Intervals for Estimating Population Mean Under Both Non-Normality and in Presence of Outliers</i>	442
Juthaphorn Sinsomboonthong, Moustafa Omar Ahmed Abu-Shawiesh, Bhuiyan Mohammad Golam Kibria	
<i>A Solution Applying the Law on Road Traffic into A Set of Constraints to Establish A Motion Trajectory for Autonomous Vehicle</i>	450
Quach Hai Tho, Huynh Cong Phap, Pham Anh Phuong	
<i>Degradation Process in Pipeline and Remaining Useful Lifetime Estimation Based on Extended Kalman Filtering</i>	457
Med Hedi Moulahi, Faycal Ben Hmida	
<i>Spot Toyota: Design and Development of a Mobile Application for Toyota's Promotion Actions to the Young Audience</i>	469
Nuno Martins, Joel Enes	
<i>Monte Carlo Estimation of Dose in Heterogeneous Phantom Around 6MV Medical Linear Accelerator</i>	478
Zakaria Aitelcadi, Mohamed Reda Mesradi, Redouane El Baydaoui, Ahmed Bannan, Abdennacer Ait Ayoub, Kamal Saidi, Saad Elmadani	
<i>5G mm-wave Band pHEMT VCO with Ultralow PN</i>	487
Abdelhafid Es-Saqy, Maryam Abata, Mohammed Fattah, Said Mazer, Mahmoud Mehdi, Moulhime El Bekkali, Catherine Algani	

- Smart Transmission Line Maintenance and Inspection using Mobile Robots* 493
Thongchai Disyadej, Surat Kwanmuang, Paisarn Muneesawang, Jatuporn Promjan, Kanyuta Poochinapan
- Measurement of Employee Awareness Levels for Information Security at the Center of Analysis and Information Services Judicial Commission Republic of Indonesia* 501
Mainar Swari Mahardika, Achmad Nizar Hidayanto, Putu Agya Paramartha, Louis Dwysevrey Ompusunggu, Rahmatul Mahdalina, Farid Affan
- The Application of Mobile Learning Technologies at Malaysian Universities Through Mind Mapping Apps for Augmenting Writing Performance* 510
Rafidah Abd Karim, Airil Haimi Mohd Adnan, Mohd Haniff Mohd Tahir, Mohd Hafiz Mat Adam, Noorzaina Idris, Izwah Ismail
- A Survey and an IoT Cybersecurity Recommendation for Public and Private Hospitals in Ecuador* 518
Maximo Giovani Tazado Espinoza, Joseline Roxana Neira Melendrez, Luis Antonio Neira Clemente
- Risk Management: The Case of Intrusion Detection using Data Mining Techniques* 529
Ruba Obiedat
- Business Process Design for Widuri Indah School Management System with the Support of Cloud Computing* 536
Yulyanty Chandra, Roy Willis, Calvin Windoro, Sfenrianto
- Efficiency Enhancement of p-i-n Solar Cell Embedding Quantum Wires in the Intrinsic Layer* 540
Nahid Akhter Jahan, M. Mofazzal Hossain
- Efficient Discretization Approaches for Machine Learning Techniques to Improve Disease Classification on Gut Microbiome Composition Data* 547
Hai Thanh Nguyen, Nhi Yen Kim Phan, Huong Hoang Luong, Trung Phuoc Le, Nghi Cong Tran
- Dynamics Model and Design of SMC-type-PID Control for 4DOF Car Motion Simulator* 557
Pham Van Bach Ngoc, Bui Trung Thanh
- Promotion of the Research Activities at the Image Processing Research Laboratory (INTI-Lab) of the UCH as Knowledge Management Strategy* 563
Avid Roman-Gonzalez, Natalia Indira Vargas-Cuentas
- A Fuzzy-PID Controller Combined with PSO Algorithm for the Resistance Furnace* 568
Trinh Luong Mien, Vo Van An, Bui Thanh Tam

Warehouse Relocation of a Company in the Automotive Industry Using P-median 576
Zarate-Zapata, Aldo Cesar, Garzón-Garnica, Eduardo Arturo, Cante-Mota,
Román, Olmos-Álvarez, Fernando, Martínez-Flores, José Luis, Sánchez-
Partida, Diana

Solutions for Building a System to Support Motion Control for Autonomous Vehicle 583
Quach Hai Tho, Huynh Cong Phap, Pham Anh Phuong

Evaluation of Uncertainty Measurement Calculation for Vector Network Analyzer From 300 kHz to 8.5 GHz

Tan Ming Hui^{*1,2}, Ahmad Yusairi Bani Hashim²

¹Radio Frequency Calibration Laboratory, National Instruments (M) Sdn Bhd, 11900, Malaysia

²Faculty of Manufacturing Engineering, Universiti Teknikal Malaysia Melaka, 76100, Malaysia

ARTICLE INFO

Article history:

Received: 04 January, 2020

Accepted: 14 April, 2020

Online: 03 May, 2020

Keywords:

Measurement Uncertainty

Network Analyzer

Reflection

ABSTRACT

Increasing the telecommunications products that allow Vector Network Analyzer is becoming more common tools to measure the S-Parameter. It will be an absolute number from the S-Parameter measurements produced in real and imaginary, other words it is also known as the product of the calculation. The calculation findings do not include the systematic and random errors. It's the reaction of the engineer to mitigate the likelihood of random and systemic errors. One of the common random error solutions is through the statistical analysis in the Vector Network Analyzer either repeated measurement or turn on high averaging measurement. The more data assessed, the greater the engineer's confidence in evaluating random errors did not contribute significant errors. Systemic Error is consistent and reproducible when the measurement is made. One way of harmonizing these errors is to evaluate uncertainty measurements in the calculation for Vector Network Analyzer to perform measurements of reflection and transmission. Transmission measurements produce the three systematic errors that were directivity, source match and frequency response reflection tracking. This paper will concentrate from 300 kHz to 8.5 GHz directivity experimental to determine the accuracy of the Vector Network Analyzer. The experimental results will check balance with the Vector Network Analyzer specification. It is a validation process to ensure the Vector Network Analyzer meets the specification in order to perform an accurate measurement. The estimation of measurement uncertainty also refers to the Metrology 100 series Joint Committee for Guide to the Expression of Uncertainty in Measurement. The uncertainty expended should apply to Student Table's confident level of 95%. It creates awareness to demonstrate the importance of measurement quality associated with the uncertainty, particularly for an ISO17025:2017 certified competence testing and calibration laboratory. Without the uncertainty associate to the measurement, it is not complying to the standard ISO17025:2017.

1. Introduction

The raise of the telecommunication products enabling Vector Network Analyzer (VNA) [1] is getting more popular instruments to making the S-Parameter measurements. The Vector Network Analyzer is measuring in Time Domain Reflection and Time Domain Transmission converting into time domain. The Vector Network Analyzer had been helped engineer to solve the Fourier Transformation calculation and giving the results in real and

imaginary number. The real and imaginary would be an absolute number, other words it is also known as measurement result. The measurement results are not including the random and systematic errors yet.

In the Vector Network Analyzer random error is statistical deviations in any direction during calculation. This error's source is unknown. It is not a mistake that the technician made during the calibration of the Vector Network Analyzer. The same issue is also encountered even through a highest traceability chain in calibration performed by national measurement laboratory. It is not possible to eradicate random error from a calculation. It is the

*Tan Ming Hui, National Instruments (M) Sdn Bhd, 8 Lebuhr Batu Maung 1, 11900 Bayan Lepas, Penang, Malaysia, Contact No : +604-3776326 & Email: P051510012@student.utem.edu.my; ming.hui.tan@ni.com

www.astesj.com

<https://dx.doi.org/10.25046/aj050301>

duty of the engineer to mitigate the probability of the random errors. One of the common random error solutions is through the statistical analysis in the Vector Network Analyzer either repeated measurement or turn on high averaging measurement. The more data taken into calculation, the greater the engineer's confidence in evaluating a random error did not contribute substantial errors to the calculation of the Device Under Test [2].

Systematic error in a Vector Network Analyzer is reproducible through the calibration experiment and unable to determine the systematic error by statistical analysis. This mistake demonstrates consistently the same direction in making measurements. Systematic error contributor may be calculated by instrument impact factors, environmental factors, incoming source frequency and others do not cause by statistical contribution.

Random and systemic errors and consequences can lead to measurement uncertainty reporting in a huge number. Calculation of measurement uncertainty often refers to the Metrology 100 series Joint Guide Committee which is the Guide to Expression of Uncertainty in Measurement. It is creating awareness to demonstrate the importance of measurement quality associate with the uncertainty.

Many aspects in our life, we are accustomed to the doubt to estimate length in Meter, weight in Kilogram, temperature in Kelvin, time in Second, electric current in Ampere, amount of substance in Mole and luminous intensity in Candela. These are the 7 units of measure defined by the International System of Units were defined with its traceability. Quantitative measurement is not complete without uncertainty being reported with the measurement. For example, estimate a weight of a kilo pack of sugar. The manufacturer always including a tolerance from the weight that they measured. The manufacturer also has a tolerance of the weight they were examining. This means the measurement is in question by the manufacturer. This is the key reason why calibration is necessary to retain its tractability to the International System of Units. Another example of new Fifth Generation of Mobile Technologies (5G) sub-6GHz low frequency spectrum antenna required calibration up to 7.125 GHz [3]. The antenna manufacturer required to purchase the Vector Network Analyzer to perform analysis of the antenna reflection and transmission. Vector Network Analyzer examinations will help determine whether the antenna produce is pass or fail. Vector Network Analyzer's accuracy is critical to ensure the antenna that is manufactured within the product's limits. In this paper the same approach definition applies in the two ports Vector Network Analyzer measurement for transmission and reflection. To calculate the measurement uncertainty, the associated contributor must be established which will affect the calibration. Such calibration would give the antenna manufacturer a source of uncertainty in the calibration report associated with an absolute measurement [4]. This uncertainty reported in the calibration reports consists of the Type A and Type B [5] uncertainty calculation from the calibration provider. Type A uncertainty consist of repeatability [6] and Type B uncertainty consists of product error such as reflection coefficient [7]. The calibration report will be used in a subsequent evaluation of uncertainty or other words names as "imported uncertainty".

This paper will implement a calibration of the 3 decibel (dB) fixed attenuator defined as a half power loss in the radio frequency

transmission line [8]. The half power loss is the best case for simulating radio frequency components as measured in the Vector Network Analyzer in terms of transmission and reflection. It is a passive device and high stability in repeating measurement compare to other active device such as antenna, amplifier, mixer and filters in the production floor. These products manufactured in the production floor would own the wider specification, tolerance and low stability. It is not a good sample for Vector Network Analyzer to perform the calibration and measurement uncertainty analysis because wider specification and tolerance contribute inaccurate measurement. At the same time, it will cause the Vector Network Analyzer measured in low accuracy, low precision and high uncertainty in calibration. Evaluation of type A uncertainty required a high stability system upon repeated measurement. This is for achieving the optimal measurement of type A uncertainty. Meanwhile, uncertainty of type B would allow a well-known good Vector Network Analyzer to perform the calibration of the transmission and reflection. This paper will demonstrate through calibration process associated with expanded uncertainty to the 3 dB fixed attenuator measurement. The key advantage is the same technique often used for the process of calibration of power sensors, radio frequency cables and connectors. The evaluation of uncertainty measurement in calculation from 300 kHz to 8.5 GHz will be applied if the same PXIe-5632 Vector Network Analyzer remains.

This paper is an extension of the work originally presented in the 2016 7th International Conference on Mechanical, Industrial, and Manufacturing Technologies [9] and Journal of Engineering Science and Technology Vol. 14, [10]. The limitation of this paper is frequency bandwidth from 300 kHz until 8.5 GHz only and frequency response transmission tracking analysis did not perform.

At the end of this paper the effects of measuring uncertainty will be compared with the other well-known manufacturer of the Vector Network Analyzer. It is to compare the estimation of the expanded uncertainty in measurement with the specific frequency bandwidth range. Improvements for future research can be established from the comparison.

2. Evaluation of Uncertainty Measurement

2.1. Type A Evaluation of Uncertainty

Evaluation of measurement uncertainty classified into two categories Type A and Type B. Type A evaluation evaluates the uncertainty resulted from the statistical study. The statistical analysis may be an undergrowth series of repeated measurements of the same process. The series of repeated measurements will fill in a standard deviation of mean. Type B evaluates uncertainty other than statistical analysis. The uncertainty is based on the manufacturer published specification.

Estimation of Type A evaluation of uncertainty can be applied when a set of measurement was recorded under the same condition with minimum 2 repeated measurement. An example of a quantity Y input with N statistically independent ($N > 1$) observed as y_j (where $j = 1, 2, 3, \dots, n$). The estimation of quantity of Y is \bar{y} , arithmetic mean wrote as equation (1) or applying the function of Average in Excel shown by equation 1 [11].

$$\bar{y} = \frac{1}{N} \sum_{j=1}^N y_j \quad (1)$$

The arithmetic mean of \bar{y} is evaluated in equation 2 or applying the function of VAR.S in Excel to estimate the variance of probability distribution define as $s^2(y)$ shown by equation 2 [11].

$$s^2(y) = \frac{1}{N-1} \sum_{j=1}^N (y_j - \bar{y})^2 \quad (2)$$

After estimation of arithmetic mean and variance had been calculated, the experimental standard deviation for Type A uncertainty was calculated in equation 3 by applying the function of STDEV in Excel shown by equation 3 [12].

$$s^2(\bar{y}) = \frac{s^2(y)}{N} \quad (3)$$

The Type A standard uncertainty $u(\bar{y})$ associated with the initial set of measurement repeated \bar{y} is the experimental standard deviation formula in equation 3. The standard uncertainty calculation is applying the same formula in equation 3. In order to differentiate the standard uncertainty and standard deviation, shown by equation 4. Type A standard uncertainty share the same formula in equation 3.

$$u(\bar{y}) = s(\bar{y}) \quad (4)$$

The main purpose of repeated measurement is to disperse the data into a statistical waveform. This study obtains 5 repeated measurements to fill in the Type A measurement uncertainty.

Table 1: PXIe-5632 Type B contributors

Frequency Range	300 kHz to <5 GHz	5 GHz to 8.5 GHz	Unit
Directivity	0.0106	0.020	dB
Source Match	0.0125	0.0195	dB
Reflection Tracking	0.1	0.1	dB
Transmission Tracking	0.12	0.12	dB
Load Match	0.0048	0.0114	dB
Power Step Resolution	0.01	0.01	dB
Trace Noise	0.0060	0.0060	dB

2.2. Type B Evaluation of Uncertainty for PXIe-5632

Type B contributor to the uncertainty was based on the product specification of the PXIe-5632 Vector Network Analyzer. It is necessary to list as many as which aspect of the system will lead to the error. Table 1 describing seven types of contributors.

Vector Network Analyzer directivity is the relation between the leakage signal and the reflected signal. The lower directional leakage signals the better measurement quality. Source match define as the internal reflection error between the Vector Network www.astesj.com

Analyzer generator and device under test. Reflection tracking is the loss from the test port, cable and connection at port 1 and port 2 independently at the network analyzer. Transmission tracking is the loss from the test port, cable and connection from port 1 to port 2 at the network analyzer. Load match is the error occur from the device under test to the receiver of network analyzer. Power step resolution refer to the most sensitivity resolution setting in the Vector Network Analyzer. Trace noise is the stability of the power when making the measurement across the bandwidth [13] swept in the network analyzer.

In the practical to perform a complete test and calibration set up shown by Figure 1 is a full two port characteristic calibration. This set up consists of a NI PXIe-1071 chassis, NI PXIe-8135 Controller, NI PXIe-5632 Vector Network Analyzer, a pair of gore cable and NI Automatic VNA Calibration Module 70 kHz until 9 GHz.

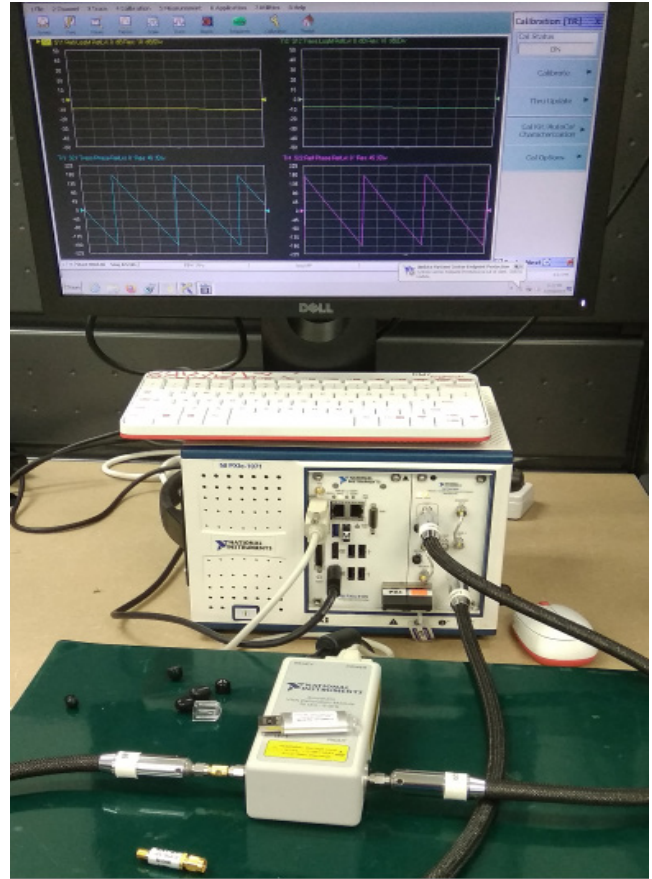


Figure 1 : A set up of PXIe-5632 Calibration

There are other variables that may contribute to the Type B uncertainty other than the 7 contributors mentioned in Table 1. The performance of the equipment would be degraded over time. In the product specification, some manufacturer may mention the aging or drift. It is up to engineers to either estimate the drift relative to the maximum or minimum specification limits or included as another Type B uncertainty component. One way of solving this problem is by calibrating the entire device shown in Figure 1. The manufacturer recommended one-year calibration interval for Vector Network Analyzer and Automatic Calibration Module. It is to ensure the absolute reading measured from this system is traceable to the International System of Units.

Once the calibration had been completed in Figure 1, this system is ready to measure a wide range of radio frequency devices such as antenna, amplifier, attenuator, splitter, filters and many more. In this paper, a 3 dB fixed attenuator manufacturer by Mini-Circuit with model VAT-3+ had been selected as shown by Figure 2.

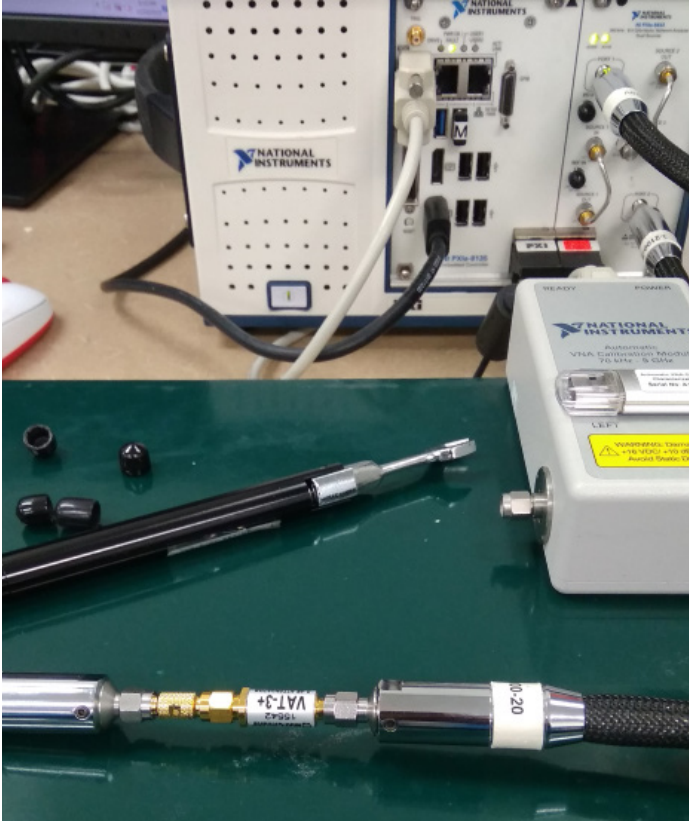


Figure 2: A 3 dB Fixed Attenuator under full two ports calibration

The 3 dB Attenuation $L(X)$ to be calibrated was obtain shown by equation 5.

$$L(X) = \delta Lm + \delta Lk + \delta Lr \quad (5)$$

Where:

δLm = Correction to mismatch loss

δLk = Correction for leakage for signal between input and output

δLr = Resolution of the network analyzer

The 3 dB Attenuation $L(X)$ will be measured in the Vector Network Analyzer from 300 kHz until 8.5 GHz are known as absolute reading. These absolute reading will associate with estimated measurement uncertainty at the end of this paper. In order to measure the 3 dB Attenuation, there are 3 factors contribute to the measurement. It is identified as correction to mismatch loss, correction for leakage for signal between input and output and resolution of the network analyzer.

Correction to mismatch loss define as the calibration system consists of NI PXIe-5632 Vector Network Analyzer, a pair of gore cable and NI Automatic VNA Calibration Module 70 kHz until 9 GHz to perform the characteristics impedance calibration. Where

the characteristics impedance calibration accuracy traceable to the NI Automatic VNA Calibration Module.

Correction for leakage for signal between input and output is define as the power dissipate along the transmission line from the Vector Network Analyzer port 2 to port 1 (forwards transmission S21) and port 1 to port 2 (reverse transmission S12).

Resolution of the network analyzer refer to least significant digit measured from the Vector Network Analyzer. The least significant digit refers to the lowest digit from the display.

3. Uncertainty Analysis for PXIe-5632 Network Analyzer

Uncertainty analysis for PXIe-5632 Vector Network Analyzer required to combine Type A and Type B contributors. The Type A uncertainty had been completed in previous study. Refer to Table 2, there are 7 Type B contributors and 1 Type A contributor in the uncertainty analysis. Each of the Type A and B contributors required to identify the distribution type. Distribution type segregate by Normal Distribution, U-shape, Triangle or Rectangular.

Table 2: Uncertainty component's unit, evaluation and distribution type

Uncertainty Component	Units	Evaluation Type	Distribution Type
Directivity	dB	B	U Shape
Source Match	dB	B	U Shape
Reflection Tracking	dB	B	Normal
Transmission Tracking	dB	B	Normal
Load Match	dB	B	U Shape
Power Step Resolution	dB	B	Rectangular
Trace Noise	dB	B	U Shape
Repeatability	dB	A	Normal

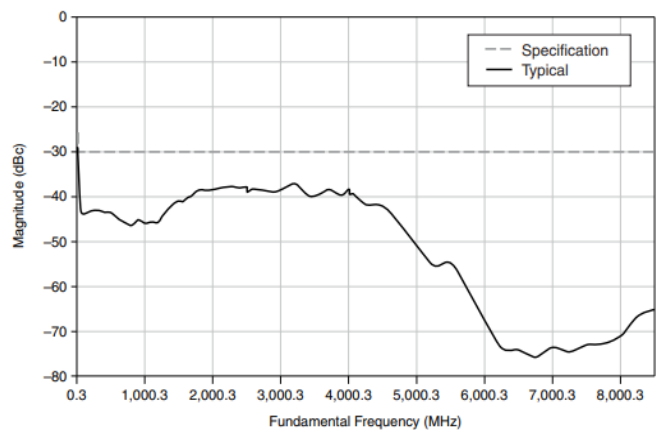


Figure 3 : PXIe-5632 Source Second Harmonics

Directivity, Source Match, Load Match and Trace Noise are identified as U-Shape distribution. The 4 contributors are mainly generated from the Vector Network Analyzer internal generators and receivers. It is systematic errors in the Vector Network Analyzer. An example PXIe-5632 generator second harmonic

recorded in the product specification [14] as shown by Figure 3. The trend of the source second harmonics does not populate in Normal Distribution, Rectangular or Triangle distribution. The source distribution type is populating close to U-Shape.

Reflection and Transmission Tracking were classified as normal distribution since uncertainty values recorded in dB were taken from the manufacturer specification. Element of repeatability uncertainty is normal distribution type. Table 3 will extend each of the uncertainty components calculation in divisor, standard uncertainty, degree of freedom and combine uncertainty.

Table 3: Uncertainty calculation from 300 kHz to <5 GHz

Uncertainty Component	Divisor	Standard Unc.	Degree of Freedom	Combine Unc.
Directivity	1.4142	0.0107	1000	0.0076
Source Match	1.4142	0.0195	1000	0.0138
Reflection Tracking	2	0.0200	1000	0.0100
Transmission Tracking	2	0.0700	1000	0.0350
Load Match	1.4142	0.0048	1000	0.0034
Power Step Resolution	3.4641	0.0100	1000	0.0029
Trace Noise	1.4142	0.0060	1000	0.0042
Repeatability	1	0.0176	4	0.0176
Combined Standard Uncertainty				0.0432
Effective Degree of Freedom				135.8664
Coverage Factor				1.96
Expanded Uncertainty				0.085
Unit				dB

For each variable, Table 3 calculates 4 contributors which were divisor, standard uncertainty, degree of freedom and combined uncertainty. Each part of the uncertainty was defined as the same element measured in dB. In this case, no conversion is necessary in order to convert to the same base unit. The sensitivity factor [15] for each contributor is equal to 1. U-Shape distribution divisor is 1.4142. The manufacturer reported standard uncertainty for reflection and transmission tracking are 0.02 dB and 0.07 dB respectively. Reflection and transmission tracking divisor are 2. Rectangular distribution divisor is 3.4641 [16].

The degree of freedom for the 7 Type B contributors are calculate as infinite. It is because the uncertainty contributors were based on the product specification. Which means the 7 Type B contributor are the worst-case [17] standard uncertainty. When a worst-case uncertainty contributor taken from the product specification, the values involved in the calculation was infinite.

Combined uncertainty [18] applied shown by equation 6.

$$U_L = \frac{\delta Us}{\delta Ud} \times \delta Uf \tag{6}$$

Where:

δUs = Standard uncertainty contributor

δUd = Distribution type of divisor

δUf = Sensitivity Coefficient

Combined Standard Uncertainty U(s) [19] applied shown by equation 7.

$$U(s) = \sqrt{\sum_{L=1}^N (U_L)^2} \tag{7}$$

Where:

U_L = Uncertainty Components

Effective Degree of Freedom V(eff) [20] applied shown by equation 8.

$$V(eff) = \frac{U(s)^4}{\sum_{j=1}^N \frac{\delta Us^4 \delta Ud^4}{Vf}} \tag{8}$$

Where:

U(s) = Combined Standard Uncertainty

$\delta U(s)$ = Standard uncertainty contributor

δUd = Distribution type of divisor

Vf = Degree of Freedom

Coverage factor refer to Effective Degree of Freedom in Table 3. It was calculated as 135.8664. Refer the value 135.8664 in Student T table. From Student T table, it was located between 100 and 1000 as shown by Figure 4. Choose the round up value at 95% confidence level as 1.962.

40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
Z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

Figure 4 : Student T Table

Expanded Uncertainty E(u) [21] applied shown by equation 9.

$$E(u) = Us \times Cf \tag{9}$$

Where:

U(s) = Combined Standard Uncertainty

Cf = Coverage Factor

From Table 3, it was concluded that the Expanded Uncertainty from 300 kHz until 5 GHz reported at 0.085 dB. The same methodology applied from 5 GHz to 8.5 GHz frequency range as shown by Table 4.

Coverage factor refer to Effective Degree of Freedom in Table 4. It was calculated as 115.2108. Check the calculated reading 115.2108 from Student T table. From the Student T table, it was located between 100 and 1000 as shown by Figure 5. Choose the round up value at 95% confident level as 1.962.

Table 4: Uncertainty calculation from 5 GHz to 8.5 GHz

Uncertainty Component	Divisor	Std Unc.	Degree of Freedom	Combine Unc.
Directivity	1.4142	0.0200	1000	0.0141
Source Match	1.4142	0.0195	1000	0.0138
Reflection Tracking	2	0.0200	1000	0.0100
Transmission Tracking	2	0.0700	1000	0.0350
Load Match	1.4142	0.0227	1000	0.0161
Power Step Resolution	3.4641	0.0100	1000	0.0029
Trace Noise	1.4142	0.0060	1000	0.0042
Repeatability	1	0.0212	4	0.0212
Combined Standard Uncertainty				0.0495
Effective Degree of Freedom				115.2108
Coverage Factor				1.96
Expanded Uncertainty				0.097
Unit				dB

40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
Z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

Figure 5 : Student T Table

From Table 4, it was concluded that the Expanded Uncertainty from 5 GHz until 8.5 GHz reported at 0.097 dB. The expanded uncertainty evaluation was reported respected to its frequency points as shown by Table 5. Forwards transmission (S21) and reverse transmission (S12) was the magnitude of absolute measurement associate with expanded uncertainty. The expended uncertainty should report in 2 decimal points.

Table 5: An Example of Calibration Report with Expanded Uncertainty

DUT Calibration Report				
Frequency (GHz)	S21	Uncertainty	S12	Uncertainty
0.0003	-2.939	0.084	-2.951	0.085
1	-3.092	0.084	-3.111	0.085
2	-3.154	0.084	-3.168	0.085
3	-3.189	0.084	-3.195	0.085
4	-3.29	0.084	-3.297	0.085
5	-3.239	0.084	-3.242	0.085
6	-3.379	0.093	-3.395	0.097
7	-3.408	0.093	-3.425	0.097
8	-3.497	0.093	-3.498	0.097
8.5	-3.553	0.093	-3.562	0.097

4. Discussion

In this paper the measurement uncertainty calculation was referred to the process flow of ISO / IEC Guide 98-1 to quantify the measurement uncertainty as Figure 6.

ISO/IEC GUIDE 98-1:2009(E)

JCGM 104:2009

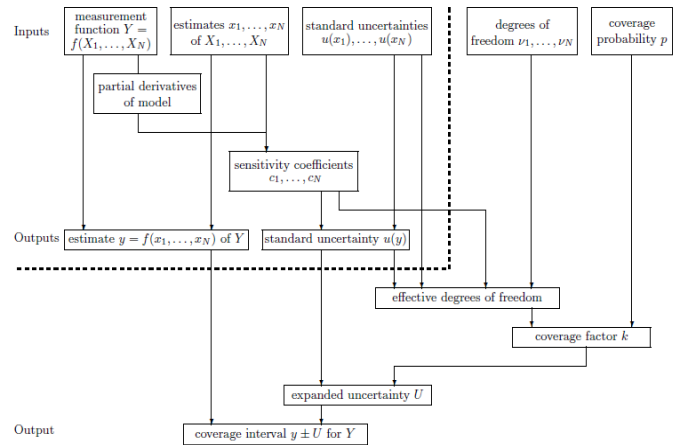


Figure 6 : ISO/IEC Guide 98-1 Uncertainty Calculation Process Flow

The flow diagram was divided into two major groups. It is input and output. In this paper, Tables 1 and 2 represent input. Whereas Tables 3, 4 and 5 reflect output. All tables describe the flows necessary to calculate the measurement uncertainty. There are total 6 systematic errors in reflection and transmission were found in the Vector Network Analyzer. It is identified as directivity, source match and frequency response reflection tracking for reflection. Transmission measurement generate another 3 systematic errors identified as isolation, load match and frequency response transmission tracking. The limitation from this paper to calculate the measurement uncertainty is without isolation taking into the consideration.

However, the vector Network Analyzer and Calibration Kits were used in the calculation of measurement uncertainty will undergo the process of validation to ensure whether the hardware meets the minimum specification as discuss in 4.1.

4.1. Vector Network Analyzer Directivity Validation by Using Anritsu Mechanical Verification Kits

The reading of Directivity and Source Match in Table 1 was released by the manufacturer as product specification. Validation is necessary to ensure the Vector Network Analyzer complies with the specification. To perform the validation, verification kits are expected using the airline, as shown in Table 6.

Table 6: Directivity and Source Match Experimental Test

Experimental	Directivity	Source Match
300 kHz to 5 GHz	Anritsu SC7953 Super Termination	Anritsu 23K50 Male Short
5 GHz to 8.5 GHz	Anritsu SC4808 20 dB Offset Termination and Anritsu SC7594 Airline	Anritsu 23KF50 Female Short and Anritsu SC7594 Airline

Directivity is load test and source match is short test. Both directivity and source match sharing the same Anritsu SC7594 airline in high frequency. Either the validation of the directivity or the source match is enough to justify the efficiency of the Vector Network Analyzer, and the calibration kits comply with the specification specified in Table 1. In this paper the validation of the directivity was chosen to compare against the product specification. This is because the vector network analyzer uses the directional couplers or bridges to make reflected measurements. When the Vector Network Analyzer calibrated with the calibration kits, a small amount of incident signal leakage through the directional couplers or bridges into the receiver. This incident happens on both ports 1 and 2 of the Vector Network Analyzer. The leakage signal will affect the accuracy and precision of the Vector Network Analyzer. The leakage signal in directivity validation could be validated using the SC7953 super termination and SC4808 20-dB offset termination attachment to Anritsu SC7594 airline as shown in Figure 7.



Figure 7 : Directivity Validation by Using Anritsu Verification Kit from 300 kHz to 5 GHz (left) and 5 GHz to 8.5 GHz (right)

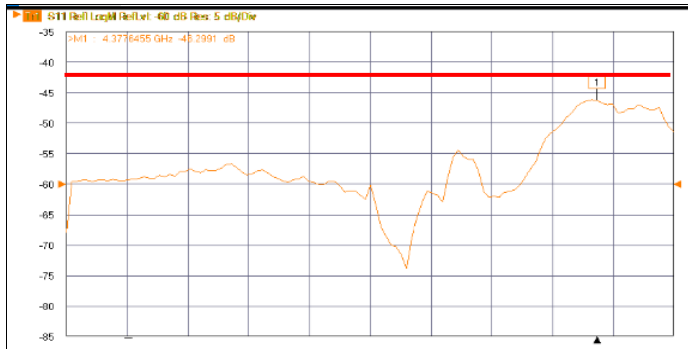


Figure 8 : PXIe-5632 Port 1 Directivity Validation Results Measured from 300 kHz to 5 GHz



Figure 9 : PXIe-5632 Port 2 Directivity Validation Results Measured from 300 kHz to 5 GHz

Once the Anritsu SC7953 super termination attach to the ports 1 and 2 at the Vector Network Analyzer, the SC7953 super termination will absorb the power transmitted from the Vector Network Analyzer generator. The results of this test shown by Figure 8 for port 1 and Figure 9 for port 2 of the Vector Network Analyzer respectively measured at -46.29 dB and -48.21 dB.

The maximum product specification is -42 dB from 300 kHz until 5 GHz. This experiment shows that the Vector Network Analyzer performed better than the specification required.

Figures 10 and 11 demonstrate the ports 1 and 2 directivity test measured from 5 GHz until 8.5 GHz. The maximum accepted ripples are 3.1dB. The Vector Network Analyzer was found within the product specification.

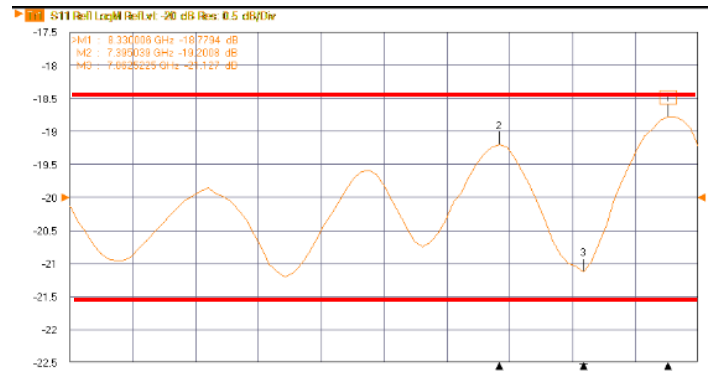


Figure 10 : PXIe-5632 Port 1 Directivity Validation Results Measured from 5 GHz to 8.5 GHz

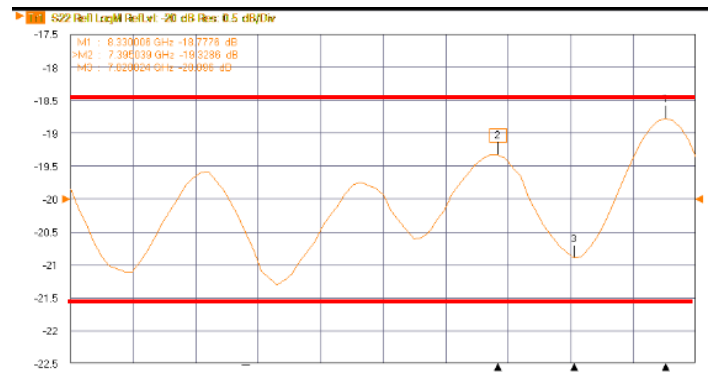


Figure 11 : PXIe-5632 Port 2 Directivity Validation Results Measured from 5 GHz to 8.5 GHz

In this validation experimental, the Vector Network Analyzer directivity measurement swept from 300 kHz to 8.5 GHz concludes the Vector Network Analyzer was performed better than the required specification. Since the directivity of the PXIe-5632 Vector Network Analyzer performed better than the warranted specification, directivity component listed in Table 1 is validated.

4.2. Vector Network Analyzer Mismatch Validation with 3 dB Fixed Attenuator Product Specification

The Anritsu Verification kit used to measure the Directivity and Source Match ripples. The ripples measured from the Vector Network Analyzer would be the Type B contributor respected to its frequency bandwidth. Meanwhile the Load Match from Table 1 was determined from below criteria as Table 7.

Table 7: Fixed Attenuator and VNA Mismatch Calculate in dB

Frequency	3 dB Attenuator	VNA	Mismatch (%)	dB
300 kHz to 5 GHz	1.15	1.016	0.1107	0.0048
5 GHz to 8.5 GHz	1.5	1.032	0.5256	0.0227

From Table 7, the voltage standing wave ratio for 3 dB attenuator measure at 1.15 and 1.5 were taken from the manufacturer specification. The voltage standing wave ratio for Vector Network Analyzer measured at 1.016 and 1.032 was conversion from the product specification Load Match 42 and 36 dB respectively. To calculate the Load Match from the Vector Network Analyzer together with the fixed attenuator VAT-3+ by applying the mismatch formula [22]. The calculated mismatch was in percentage. It is required to convert the mismatch in percentage to dB [23].

Table 2 used to determine the distribution type for each contributor. The main reason why reflection and transmission tracking did not category as U-Shape distribution type because from the product datasheet it was provided the info as uncertainty in typical. The uncertainty value reported in typical is identical to 2 sigma uncertainty. In order to calculate the Type B contributor for reflection and transmission tracking correctly, it is normal distribution type. If both reflection and transmission tracking identified as U-Shape distribution type, the expanded uncertainty will be increased as Tables 8 and 9. The uncertainty will be increased from 0.085 to 0.111 and 0.098 to 0.122. The expended uncertainty in Tables 8 and 9 could not classified incorrect expended uncertainty calculation because it is meet the ISO/IEC Guide 98-1 method. The only disadvantage for Tables 8 and 9 is the calculated expanded uncertainty approximately 30% increase from the initial calculation in Tables 3 and 4 respectively.

Table 8: Effect of Reflection and Transmission Distribution Type Changed from 300 kHz to 5 GHz

Uncertainty Component	Distribution Type	Divisor	Std Unc.	Combined
Directivity	U	1.4142	0.0107	0.0076
Source Match	U	1.4142	0.0195	0.0138
Reflection	U	1.4142	0.0200	0.0141
Transmission	U	1.4142	0.0700	0.0495
Load Match	U	1.4142	0.0048	0.0034
Power Step	Rectangular	3.4641	0.0100	0.0029
Trace Noise	U	1.4142	0.0060	0.0042
Repeatability	Normal	1	0.0176	0.0176
Combined Standard Uncertainty				0.0565
Effective Degree of Freedom				337.94
Coverage Factor				1.96
Expanded Uncertainty				0.111
Unit				dB

4.3. Vector Network Analyzer Reflection and Transmission Improvement

Distribution type identification for each contributor in table 2 is very important. We need to understand and identify the

distribution type for each contributor correctly in order to compute the lowest expended uncertainty. The lower expended uncertainty calculated, the better system it is.

Table 9: Effect of Reflection and Transmission Distribution Type Changed from 5 GHz to 8.5 GHz

Uncertainty Component	Distribution Type	Divisor	Std Unc.	Combined
Directivity	U	1.4142	0.0200	0.0141
Source Match	U	1.4142	0.0195	0.0138
Reflection	U	1.4142	0.0200	0.0141
Transmission	U	1.4142	0.0700	0.0495
Load Match	U	1.4142	0.0227	0.0161
Power Step	Rectangular	3.4641	0.0100	0.0029
Trace Noise	U	1.4142	0.0060	0.0042
Repeatability	Normal	1	0.0212	0.0212
Combined Standard Uncertainty				0.0614
Effective Degree of Freedom				251.7565
Coverage Factor				1.98
Expanded Uncertainty				0.122
Unit				dB

4.4. PXIe-5632 Vector Network Analyzer Limitation and Improvement for Future Work

A comparison against the calibration and measurement capability for transmission S21 with other manufacturer had been completed in Table 10. The purpose of the comparison is to measure the gap analysis between other Vector Network Analyzer manufacturers and this study.

Table 10: Transmission S21 Comparison With other Manufacturer

Transmission S21	Study	OEM1	OEM2
Calibration and Measurement Capability in dB	0.086 to 0.094	0.057	0.029 to 0.056
VNA system in use	NI PXIe-5632	HP 8753ES	MS 462xx

The calibration and measurement capability for OEM1 and OEM2 were downloaded from Accredited body website. The manufacturer published the best capability in Transmission S21 measurement. The differences between this OEM1 and this study is approximate 33.7% better in measurement capability. Meanwhile the gap analysis between OEM2 and this study was approximate 66.2% better in measurement. Both OEM1 and OEM2 delivered the better number in expanded uncertainty calculation because both OEM1 and OEM2 select the higher accuracy of the Vector Network Analyzer. To improve this study, we need to focus at the main contributors is Tables 3 and 4 in order to enhance the system.

From Tables 3 and 4, it was found that Transmission Tracking and Repeatability contribute more than half of the expended uncertainty calculation. To improve the Transmission Tracking uncertainty contributor, we need a calibrated 3 dB attenuator from National Measurement Laboratory becomes a reference attenuator to calibrate the NI PXIe-5632 Vector Network Analyzer. A calibrated 3 dB attenuator from United Kingdom National

Measurement Institute, National Physical Laboratory (United Kingdom Accreditation Service number 0478) could carry the reflection and transmission uncertainty as low as 0.01 dB and 0.0012 dB respectively. The known attenuation value calibrated from National Physical Laboratory could overwrite the NI PXIe-5632 Vector Network Analyzer product specification in reflection and transmission Tracking accuracy by 2 time and 58 times improved respectively calculated in Tables 11 and 12.

Table 11: Reflection and Transmission Tracking Uncertainty With Reference Attenuator from 300 kHz to 5 GHz

Uncertainty Component	Standard Uncertainty	Combined
Directivity	0.0107	0.0076
Source Match	0.0195	0.0138
Reflection	0.0100	0.0071
Transmission	0.0012	0.0008
Load Match	0.0048	0.0034
Power Step	0.0100	0.0029
Trace Noise	0.0060	0.0042
Repeatability	0.0176	0.0176
Combined Standard Uncertainty		0.0243
Effective Degree of Freedom		14.4105
Coverage Factor		1.96
Expanded Uncertainty		0.048
Unit		dB

Table 12: Reflection and Transmission Tracking Uncertainty With Reference Attenuator from 5 GHz to 8.5 GHz

Uncertainty Component	Standard Uncertainty	Combined
Directivity	0.0200	0.0141
Source Match	0.0195	0.0138
Reflection	0.0100	0.0071
Transmission	0.0012	0.0008
Load Match	0.0227	0.0161
Power Step	0.0100	0.0029
Trace Noise	0.0060	0.0042
Repeatability	0.0212	0.0212
Combined Standard Uncertainty		0.0343
Effective Degree of Freedom		27.2742
Coverage Factor		1.98
Expanded Uncertainty		0.068
Unit		dB

Nevertheless, the significant improvement in reflection and transmission tracking accuracy also will help to recalculate the Type A repeated uncertainty contributor to a smaller number in uncertainty are based on theory and for future study.

4.5. Summary of Discussion

Table 13 is the summary of improvement calibration technique applied the 3 dB reference attenuator into the NI PXIe-5632 Vector Network Analyzer calibration system. When this technique applied in the NI PXIe-5632 calibration system, the calibration and measurement capability performance is like other manufacturer Vector Network Analyzer and the investment is the lowest among the others.

Table 13: Summary of Improvement Calibration Technique NI PXIe-5632 With Reference Attenuator

Transmission S21	Study	OEM1	OEM2
Calibration and Measurement Capability in dB	0.086 to 0.094	0.057	0.029 to 0.056
VNA system in use	NI PXIe-5632 With 3dB Reference Attenuator	Agilent 8753ES/E5071C	Anritsu MS462x
Investment	Low - USD 20k	High - USD 50k	Medium- USD 35K

5. Conclusion

In this paper, an evaluation of measurement uncertainty for NI PXIe-5632 Vector Network Analyzer from 300 kHz to 8.5 GHz is calculated. This frequency bandwidth is widely used in telecommunications particularly in the calibration system for the 5 G sub-6GHz. This paper clarified the extended uncertainty measurement applied in the Vector Network Analyzer, valued to specific calibration techniques. A complete mathematical review shows the effects in each step involve changes to the contributor of uncertainty. This paper could help improve the scattering of parameters for commercial laboratories to extend the calibration potential in the new sector. The calibration technique of the 3 dB fixed attenuator also helps the commercial calibration laboratory explore existing measuring capacities and increase the degree of competence in the laboratory. In the discussion, it is explained different Vector Network Analyzer techniques applied in the NI PXIe-5632 could generate different expanded uncertainty in calculation. Based on the laboratory budget planning, the laboratory is free to choose which calibration technique would apply in the new scope. A part of the limitation of this paper is that isolation does not take into the estimation of measurement uncertainty calculation. It is because the manufacturer did not specify the specification for the isolation, but the systematic error of isolation could be established by validation or experimental. As future of this work, the first development is to increase the frequency from 8.5 GHz to 50 GHz. The second development is extending to Vector Network Analyzer Scattering Parameters port 1 (S11) and port 2 (S22). The calculation of the reflections S11 and S22 involved only port 1 or port 2. In addition, a measure of the test uncertainty ratio may also be added in future to satisfy the requirements of ANSI / NCSL Z540.

6. Acknowledgements

This work was supported in part by the Malaysian Ministry of Higher Education under the MyBrain15 program and Universiti Teknikal Malaysia Melaka.

References

- [1] National Instruments, Calibration Procedure PXIe-5632 8.5 GHz Vector Network Analyzer, National Instruments USA, 2018.

- [2] Agilent, Fundamentals of RF and Microwave Power Measurements (Part 1), Agilent Technology USA, 2003.
- [3] M. Andreas, Chandramouli, M. Baker, Devaki, "5G Evolution: A View on 5G Cellular Technology Beyond 3GPP Release 15," in IEEE Access PP(99):1-1, 2019 <https://doi.org/10.1109/ACCESS.2019.2939938>
- [4] Mini-Circuits, Fixed attenuators help minimize impedance mismatches, Mini-Circuit USA, 2015.
- [5] Metrology, ISO/IEC Guide 98-1 Uncertainty of Measurement, Switzerland: ISO, 2009.
- [6] G. G. Jing, "How to measure test repeatability when stability and constant variance are not observed" International Journal of Metrology and Quality Engineering, vol. 9, Article 10, 2018. <https://doi.org/10.1051/ijmqe/2018007>
- [7] Agilent, Fundamentals of RF and Microwave Power Measurements (Part 4), Agilent Technologies USA, 2006.
- [8] T. J. Roupael, Wireless Receiver Architectures and Design Antennas, RF, Synthesizers, Mixed Signal, and Digital Signal Processing, Academic Press, 2015
- [9] M.H. Tan, B.H Ahmad Yusairi, S. Mohd Rizal, "An Analysis for 2.4mm-2.4mm RF Connector Insertion Loss Measure From 45MHz Until 50GHz by Using Electronic Calibration Module and Mechanical Calibration kits in a Network Analyzer," in MIMT 2016, Cape Town South Africa. <https://doi.org/10.1051/mateconf/20165403006>
- [10] M.H. Tan, B.H Ahmad Yusairi, "Radio Frequency Connector Insertion Loss Measured From 300 kHz Until 8.5 GHz By Using Network Analyzer And Mechanical Calibration Kits," in Journal Of Engineering Science and Technology, 1587-1600, 2019. <https://doi.org/10.5281/zenodo.3590276>
- [11] D.A. Nix, A.S. Weigend, "Estimating the mean and variance of the target probability distribution," in ICNN'94, 374138, 1994. <https://doi.org/10.1109/ICNN.1994.374138>
- [12] X. Wan, W. Wang, J. Liu, T. Tong, "Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range," in BMC Medical Research Methodology, vol. 14, Article 135, 2014. <https://doi.org/10.1186/1471-2288-14-135>,
- [13] Agilent, Fundamentals of RF and Microwave Power Measurements (Part 2), Agilent Technologies USA, 2006.
- [14] National Instruments, Specifications PXIe-5632 8.5 GHz Vector Network Analyzer, National Instruments USA, 2018.
- [15] Mills et. Al, Study of Proposed Internet Congestion Control Mechanisms, NIST Special Publication 500-282 USA, 2010.
- [16] S. Bell, Beginner's Guide to Uncertainty of Measurement, National Physical Laboratory UK, 2001.
- [17] Agilent, Fundamentals of RF and Microwave Power Measurements (Part 3), Agilent Technologies USA, 2003.
- [18] S. Bell, Good Practice Guide No. 11 The Beginner's Guide to Uncertainty of Measurement, National Physical Laboratory Issue 2 UK, 2001.
- [19] Kuyatt, B. N. Taylor, E Chris, Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, National Institute of Standards and Technology USA, 1994.
- [20] G. L. Harris, Selected Laboratory and Measurement Practices and Procedures to Support Basic Mass Calibrations, National Institute of Standards and Technology USA, 2019.
- [21] D. Flack, Measurement Good Practice Guide No. 130 Co-ordinate measuring machine task-specific measurement uncertainties, National Physical Laboratory UK, 2013.
- [22] R. Lindsay, Understanding Measurement Uncertainty in Power Measurement, Anritsu Company USA, 2015.
- [23] Jones, G. Davis, Ralph, Sound Reinforcement Handbook Second Edition, Hal Leonard Corporation USA, 1989.

Parameter Estimation on Two-Dimensional Advection-Dispersion Model of Biological Oxygen Demand in Facultative Waste Water Stabilization Pond: Case Study at Sewon Wastewater Treatment Facility

Sunarsih Sunarsih*, Dwi Purwantoro Sasongko, Sutrisno Sutrisno

Department of Mathematics, Diponegoro University, 50275, Indonesia

ARTICLE INFO

Article history:

Received: 20 March, 2020

Accepted: 16 April, 2020

Online: 03 May, 2020

Keywords:

Advection-Dispersion

Parameter Estimation

Quadratic programming

Waste Water treatment

ABSTRACT

To build a precise mathematical model describing a natural phenomenon, parameter estimation is needed to achieve the best parameter value. In this paper, we have calculated the best parameter value for a two-dimensional advection-dispersion differential equation of the biological oxygen demand degradation process in a facultative wastewater stabilization pond. This research was conducted with case study data collected from Sewon, Bantul facultative wastewater treatment facility located in Yogyakarta, Indonesia. The method employed in this research is based on the least square value by minimizing the difference between the observed data and the simulated data via quadratic programming using the interior point algorithm. This method gave the best value for the parameters observed in the model i.e. dispersion constant and the flow rate velocity. From the results, we have achieved that the best value for dispersion constant is 0.25, the velocity of the flow rate in the x-direction is 0.1, and the velocity of the flow rate in the y-direction is 0.15 whereas the relative error of this parameter estimation was 11.5% that is acceptable.

1. Introduction

People who dispose of their domestic wastewater directly to the river are still found around the world. It has a high potentiality to increase pollution in the river. Then, the water quality in the river will be decreased. Hence, wastewater treatment is needed to reduce the pollutant in the wastewater before it is disposed of. Many countries have been developing wastewater treatment plants (WWTP) to reduce the negative impacts of the wastewater they produced. Commonly, a WWTP was built with components of an inlet section at the beginning of the process, facultative ponds, maturation ponds, and outlet section [1]. To optimize the pollutant reducing on facultative ponds, many researchers have been developed some mathematical models to analyze the physical, biological and chemical processes on these ponds during the treatment. In the biological process, the organic material is decreased during a natural process which utilizes the bacteria and algae in the wastewater [2]–[4].

In mathematical theories, many mathematical models are useful to analyze some phenomena which will produce some results to be used for mitigation, optimization, evaluation, etc. One of the most useful mathematical models is the partial differential equation. For example, an elliptic partial differential

equation was used to solve a production planning problem [5], a partial differential equation model was formulated for infrared image enhancement [6], and a mathematical model was applied for inflammatory edema formation [7]. The more special mathematical model in a class of partial differential equations is an equation for the advection-dispersion phenomenon. There were many pieces of researches conducted to solve and to apply the advection-dispersion model. For examples, an advection-dispersion rule was applied for analyzing of transport of leaking CO₂-saturated brine along a fractured zone [8], a fractional advection-dispersion model was applied for hillslope tracer analysis, and a 3D advection-dispersion model was used to analyze the distribution of dissolved oxygen in a facultative pond [9]. The analytical solution of a partial differential equation is commonly not easy to find. Therefore, many researchers are more prefer to the numerical method to solve. Special for the advection-dispersion model, some numerical methods were developed such as random lattice Boltzmann method [10], Haar wavelets coupled with finite differences [11], unified transform/Fokas method [12], a numerical method based on Legendre scaling functions [13] and many more. Each of these methods had some superiority and weakness. Furthermore, a mathematical model contains some parameters in the equation. To determine the value of these parameters, parameter estimation is needed to be performed.

*Corresponding Author: Sunarsih, narsih_pdl@yahoo.com

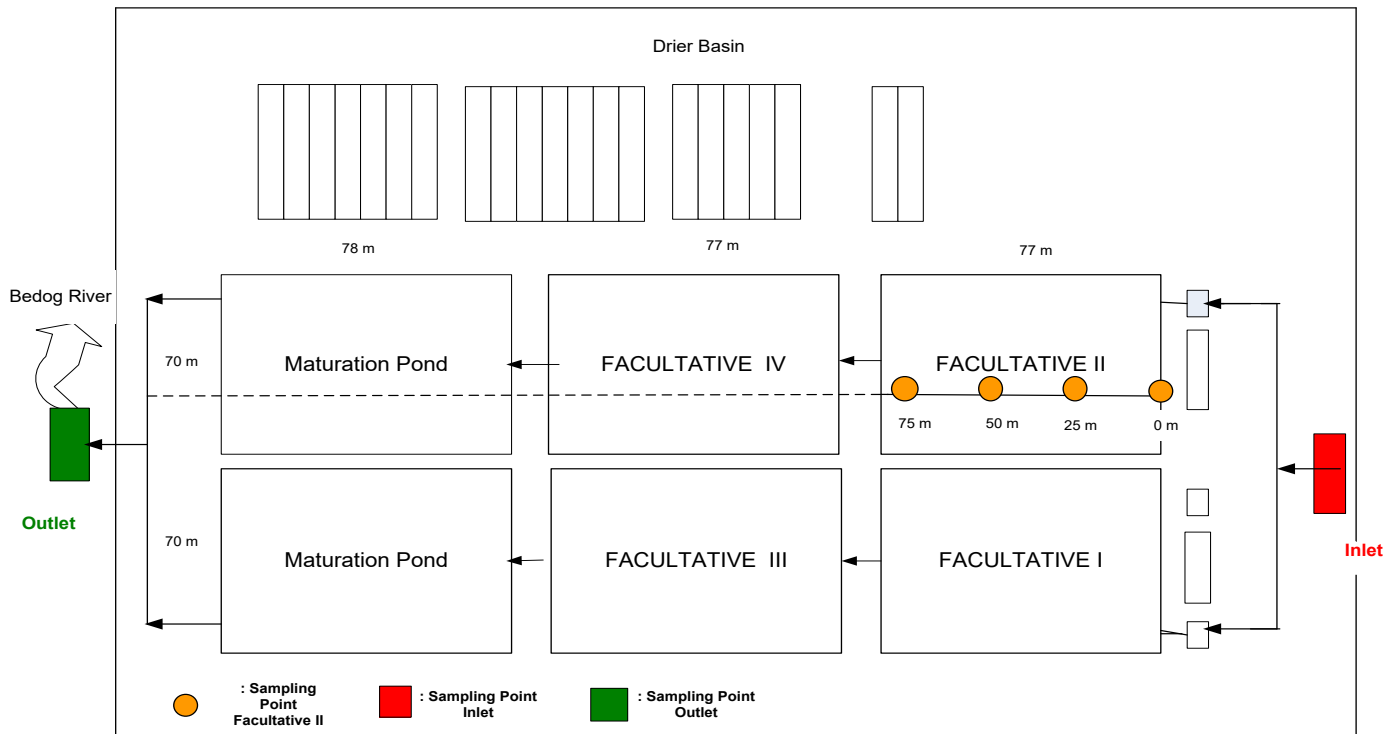


Figure 1: Sewon Bantul wastewater treatment plant [19]

Special for a partial differential equation, to be considered well to observe the modeled phenomenon, a parameter estimation process is needed to obtain the best parameter value to represent the phenomenon. There are many methods to be able to perform parameter estimation from classic methods like least square to some new methods like multi-parametric programming [14] and novel mixed artificial neural network [15]. Many researchers have been conducted some researches for parameter estimation process which have been applied for many problems like Ornstein–Uhlenbeck process [16], plasmonic QED problem [17], and gravitational waves problem [18].

In this research, we conduct a study of parameter estimation on a BOD degradation mathematical model for a two-dimensional differential equation based on the advection-dispersion model for a facultative wastewater stabilization pond. To obtain the best parameter, we use the BOD observation data taken in the Sewon wastewater stabilization pond located in Bantul, Yogyakarta, Indonesia.

2. Material and Method

The methodology used in this research is explained as follows. First, we explain the WWTP that we observed. The sample data of the BOD concentration used in the parameter estimation process were observed in Sewon WWTP located in Bantul, Yogyakarta province, Indonesia. The scheme of this WWP is illustrated in Figure 1.

This WWTP’s layout can be separated into three main parts that are inlet section, facultative ponds section, maturation ponds section, and an outlet section. These facultative ponds are connected as illustrated in the figure where the size of each facultative pond is around $(77 \times 70 \times 4) \text{ m}^3$. The sampling point of this research is explained in Figure 2. The used method to test

the BOD concentration in the wastewater sample is APHA.AWWWA.WEF 5190 B-2012.

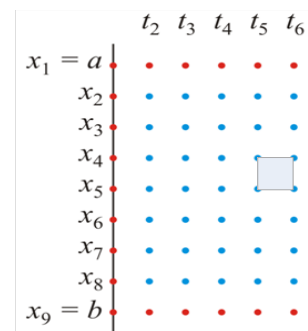


Figure 2: BOD sampling points

2.1. Assumptions

The governed differential equation in this research is working under the following assumptions:

- The observation data were collected in Sewon WWTP in the sunny season. We assume the achieved parameter is suitable for the sunny season.
- The governed differential equation contains only two dimensions i.e. x -direction and y -direction. We omit the depth parameter and assume the BOD concentration is uniformly distributed in the depth dimension.
- Dispersion constant in the x and y direction are assumed to be identical in the horizontal direction.
- We observe only the BOD degradation in the facultative ponds. We omit the treatment on the maturation pond.

The governed differential equation used the following symbols:

- i : index of sampling point coordinate in the x -direction

- j : index of sampling point coordinate in the y -direction
- x : sampling point coordinate in the x -direction
- y : sampling point coordinate in the y -direction
- $D_{x,y}$: dispersion constant in the x or y direction
- $C_{i,j}^n$: BOD concentration on sampling point (i,j) at observation time instant n (ML^{-3})
- Δx : step length of observation sampling point in the x -direction
- Δy : step length of observation sampling point in the y -direction
- Δt : step length step of observation time instant
- $\hat{C}_{i,j}^n$: observation data value of BOD concentration at sampling point (i,j) at observation time instant n (ML^{-3})
- u : velocity of the flow rate of the wastewater in the x -axis direction (LT^{-1})
- v : velocity of the flow rate of the wastewater in the y -axis direction (LT^{-1}).

$$\begin{aligned} \frac{C_{i,j}^{n+1}}{\Delta t} = & \left[-\frac{2D_h}{(\Delta x)^2} - \frac{2D_h}{(\Delta y)^2} + \frac{1}{\Delta t} \right] C_{i,j}^n \\ & + \frac{D_h}{(\Delta x)^2} C_{i+1,j}^n + \frac{D_h}{(\Delta x)^2} C_{i-1,j}^n \\ & + \frac{D_h}{(\Delta y)^2} C_{i,j+1}^n + \frac{D_h}{(\Delta y)^2} C_{i,j-1}^n \\ & - \frac{u}{2\Delta x} C_{i+1,j}^{n+1} + \frac{u}{2\Delta x} C_{i-1,j}^n \\ & - \frac{v}{2\Delta y} C_{i,j+1}^{n+1} + \frac{v}{2\Delta y} C_{i,j-1}^n. \end{aligned} \quad (2)$$

The initial value of the model is obtained by sampling with the following method. Firstly, by using the observation BOD concentration data shown in Table 1, we have calculated the curve fitting process to obtain a BOD concentration function of the grid based on these data.

2.2. Mathematical Model

The BOD degradation process can be illustrated in Figure 3. The velocity of the flow rate of the wastewater in the x -axis direction is u (LT^{-1}) whereas notation v is the velocity of the flow in the y -direction (LT^{-1}).

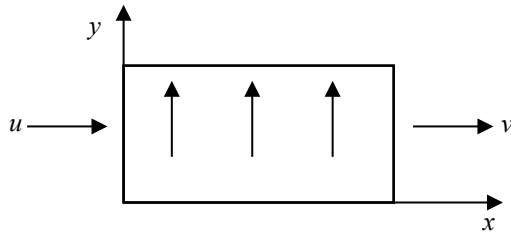


Figure 3: The BOD degradation process

Let $D_{x,y}$ be the dispersion constant in the x or y direction (L^2T^{-1}), the parameter estimation is conducted for the following BOD degradation mathematical model in the two-dimensional equation based on the advection-dispersion process [20]:

$$\frac{\partial C}{\partial t} = -\frac{\partial(uC)}{\partial x} - \frac{\partial(vC)}{\partial y} + D_x \frac{\partial^2 C}{\partial x^2} + D_y \frac{\partial^2 C}{\partial y^2} \quad (1)$$

where C is the BOD concentration (ML^{-3}). We assume that $D_x = D_y = D_h$ which means that the dispersion constant in the x and y direction are identical as the horizontal direction denoted by D_h . By applying the finite difference method to solve the model numerically, we have the following solution

$$\begin{aligned} \frac{C_{i,j}^{n+1} - C_{i,j}^n}{\Delta t} + u \left[\frac{C_{i+1,j}^{n+1} - C_{i-1,j}^n}{2\Delta x} \right] + v \left[\frac{C_{i,j+1}^{n+1} - C_{i,j-1}^n}{2\Delta y} \right] \\ = D_h \left[\frac{C_{i+1,j}^n - 2C_{i,j}^n + C_{i-1,j}^n}{(\Delta x)^2} \right] + D_h \left[\frac{C_{i,j+1}^n - 2C_{i,j}^n + C_{i,j-1}^n}{(\Delta y)^2} \right]. \end{aligned}$$

By simple algebraic manipulation, we have the following solution

Table 1: BOD Concentration Data for initial condition curve fitting

Sampling point	y							
	1	2	3	4	5	6	7	
x	1	64.06	66.12	41.79	34.92	64.06	66.12	41.79
	2	34.92	64.06	66.12	41.79	34.92	37.44	40.64
	3	36.51	39.43	63.85	37.87	49.68	69.84	64.06
	4	66.12	41.79	34.92	37.44	40.64	36.51	39.43
	5	63.85	37.87	49.68	69.84	64.06	66.12	41.79
	6	34.92	37.44	40.64	36.51	39.43	63.85	37.87
	7	49.68	69.84	69.78	55.35	69.86	42.68	42.68

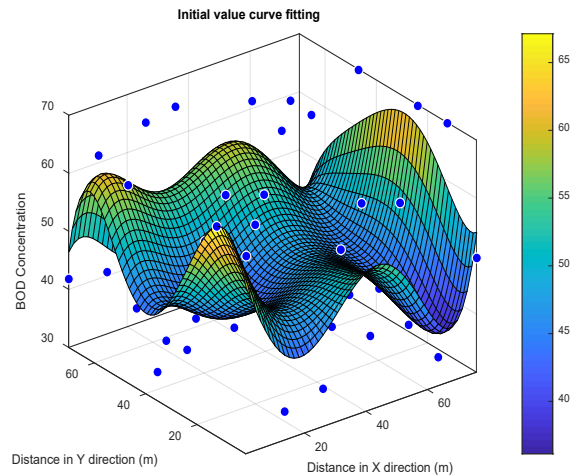


Figure 4: Initial value curve fitting

$\Delta x = 0.077$, $\Delta y = 0.07$, curve fitting is resulting the following polynomial of degree 5:

$$\begin{aligned} f(x,y) = & 58.11 - 1.454x + 2.056y - 0.007594x^2 \\ & - 0.01097xy - 0.1163y^2 + 0.003586x^3 \\ & - 0.003701x^2y + 0.004778xy^2 + 0.0002471y^3 \\ & - 8.68 \cdot 10^{-5}x^4 + 6.926 \cdot 10^{-5}x^3y \\ & + 4.355 \cdot 10^{-5}x^2y^2 - 9.643 \cdot 10^{-5}xy^3 \end{aligned}$$

$$\begin{aligned}
 &+4.816 \cdot 10^{-5} y^4 + 5.808 \cdot 10^{-7} x^5 \\
 &-3.311 \cdot 10^{-7} x^4 y - 2.785 \cdot 10^{-7} x^3 y^2 \\
 &+2.037 \cdot 10^{-7} x^2 y^3 + 5.489 \cdot 10^{-7} xy^4 \\
 &-4.927 \cdot 10^{-7} y^5
 \end{aligned}$$

which is illustrated in Figure 4. The derived initial value curve $f(x, y)$ is then used as the initial value of the BOD concentration for all sampling points in the parameter estimation calculation.

3. Results and Discussions

Let $\hat{C}_{i,j}^n$ denotes the observation data value of BOD concentration at sampling point (i, j) at observation time instant n . The governed optimization is formulated as

$$\min \sum_i \sum_j \sum_n \left(\hat{C}_{i,j}^n - C_{i,j}^n \right)^2 \tag{3}$$

subject to:

$$\begin{aligned}
 \frac{C_{i,j}^{n+1}}{\Delta t} = & \left[-\frac{2D_h}{(\Delta x)^2} - \frac{2D_h}{(\Delta y)^2} + \frac{1}{\Delta t} \right] C_{i,j}^n + \frac{D_h}{(\Delta x)^2} C_{i+1,j}^{n+1} \\
 & + \frac{D_h}{(\Delta x)^2} C_{i-1,j}^n + \frac{D_h}{(\Delta y)^2} C_{i,j+1}^n + \frac{D_h}{(\Delta y)^2} C_{i,j-1}^n \\
 & - \frac{u}{2\Delta x} C_{i+1,j}^{n+1} + \frac{u}{2\Delta x} C_{i-1,j}^n - \frac{v}{2\Delta y} C_{i,j+1}^{n+1} + \frac{v}{2\Delta y} C_{i,j-1}^n.
 \end{aligned}$$

This mathematical model is explained as follows. The objective is find the best value for the parameters that will be estimated i.e. the flow rate in the x -direction & y -direction, and the dispersion constant. The term “best” means that we want to find the value for the parameters so that the simulated data derived from (2) will be closest to the observation data. This can be achieved by adopting the least square theory i.e. minimizing the difference between the simulated data and the observation data of the BOD concentration value for all time instants (observation time) and for all sampling points in the quadratic form as formulated in (3) subject to the governed equation of the BOD concentration.

Table 2: BOD concentration (mg/L) data for parameter estimation at time sampling 2 PM

Sampling point	y				
	1	2	3	4	
x	1	37.44	40.64	36.51	39.43
	2	38.32	44.24	38.38	50.97
	3	46.10	23.32	47.06	53.86
	4	55.79	47.05	23.98	64.44
	5	51.02	36.40	35.56	54.71

Table 3: BOD Concentration (mg/L) data for parameter estimation at time sampling 7 PM

Sampling point	y				
	1	2	3	4	
x	1	63.85	37.87	49.68	69.84
	2	48.33	33.84	66.78	76.40
	3	49.25	26.05	43.79	42.56
	4	50.28	47.43	40.66	49.90
	5	87.05	55.08	51.34	52.44

By using the MATLAB programming language and interior-point algorithm solver, we run the parameter estimation by using the observation data. The results are the value for parameters the flow rate, dispersion coefficient in the x -direction, and dispersion coefficient in the y -direction.

Tables 2 and 3 show the BOD concentration observation data at sampling time 2 PM and 7 PM that we used to perform the parameter estimation. The optimization process or solving (3) was performed in MATLAB R2017b on a daily used personal computer with 3.2 GHz of processor, 8 GB of memory and Windows 10 of the operating system. From the solution for (3), the parameter value for dispersion constant is 0.25 and the velocity of the flow rate is 0.1 for the x -direction and 0.15 for the y -direction. These values can be used in the model so that the model will give the best fit to the data with a relative error of 11.5% of the simulated data from the observed data. We argue that this relative error is acceptable which means that these results are implementable. These parameter estimation results, then, can be used to estimate the dynamics of the BOD concentration in the corresponding facultative domestic wastewater ponds via the governed differential equation (1).

According to the fact that the more the observation data are used, the better the estimated parameters' value are achieved. This means that derived parameters' value in this research might be improved by the policy maker if more observation data are available to compute although it will be costly. Moreover, time of sampling is also affecting the estimated parameter. Then, the time for collecting the observation data would be better with different conditions such as various weather. Hence, the policy maker is suggested to collect the data with multiple observations so that the estimated parameters would be better.

4. Concluding Remarks

We have considered the parameter estimation for a two-dimensional advection-dispersion mathematical model using BOD concentration data at Sewon wastewater facultative ponds located in Bantul, Yogyakarta, Indonesia. The parameters that we estimated are the flow rate in the two directions (x -direction and y -direction) and the dispersion constant. The parameter estimation process was modeled as a quadratic optimization problem in which the objective is finding the best parameter for the model and observation data using the least square scheme. We have achieved the best parameter for them with a relative error of 11.5%.

In our future works, we will develop the model for the three-dimensional case so the depth of the water will be included in the model. Furthermore, the other models for other biological/chemical processes such as chemical oxygen demand (COD), dissolved oxygen (DO), plankton, and sediment analysis are interesting to study.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research is supported by DRPM KEMENRISTEKDIKTI INDONESIA under PDUPT research grant 2018.

References

- [1] S. Sunarsih, P. Purwanto, and W. S. Budi, "Mathematical modeling regime steady state for domestic Wastewater Treatment facultative stabilization ponds," *J. Urban Environ. Eng.*, vol. 7, no. 2, pp. 293–301, 2013, doi: 10.4090/juee.2013.v7n2.293301.
- [2] S. Kayombo, T. S. A. Mbvette, A. W. Mayo, J. H. Y. Katima, and S. E. Jorgensen, "Diurnal Cycles of Variation Physical-Chemical Parameters in Waste Stabilization Ponds," *Ecol. Eng.*, vol. 18, no. 1, pp. 287–291, 2002.
- [3] Beran and Kargi, "A dynamic mathematical model for waste water stabilization ponds," *Ecol. Modell.*, vol. 181, no. 1, pp. 39–57, 2005, doi: 10.1016/j.ecolmodel.2004.06.022
- [4] L. Puspita, R. E., and S. I. N.N., "Lahan Basah Buatan di Indonesia. Wetlands International Indonesia Programme," 2005.
- [5] D. Covei and T. A. Pirvu, "An elliptic partial differential equation and its application," *Appl. Math. Lett.*, vol. 101, p. 106059, 2020, doi: 10.1016/j.aml.2019.106059.
- [6] D. N. Liu, R. Hou, W. Z. Wu, J. W. Hua, X. Y. Wang, and B. Pang, "Research on infrared image enhancement and segmentation of power equipment based on partial differential equation," *J. Vis. Commun. Image Represent.*, vol. 64, p. 102610, 2019, doi: 10.1016/j.jvcir.2019.102610.
- [7] R. F. Reis, R. W. dos Santos, B. M. Rocha, and M. Lobosco, "On the mathematical modeling of inflammatory edema formation," *Comput. Math. with Appl.*, 2019, doi: 10.1016/j.camwa.2019.03.058.
- [8] N. Ahmad, A. Wörman, X. Sanchez-Vila, and A. Bottacin-Busolin, "The role of advection and dispersion in the rock matrix on the transport of leaking CO₂-saturated brine along a fractured zone," *Adv. Water Resour.*, vol. 98, pp. 132–146, 2016, doi: 10.1016/j.advwatres.2016.10.006.
- [9] Sunarsih, D. P. Sasongko, and Sutrisno, "Numerical Solution of a 3-D Advection-Dispersion Model for Dissolved Oxygen Distribution in Facultative Ponds," *E3S Web Conf.*, vol. 31, 2018, doi: 10.1051/e3sconf/20183103006
- [10] A. A. Hekmatzadeh, A. Adel, F. Zarei, and A. Torabi Haghghi, "Probabilistic simulation of advection-reaction-dispersion equation using random lattice Boltzmann method," *Int. J. Heat Mass Transf.*, vol. 144, p. 118647, 2019, doi: 10.1016/j.ijheatmasstransfer.2019.118647.
- [11] S. Haq, A. Ghafoor, and M. Hussain, "Numerical solutions of variable order time fractional (1+1)- and (1+2)-dimensional advection dispersion and diffusion models," *Appl. Math. Comput.*, vol. 360, pp. 107–121, 2019, doi: 10.1016/j.amc.2019.04.085.
- [12] F. P. J. de Barros, M. J. Colbrook, and A. S. Fokas, "A hybrid analytical-numerical method for solving advection-dispersion problems on a half-line," *Int. J. Heat Mass Transf.*, vol. 139, pp. 482–491, 2019, doi: 10.1016/j.ijheatmasstransfer.2019.05.018.
- [13] H. Singh, R. K. Pandey, J. Singh, and M. P. Tripathi, "A reliable numerical algorithm for fractional advection–dispersion equation arising in contaminant transport through porous media," *Phys. A Stat. Mech. its Appl.*, vol. 527, p. 121077, 2019, doi: 10.1016/j.physa.2019.121077.
- [14] E. Che Mid and V. Dua, "Parameter estimation using multiparametric programming for implicit Euler's method based discretization," *Chem. Eng. Res. Des.*, vol. 142, pp. 62–77, 2019, doi: 10.1016/j.cherd.2018.11.032.
- [15] X. Hou, J. Yuan, C. Ma, and C. Sun, "Parameter estimations of uncooperative space targets using novel mixed artificial neural network," *Neurocomputing*, 2019, doi: 10.1016/j.neucom.2019.02.038.
- [16] Y. A. Kutoyants, "On parameter estimation of the hidden Ornstein–Uhlenbeck process," *J. Multivar. Anal.*, vol. 169, pp. 248–263, 2019, doi: 10.1016/j.jmva.2018.09.008.
- [17] H. R. Jahromi, "Parameter estimation in plasmonic QED," *Opt. Commun.*, vol. 411, no. August 2017, pp. 119–125, 2018, doi: 10.1016/j.optcom.2017.11.020.
- [18] A. Blaut, "Parameter estimation accuracies of Galactic binaries with eLISA," *Astropart. Phys.*, vol. 101, pp. 17–26, 2018, doi: 10.1016/j.astropartphys.2018.04.001.
- [19] Sunarsih, Widowati, Kartono, and Sutrisno, "Mathematical Analysis for the Optimization of Wastewater Treatment Systems in Facultative Pond Indicator Organic Matter," *E3S Web Conf.*, vol. 31, no. 05008, pp. 1–3, 2018, doi: 10.1051/e3sconf/20183105008
- [20] Sunarsih, S. Dwi P, and Sutrisno, "Mathematical Model of Biological Oxygen Demand in Facultative Wastewater Stabilization Pond Based on Two-Dimensional Advection-Dispersion Model," *Am. J. Eng. Res.*, vol. 5, no. 11, pp. 1–5, 2016.

A Novel Hybrid Method for Segmentation and Analysis of Brain MRI for Tumor Diagnosis

Kapil Kumar Gupta^{*1,2}, Namrata Dhanda¹, Upendra Kumar³

¹Department of Computer Science and Engineering, Amity University, Lucknow, India

²Department of Computer Science and Engineering, Shri Ramswaroop Memorial University, Barabanki, India

³Department of Computer Science and Engineering, Institute of Engineering and Technology, Lucknow, India

ARTICLE INFO

Article history:

Received: 01 January, 2020

Accepted: 11 April, 2020

Online: 03 May, 2020

Keywords:

Brain Tumor

Image Segmentation

Fuzzy C-Mean Clustering

Adaptive mean and variance based filter

Noise Removal

ABSTRACT

It is difficult to accurately segment brain MRI in the complex structures of brain tumors, blurred borders, and external variables such as noise. Much research in developing as well as developed countries show that the number of individuals suffering tumor of the brain has died as a result of the inaccurate diagnosis. The proposed article, a novel hybrid method improves segmentation accuracy. The proposed research includes three basic steps. In the first step, the adaptive filter based on mean and local variance is utilized for noise removal in the input images. It helps in de-noising to a different orientation and scale, creates numerous responses for all components in the medical images while preserving the edges. In the second step, the development of a hybrid method takes place. It is the combination of extended K-mean clustering and fuzzy C-mean clustering. The purpose of the research is to develop a hybrid segmentation structure of single-channel T1 MR Images for multiform benign and malignant tumors. It removes the limitation of prefixed cluster size which helps in improving the segmentation accuracy by reducing the sensitivity of the clustering parameters. In the third step, the morphological non-linear operation performed for the removal of the non-tumor part. The proposed approach is evaluated against various statistical parameters such as mean, standard deviation, entropy, correlation, homogeneity, smoothness and variance. The parameters result predicts a greater balance between the automated tumor areas extracted by radiologists with the tumor areas extracted by the proposed method. The findings show that the proposed hybrid method achieves a 98% level of segmentation accuracy.

1. Introduction

Advanced 3-D medical imaging techniques produce highly efficient and reliable images [1]. Segmentation of the medical image is the method of automatically or semi-automatically detecting boundaries within a 2D or 3D image. A significant challenge in segmenting medical images is the increased variation of medical images. First and foremost, there are significant modes of variation in human anatomy. Furthermore, a number of different modalities like X-ray, CT, MRI, Microscopy, PET, SPECT, Endoscopy and OCT are used in the processing of medical images [2]. Normal cells develop regulated in the human brain, whereas new cells replace damaged cells, tumor cells produce uncontrolled in the brain, which is not yet understood. Benign or malignant can

be primary brain tumors. Brain tumors are heterogeneous and extremely variable in size, place, shape, and appearance. Automatic benign and malignant brain tumor segmentation is a difficult task. Some MR imaging artifacts also increase the complexity level in the tumor segmentation [3]. The slow growth of a benign brain tumor is certain and rarely diffuses, while its cells are not malignant. Figure 1 display below containing different types of benign MRI tumor images.

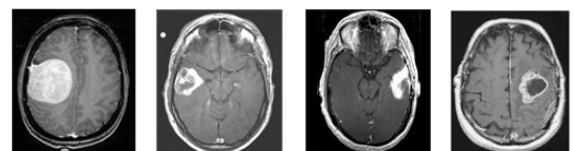


Figure 1: Brain MRI images containing different types of benign brain tumor

*Kapil Kumar Gupta, Amity University, Shri Ramswaroop Memorial University, kapilkumargupta2007@gmail.com

Figure 2 display containing different types of malignant MRI tumor images.

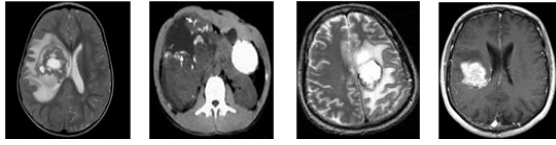


Figure 2: Brain MRI images containing different types of malignant brain tumor

The tumor, causing normal tissue damage, compression and elevated intra-cranial pressures can have an effect on the brain [4]. Symptoms like headache, dizziness and visibility problems can vary by type, size, and location of the tumor in the brain.

In this paper, Image segmentation relates to pixel classification or grouping, so that each class or group represents an object that is viewed differently. Different characteristics are used for this purpose to discriminate between one item and another [5]. Texture, boundaries, edges, and color are some of the characteristics most commonly used to differentiate between different objects [6, 7]. Edge is a significant image feature, so edge detection plays an important role in the segmentation. A challenging issue for edge detection of digital medical images is noise. De-noising images is a classic yet active subject as it is an essential step in many applications. Since the pixel characteristics are discontinuous, there will be comparatively evident variations between the pixel characteristics on both sides of the edges. Therefore, the fundamental concept of the edge-based segmentation algorithm is to use some techniques to discover the limits. On one side of the corners the pixels are divided into one sub-image and on the other side, the pixels are considered to belong to another sub-image. It is sensitive to noise and generally gets incomplete data [8]. Linear filters that transform the image into a constant matrix, in the presence of additive noise, will generate the image that is blurred with the poor position of features and incomplete noise reduction [9]. A lot of research has taken into consideration and conclude that the adaptive mean and variance based filter is better than other filters for de-noising medical images.

Clustering techniques are used to distinguish groups of similar objects in a multivariate data set gathered from areas such as Biomedical and Geospatial etc.. The Fuzzy C- Mean (FCM) clustering technique applies to a broad range of data in the assessment of biomedical information. For any set of numerical information, this program produces fuzzy partitions and prototypes. These partitions are helpful to support recognized substructures or to suggest substructure in unexplored information. A generalized minimum-square objective function is the clustering criterion used to aggregate subsets [10]. Extended K-means algorithms are used to separate the area of concern from the context [11].

For the assessment of Brain Tumor, the statistical parameters are implemented. Under the average data technique, statistical and mathematical parameters such as Entropy, Standard Deviation, Mean and Variance are introduced [12]. For the number of iterations, the statistical range of each parameter is calculated. The individual statistical parameter assessment is performed with its impact on MR images of the brain tumors. Factors such as intensity, artifact presence and similarities in gray-level images are also taken into account in the segmentation process.

2. Background Study

Segmentation of the magnetic resonance images relates to the pixel clusters. Each cluster represents an object that is viewed differently. Distinct characteristics are used to discriminate between one object to another. Medical image processing is always a focal point for researchers. The challenge is to predict the accurate areas of brain tumors within the given time frame. Many studies suggested in the field of brain tumor identification have been released over the past 30 years. Research has grown exponentially over the past decade. An article was written by A. Kharrat et. al. on the brain tumor recognition using the Wavelet Transform segmentation approach to break up images of MRI. In this proposed method k-means are implemented to remove the locations or tumors that are suspect. The wavelet transformation functions $\Psi_{x,y}(s)$ called the wavelet family in equation 1. Equation 2 evaluates many wavelets, which are derived from the scaling of a solitary function $\Psi(s)$ called the mother wavelet [13].

$$B_{x,y} = \int_{-\infty}^{\infty} a(s) \Psi_{x,y}(s) ds \quad (1)$$

$$\Psi_{x,y}(s) = \frac{1}{\sqrt{x}} \Psi\left(\frac{s-y}{x}\right) \text{ with } x \neq 0 \quad (2)$$

The method displays images at different levels of resolution. The wavelet test also allows the images to be compact or to de-noise without any noticeable degradation. In this study, researchers achieved 21.6272 PSNR value and 17.7768 MSAD value. R. Ratan and. al. suggested a detection method for brain tumors based on multi-parameter analyzes of MR images. Segmentation of the Watershed was implemented in this process. By deciphering a force picture's inclination guide as a tallness appreciation, we get lines that seem to be edged by all accounts. The journey from the dividing lines to the corresponding catchment bowl would be discovered if the forms were a precipitation dropper. These lines of isolation are called watersheds. The shift in the watershed can be created on a grayscale by flooding method. It was used to detect brain tumors in 2D and 3D medical images [14]. A properly organized quick diagnostic of brain tumor technique was suggested by Qurat-Ul-Ain et al. In the first stage extraction of texture, features takes place while the second stage consists of the classification of brain pictures on the basis of texture feature using the hybrid base classifier. The extracted tumor region is categorized as malignant by using two-phase segmentation processes. The segmentation process in this technique involves stages of removal of the skull and extraction of tumors. The researchers achieved 99% classification accuracy [15]. M. U. Akram et al. suggested an automatic diagnostic system technique for MR images. In this approach, segmentation and fragmentation of the images are obtained in three phases. In the first phase, preprocessing is performed to evacuate the blurring and to sharpen the picture. In the second phase, thresholding based segmentation takes place and in the third phase, fragmented images are processed with morphology, and tumor cover is carried out to remove the falsified pixels. The authors achieved an average of 97% segmentation accuracy in the experiment [16]. S. Baurer et al. placed forward a distinctive approach in which tumor patients can select a healthy brain atlas for medical images. The simulation of tumor growth along with trained algorithms is used to equate a safe image of the brain with that of the patient's brain. Finally, the non-stringent identification between the mutated atlas and the patient

image is used to allow a better contrast. It provides flexibility in the segmentation and enhanced tumor simulation and tumor enhancement treatment of atlas-based brain tumor images. In this approach, iterations are repeated until 90% of the tumor volume achieved [17]. A multiform brain tumor segmentation proposed by J. Huo et. al. used a hybrid approach. In this approach, they adopted three different methods: the first one is fuzzy connectedness, the second one is grow cut, and last is a classification of voxel using the vector supporting machine. As a hybrid rule, a confidence map averaging (CMA) technique has been used. In order to capture the global "joining together" of the voxels, a fuzzy correctness frame (FC) assigns the target object flipped affinities during classification. A paired t-test was found to be greater than 5 in the results of the experiments. A. Tom et. al. Suggested a technique of segmentation using invariant geometric transform analysis to detect brain tumors. The authors proposed a unique algorithm which is the combination of translational, rotational and scaling. The feature vector is calculated using the combination of shape, position, and texture. After that, the calculation of the Euclidean distance between the input vector and the preciously stored vector takes place. Images with minimum Euclidean distance are considered whose result is above a certain level i.e. threshold level. Experimental results achieved 90% of segmentation accuracy [18]. R. Vijayarajan et.al. explain about the medical image segmentation using fuzzy C-mean clustering technique where averaging is done by fusion of principal components. Image fusion is a technique of transmitting all appropriate and complementary image information into a single composite image from multiple MRI from the same source or from different sources. In this method, source images are segmented by the FCM clustering algorithm into K-number of clusters and the segmented regions are sorted specifically for fusion regions. If the principal components for the k-clusters is $p1^i$ and $p2^i$ where, i is from 1 to K then the average of all the parts is assessed by equation 3 and 4 [19].

$$p1 = (1/K) \sum_{i=0}^k p1^i \quad (3)$$

$$p2 = (1/K) \sum_{i=0}^k p2^i \quad (4)$$

K is subjective to various image inputs. And $p1$ and $p2$ are fusion weights. The fused image is obtained by the following equation 5 :

$$I = p1 * i1 + p2 * i2 \quad (5)$$

Where, $i1$ and $i2$ are two images which are going to be fused. The average quality index is around 85% which better than the other algorithm as given in the experiments.

The distinctive local, independent transformation based segmentation and classification of CT and MRI images have been submitted by M. Huang et al. In their study, the LIC method is employed to identify the segments into separate classes. Locality plays an important part in LITC's autonomous urban development, which solves problems. The locality is used to bind the anchor code to solve linear transformation issues instead of other coding techniques. On the basis of the actual patient results, the average dice similitude of the proposed whole tumor, tumor center, and tumor section are 0.84, 0.685 and 0.585, respectively [20]. The

method was suggested by N. Dhanachandra et. al. for the solution of the problem of segmentation using the K-mean algorithm of clustering and subtractive clustering algorithm. In this approach, the enhancement in the image quality, partial stretching improvement is applied to the picture. The clustering subtractive approach is an information clustering technique. Therefore, the subtractive cluster is used in k-means algorithms for the division of the images. The median filter is then placed in the segmented image to eliminate any undesired part from the image. Experimental results showed that RMSE is 0.0017 and PSNR is 35.77 which is a good result for segmentation. The ACM segmentation and ANN-LM classification methods used in brain tumor MRI analysis were used by A. Shenbagarajan and others. The proposed image analysis of MRIs using the ACM method regionally used to segment and the ANN algorithm based on the LM to efficiently identify MRIs as normal brain and timorous brain. The proposed MRI image-based brain tumor evaluation should effectively address the segmentation process and the classification method for brain tumor evaluation using feature removal techniques so that the technique would lead to a better result of brain tumor diagnostics before use in medical areas [21]. M. Sornam and others focus on establishing an independent diagnostic tumor program using pictures of MR weighted T1 and T2. The beginning step is to separate the brain MR images into benign and malignant by filtering images and the k-mean algorithm. The textural and form-based extraction process is performed with Wavelet and Zernike methods in the intermediate step. The final step is to determine the difference entre benign and malignant tumors using the ELM (Extreme Learning Machine) algorithm. Researchers found that the proposed algorithm achieved 77% of segmentation accuracy [22]. Figure 3 depicts an example of wavelet and Zernike based segmentation results.

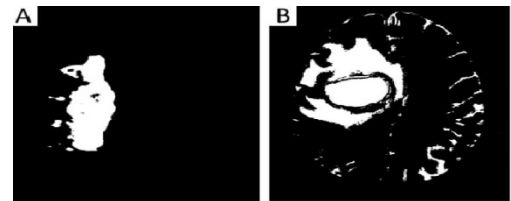


Figure 3: A – Benign, B – Malignant are the images after segmentation using k-mean [22].

In order to classify images of the MR of brain tumor, R.Ahmed et al. suggested the use of a support vector machine and ANN. For the segmentation of MRI images, the combination of tempered K-means and modified Fuzzy C-means (TKFCM) is used. TKFCM assimilates with few improvements K-means and Fuzzy C-means. A combination of Temper with K-means has identified in the brain MR images by gray level intensity. Equation 6 shall assess the gray based strength and picture complexity.

$$A(a_i, b_j) = \sum_{i=m+1}^{p+m} \sum_{j=m+1}^{q+m} R(a_i, b_j) S_{PQ}(resize) \quad (6)$$

Equation 7 gives the Temper window.

$$S_{PQ} = \sum_{i=m+1}^{p-m} \sum_{j=m+1}^{q-m} R(a_i, b_j) \quad (7)$$

Tempe-based image matrix is given with G gray-level intensity number and B bin numbers used for image R (a_i, b_j) temper detection. Where, m is defined as $m = (winsize - 1)/2$. Application

of exact temper values, rank and column is obtained by placing the desired temper.

$$L_g = \sum_{i=1}^D \sum_{j=1}^G A(a_i, b_j) || a_i - d_j ||^2 \quad (8)$$

$$L_p = \sum_{j=1}^Q \sum_{i=1}^G (V_{ij}) * (h_{ij})^2 \quad (9)$$

The description with which the tumor area could be located as follows (8) and (9) equation:

$$L_{gp} = \phi_d(L_g, L_p) \quad (10)$$

P and Q are row and column of A (a_i, b_j) and d is contour value. Cluster number, cluster center, and data point's number are identified by G, Q, and D respectively. The proposed study evaluates the sample data with 97.37% classification accuracy [23]. X. Ma et.al. proposes a noise removal method in SAR images. In this method, shift-invariant K-means singular value decomposition (SVD) with guided filter is used. The entire technique is made up of two steps. The first stage is to deal with the noisy image, the K-SVD shift, and the original de-noisy images. In the second step, the guided filtration is performed for the initial de-noisy image. Researchers achieved an average of 31.4040 PSNR over 6 images in the given method [24]. The automated brain tumor identification technique suggested by A. Mukaram et. al. is composed of the picture and separating two principal sections. Smoothing of the image can be accomplished through the use of a median filter, followed by an image enhancement method that can be accomplished through Sobel edge detection. For the Segmentation purpose, the K-means pillar algorithm is used. Pillar K-mean algorithm involves pillar pixel selection for efficient segmentation [25]. Hanafy M. Ali presented a study on MRI images for noise removal. MRI images are usually susceptible to noise such as the Gaussian noise, salt, pepper, and spatula noise. Therefore, having a precise brain picture is an extremely severe activity. A precise brain image is very important for further diagnosis. To remove noise, a median filter algorithm is used. Input images are combined with Gaussian noise and salt and pepper noise at various levels. To solve the problem of noise reduction it will be introduced a combination of the proposed Middle filter (MF), adaptive Middle Filter (AMF) and adaptable Wiener Filter (AWF). The filters are used to eliminate the additive noise from images of the MRI [26]. S. Riaz et. al. applied a unique technique on images having large scale data. This incorporates rough Fuzzy C-mean clustering with the methodology of the neural network. The main concept is to first create an original cluster core, then update image clusters and representations jointly during preparation, using a single layer of clustering by multi-layers of CNN. Fuzzy rough C-mean (FRCM) will be used for upgrading the forward pass cluster centers, while Stochastic Gradient Decent (SGD) will be used to upgrade CNN parameters. This rough collection of low and limited approaches addresses uncertainty, vacuity, and incompleteness in the definition of clusters and fuzzy sets which make it possible to effectively manage overlapping partitions in a noisy environment. The experimental result predicts that the given method has an average of 91% clustering accuracy among 4 different datasets [27]. Exhaustive research carried by A. Zotin et. al. used in brain tumor pictures to detect edges. In this technique, you look for a set of the fuzzy cluster by searching iteratively. A related fuzzy membership matrix specifying the

membership level of the kth element is prepared in the first set of vectors of the ith cluster. Equation 11 calculates the objective function.

$$f(D, C) = \sum_{i=1}^n \sum_{k=1}^m d \cdot (x_m - c_n)^2 \quad (11)$$

Where, d is the membership value, x_m is the mth cluster and c_n is the centre of the nth cluster. After dividing the image into a series of homogenous clusters with the fuzzy C-mean (FCM) algorithm the canny edge detector is implemented. The authors achieved 98.4% segmentation accuracy in 2 images [28, 29]. Proposed A Robust segmentation on MRI and field correction, incorporated into a clustering model through local contextual information by Z. Zhang and J. Song. In this study, the precise segments of brain MRI, which is corrupted by noise and intensity homogeneity, are used to produce a new robust clustering of local context information (RC-LCI). In order to produce the respective anisotropic weight to update the present core pixel and eliminate noisy pixels, the weighting method incorporating local contextual data was used to construct a pixel in the principal pixel region. A multiplying structure consisting of a real image product and a bias field could then efficiently segment the brain RMI and analyze the bias domain. In terms of the coefficient of Jaccard similarity, the results obtained by RC_LCI were increased by 0.195 +/- 0.125 [30]. A unique algorithm proposed by F. Han for the detection of high-speed moving locations in a noisy environment. In conjunction with the Wavelet Coefficients tree and Lipschitz Expertise Noise Property, the definition of the neural network activation function is agreed on for the activation evaluation technique. Then it is possible to retain the important wavelet coefficients. The non-important coefficients were also excluded by the process of testing the independent coefficients. High-frequency data can be preserved during the wavelet transformation by raising the disappearance of wavelet filters which is useful for edge sensing [31]. A. Nyma et. al. presented a research paper on a hybrid technique for MRI segmentation. In this study, they adopt the vector median filter for the removal of noise. Thereafter, the Otsu threshold is used to define the homogenous regions of the input image in the first coarse segmentation process. An improved FCM is used to separate brain MR pictures in several segments that use the optimal deletion factor in the given data set for perfect clustering. The experimental result showed that it achieves around 96% of segmentation accuracy [32]. M.Z. Abderrezak et. al. discuss a new hybrid approach for the analysis of brain MRI. In this approach, the nonlocal median filter is used for the preprocessing. For the post-processing, authors propose a new model that is based on geometrical transformation. The segmentation efficiency is up to the mark for the tumor and the multiple detections of sclerosis [33]. R.Ahmad et. al. discussed a different approach for solving brain tumor problem. The combined SVM and ANN classification techniques were adopted by the researchers. In order to segment an image, temper based KMFCM is used. The numbers of clusters are assigned more than the standard K-mean algorithm. Automatically modified FCM membership eradicates tumor field contouring issues. The statistical characteristics obtained from segmented images are used to identify and distinguish tumors from normal SVM brain MRI images. The classification of benign tumors is based on 2 sets of features and ANN is used for four stages of malignant. In this method, 97.37% classification accuracy achieved.

3. Proposed Hybrid Method

On the basis of the historical research and result analysis of the different segmentation methods for brain tumor diagnosis, the authors proposed a unique hybrid method for automatic brain tumor detection. The framework of the proposed method is given below:

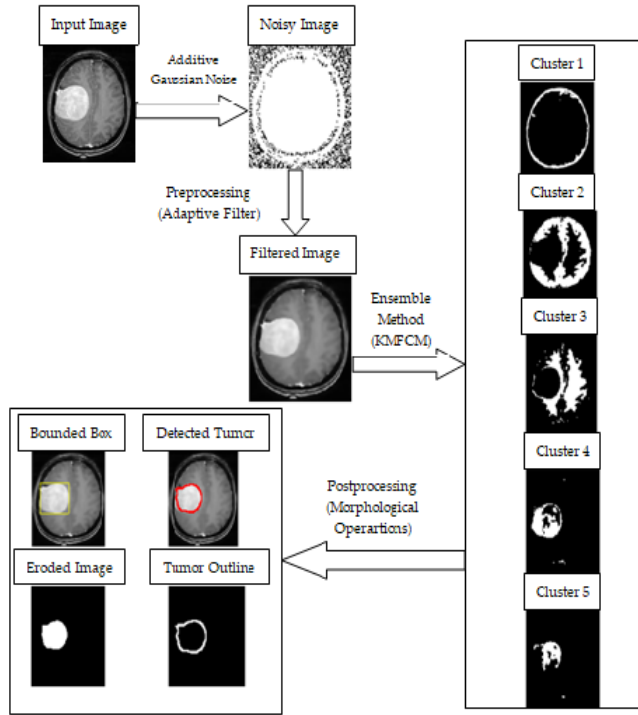


Figure 4: The framework of the proposed novel hybrid method for brain tumor segmentation

As shown in Figure 4, this article primarily divides the segmentation algorithm into three components.

Step 1: Preprocessing: Input the brain image from the Kaggle database [34]. Add Gaussian noise with different levels like 0.005, 0.007, etc. In this step noise removal is takes place with the help of adaptive mean and variance bas filter. By smoothing the MR images, it eliminates noise. This filter also reduces the variability in the intensity of MR images between one pixel to another.

Step 2: Applying Hybrid Method: In this step, a hybrid method which is the combination of modified K-mean clustering and fuzzy C- mean clustering (KMFCM) is applied. Segmented image with 5 different clusters are shown in figure 4. Brain tumor clustering and segmentation is done according to a threshold value.

Step 3: Post-processing: In the post-processing morphological operations are applied to the finding exact tumor region. For this purpose, the bounded box is used to show closed tumor areas. Also, the exact prediction of brain tumor some more eroded images and tumor outline images is added.

3.1. Preprocessing

On the degraded images containing Gaussian noise, the adaptive filter is applied. The mean and variance are the two statistical measures depending on a locally adaptive filter with a

specified region of $m \times n$ window. The Function of adaptive filter is given below:

$$f(x,y) = g(x,y) - (\sigma_{noise}^2 / \sigma_{local}^2)(g(x,y) - \mu) \quad (12)$$

Where, $g(x, y)$ is the pixel value of the image at position (x, y) , σ_{noise} is the variance of the overall noisy image, σ_{local} is the variance of the local region and μ is the local mean. The overall noise removal process is explained below with the help of an example:

- Define a 3×3 window size.
- Consider an input image B with additive Gaussian noise level=0.005. Noisy image is a 218×180 matrix having values from 0 to 255.
- Appending 0's to all four sides of the matrix B. The resultant matrix is C of order 222×184 .
- Calculate the local mean (μ) and variance by sliding 3×3 window on the entire matrix C. Local mean is the mean of window of a matrix of 3×3 .
- Calculate local variance (σ_{local}) with the given below equation:

$$\sigma_{local} = \text{mean}(\text{window}^2) - (\text{mean}(\text{window}))^2 \quad (13)$$

- The noise variance (σ_{noise}) of the whole image is calculated by taking the average of all local_variance.
- If ($\sigma_{noise} > \sigma_{local}$) then $\sigma_{local} = \sigma_{noise}$.
- The Filtered image is calculated by the following equation:

$$I_{filtered} = B - (\sigma_{noise}^2 / \sigma_{local}^2) \times (B - \mu) \quad (14)$$

Noisy image and the filtered image is shown in the figure below:

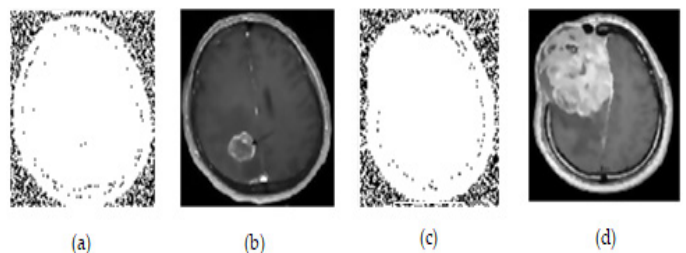


Figure 5: (a) Noisy Benign brain tumor image (b) Filtered image of (a) after processing adaptive filter (c) Noisy Malignant brain tumor image (d) Filtered image of (c) after processing adaptive filter

3.2. Applying Hybrid KMFCM Method

The proposed hybrid method is the combination of the modified K-mean clustering technique and Gaussian based fuzzy C- mean clustering technique. In the first stage, the extended K-mean is applied for deterministic cluster centroid initialization to prevent overfitting and then use the Fuzzy C-means algorithm for enhancing the classification capability. The extended K-mean algorithm is based on the standard K-mean algorithm having the capability of deterministically initializing the centroids. Maximizing the distance between the original cluster centroids is the fundamental principle of the extended K-means algorithm for initializing cluster centroids. This technology allows cluster

centroids to be initialized deterministically and resolve the shortcomings of the K-means algorithm linked to their initialization instability. Given below is the method of initialization of an extended K-mean algorithm:

- Select a sample pixel as the first initialized cluster centroid from the information set randomly.
- Select the centroids of the all k- clusters and compute the distance between sample pixel (i.e. the point within the cluster) and cluster centroid.

$$d_i = \sqrt{((x_c - x_i)^2 + (y_c - y_i)^2)} \quad (15)$$

Where, (x_c, y_c) is the centroid of the cluster and (x_i, y_i) is the i^{th} sample pixel.

- After finding the distances between cluster points with the remaining pixel values, extract the point with minimum distance.
- Compute the probability of remaining sample pixel by using equation 16

$$P(\text{pixel}) = d(x')^2 / \sum_{x \in R} d(x)^2 \quad (16)$$

Where, x is the sample pixel within the set of R pixels. $d(x)$ is the distance between the sample pixel x to closest cluster center and $d(x')$ is the distance between the nearest pixels. By probability, select the sample with the biggest probability as the new centroid cluster.

- Repeat the above method until the determination of k cluster centroids.
- Now applying, fuzzy C-mean clustering for further updating in the cluster center.
- Let, X_i is the input cluster centers obtained from extended k-mean clustering.
- FCM partitions input information set X_i , where $i=1$ to R , is assigned to different clusters by assigning membership values to the set of points.
- Calculate the membership values with the help of objective function given below:

$$F_m = \sum_{i=1}^R \sum_{j=1}^c \mu_{ij}^m (x_i - c_j)^2 \quad (17)$$

Where, $m > 1$, μ_{ij} is the degree of membership of cluster input data set x_i in the j^{th} cluster having cluster centre c_j .

- k is the number of steps. Iteratively optimizes the objective function and updates the degree of membership μ_{ij} and the c_j cluster centers.

$$\mu_{ij} = \left(\sum_{k=1}^c \left(\frac{x_i - c_j}{x_i - c_k} \right)^{2/(m-1)} \right)^{-1} \quad (18)$$

$$c_j = \left(\frac{\sum_{i=1}^R (\mu_{ij}^m \times x_i)}{\sum_{i=1}^R \mu_{ij}^m} \right) \quad (19)$$

- Calculate the value of $\mu_{\text{diff}} = \mu_{ij}^{(k+1)} - \mu_{ij}^k$. If the μ_{diff} is less than ϵ then terminates. The value of the ϵ is the fixed value between 0 and 1 throughout the computation of the hybrid method.

The results after applying the hybrid method are given below in figure 6:

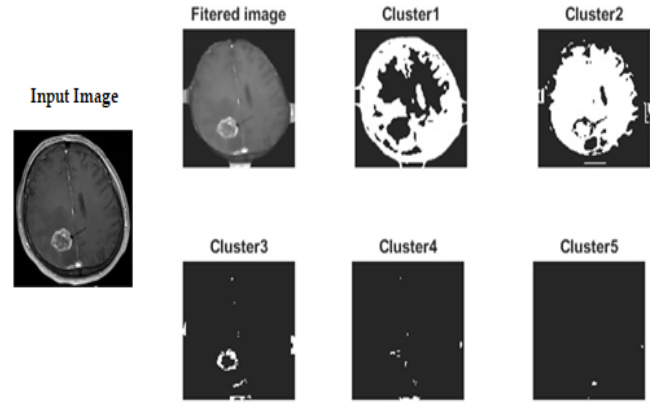


Figure 6: Clusters numbered 1 to 5 are shown after processing input image to the proposed KMFCM method

3.3. Post-processing

Morphological operations such as erosion and dilations applied to the results to improve the accuracy of segmented areas.

- Erosion is a method in which the structuring element is translated across the image domain. In a binary picture, erosion shrinks or thins items. A structuring element controls the way and magnitude of shrinking.
- The erosion by a structuring factor S of the binary image I generates a new binary image $I_{\text{eroded}} = I \ominus S$ with those positions (x, y) where the structuring element S matches the input image I originated, i.e. $I(x,y) = 1$ suited I and 0 otherwise, this process will be repeated for all the image pixels.
- Dilation is an operation in the binary image that develops or thickens items. A structuring component controls the way and magnitude of shrinking. Dilation is a method in which the structuring element is translated through the image domain.

\oplus

The results after applying the morphological operations are given below in figure 7:

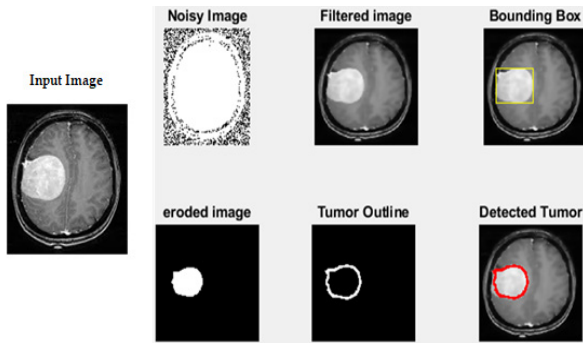


Figure 7: Eroded image, dilated image (Tumor Outline) and detected tumor are shown after processing input image to the morphological operation.

4. Evaluation Results

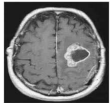
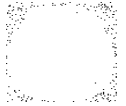












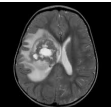













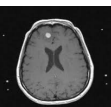













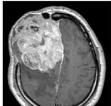













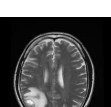













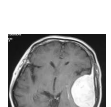













We randomly select 250 images from the kaggle database for the purpose of testing the proposed method [34]. These images are collected from different patients to check the reliability of the proposed method. Evaluation results of 10 images are given in table 1.

5. Result Analysis

The result is assessed on various parameters such as mean, standard deviation, Entropy, RMS, correlation, variance, smoothness, kurtosis, and skewness in order to inspect various aspects of this proposed method.

Table 1. Evaluation results of 10 different images containing benign and malignant tumor

Input Image	Noisy Image	Results after Applying Hybrid Method						Types of Tumor
								Benign
								Malignant
								Benign
								Benign

								Malignant
								
								Malignant
								
								Malignant
								
								Malignant
								
								Malignant
								
								Malignant
								

5.1. Platform for Evaluation

MATLAB implements the proposed model with the specific system requirements as shown in the table 2.

5.2. Result Analysis after Applying Proposed Method

Result analysis based on the statistical parameters is given in table 3. The statistical parameters are mean, standard deviation, entropy, RMS, correlation, variance, smoothness, kurtosis, and skewness. Table 4 predicts the detected tumor area, elapsed time. It is also able to classify the tumor is benign and malignant categories. Tables 5 show the comparative analysis between the size of the image and the elapsed time for brain tumor detection.

5.3. Histogram Analysis of Size of the Image vs. Elapsed Time

Figure 8 shows the histogram analysis of the result shown in table 5. Histogram analysis of the images shown that as size increase the time of brain tumor detection also increases.

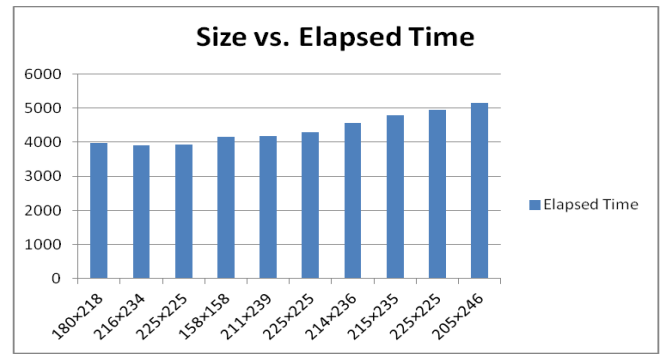


Figure 8: Histogram analysis of 10 different images based on the sizes in pixels and elapsed time in the Matlab

5.4. Comparison Based on the Segmentation Accuracy

Comparative result analysis based on the segmentation accuracy is shown in table 6. The proposed method contains 98% segmentation accuracy.

Table 2. System requirements of the proposed hybrid method

Operating System	Processor	Disk Space	RAM	Graphics
Windows 10	Any Intel or AMD x86-64 processor	2 GB	2 GB	It is suggested to have a hardware accelerated OpenGL 3.3 graphics card with 1 GB GPU RAM.

Table 3. Result evaluation of 10 images with different parameters like mean, standard deviation, entropy, RMS, correlation, variance, smoothness, kurtosis, and skewness

Image	Mean	Standard Deviation	Entropy	RMS	Correlation	Variance	Smoothness	Kurtosis	Skewness
1	0.00324997	0.0944484	3.25081	0.094491	0.149077	0.0089188	0.920936	6.31294	0.346882
2	0.00262226	0.087009	3.31088	0.087039	0.124667	0.0075712	0.921683	7.78956	0.662543
3	0.00311785	0.0857021	3.62058	0.085749	0.091271	0.0073174	0.933302	6.71334	0.514491
4	0.00260813	0.0999868	2.77989	0.100000	0.131878	0.0099708	0.86225	7.30026	0.505607
5	0.00287338	0.0883515	3.65058	0.088388	0.079195	0.0077465	0.927916	5.70919	0.385260
6	0.003997	0.08567	3.21550	0.085750	0.114946	0.007309	0.9472	10.49700	0.829200
7	0.002789	0.08835	3.08390	0.08839	0.106534	0.007804	0.9239	11.15900	0.972100
8	0.002625	0.08701	3.4965	0.08704	0.039998	0.007511	0.9218	6.8506	0.4296
9	0.003545	0.08569	2.6891	0.08575	0.136652	0.007345	0.9409	13.195	1.1142
10	0.004851	0.08968	3.6828	0.0898	0.065291	0.008028	0.9547	5.3377	0.3763

Table 4. Comparative result analysis of 10 images with different parameters like detected tumor area, elapsed Matlab time and type of tumor.

Image	Detected Tumor Area	Elapsed Time(mili seconds)	Types of Tumor
1	753.925	3967.4151	Benign
2	15.9	3898.48409	Malignant
3	56.18	3934.035306	Benign
4	55.385	4159.82001	Benign

5	23.85	4187.05282	Malignant
6	120.575	4283.705	Malignant
7	19.345	4565.229	Malignant
8	563.655	4777.548	Malignant
9	55.385	4952.883	Malignant
10	1444.52	5151.964	Malignant

Table 5. Comparative result analysis based on the size of the image and elapsed time

Image	Size of the Image (Pixels)	Elapsed Time (mili seconds)
1	180×218	3967.4151
2	216×234	3898.48409
3	225×225	3934.035306
4	158×158	4159.82001
5	211×239	4187.05282
6	225×225	4283.705
7	214×236	4565.229
8	215×235	4777.548
9	225×225	4952.883
10	205×246	5151.964

Table 6. Comparative result analysis based on segmentation accuracy

S. No.	Segmentation Technique	Imaging Modalities	Image Task(s)	Segmentation Accuracy
1	Proposed Hybrid Method	MRI	Noise Removal and Segmentation	98%
2	Wavelet transform with morphology based segmentation technique	MRI	Segmentation	80%
3	Computer Aided method that includes the segmentation of the morphological erosion and dilation	MRI	Segmentation	97%
4	Multistage modeling using Eulerian approach based segmentation technique	MRI	Registration and Segmentation	90%
5	Probabilistic neural network based segmentation technique	MRI	Segmentation and Classification	79%
6	Watershed & Thresholding based segmentation technique	CT Images and MRI	Segmentation	92%
7	Geometric transform invariant based segmentation technique	CT Images and MRI	Segmentation	90%
8	Watershed based segmentation technique	MRI	Segmentation	97%
9	SVM and ANN based segmentation technique	MRI	Segmentation and Classification	97.37%

6. Conclusion

The brain images in this study are different in tumor sizes. Therefore, the tumor region is calculated by the proposed process. A variety of statistical and machine learning algorithms such as extended K-mean and modified fuzzy c-means are combined with noise removal adaptive filter technique for implementing such an efficient algorithm. The non-linear morphological erosion and dilation operations carried out to remove the non-tumor portion. As a result, segmentation accuracy and time of brain tumor detection improved with previously discussed techniques in this era. The proposed hybrid method achieves 98% segmentation accuracy. Also, an average of 4 to 5 seconds is the time taken to detect brain tumors depending upon the size of the image. Depending on the outcome review of the various techniques, segmentation level accuracy is higher than other approaches in the background study. It is also found that the statistical parameters such as standard deviation and image are small, which indicates that the statistical analysis outcomes of the proposed method are better than previously discussed algorithms.

Conflict of Interest

So far the knowledge, the authors declare no conflict of interest regarding this article.

Acknowledgment

This work is supported by the director, ASET, Amity University, Lucknow. They would also like to thank the anonymous readers and reviewers for their important and constructive feedback and suggestions, which significantly improved this paper's content.

References

- [1] N. Goel, Dr. A. Yadav and Dr. B.M. Singh, "Medical Image Processing: A Review" IEEE International Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity (CIPECH), 57-62, 2016.
- [2] Sivakumaran, "Identifying Bone Cancer Using Markov Random Field Segmentation", 4 June 2018, Available online: <https://electronicsforu.com/electronics-projects/prototypes/cancer-markov-random-field-segmentation> . [Accessed: 21-August-2019]
- [3] Huo, Jing & Okada, Kazunori & van Rikxoort, Eva & Kim, Grace Hyun & R Alger, Jeffry & B Pope, Whitney & Goldin, Jonathan & S Brown, Matthew, "Hybrid segmentation for GBM brain tumors on MR images using confidence-based averaging". Medical physics, 2013. 40. 093502. 10.1118/1.4817475.
- [4] A. Shenbagarajan, V. Ramalingam, C. Balasubramanian and S. Palanivel, "Tumor Diagnosis in MRI Brain Image using ACM Segmentation and ANN-LM Classification Techniques", Indian Journal of Science and Technology, Vol 9(1), 1-12, Jan 2016.
- [5] Aiman Badawi and Muhammad Bilal, "High-Level Synthesis of Online K-Means Clustering Hardware for a Real-Time Image Processing Pipeline", J. Imaging 2019, 5, 38; doi:10.3390/jimaging5030038
- [6] Yuheng, S.; Hao, Y. Image Segmentation Algorithms Overview. arXiv, 2017; arXiv:1707.02051.
- [7] Cardoso, J.S.; Corte-Real, L., "Toward a generic evaluation of image segmentation", IEEE Trans. Image Process. 2005, 14, 1773–1782.
- [8] Chong Zhang , Xuanjing Shen, Hang Cheng, and Qingji Qian, "Brain Tumor Segmentation Based on Hybrid Clustering and Morphological Operations", Hindawi, International Journal of Biomedical Imaging, Volume 2019, Article ID 7305832, 11 pages <https://doi.org/10.1155/2019/7305832>.
- [9] N. Nezamoddini-Kachouie and P. Fieguth, "A Gabor based technique for image denoising", Canadian Conference on Electrical and Computer Engineering, 2005. Saskatoon, Sask., 2005, pp. 980-983. doi: 10.1109/CCECE.2005.1557140
- [10] James C.Bezdek, Robert Ehrlich, William Full, "FCM: The fuzzy c-means clustering algorithm", Computers & Geosciences volume 10, issues 2–3, 1984, Pages 191-203.
- [11] Nameirakpam Dhanachandra, Khumanthem Manglem, Yambem Jina Chanu, "Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm", Procedia Computer Science volume 54, 2015, pages 764-771. <https://doi.org/10.1016/j.procs.2015.06.090>
- [12] Kale Vaishnav, Vandana B. Malode, "A Novel Approach based on Average Information Parameters for Investigation and Diagnosis of Lung Cancer using ANN", Pattern Recognition and Image Analysis april 2018, volume 28, Issue 2, pp 301–309.
- [13] Ahmed Kharrat, Mohamed Ben Messaoud, Nacéra BENAMRANE and Mohamed ABID, "Detection of Brain Tumor in Medical Images", International Conference on Signals, Circuits and Systems, PP 1-6, 2009.
- [14] Rajeev Ratan, Sanjay Sharma and S. K. Sharma, "Brain Tumor Detection based on Multi-parameter MRI Image Analysis", ICGST-GVIP Journal, Volume (9), Issue (III), PP- 9- 17, June 2009.
- [15] Qurat-Ul-Ain, Ghazanfar Latif, Sidra Batool Kazmi, M. Arfan Jaffar and Anwar M. Mirza, "Classification and Segmentation of Brain Tumor using Texture Analysis", Recent Advances in Artificial Intelligence, Knowledge Engineering and Data Bases, PP 147- 155, Jan 2010.
- [16] M. Usman Akram and Anam Usman, "Computer Aided System for Brain Tumor Detection and Segmentation", International Conference on Computer Networks and Information Technology, PP 299-302, July 2011.
- [17] Stefan Bauer, Christian May, Dimitra Dionysiou, Georgios Stamatakos, Philippe B'uchler, and Mauricio Reyes, "Multiscale Modeling for Image Analysis of Brain Tumor Studies", IEEE Transactions on Biomedical Engineering, Vol. 59, No. 1, ,PP 25-29, January 2012.
- [18] Arun Tom, P. Jidesh, "Geometric transform invariant Brain-MR Image Analysis for Tumor detection", IEEE conference on Circuits, Controls and Communications (CCUBE), 1-6, Dec 2013.
- [19] R. Vijayarajan and S. Muttan, "Fuzzy C-Means Clustering Based Principal Component Averaging Fusion", International Journal of Fuzzy Systems, Vol. 16, No. 2, June 2014 PP-153-159.
- [20] Meiyuan Huang, Wei Yang, Yao Wu, Jun Jiang, Wufan Chen and Qianjin Feng , "Brain Tumor Segmentation Based on Local Independent Projection-based Classification", IEEE Transactions on Biomedical Engineering , 2633 - 2645 , Vol. 61, Issue 10, Oct 2014.
- [21] K. K. Gupta, N. Dhanda and U. Kumar, "A Comparative Study of Medical Image Segmentation Techniques for Brain Tumor Detection", 2018 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2018, pp. 1-4.
- [22] M. Sornam , Muthu Subash Kavitha , R. Shalini, "Segmentation and Classification of Brain Tumor using Wavelet and Zernike based features on MRI", IEEE International Conference on Advances in Computer Applications (ICACA), Coimbatore, 166-169, Oct 2016.
- [23] Rasel Ahmmmed, Anirban Sen Swakshar, Md. Foisal Hossain, and Md. Abdur Rafiq, "Classification of Tumors and It Stages in Brain MRI Using Support Vector Machine and Artificial Neural Network", In Proc. IEEE International Conference on Electrical, Computer and Communication Engineering (ECCE), Bangladesh, 229-237, Feb 2017.
- [24] Xiaole Ma, Shaohai Hu and Shuaiqi Liu, "SAR Image De-Noising Based on Shift Invariant K-SVD and Guided Filter", remote sensing, November 2017.
- [25] Arbaz Mukaram, Chidananda Murthy and M.Z.Kurian, "An Automatic Brain Tumour Detection, Segmentation and Classification Using MRI Image", International Journal of Electronics, Electrical and Computational System, Volume 6, Issue 5 May 2017.
- [26] Hanafy M. Ali, "MRI Medical Image Denoising by Fundamental Filters", in High-Resolution Neuroimaging - Basic Physical Principles and Clinical Applications, Ahmet Mesrur Halefoğlu, Intechopen, pp-111-124, March, 2018.

- [27] Saman Riaz, Ali Arshad and Licheng Jiao, “Fuzzy Rough C-Mean Based Unsupervised CNN Clustering for Large-Scale Image Data”, *applied sciences* October 8(10), 1869 , 2018.
- [28] Alexander Zotin, Konstantin Simonov, Mikhail Kurako, Yousif Hamad, Svetlana Kirillova, “Edge detection in MRI brain tumor images based on fuzzy C-means clustering”, *22nd International Conference on Knowledge-Based and Intelligent Information & Engineering Systems*, PP-1262-1270, 2018.
- [29] Gupta K.K., Dhanda N., Kumar U., “Depth Analysis of Different Medical Image Segmentation Techniques for Brain Tumor Detection”, in: Jain L., Virvou M., Piuri V., Balas V. (eds) *Advances in Bioinformatics, Multimedia, and Electronics Circuits and Signals. Advances in Intelligent Systems and Computing*, vol 1064, 2020. Springer, Singapore.
- [30] Zhe Zhang and Jianhua Song, “ A Robust Brain MRI Segmentation and Bias Field Correction Method Integrating Local Contextual Information into a Clustering Model”, *Recent Advances on Signal Processing and Deep Learning for Public Security and Engineering Applications*, *Appl. Sci.* , 9(7), 1332, March, 2019.
- [31] Fangfang Han , Bin Liu , Junchao Zhu and Baofeng Zhang, “Algorithm Design for Edge Detection of High-Speed Moving Target Image under Noisy Environment”, *Sensors* 2019, 19(2), 343; <https://doi.org/10.3390/s19020343>
- [32] Alamgir Nyma, Myeongsu Kang, Yung-Keun Kwon, Cheol-Hong Kim and Jong-Myon Kim, “A Hybrid Technique for Medical Image Segmentation”, *Hindawi Publishing Corporation, Journal of Biomedicine and Biotechnology*, Volume 2012, Article ID 830252, 7 pages, doi:10.1155/2012/830252.
- [33] Mohamed Zaki Abderrezak, Mouatez billah Chibane, Prof. Karim Mansour, “A New Hybrid Method for the Segmentation of the Brain MRIs”, *Signal & Image Processing: An International Journal (SIPIJ)* Vol.5, No.4, 77-84, August 2014.
- [34] Navoneel Chakrabarty, “Brain MRI Images for Brain Tumor Detection Dataset”, *Appl.* 2019. Available online: <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>. [Accessed: 21-August-2019]

Centralized System of Universities Learning Materials

Ruslan Vynokurov*, Volodymyr Tigariiev, Oleksii Lopakov, Kateryna Kirkopulo, Olena Pavlyshko

Odessa National Polytechnic University, Institute of Industrial Technologies Design and Management, 65044, Ukraine

ARTICLE INFO

Article history:

Received: 15 January, 2020

Accepted: 11 April, 2020

Online: 03 May, 2020

Keywords:

Distance Learning

Education

Studying

Website

VR

AR

ABSTRACT

This article considers the creation of a program / website as an example of a centralized system for all universities in order to more easily familiarize an applicant and / or student with the internal politics and capabilities of the university. Existing problems in the training system are examined, why distance learning is relevant, what problems exist in it and how they can be solved. Based on the proposed distance learning projects, an argument is built on its meaninglessness without further advancement, and how a website concept can help solve this problem. The concept of the website and its possible implementation using existing analogues in other areas that successfully complete the tasks are considered. Theoretically, using the proposed methods, the concept will simplify the choice of a further place of study for schoolchildren and applicants. The method consists in parsing the desired university website into the necessary categories: a number of specialties, disciplines, foreign programs, distance learning materials and so on. This will allow you to create a library of universities, their materials and open data (specialties, disciplines), without coming to excesses, such as creating a video hosting service for broadcasting recorded disciplines, less data to be stored, and more.

1. Introduction

This paper is an extension of work originally presented in IEEE PICS&T 2018 [1].

The purpose of this article is to consider the possibility of creating a website that will combine higher education institutions with the goal of easier navigation in higher education.

The original article examined why the decision to use the Unreal Engine 4 (<https://www.unrealengine.com/en-US/>) game engine is relevant to create a platform. Thanks to blueprint technology, which is a system of nodes with built-in code scripts, you can easily perform many tasks without knowing the syntax of a programming language, in this case C++. Saving time for studying the syntax and focusing forces on the program logic, it becomes possible to devote more time directly to the program itself, which positively affects the quality.

Also in the original article, the creation of a workflow for filling the platform with materials was considered: what components are needed to create a working system, what programs and their bounds were proposed for compiling the content of the proposed lessons, a conceptual example used at the Odessa

National Polytechnic University was considered. All this was done in order to describe one of the possible options for creating such a platform/hub, where students could find the materials they need and navigate the relevance of knowledge that the university has been presenting during the studying.

This article focuses on the consideration of the previously described problem, its expansion and argumentation. An approach was taken to collect information from third-party sources and combine them to provide more specific structured information to applicants through the website. The process of the platform creation was partially considered [1,2].

At the moment, students are often not aware of what universities are like. According to the experience of the authors who live in Ukraine, at the start of student enrollment, different universities arrange open house days when you can go in and look at the place where the applicant thinks to continue his studies. According to this event, the student receives a plot of information about who he will study with, who will teach him and what he will be able to achieve. Often, since in Ukrainian universities a shortage of students in certain specialties is a frequent occurrence, it is customary to embellish a university story and indicate only the advantages of the specialty and/or specialization, without informing applicants about possible shortcomings. The

*Ruslan Vynokurov, Email: ruslan.vinokurov14@gmail.com

disadvantages can be in different areas: poor teaching staff, the relevance of the teaching materials presented, poor learning conditions and more.

There is also the problem of choice, which should solve the confines between graduation and the beginning of a university.

The problems that prevent students from choosing their future profession will be considered below.

2. Main Problem

There are a number of problems that do not allow obtaining the necessary knowledge for a person. It is also necessary to indicate that these problems are considered in several areas: preparation of the applicant for admission to the university and student gaining knowledge in the university environment. According to the authors, there are two most common problems:

- The problem of awareness.
- The problem of the existing education system.

2.1. Awareness problem

The authors' experience is based on a model of the Bologna system, as well as on the basis of statistics from some universities in the Russian Federation. The moment of transition from school to a higher educational institution often raises many questions, since this stage is constantly updated and supplemented by various acts. All these updates are briefly explained, but the busyness of schoolchildren with the upcoming exams does not make it possible to immerse themselves in the question completely.

The exam procedure itself is relatively simple, but preparing for them can entail enormous costs in time, which limits students in reflecting on who they want to become in the future. Short time for studying the material and pressure from the family and the teaching staff leads to procrastination among schoolchildren, which affects their general level of anxiety, productivity and academic performance [3]. Such employment and a psychologically difficult situation do not allow the student to adequately assess their strengths and weaknesses, which would allow to come to the correct decision regarding the choice of specialty and specialization.

According to the considered statistics, only 11% of schoolchildren in grades 8-9 are determined with the place of forthcoming education. 54.1% is determined only by the end of grade 11. 5.2% of students initially knew where they would go. 19.7% are determined either immediately before admission, or do not do it at all. The choice is complicated by a large number of criteria and a variety of conditions, thanks to through which the choice of university is made: parents' advice, teachers' advice, print university advertising, television and radio advertising, open day at the university, conversations with university students, speeches, meetings, conversations with teachers, job fairs, university rankings in the media and more [4]. Due to the fact that applicants cannot decide on the choice of a profession at school, they are trying to find the specialty that is relevant and necessary at the time of selection — the trend.

Often, professions are selected on the basis of superficial knowledge about this type of activity. Each specialty has its own

requirements for a person in different fields: level of knowledge, stress tolerance, responsibility, level of emotional intelligence, sociability, teamwork and other human qualities. This can be called a characteristic of a single profession, which for the most part should be observed in candidates. Many people make an unconscious choice of their future profession, since the characteristic is a huge amount of details and, often, the only way to understand whether you can relate to this specialty is an internship, which can be completed with an average level of knowledge in the specialty, which suggests studying in this area. Therefore, those 54.1%, as well as 19.7% of schoolchildren who make decisions at the last moment, cannot have a complete picture of what they need to study in order to become a specialist in a certain kind of activity, since this requires a lot of amount of time. This leads to poor-quality training of students due to the lack of their motivation to study further. And this, in turn, to work not in the specialty and lack of qualifications in another kind of activity.

2.2. Education system problem

Countries should try to develop in the field of education as fast as countries with similar capabilities do [5]. The main problem in the development of distance learning is an incorrect education system. According to statistics from the American Federation's National Policy Summit for Children, online schools in their early days did not achieve proper exams due to funding for a non-working education system. From the very beginning, virtual schools were suitable only for some disciplines [6], but with the advent of new mechanics it became stronger in others, which gradually leads to the popularity of this approach for learning.

In the CIS countries, this situation lies in the lag of the teaching methodology, which is manifested in the students' lack of interest, since the training system is not clear and complex, which leads to a lack of understanding of the need for a particular discipline and/or specialty as a whole. This is due to the formalism of the materials presented. It is expressed in a mechanical, passive, insufficiently meaningful assimilation by students of teaching material, in memorizing verbal formulas by them, devoid of specific content, in the inability to connect the knowledge gained with life [7]. This educational system was used in the USSR and developed to the maximum possible to meet the requirements of the ruling party. Subsequently, the education system was not changed and only updated the knowledge provided, which led to the depletion of the system and its inability to improve further, namely, to include modern technologies as part of the learning process, such as AR/VR/MR and so on [8].

3. Information About Distance Learning

At the moment, distance learning is becoming more and more popular in the world, which helps many students to gain knowledge without being in an educational institution. Distance learning is gradually becoming the standard throughout the world. This can be seen by the number of queries in popular search engines. For example, according to the data from the multi-functional SEO platform Serpstat, on google.com the query "distance learning" is entered into the search line 133900 times a month (on average for 2019), and the phrase "online school" is 255300 times under the same conditions. Now the term "online school" is becoming popular, which consists in distance learning through various communication platforms and video conferencing

of the teacher. Special business accelerators are being created, which are aimed at helping start-ups and companies related to this type of studying and helping to transform a successful business out of it through templates and promotion schemes.

The largest universities in the world have long possessed this type of education, for example: Boston, Florida and Arizona State Universities, the University of Wisconsin at Madison, as well as the State University of Pennsylvania in the USA. With the help of this system, the University of Florida allowed 400000 students to get education in 135 countries of the world. There are examples in the UK: University of Liverpool, Suffolk University Campus, Ruskin University England, School of Oriental and African Studies, University of Manchester. The latter currently provides knowledge to about 40000 students only through distance learning in 154 countries around the world [5].

3.1. What is wrong with distance learning organization

Despite such local and international success, it is not possible to refer to the current distance learning - there is no primary source that would immediately tell about many available options from different universities and at the same time give access and description to it.

The above examples are well-known mastodons of science, which in any case will be found by applicants if they wish, because you can hear about them at school age as the most desirable places to study, but applicants absolutely do not know anything about other places, which also may be good for learning. Examples show that distance learning is valid, but it is impossible to confine oneself only to these mastodons - on the contrary, it is necessary to expand knowledge, it is necessary to make it clear that the world of education is much wider and more spacious, and does not close to the top ten most popular universities.

For example, the YouTube website (<https://www.youtube.com/>) is a great place to store distance learning videos, but it was not originally created for this. You can restrict access to video, as well as collect video playlists, but adding a description to each playlist to make it convenient is impossible due to lack of functionality. And it would be possible to create an analogue of Instant View (hereinafter referred to as the IV) from Telegram (<https://telegram.org/>) and still store the video on a service from Google, but at the same time pre-set and display it with the necessary data on the website.

The point is that it is necessary to fundamentally change the existing system and introduce a core that can connect all that is different that exists at the moment. Each person should have the opportunity to look at different universities and decide which one will be most suitable for him. This implies the lack of centralization of the final result, in this case, the showcase of distance learning and the university as a whole.

3.2. Why is this a problem

At this stage, an applicant who has the opportunity to enter any university should find universities on the Internet and check the availability of specialties, specializations, what he will be taught, who he will become in the end, what perspectives the industry has and much more. It takes time to surf on the Internet, which students do not have, as it was described earlier. The problem is that the

more responsible the student approaches the question of exams, the more irresponsible he is to his future. It is quite difficult to find a balance, since both questions are complex and difficult to study, so the idea of centralizing information about universities from one point seems quite reasonable.

3.3. How a solution to this problem affects on progress

When a website appears with an extensive list of universities, it becomes possible to compare their statistics: popular specialties, the relevance of discipline programs, teacher qualifications, etc.

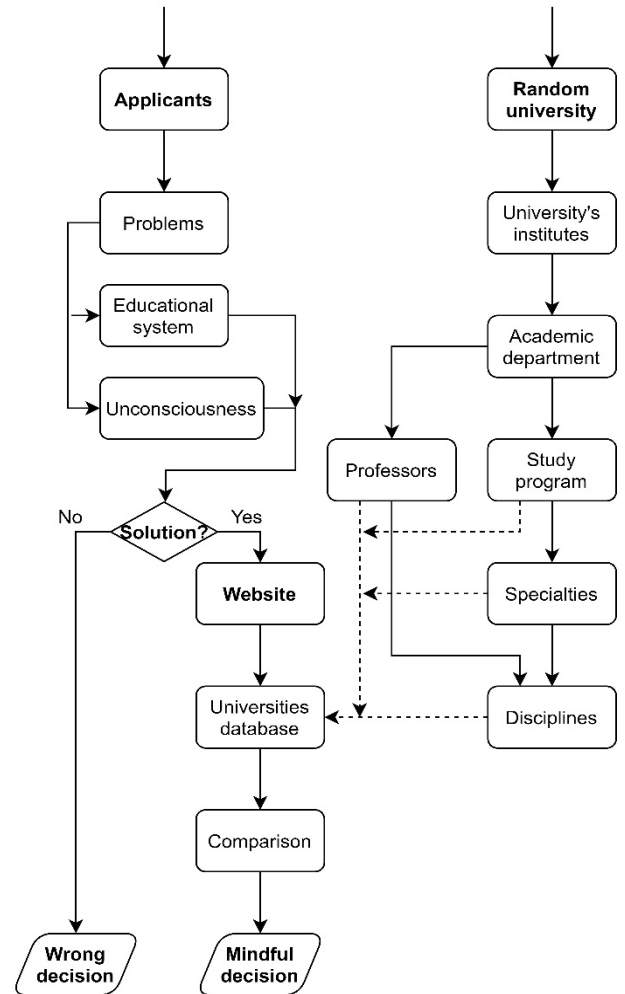


Figure 1: Scheme of each applicant choice and how the solution should be implemented.

Thanks to this website, students can easily compare and contrast universities: see different and general, study programs, with which foreign partners there is cooperation, foreign programs, opportunities for graduation, with which companies the educational institution cooperates. Also, applicants gain the opportunity to see other universities, the existence of which they might not know. Since the website will also have distance learning materials, this allows applicants to review the teaching methodology, compare with other universities and make an informed decision about admission to a particular educational institution.

This implies a great knowledge of the choice of a university, as well as a dry balance of the necessary information without frills,

which allows a sober assessment of the proposed educational institution. As a result, students study in more relevant specialties for them at the universities most suitable for this, which is what the research of this article seeks. The overall idea is described on the Figure 1.

4. Solution

The examples of solutions that allow you to learn much more efficiently and faster are shown below, but due to the lack of information about such solutions, few people think about their existence. At the same time, they can provide an incentive to create competitive software, or make additions to existing code, which will improve the UX and functionality of existing programs. This approach is widespread in the IT field and this is evidenced by such communities as GitHub, Stack overflow, the open source Unreal Engine 4, which can be supplemented by various plug-ins and the Blender 3D modeling program, which also has open source code and can be supplemented by functionality that users deem necessary.

All this speaks of the popularity of the approach when people, specialists in their field, are aware of this. That is why it is necessary to show and disseminate new techniques and show them to a wide audience, in which the proposed website concept can help.

4.1. Solution examples

There are different AR applications and many of them have very different goals. AR technology allows you to reproduce objects that are difficult to read in 3D models, which makes it easier to accept abstract and complex content. This is useful for those who remember visually and who need a translation of the theory into a tangible concept. For example, Polytechnic University of Leiria, in Portugal introduced AR to math classes and students speak of it as a very useful, easy and interesting solution.

Nvidia Holodeck may be an example of VR technology, whose technology allows people from different parts of the planet to work together in virtual space, makes it possible to develop design, assembly lines, technological processes and much more, having their 3D counterparts in a virtual environment [1].

A team from Stanford University created a project called STRIVR. It began as a VR training tool for their university football team. Subsequently, they found out about the project, and it turned into a platform for staff training and works with companies such as Walmart. United Rentals (hereinafter referred to as UR), which is the largest equipment rental company in North America, trains third-party sales representatives (hereinafter referred to as TPSR) using STRIVR technology. Using photos or videos in the classroom does not allow students to gain practical experience, because UR used the STRIVR immersion platform to create the next generation TPSR tutorial that uses VR technology. It includes 5 practical steps in VR, so that training takes place with a sense of presence directly at the workplace, but actually not there. Thus, they achieved an increase in efficiency and reduced training time by 40% [1].

4.2. Conclusions from the proposed concepts

Above are local, point solutions that allow you to gain an increase in the effectiveness of training using practical experience.

This decision is relevant, as the current education system in the CIS has reached its maximum. At the moment, it cannot fully compare the modern requirements of social expectations and real educational results [8], which leads to a lag and less interest of students to study.

Also, these examples are relevant, since AR/VR/MR technologies are considered one of the most promising technologies of the 21st century, and also create new ways of presenting information and studying.

The disadvantage of this system is the lack of cooperation between such solutions, which does not allow to switch from one technology to another within the framework of one application. Having created a common basis that opens up access to any of these solutions, you can think about implementing a technology that will demonstrate different disciplines (where applicable), allowing students to understand what a single discipline is in a particular specialty and what it manifests itself.

This technology allows to solve the problem of awareness in view of a more affordable solution for students and the opportunity to get some practical experience immediately before studying in this area.

Also, this system involves either a complete change in training programs, or their adaptation to a new approach, which solves the problem of the education system and allows to increase the effectiveness of training, a more relevant choice by students of future specialties and areas in which they will study.

4.3. Suggestion of solution

It is proposed to create a website. It will serve as the basis for comparing universities, searching for information about educational institutions, study programs in them. It is proposed to minimize the work and use what already exists, namely, materials and approaches.

Telegram messenger has IV technology, which was created specifically to display the content of other websites inside the messenger. When a user shares a website link inside Telegram, a bot instantly fires, which checks for the presence of a template for the website. A template is a set of rules that interprets a website as necessary, removing all unnecessary and leaving a purely necessary context in the form of text, links, tables, images. If such a template is found, the website opens inside the messenger without loading the page - instantly. It is noteworthy that any user can create a template and send it to the messenger developers for verification and subsequent approval. After confirming the suitability of the template, any page of the website will open according to the rules established in the template.

The idea of the proposed website is a similar approach. Instead of re-creating the content, you can duplicate it and interpret it in the right way, managing the content as you need.

This solution allows you not to create your own video hosting services, like YouTube, for storing university video data. Instead, you can simply link to existing pages and present information from the websites of universities and their individual projects, whether it's distance learning, or international programs, in one light and together in a user-friendly way.

This will allow you to view from one place all those existing lectures and materials, regardless of what they will be: an article, video, or the AR / VR / MR application. In the case of the latter, most likely, it will be necessary to install additional software recommended for their reproduction, but such a question is not included in the scope of this article.

The uniqueness and unification of the website is a guarantee of freedom in the implementation of projects (websites) of other universities. The existence of such a website will not limit the creators of university websites to change anything in the design or development approach.

5. Implementation

As mentioned earlier, the concept is a website that uses a similar IV technology created by Telegram.

The proposed website concept can solve the problems indicated in section 2. Since the concept collects a database of different places for higher education, having the ability to expand and modify the list of institutions and their characteristics according to the method below, the problem of awareness can be overcome with a single condition: the provision of such a basis to applicants in advance and with an explanation.

The system works like an online store with a large number of filters - what interests the applicant has selected: direction, specialty (final profession), country, tuition fees, other filters that will allow you to get a selection of final results. Provides a description of the specialty, possible previews of training materials for better reference. Applicants have a chance to look at a specialty and form an opinion about it, which makes it easier to make a decision.

The final result of solving the problem of awareness remains behind the desire of applicants to understand this issue, but the presented concept will simplify the task for them by collecting everything they need in a single database.

Initially, existing university materials will be presented, both on distance learning and lecture materials from various platforms, for example, YouTube. The website will initially focus on downloading various additional materials from external sources. This is due to the lack of servers and capacities that are required to service such networks. The website will only target educational materials from trusted sources that are either presented at the university, or distance learning materials. This is also the main difference from YouTube, Vimeo (<https://vimeo.com/>) and other similar services, where in addition to educational content, you can also find entertaining, which can distract the user. Thus, having the opportunity to accept materials from a variety of resources, the website acts as a place where only training materials will be collected.

5.1. Functionality and purpose of the site

To be able to open articles from a single resource, a template is created that transforms the data into the desired form and provides it in an IV which is a simplified form without such additional aspects as a list with other articles, website menu, advertising, and so on.

The transformation of the page of any resource is carried out in the following 4 steps:

- Determining which parts of a particular resource will be better suited to the conditions for creating the template;
- Parses the basis of the website for the main tags of HTML code to present them in accordance with the requirements of the template. In case an information element is found that does not meet the rules of the template, it is necessary to use the functions provided by Telegram that transform them into a suitable element;
- Remove all unnecessary content and provide a clean IV page for viewing using a dedicated feature;
- Check the template using 5-10 links to other articles of the same resource to make sure that all applied functions work correctly and there is no need to use others to display additional content (for example, replace one HTML tag with another, insert a paragraph tag inside text structures and so on).

According to a case study that was provided by Telegram based on an analysis of an article from medium.com, you can understand and take as a basic principle how the website content is transformed and its new look is compiled.

This approach is quite flexible and scalable. Also, the Telegram code is open and exposed on the GitHub resource, which allows you to use the acquired technologies as needed.

6. Related Work

At the moment, work is underway on the development of application authentication and the functionality of the hub start page, the second component of the proposed solution, which was presented in the original article [1]. To create authentication, a set of libraries is used to implement standard scenarios (hereinafter referred to as the SDK) of GameSparks (<https://bit.ly/3besv6D>), which allows us to develop user authorization logic in detail. The start page at the moment is a block that contains video, the main control panel and news.

For authentication, blueprint technology was used, which allowed us to create user registration and authentication. Figure 2 shows the blueprint used.

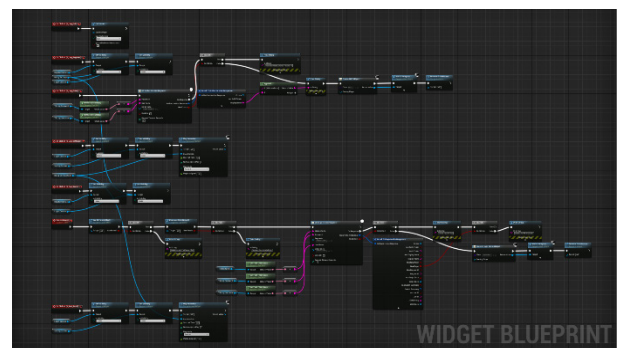


Figure 2: Authentication blueprint.

At this stage, the website has not yet been developed, since the final goal is constantly updated and optimized.

7. Future Work

As the upcoming works, the authors of this article see the development of the hub application, namely the creation of an application for playing their own AR/VR/MR projects and interaction with them. It will be considered creating a user account and saving information about it inside the application, the ability to edit it. Creating basic templates for learning materials, for example: video lecture, video with auxiliary files, AR project, VR project and more, in which all the necessary variables and components that create this template will be explained and recorded. Thanks to it, content of the same type will be able to collect in arrays and automatically load.

Viewing videos from YouTube, Vimeo and similar websites will be implemented, when creating the subsequent pages of the application.

Creation of the website described in this article where the functionality for comparing universities will be created. A bundle of a website and an application for synchronizing databases, since, for example, the GameSparks SDK stores all accounts inside its server.

The application will be created using Unreal Engine 4, which will fully realize the main goal of the article: quick familiarization of applicants with the specialties and professions taught by universities in order to make a more informed and rational choice of further education.

References

- [1] V. Tigariyev, R. Vinokurov, A. Pavlyshko, I. Sinko, "Informational Training System Mindgate" in 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), IEEE, 2019.
<https://ieeexplore.ieee.org/document/8632104>
- [2] V. Tigariyev, R. Vynokurov, "Creation Of a Learning System Based On The Game Engine Unreal Engine 4 (article in Russian with an abstract in English)" in 2018 ICT 7th International Scientific and Technical Conference, Odessa, Ukraine, 2018.
http://immm.opu.ua/files/archive/n4_v8_2018/immm_n4_v8_2018.pdf
- [3] J.R. Ferrari "Self-handicapping by procrastinators: Effects of task importance and performance privacy" in 1990 Doctoral dissertation, Adelphi University, Garden City, NY, USA, 1990.
- [4] G. Evdonin, S. Mamedova, "Modeling of choosing of university for entrants (article in Russian with an abstract in English)" in 2014 Управленческое консультирование, 2014.
<https://cyberleninka.ru/article/n/modelirovanie-protsesssa-vybora-vuza-abiturientom-pri-postuplenii>
- [5] R. Vynokurov, "Расстояние Как Основная Проблема Информационного Обеспечения Образовательных Процессов" in 2019 TIRMIA/TI 5th International Scientific and Practical Conference of students, postgraduates and junior researches, Odessa, Ukraine, 2019.
- [6] A. Molnar, G. Miron, C. Gulosino, C. Shank, C. Davidson, M. K. Barbour, L. Huerta, S. R. Shafer, J. K. Rice, D. Nitkin, "Virtual Schools in the U.S. 2017" in 2017 Boulder, CO: National Education Policy Center, 2017.
<https://nepc.colorado.edu/publication/virtual-schools-annual-2017>
- [7] O. Lebedev, "The End of Compulsory Education? (article in Russian with an abstract in English)" in 2016 Вопросы образования, 2016.
<https://cyberleninka.ru/article/n/konets-sistemy-obyazatel'nogo-obrazovaniya>
- [8] K. Gusev, "XXIV съезд КПСС и отечественная историография" in 1971 Издательство Академии наук СССР, 1971.
<https://bit.ly/2uCNSxQ>

Dynamic Objects Parameter Estimation Program for ARM Processors Based Adaptive Controllers

Vasily Olonichev*, Boris Staroverov, Maxim Smirnov

Department of Automation and Microprocessor Technology, Federal State budgetary Educational Institution of Higher Education Kostroma State University, Kostroma, Russia

ARTICLE INFO

Article history:

Received: 08 January, 2020

Accepted: 22 April, 2020

Online: 03 May, 2020

Keywords

Adaptive control

LSQ parameter estimation

ARM_CORTEX processors

Program optimization

ABSTRACT

Modern microcontrollers are capable to realize not only traditional PID-regulators but also adaptive ones. Object of control parameters estimation is the biggest part of adaptive control from the point of view of time consumption. The ways to reduce this time for digital control systems based on ARM-CORTEX 32-bit and 64-bit processors are shown in the article. These ways include source code refactoring, using vector registers and parallelism of code. As result of program improvement, a new algorithm for least squares method was suggested. Intrinsic for vector operations and OMP directives were added to the program to realize data and code parallelism. All options were tested for time consumption in order to find out the best decision. The program suggested may be useful while realizing adaptive controller based on single-board mini-computers and microcontrollers

1. Introduction

This work is an extension of conference paper "Optimization of the Program for Run Time Parametrical Identification for ARM Cortex Processors" originally presented in "2018 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)" [1]. Conference paper has the results obtained only with the 32-bit armv7 processor. The results obtained with the 64-bit armv8 processor are added in the current paper. Also, the questions of alignment data in memory and leftovers processing during vectorization are considered.

Nowadays most of the micro-controllers are based on inexpensive but at the same time powerful ARM processors. High computing abilities of these processors allow to realize not only simple PID-controllers but more sophisticated adaptive controllers. Adaptive digital controllers are indispensable for technological processes which require high quality of control. In this case oscillation and overshooting are inadmissible and setting time must be minimal. And digital controllers are able to improve process control performance significantly. The theory of digital control systems was developed in 70-80 years of the last century. In particular, K. Astrom and B. Wittenmark [2] and R. Iserman [3] showed that digital controllers are the best when aperiodic transient processes are wanted and described how to make state variable modal digital controller capable to provide any in advance known characteristics of the transient process. When an object of

control is timeinvariant it is possible to use experimental data and find out the coefficients of the digital transfer function of this object and the coefficients of digital modal controller and digital observer preliminarily. These calculations are carried out only once and their results may be used as the constants in the program for direct digital control. But when the object of control is unstable i.e. its characteristics drift with the time, or its characteristics are non-linear and its linear approximation depends on operating point, all above mentioned calculations must be made repeatedly at run time within controller itself with the pace of technological process. And in this case time consumption for such calculations may be crucial. In other case quality of control may decrease drastically.

2. Using least squares method for parameter evaluation

The most substantial part of calculations in discussion is the object's parameter estimation, i.e. the process of finding out its digital transfer function coefficients. And the task of this article is to show how time consumption of the program for parameter estimation may be reduced. The novelty of this work is in making new program realization of the least-squares method that includes refusing of function decomposition of the code. As a result, some intermediate matrices may be dropped and outer loops of sequential stages of calculations may be linked. Such a decision allows reducing the program's time and memory consumption that is very significant in the case of real-time applications based on the microcontrollers. Also, for further optimization, the code

*Corresponding Author: Vasily Olonichev, v_olonichev@ksu.edu.ru

parallelism and different ways of data parallelism are added to the program and a comparison of the obtained results is made.

The general method of the object of control parameter estimation is the least-squares method or LSQ [4]. Using LSQ, one can get a transient function of the object in a discrete form. The coefficients of this transfer function are used to find out parameters for tuning adaptive dynamic regulators [3].

Despite the fact that LSQ is a rather old method, there are still a lot of publications connected with it. Each of them is devoted to special issue. For example, identification of nonlinear systems [5], time-varying magnetic field analysis [6], signal processing [7], lines approximation in multidimensional space [8]. But among them, researches connected with LSQ utilization at microcontrollers when amount of RAM is limited and time constraint are strict are not found.

LSQ is widely used in many fields of applied calculations and it's realization is available in many program libraries. First of all must be mentioned lapack library (<https://www.netlib.org/lapack/>) - standard library that may be found in many Unix-like operating systems. Another widely used library is gnu scientific library or gsl (<https://www.gnu.org/software/gsl/>). These libraries use QR-decomposition [9] for solving LSQ problem. This method is considered to be one of the fastest for the big dimension tasks.

But the objects of control usually are described with the models that have order between 2 and 6. So it is reasonable to suggest that program based on QR-decomposition will spend more time making preliminary computing than useful one when applied to such small systems. And with the micro-controllers situation when it will be not enough memory for such libraries is also possible.

To check this hypothesis the program was written that applied matrix operations as LSQ suggests [9] and as is shown in (1):

$$\begin{bmatrix} A \\ B \end{bmatrix} = [X^T \cdot X]^{-1} \cdot X^T \cdot Y \quad (1)$$

where X – matrix with the dimension of Nx2M filled with values taken from input and output of the object with the regular intervals of time; Y – vector with the dimension of N filled with the values taken from the output of the object; A – output vector with the dimension of M having coefficients of transfer function denominator; B – output vector with the dimension of M having coefficients of transfer function nominator; L – number of experimental points; M – order of the object of control; N = L – M – number of equations.

Covariance matrix is represented in (2):

$$C = [X^T \cdot X]^{-1} \quad (2)$$

Equation (2) that is the part of (1) plays special role and is used for evaluation of parameters dispersion [10].

How to fill matrix X and vector Y with the experimental values is shown in [3].

Other reasons to write such program are the following:

At the first, for correct parameter estimation of the object of control the trace of covariance matrix must be minimal and the

value of this trace depends mainly on sampling period [11]. This means that at the first stage of parameter estimation time step may vary and alongside with the model's parameters the trace of covariance matrix must be computed. Getting the covariance matrix is the part of LSQ method, so it's trace may be computed when direct matrix operations are involved. While using library functions this data are hidden inside them. And to get covariance matrix one needs to repeat a bigger half of computations already made.

At the second, during LSQ parameter evaluation some dynamically allocated matrices are used for storing intermediate data. Each time when library function is called the memory is allocated for them and then released. While using our own code we can allocate memory for intermediate matrices only once and then use them with every next time step.

As equation (1) shows LSQ consists of matrices multiplication and matrix inversion. Matrices multiplication may be written from scratch. An example of matrix inversion can be found in the Internet (<http://www.programming-techniques.com/2011/09/numerical-methods-inverse-of-nxn-matrix.html>). This program realizes the Gauss-Jordan method [12].

3. Program Testing and Optimization

3.1. General Information about Testing

For realization and testing of the programs singleboard mini-computers CubieBoard-3 and Odroid-C2 were used. CubieBoard-3 has two core CPU CORTEX-A7 (ARMv7) with the frequency 1GHz and 2G of RAM. Operating system Linux Ubuntu 18.04.1 with the kernel 4.19.57 is installed on this computer alongside with gcc compiler v 7.4.0. Odroid-C2 has quad core CPU CORTEX-A53(ARMv8) with the frequency 1.5GHz and 2G of RAM, operating system Linux Ubuntu 16.04.09 with the kernel 3.14.79 and gcc compiler 5.4.0. For the conference paper [1] programs were made with the gcc v 4.6.3, so results presented in this article may slightly differ, first of all due to the fact that the realization of optimization in these compilers is not identical.

For program realization, the C++ language was chosen. It's a common practice nowadays even for embedded systems. If you do not use classes with the virtual functions, the productivity of the result code is almost the same, and at the same time, the full power of C++ as a language of generic programming is available.

All programs were compiled with the -O2 level of optimization.

Working with the matrices, one must decide how to store them in memory. First way suggests using dynamic one dimensional flat vectors. Matrix elements in this case are accessible with the function or overloaded operator () taking as arguments row and column number. E.g. getElem(X,i,j) or X(i,j). Second way suggests using dynamic two dimensional arrays. In this case X[i] is a vector of pointers containing the addresses of matrix's rows and X[i][j] is an element of matrix. First way is considered to be slightly faster as the data occupy continues space in memory. But for an adaptive controller second way is more suitable, because in this case matrix X must be renewed with every next step in time. Old data must be removed and new once added. With the flat vector all elements counted with thousands will be moved within

this vector from the end to the beginning. But with the two-dimensional array only pointers counted with tens will be moved.

3.2. Comparison of Code Using Library Functions and Code Using Direct Operation with the Matrices

To compare time consumption for LSQ 3 programs were written. The first used function `dgel` from the `lapack` library, the second used function `gsl_linalg_QR_ksolve` from the `gnu scientific library` and the third realized equation (1) using self-written functions for matrix multiplication and inversion. Time consumption for the calculations was found out as the difference between the time measured before and after the calculations. To measure time the function `clock_gettime` was used.

Time consumption were determined with the matrices of the following dimensions $M \times N$: 2×40 , 3×60 , 4×80 , 5×100 , 6×120 , 7×140 , 8×160 and 10×200 . Objects of higher orders require more experimental points. But this does not mean that to evaluate parameters of the for example 4 order object, one must use exactly 80 experimental points. It is just an average value.

The matrix X and vector Y were filled with the random numbers because we need to get not the results of LSQ-evaluation but only time consumption for getting them.

Calculations for each program and for each dimension were repeated 5 times. Maximum and minimum were removed and out of the rest three measures, an average value was calculated. To run programs, calculate time consumption and to make plots the special script in Python language was written. The resulting plots with the time consumption against object order are shown in Figure 1 for CubieBoard3 and Figure 2 for Odroid-C2.

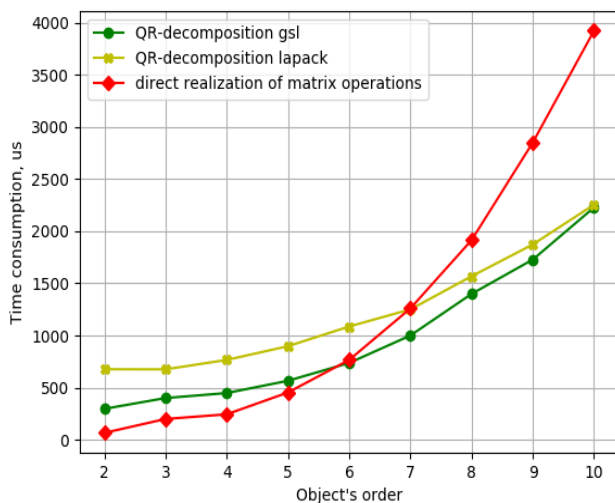


Figure 1: Time consumption for LSQ parameter estimation against object's order by the programs with the double precision numbers on CubieBoard-3.

The plots in Figures 1 and 2 show that the program that uses matrix operations takes less time than library functions for the tasks of small dimensions. But this program must be rearranged in order further to improve its efficiency.

First, decision must be made is there any sense to replace the double precision variables with the single precision variables or with the fixed-point ones.

3.3. Time consumption for Carrying out Arithmetic Operations

For this purpose, it will be useful to find out how many time take arithmetic operations with the operands of the different types. And proper program with the three operands expression was written and run on CubieBoard-3 and Odroid-C2 computers:

$$a = b \# c$$

where $\#$ in turn is $+$ - $*$ and $/$; $b=2$; $c=3$.

The results are shown in Table 1.

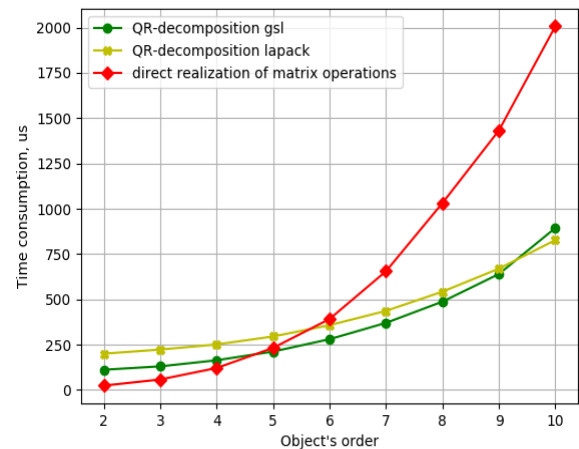


Figure 2: Time consumption for LSQ parameter estimation against object's order by the programs with the double precision numbers on Odroid-C2.

Table 1: Time Consumption in μs for Carrying out 1000 Operations on Single-Board Computers CubieBoard 3 and Odroid-C2

Variable type	Operation			
	Addition	Subtraction	Multiplication	Division
CubieBoard-3				
int	13,5	13,5	14,6	30,2
long	13,5	13,5	14,6	30,3
float	14,6	14,6	14,6	29,2
double	14,6	14,6	17,7	43,8
Odroid-C2				
int	7,1	7,1	8,5	8,5
long	7,1	7,1	9,1	8,5
float	9,8	9,8	9,8	15,6
double	9,8	9,8	9,8	21,5

As Table 1 shows, the time consumption for processing integer variables is less than the time for processing floating point variables by 8% for `armv7` and 28% for `armv8`. But realization of LSQ in adaptive controller requires support of numbers in wide range of values. For example, parameter evaluation of the 3 order object using the experimental results where input varies between 0.0 and 1.0 and output between 0.0 and 200.0 will give values in the intermediate matrices varying from $1.0e-3$ to $1.0e6$. Obviously, integer values can't be used for such computing.

Time consumption for single precision and double precision floating point variables is identical for addition, subtraction and multiplication at both platforms. But division takes considerably more time especially for double precision variables.

3.4. Using Single Precision Variables and Parallelism of Code

Time consumption of the program depends not only on time required for the computing but also on cache misses [13]. Float variable takes 4 bytes and double takes 8 bytes. That means that the processor's cache can hold more data in case of float variables and so, cache misses will be met less often.

There is one more reason to use variables of single-precision:

- The vector unit of ARM-CORTEX-A (ARMv7) processes only single-precision floating-point numbers, and vectorization is a significant source of increasing program efficiency.
- Microcontrollers STM32 has hardware support only for single precision, and software emulation of double precision is rather slow.

Usually, it's recommended to avoid single precision variables in calculations [14]. But in our case, the dimension of the task is not big. And if to use in the experiments optimal sampling period, well-conditioned matrices will be obtained [11]. LSQ identification of the same object made with double precision and single precision numbers is almost identical. For these calculations, real experimental results obtained with the object with orders 3 and 4 were used.

There is also one more source of program efficiency improving. It is parallelism of code. And gcc compiler supports OpenMP specification, which provides parallelism or multi-threading. The results obtained are presented in Figures 3 and 4.

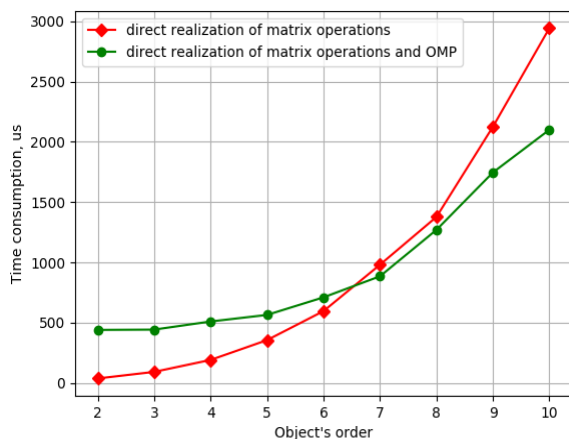


Figure 3. Time consumption for LSQ parameter estimation against object's order by the programs with the single precision numbers on CubieBoard-3.

The CubieBoard3 computer has 2 core processor and Odroid-C2 has 4 core processor.

Figures 3 and 4 show that using single precision variables instead of double has given about 25% increase in productivity. At the other hand on CubieBoard-3 parallelism has given expected results. I.e. while working with the small matrices synchronization between threads takes more time then parallelism saves it. And for

objects with order 7 and higher parallel program becomes faster. But the results obtained on Odroid-C2 show that parallelism instead of increasing productivity reduces it in the whole range. The explanation of this fact may be following. The program calls several functions, and directives for parallel code are placed within them. So OMP preprocessor generates code that creates and cancels threads within each function. At CubieBoard-3 more modern gcc compiler was used with this problem fixed.

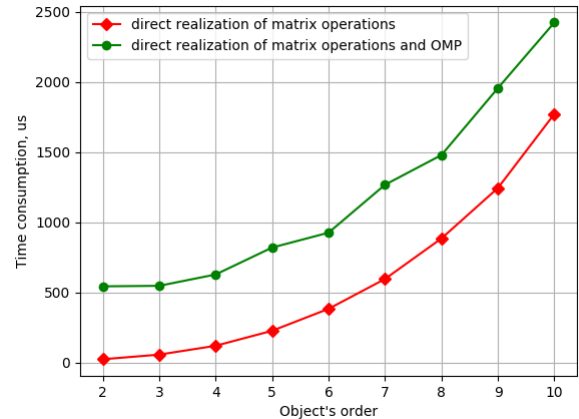


Figure 4. Time consumption for LSQ parameter estimation against object's order by the programs with the single precision numbers on Odroid-C2.

3.5. Refactoring of the Code

Next stage of optimization is concerned the code itself.

The first stage is also connected with the problem of cache-missing. In the programs written in C, matrices are allocated in memory row-wise. During multiplication one of the matrices is scanned row by row, and a big piece of data is loaded into cache. But another matrix is scanned column by column, and getting the next element may come to a cache-miss. If preliminary to transpose another matrix it will be also scanned row by row [13].

There is a division in the inner loops of the function that makes matrix inversion. If to calculate the inverse number in the outer loop and replace division with the multiplication we can get another source of the productivity raising.

From the point of view of the structural programming the code must be divided into functions each of them makes logically complete operation. In our case these are matrix multiplication and inversion. But its also known that such structural decomposition may reduce productivity of the program.

If to make one function that makes all calculations in one step it will be possible to take into account specific properties of computing. Informational matrix $X^T \cdot X$ is symmetric and it's possible to calculate only half of it. And also, it's possible to get away two intermediate matrices with the sizes $[2 * M][2 * M]$. And intermediate matrix $[2 * M][N]$ may be replaced with a vector with the size $[N]$. As a result, the number of cache-misses and total consumption of memory will be reduced. The last is especially significant for STM32 microcontrollers with limited RAM. In this one function, it's also possible to improve an algorithm. In this new program, the next stage of calculation will be started within an outer loop of the previous stage.

Alongside with the algorithm improvement, OMP directives were added to the code in order to apply parallelism. The resulting program was compiled twice. First time with an OMP flag to get parallel code. And second time without this flag to get code without multithreading.

Calculations were repeated with the same data. The results are presented in Figures 5 and 6.

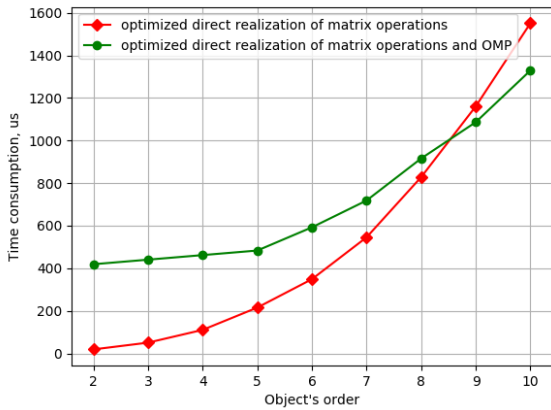


Figure 5: Time consumption for LSQ parameter estimation against object's order by the programs with the optimized function on CubieBoard-3.

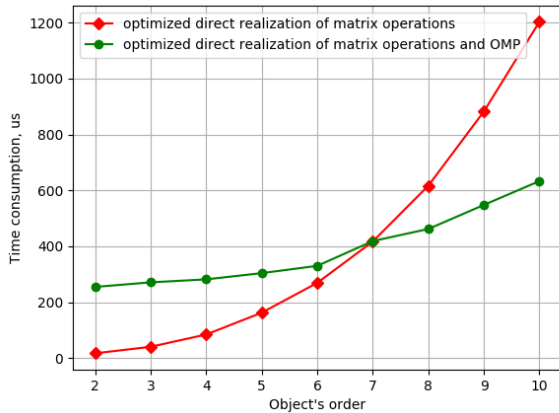


Figure 6: Time consumption for LSQ parameter estimation against object's order by the programs

Plots in Figures 5 and 6 show that optimization of code gave increasing in speed about 80% for CubieBoard-3 and 50% for Odroid-C2. And using parallelism in one function has given results on Odroid-C2. But also as one can see parallel code is faster only with 7 and higher-order object models. For the object's model with the order from 2 to 6 nonparallel code is faster and it may be recommended for practical utilization.

And more significant is the fact that the optimized program is faster than programs using lapack and gsl libraries in the whole range from 2 to 10 at both mini-computers.

3.6. Using Vector Operations

Vectorization is another way to increase program efficiency. ARM-CORTEX-A processors have the vector unit named NEON. This unit has 128-bit registers that enough to keep 4 floating-point numbers. The vector instructions process all numbers in the vectors at one time. Theoretically, it can increase efficiency by 4

times. But usually, this value is less because switching processor's pipeline between vector and regular registers takes a lot of time. (<https://developer.arm.com/products/processors/cortex-m/>).

There are many options to use vectorization: special libraries, auto-vectorization of compiler, OMP directives, NEON intrinsics, and assembly code. In our case, the best decision is to use intrinsics. They provide access to all vector instructions and allows them to apply total control of instruction flow comparing with the auto-vectorization. The efficiency of such code is close to the assembly one.

For matrices processing, data parallelism may be used in two different ways. In the first case, the elements of the matrix are loaded into the vector register horizontally, first elements of the row with the indexes from 0 to 3, then from 4 to 7 and so on. For each subset of data vector instruction multiplication with accumulation is used. After the loop is finished, the dot product is obtained as a sum of four elements of the vector register. This method may be called horizontal vectorization and is simple for realization both for matrices multiplication and matrix inversion.

The second approach that may be called vertical vectorization requires source matrices to hold data in a vector format of type float32x4_t in columns (<https://community.arm.com/processors/b/blog/posts/coding-for-neon---part-3-matrix-multiplication>). In this case, code for matrices multiplication is very simple and efficient. But during matrix inversion, non-vector variables are used rather often.

To compare these variants of vectorization, the programs were made for both of them. These programs run on CubieBoard3 and Odroid-C2 computers with the same initial data as previous programs. Resulting plots of these tests are shown in Figure 7 and 8.

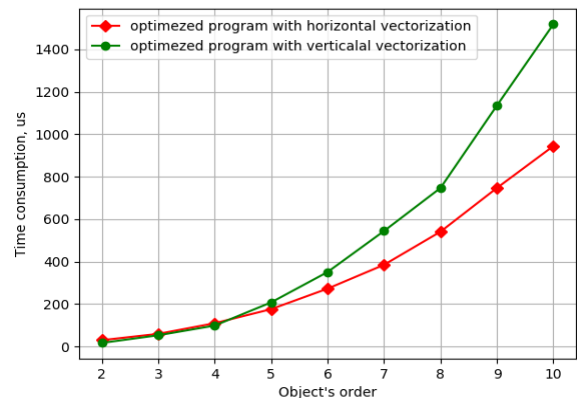


Figure 7. Time consumption for LSQ parameter estimation against object's order by the programs using vector unit NEON on CubieBoard-3.

As one can see in both cases horizontal vectorization gives better results. And further improvements and checks will be connected only with it.

3.7. Data Alignment

The next problem is data alignment. Old gcc compilers for armv7 required the directives explicitly showing that elements of the matrix rows and vector variables are aligned in memory at the boundaries multiple for 64 (<http://infocenter.arm.com/help/index.jsp?topic=/com.arm.doc.dd>

i0344k/Cihejdic.html). To find out the effect of the data alignment program with horizontal vectorization was modified. Compiler's directives `__attribute__((aligned(64)))` and `builtin_assume_aligned` were added to it.

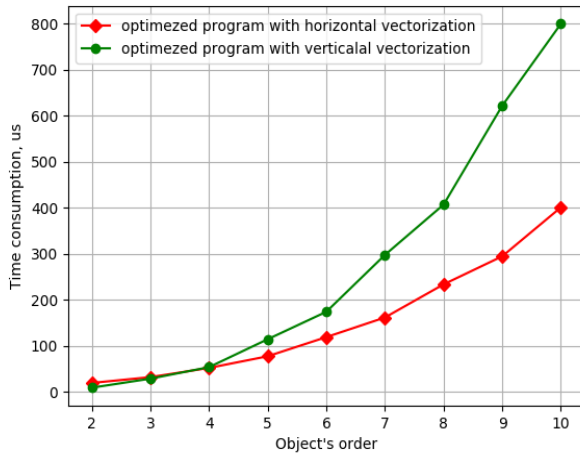


Figure 8. Time consumption for LSQ parameter estimation against object's order by the programs using vector unit NEON on Odroid-C2.

For comparison, together with 64, the numbers 8, 16, 32 and 128 were used for alignment. Running these programs along with the program without alignment directives has shown that there is no any difference between their time consumption. That means that modern gcc compiler makes data alignment without the additional directives.

3.8. Working with the Leftovers

The next problem one always meets while working with the vectorization is leftovers. ARM NEON vector register holds 4 single precision floating point numbers. The matrices sizes in real tasks are not multiple to 4, so the leftovers which can't be loaded into the vector register directly must be processed in some manner. Two ways to solve this problem are suggested (<https://community.arm.com/developer/ip-products/processors/b/processors-ip-blog/posts/coding-for-neon--part-2-dealing-with-leftovers>). The first is to process the leftovers as non-vector data. And the second is to extend matrices to the sizes multiple to 4 and fill the edges with the zeros. The second way is considered to be faster because the vector's processing is not interrupted with the non-vector operations.

In our case using extended matrices makes the code more sophisticated. Function for parameter evaluation takes two additional parameters and intermediate matrix S, holding informational and covariance square matrices side by side must be filled and processed in a not obvious way. As a result, the code of the function is tightly coupled with the rest of a program.

To avoid such a problem, the third way to solve leftovers problem were suggested. In this case function for parameter evaluation has local floating point arrays with the sizes equal to 4. The leftovers are loaded into these arrays before the main cycle of processing and the arrays are processed after the main cycle. These local vectors are extending each row of the matrix in a turn. As a result, all specific features connected with the vectorization are hidden within the function.

All three programs with different ways of solving the leftovers problem were checked for time consumption. The results are presented in Figures 9 and 10 for CubieBoard-3 and Odroid-C2 correspondingly.

Plots in Figures 9 and 10 show that variant with the local vectors is the worst from the point of view of time consumption, and the best is variant with the extended matrices. The difference for the object's order from 4 to 6 is about 50% for armv8 processor and 75% for armv7. This difference is significant, so using extended matrices for the solving of leftovers problem must be recommended.

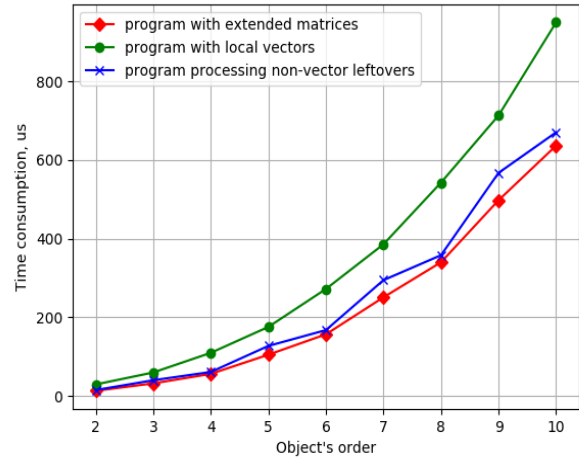


Figure 9. Time consumption for LSQ parameter estimation against object's order by the programs using vector unit NEON and different ways of leftovers handling on CubieBoard-3.

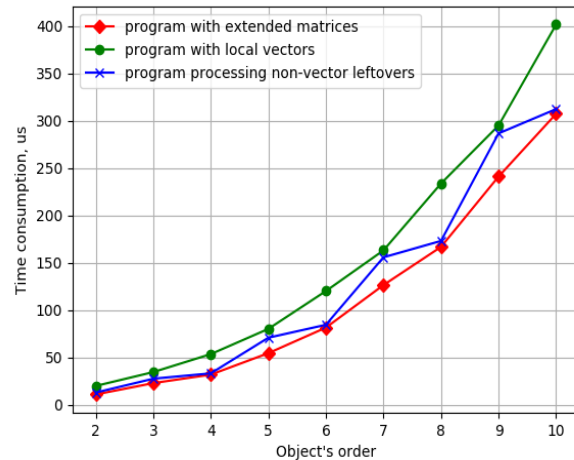


Figure 10. Time consumption for LSQ parameter estimation against object's order by the programs using vector unit NEON and different ways of leftovers handling on Odroid-C2.

3.9. Compare of all Results Obtained

All experimental data received in this work are presented in Table 2.

Types of the programs: 1 – QR-decompozition and dgels form lapack library; 2 – QR-decompozition and lsqsolve form gsl library; 3 – direct realization with the double precision numbers; 4 - direct realization with single precision matrices; 5

- direct realization with single precision matrices and multithreading; 6 – optimized function code; 7 – optimized function code and multithreading; 8 optimized function code with horizontal vectorization; 9 – optimized function code with vertical vectorization; 10 – optimized function code with horizontal vectorization and extended matrices; 11 – optimized function code with horizontal vectorization and non-vector leftovers.

Table 2. Time Consumption in μ s for LSQ Parametrical Identification on Single-Board Computers CubieBoard-3 and Odroid-C2 for different Types of the Programs

Type	Object's order								
	2	3	4	5	6	7	8	9	10
CubieBoard-3									
1	675,8	675.0	765.4	897,4	1086,1	1250,7	1568,6	1873,9	2255,9
2	295.6	399.3	447.3	565.0	736.3	999.8	1399.4	1728.7	2228.7
3	64.7	198.3	243.2	452.7	765.3	1259.5	1915.6	2852.7	3925.8
4	37.5	92.3	191.4	355.8	596.5	981.2	1377.6	2125.9	2943.5
5	440.3	442.8	510.3	564.7	710.7	885.3	1270.5	1746.3	2100.6
6	18.8	50.5	110.5	215.2	348.6	545.3	828.2	1160.9	1551.4
7	419.0	440.4	461.9	483.3	592.3	718.2	915.9	1086.5	1329.3
8	29.5	59.8	109.5	175.7	272.3	384.7	540.6	746.7	944.2
9	16.5	52.7	98.3	207.5	350.6	543.8	746.6	1135.2	1518.2
10	13.02	31.8	56.0	104.8	156.5	250.7	339.8	497.4	636.3
11	14.9	39.8	60.8	127.1	167.3	294.3	357.7	567.0	670.1
Odroid-C2									
1	200.0	222.7	251.0	295.0	357.3	436.3	542.3	669.7	827.0
2	111.3	130.0	163.7	211.7	280.0	370.0	487.3	639.7	894.0
3	23.7	57.6	121.0	231.6	391.3	655.0	1031.7	1432.3	2007.3
4	24.0	57.0	120.0	226.7	383.0	595.7	884.3	1243.7	1771.0
5	543.3	547.0	627.7	819.3	925.3	1266.7	1478.3	1955.0	2422.3
6	17.0	40.3	84.3	162.3	269.0	417.0	616.7	882.4	1203.4
7	254.6	271.0	281.7	304.0	330.0	418.7	462.0	548.0	633.4
8	19.3	32.0	52.0	77.3	119.0	161.7	233.7	294.7	401.0
9	9.0	28.7	53.7	114.3	174.0	297.7	407.0	622.0	800.0
10	10.7	22.7	31.7	54.3	81.3	126.3	166.7	241.0	307.5
11	12.7	27.33	33.0	70.7	84.3	155.7	173.0	286.6	312.5

As Table 2 shows, the measures taken to optimize the code of function for LSQ parameter estimation of the object of control have allowed to reduce time consumption in 4.7 times for the processor armv7 and in 3.5 times for armv8 for the objects with the order from 4 to 6. Such significant productivity-increasing may allow to widen substantially the area of utilization of the adaptive digital controllers.

The source code of the programs tested in this article is available for free downloading under GPL license (https://github.com/basv0/lq_armv7).

4. Conclusion

Wide using of the microprocessors systems that may provide high quality of technical objects control is restrained with the complexity of parameters estimation of these objects. As a result, the adaptive control may take more time than the technological process allows in hard real-time systems. This conclusion follows from the results of time consumption comparison for LSQ parameter estimation by the programs using function `dgel` from the linear algebra library `lapack`, function `lsqsolve` from scientific library `gsl` and functions for matrix multiplication and inversion. The tests show that direct realization of the matrix operations is

more preferable for the tasks of not big dimensions and that using a single-precision floating-point variable instead of double precision ones does not decrease the calculations accuracy for the objects with the order less than 10.

Applying multi-threading showed that it gives productivity-increasing only for the objects with the order higher than 7 while in practice objects with the order between 2 and 6 are mainly met. Increasing productivity in the whole range from 2 to 10 may be achieved by the code refactoring and using intrinsic functions for the vector computations.

An optimized function using horizontal vectorization and extended matrices with the dimensions multiple by 4 has shown the best results. And this function is recommended for practical utilization despite the fact that matrix extension makes the code more sophisticated.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] V.V. Olonichev, B. A. Staroverov and M. A. Smirnov, "Optimization of Program for Run Time Parametrical Identification for ARM Cortex Processors," *2018 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, Moscow, Russia, 2018, pp. 1-5. doi: 10.1109/ICIEAM.2018.8728595.
- [2] K. Astrom and B. Wittenmark, *Computer Controlled Systems*. Prentice-Hall, Inc., 1984.
- [3] R. Izerman. *Digital Control Sytems*. Springer-Verlag, Berlin, Heidelberg, New York, 1981.
- [4] L. Ljung, *System Identification: Theory for the User*. Prentice-Hall, 1987.
- [5] Z. Tan, H. Zhang, J. Sun et al., "Research on Identification Process of Nonlinear System Based on An Improved Recursive Least Squares Algorithm", *Proceedings of the 31st Chinese Control and Decision Conference, CCDC 2019 8832530*, pp. 1673-1678.
- [6] M. Arehpanahi, H.R. Jamalifard, "Time-varying magnetic field analysis using an improved meshless method based on interpolating moving least squares", *IET Science Measurement Technology*, vol. 12, no. 6, pp. 816-820, May. 2018.
- [7] H. Li, J. Zhang, J. Zou, "Improving the bound on the restricted isometry property constant in multiple orthogonal least squares", *IET Signal Processing*, vol. 12, no. 5, pp. 666-671, Apr. 2018.
- [8] V. Skala, "A new formulation for total Least Square Error method in d-dimensional space with mapping to a parametric line ICNAAM", *2015 AIP Conf. Proc.1738*, pp. 480106-1-480106-4, 2016.
- [9] H. Leslie, *Handbook of Linear Algebra*. – CRC Press, 2013.
- [10] C.F. Jeff Wu, Michael S. Hamada *Experiments: Planning, Analysis, And Optimization*. Wiley, New Jersey, 2009.
- [11] B.A. Staroverov, V.V. Olonichev and M.A. Smirnov, "Optimal sampling period definition for the object identification using least squares method", *Vestnik IGEU*, vol 1, pp 62-69, 2014. (article in Russian with an abstract in English)
- [12] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, New York, Port Chester, Melbourne, Sydney, 1992.
- [13] U. Drepper. *What Every Programmer Should Know About Memory*. <https://www.akkadia.org/drepper/cpumemory.pdf>.
- [14] B. Stroustrup, *The C++ Programming language*. Addison-Wesley, 1997.

Design, Implementation and Performance Analysis of a Dual Axis Solar Tracking System

Ba Thanh Nguyen¹, Hong-Xuyen Thi Ho^{2,*}

¹*Faculty of Engineering and Technology, Thu Dau Mot University, Thu Dau Mot City 75109, Vietnam*

²*Faculty of Economics, Ho Chi Minh City University of Technology and Education, Ho Chi Minh City 71307, Vietnam*

ARTICLE INFO

Article history:

Received: 09 February, 2020

Accepted: 22 April, 2020

Online: 03 May, 2020

Keywords:

Solar Tracking System

Servo Motors

Arduino

ABSTRACT

This study presents the design and construction of the Dual Axis Solar Tracking System to ensure maximum energy gain. The solar tracking system will automatically follow the sun's position to maximize the intensity of the light emitted from the sun. When the light intensity decreases, the system automatically changes its direction to get the maximum light intensity. Light Dependent Resistor (LDR) is used to track the coordinates of the sun. The two servo motors that receive signals from the central processing unit will turn the solar panel to the appropriate location for optimum performance. The energy results obtained by the dual-axis solar system are compared with single and fixed solar systems. This research provides optimal solar energy usage.

1. Introduction

Nowadays, environmental pollution is getting more serious, traditional energy sources such as oil, coal, etc. are gradually exhausted. However, the global demand for energy is forecasted to increase in the upcoming decades. Thus, many countries are promoting the development of alternative energy sources [1].

Renewable energy is an energy source that does not consume the limited resources of the earth. It can be easily and quickly replenished. Renewable energy plays an important role in meeting energy demand and contributing to climate change control [2].

Among renewable energy sources such as solar energy, wind energy, biomass energy, tidal energy, etc., Solar energy can be easily exploited in Vietnam because Vietnam has a high level of radiation, and advances in solar technology have made production costs more and more affordable [3] [4].

Because the earth is always orbiting its axis, the angle of sunlight is constantly changing, affecting the amount of solar radiation that is projected on the solar cell. To increase the efficiency of absorbing light from the sun, we need a system of devices called the solar tracking system that help the photovoltaic panels orientate to the sun.

The solar tracking system includes tracking installation system, solar panels and tracking control system. The mounting structure

is the body of the system, and the control system is its brain, which controls the movement of the solar panel system.

Many authors have studied solar tracking. Mayank Kumar Lokhande [5] presented an automatic solar tracking system. The study designed a microcontroller-based solar panel tracking system and observed that the single-axis solar tracking unit increased efficiency by 30% compared to the fixed panel.

Guiha Li, Runsheng Tang and Hao Zhong [6] experimented with a single horizontal solar tracking system. The study found that the east-west-oriented system was not significantly improved while the south-north oriented system was the best. The efficiency increase for the east-west axis is below 8% while for the south-north axis it increases 10-24%.

Chaiko and Rizk [7] have developed an effective solar tracking system. The authors designed a simple axis tracking system by using a stepper motor and light sensor. They observed that this system increased the efficiency of collecting energy by keeping a solar panel perpendicular to the sun's rays. The study also found that the electrical power increased by 30% compared to static systems.

Imam Abadi, Adi Soeprijanto and Ali Musyafa [8] designed a single-axis solar tracking system based on fuzzy logic. They performed on ATMEGA 8353 microcontroller to improve the power of the panels. The study presented that the performance of solar tracking systems with fuzzy control increased by 47% compared to static systems.

*Corresponding Author: Hong-Xuyen Thi Ho, xuyenhth@hcmute.edu.vn

Ashwin R, Varun A.K et al. [9] showed a single-axis solar tracking system based on sensors to achieve the highest energy level. The system automatically changes in the direction of solar panel to get maximum light energy. Therefore, the test results show the robustness and productivity of the proposed method.

In 2013, Anusha, Chandra and Reddy [10] designed a solar tracking system based on a real-time clock. They compared a static system and a real-time system using an ARM processor. Experiments have proven that the new system is about 40% more efficient.

Hussain S. Akbar [11] designed a single-axis solar tracking unit using the AVR microcontroller. The results show that the designed solar tracking system has improved the output power by 18-25% compared to the static panels in Kirkuk, Iraq.

In 2016, H. Fathabadi [12] tested that the energy level obtained from the dual-axis solar tracking system was 28.8% - 43.6% higher than static systems depending on the season.

Jing-Min Wang and Chia-Liang Lu [13] implemented a dual-axis solar tracking system in New Taipei City, Taiwan. Experiments show that their system boosts energy levels to 28.31% for a cloudy day.

Munna, M. S et al. [14] constructed and evaluated the performance of dual-, single- and fixed-axis solar tracking systems, the data showed that the dual-axis system is more optimal than the two systems. again.

Mustafa, F. I. et al. [15] implemented a dual-axis solar system, data collected in one day showed that the efficiency of this system is 35% higher than that of fixed systems.

The authors Chhoton, AC, and Chakraborty, N. R [16] performed a performance evaluation of the dual-axis solar tracking system, the data collected during the day showed that the performance of this system was higher than the fixed system. 40%.

In this study, the authors designed and constructed a dual-axis solar tracking system with a simple and low-cost structure that still meets the requirements of increasing the efficiency of solar energy exploitation. The automatic operating system is controlled by Arduino microcontroller. The results of this study provide optimal solar energy solutions. Moreover, the results of this study provide an optimal solution for solar energy, helping to improve the efficiency of the solar power system, effectively serving the energy production process.

This paper is organized as follows. Part 2 discusses the operating principles of a solar tracking system. Part 3 presents hardware and software design. Part 4 presents the test results, compared with the static system. The last part is the conclusion.

2. Overview of Solar Tracking System

2.1. A Solar Tracker

The solar tracking system was created to take advantage of solar radiation as efficiently as possible, increasing the amount of energy radiated to the solar rig. Thus, the amount of electricity generated will be greater than the fixed devices [17].

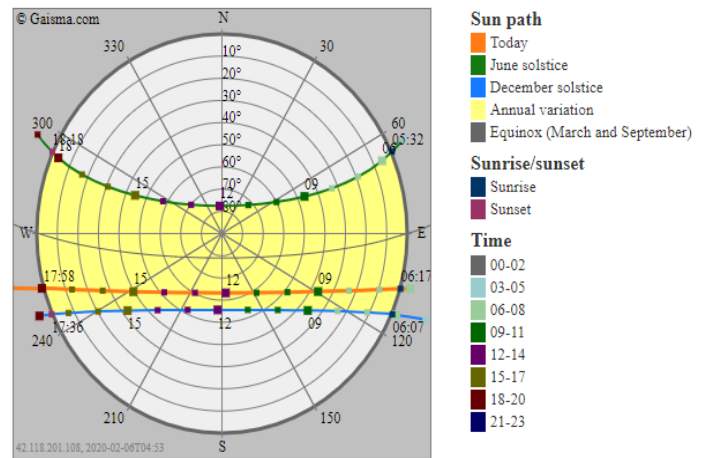


Figure 1: Sun path in Thu Dau Mot City, Vietnam [18]

A solar tracking system consists of three components: a mechanical mechanism, a solar panel and a control system.

2.2. Types of solar tracking

Solar tracking system plays an important role in the development of solar energy applications, in order to improve the efficiency of solar power systems.

According to the mechanical structure, the solar tracking system has 2 types: single and dual-axis systems [12], [19]. Of these, dual shafts are often used because it provides higher accuracy and is known to improve solar power captured capacity compared to single-axis tables [13]. In addition, there is a way to classify according to control with a positive and passive system [7-9].

2.2.1. One-Axis Trackers

The single-axis solar tracking system uses an inclined PV mount bracket and an electric motor to move the board in orbit closer to the position of the sun. Spindle can be horizontal, vertical or inclined. Figure 2 shows a general diagram of a single-axis tracker that shows both the axis of rotation (unit vector e) and the collector plane (the unit vector is normal for the collector plane). The angle between these two unit vectors is usually kept constant in this type of solar tracker.

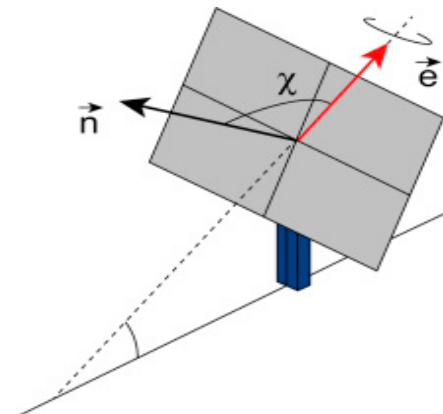


Figure 2: Characteristic vectors in a one-axis tracker [19].

2.2.2. Two-Axis Trackers

The two-axis solar tracking system can achieve the maximum power level because, because it is completely free to move in two directions, it is capable of tracking the sun anywhere.

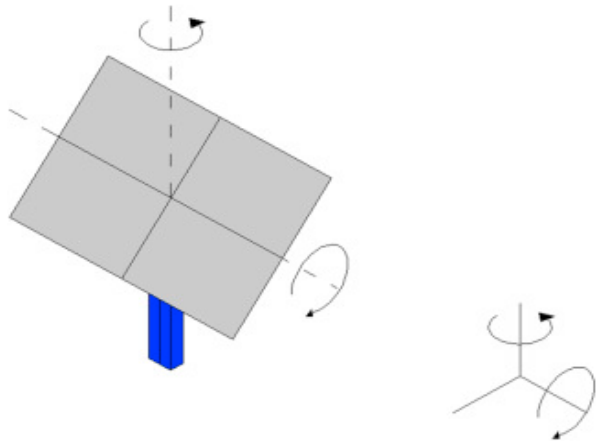


Figure 3: Characteristic movements in a two-axis tracker [19].

3. Proposed Solar Tracker System

3.1. Architecture of The Overall System

Our proposed design is a dual-axis solar tracking system, based on feedback loops. The system consists of optical resistors acting as sensors, servo motors, actuators, and Arduino microcontrollers. The whole system is divided into two main parts: hardware and software. The main equipment is listed in Table 1 and the system has diagrams as shown in Figure 4.

Table 1: List of devices

No.	Item Name	Quantity
1	Arduino Uno R3	1
2	MG996R Servo Motor	2
3	Light Dependent Resistor	4
4	Module LM2596	2
5	Solar Panel 10 W	1
6	Solar Charge Controller	1
7	Battery 12V 5Ah/10HR	1
8	DSN-VC288 DC 100V 10A Voltmeter Ammeter	1

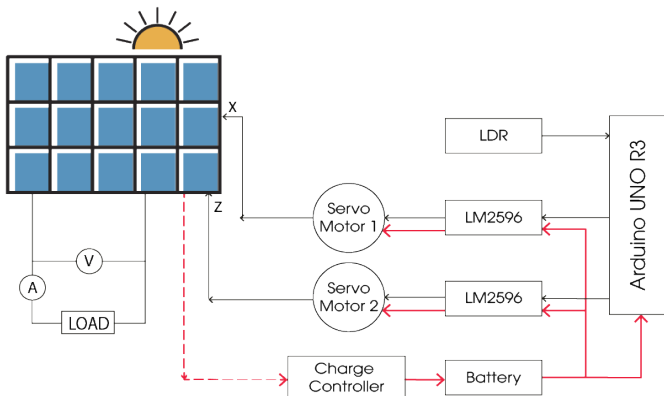


Figure 4: Block diagram of overall system

3.2. Hardware Design

Figure 3 depicts the structure of the model including 2 main components: fixed and mobile parts. Fixed parts are the base of the system; the movable joint is attached to the servo motor, placed above and below to rotate the solar panels in two directions.



Figure 5: Complete hardware setup of a solar tracking system

3.3. Software Design

The sensor system consists of 4 resistors (R1, R2, R3, R4) that will receive light from the light source. Between these 4 resistors, there will be a cross-shaped partition dividing the 4 optical sensors into 4 different directions. This partition will serve as a guide for the 4 optical barriers always towards the strongest light source, namely the sun. When the sensor assembly is perpendicular to the radiation of the sun, the values of the 4 sensors will be equal.

The signal from the sensor assembly will be transmitted directly to the Arduino control center and converted into a digital signal. Here, the Arduino compares the average of the two adjacent optical sensors to the average of the two opposite sensors.

$$A = (R1 + R2)/2 \tag{1}$$

$$B = (R3 + R4)/2 \tag{2}$$

$$C = (R1 + R3)/2 \tag{3}$$

$$D = (R2 + R4)/2 \tag{4}$$

For R1, R2, R3, R4 are the resistance values of the 4 optical resistors LDR1, LDR2, LDR3, LDR4, respectively. These values are inversely proportional to the values of voltage transferred to the Arduino.

If A = B, the servo motor controls the X-axis.

If C = D, then the Z-axis servo motor is stationary.

If A > B, it means that the light is more concentrated on R3, R4 side, 1st servo motor rotates the panel downwards until the light is perpendicular to the panel then stops.

If $A < B$, which means more light is concentrated on R1, R2 side, 1st servo motor rotates the panel upwards until the light is perpendicular to the panel then stops.

Similarly, if $C > D$, which means more focused light on side R2, R4, servo motor 2 will rotate the solar panel to the right until the light is perpendicular to the solar panel.

If $C < D$, it means that the light on the side of R1 and R3 is more concentrated, the servo motor 2 will rotate the battery plate to the left until the light is perpendicular to the solar panel, then stop.

The rotation angle of servo motors is within the limit of 0 - 180 degrees.

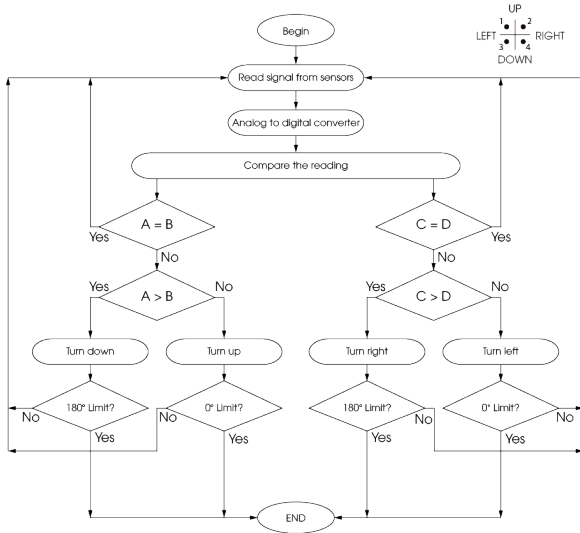


Figure 6: Flowchart of Solar Tracking System

4. Experimental Setup and Results

This study has implemented experiments for three systems: fixed solar panels, single-axis solar tracking system, and dual-axis solar tracking system. The experiment was conducted on December 12, 2019, from 6:30 to 18:30, the weather was sunny and cloudy at the location $10^{\circ} 58'49.8''N$ $106^{\circ} 40'26.4''E$.

In this experiment, we use the DSN-VC288 DC 100V 10A Voltmeter Ammeter to measure the amperage and voltage produced by the solar panel for a load (9W LED). From the recorded data, the authors use the Excel software to calculate the power output (Table 2) and plot the chart as Figure 7.

Table 2 shows the statistical results for the three systems, and these results are plotted as shown in Figure 7. Experimental results in the dual-axis tracking system have a total output of more than 14.28W (20, 77%) compared to a single-axis tracking system and 20.31 W (32.39%) more when the solar panel is left standing. Thus, the use of the dual-axis solar battery tracking system will have a greater efficiency than leaving the battery in place and tracking a single axis.

As shown in Figure 7, the performance of the dual axis system is more optimal than that of a single and fixed system. The graph descended at 11 o'clock because of the cloud cover, at 11 o'clock a cloud appeared to cover the sun, so the output power of all 3 systems decreased.

Table 2: Data for different solar tracker

Time (h: mm)	Fixed Panel (Watt)	Single Axis Tracker (Watt)	Dual Axis Tracker (Watt)
6:30	0	0	0
7:00	0,16	0,16	0,32
7:30	0,415	0,24	0,67
8:00	0,765	0,32	2,27
8:30	1,2	0,49	4,12
9:00	4,6	5,04	6,26
9:30	4,7	5,04	6,95
10:00	6,3	7,27	7,84
10:30	6,84	7,08	8,06
11:00	1,58	1,69	1,8
11:30	6,25	7,48	8,1
12:00	6,25	7,48	8,1
12:30	6	6,25	7,02
13:00	6,8	7,12	7,52
13:30	3,6	5,15	5,61
14:00	1,8	1,89	2,18
14:30	0,9	1,09	1,27
15:00	1,26	1,13	1,39
15:30	0,99	1,11	1,39
16:00	0,8	0,91	1,1
16:30	0,57	0,729	0,9
17:00	0,45	0,49	0,59
17:30	0,24	0,3	0,37
18:00	0,236	0,29	0,29
18:30	0	0	0
Total	62.706	68.74	83.02

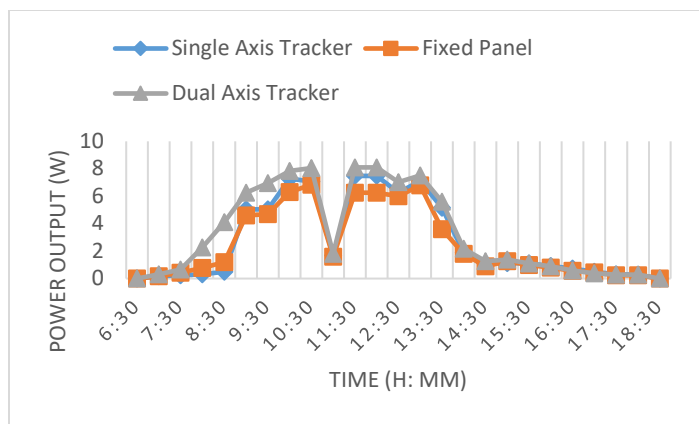


Figure 7: Graphical comparison of experimental data

5. Conclusion

Solar energy is an important renewable energy source. To get maximum solar energy, the solar tracking system is needed. In this study, the authors designed a solar tracking system that uses photoresists, Arduino microcontrollers, and servomotors. The research results show that the dual-axis solar tracking system is more optimal than the other systems, with an outstanding performance of 20.77% compared to the single-axis and 32.39% compared to fixed solar panels. The results show that this solar tracking system is a low cost, efficient and easily implementable. In the future, the study will improve the hardware with the display of time, voltage, current, power via LCD and export data via wifi network.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors appreciate the support from Thu Dau Mot University and Ho Chi Minh City University of Technology and Education in Vietnam.

References

- [1] Rezvani, A., Gandomkar, M., Izadbakhsh, M., & Ahmadi, A. (2015). Environmental/economic scheduling of a micro-grid with renewable energy resources. *Journal of cleaner production*, 87, 216-226.
- [2] Mallick, T. C., Munna, M. S., Barua, B., & Rahman, K. M. (2014, October). A design & implementation of a single axis solar tracker with diffuse reflector. In 2014 9th International Forum on Strategic Technology (IFOST) (pp. 289-293). IEEE.
- [3] Polo, J., Bernardos, A., Navarro, A. A., Fernandez-Peruchena, C. M., Ramirez, L., Guisado, M. V., & Martínez, S. (2015). Solar resources and power potential mapping in Vietnam using satellite-derived and GIS-based information. *Energy Conversion and Management*, 98, 348-358.
- [4] The World Bank. Global Solar Atlas version 2.0 (on 23-Oct-2019) <https://globalsolaratlas.info/download/vietnam>
- [5] Lokhande, M. K. (2014). Automatic solar tracking system. *International Journal of Core Engineering and Management*, 1(7), 122-133.
- [6] Li, G., Tang, R., & Zhong, H. (2012). Optical performance of horizontal single-axis tracked solar panels. *Energy Procedia*, 16, 1744-1752.
- [7] Rizk, J. C. A. Y., & Chaiko, Y. (2008). Solar tracking system: more efficient use of solar panels. *World Academy of Science, Engineering and Technology*, 41, 313-315.
- [8] I. Abadi, A. Soeprijanto and A. Musyafa (2014). Design of single axis solar tracking system at photovoltaic panel using fuzzy logic controller. 5th Brunei

- International Conference on Engineering and Technology (BICET 2014), Bandar Seri Begawan, pp. 1-6.
- [9] Ashwin R, Joshualar Immanuel K, Lalith Sharavn C, Ravi Prasad P.S, Varun A.K (2014). Design and Fabrication of Single Axis Solar Tracking System *Journal of Mechanical and Production Engineering* ISSN: 2320-2092, Volume- 2, Issue-12.
- [10] Anusha, K., & Reddy, S. C. M. (2013). Design and development of real time clock based efficient solar tracking system. *International Journal of Engineering Research and Applications*, 3(1), 1219-1223.
- [11] Hussian S. Akbar, Muayyad N. Fathallah, Ozlim O. Raouf (2017). Efficient Single Axis Tracker Design for Photovoltaic System Applications. *IOSR Journal of Applied Physics* 09(02):53-60
- [12] Fathabadi, H. (2016). Novel high accurate sensorless dual-axis solar tracking system controlled by maximum power point tracking unit of photovoltaic systems. *Applied Energy*, 173, 448-459.
- [13] Wang, J. M., & Lu, C. L. (2013). Design and implementation of a sun tracker with a dual-axis single motor for an optical sensor-based photovoltaic system. *Sensors*, 13(3), 3157-3168.
- [14] Munna, M. S., Bhuyan, M. A. I., Rahman, K. M., & Hoque, M. A. (2015, September). Design, implementation and performance analysis of a dual-axis autonomous solar tracker. In 2015 3rd International Conference on Green Energy and Technology (ICGET) (pp. 1-5). IEEE.
- [15] Mustafa, F. I., Shakir, S., Mustafa, F. F., & Naiyf, A. T. (2018, March). Simple design and implementation of solar tracking system two axis with four sensors for Baghdad city. In 2018 9th International Renewable Energy Congress (IREC) (pp. 1-5). IEEE.
- [16] Chhoton, A. C., & Chakraborty, N. R. (2017, September). Dual axis solar tracking system-A comprehensive study: Bangladesh context. In 2017 4th International Conference on Advances in Electrical Engineering (ICAEE) (pp. 421-426). IEEE.
- [17] Jovanovic, V. M., Ayala, O., Seek, M., & Marsillac, S. (2016, March). Single axis solar tracker actuator location analysis. In SoutheastCon 2016 (pp. 1-5). IEEE.
- [18] Thu Dau Mot Vietnam - Sunrise, sunset, dawn and dusk times: <https://www.gaisma.com/en/location/thu-dau-mot.html>
- [19] Reca-Cardena, J., & López-Luque, R. (2018). Design Principles of Photovoltaic Irrigation Systems. In *Advances in Renewable Energies and Power Technologies* (pp. 295-333). Elsevier.

Using Leader Election and Blockchain in E-Health

Basem Assiri*

Computer Science Department, Jazan University, 45142, Jazan city, Saudi Arabia

ARTICLE INFO

Article history:

Received: 13 March, 2020

Accepted: 20 April, 2020

Online: 03 May, 2020

Keywords:

Electronic Personal Health Record

Parallelism

Distributed Systems

Leader Election Algorithms

Blockchain Technology

ABSTRACT

The development of electronic healthcare systems requires to adopt modern technologies and architectures. The use of Electronic Personal Health Record (E-PHR) should be supported by efficient storage such as cloud storage which enables more security, availability and accessibility of patients' records. Actually, the increase of availability of E-PHR enhances parallel access, which improves the performance and the throughput of the system. Using distributed systems, users are able to communicate and to share resources to achieve specific goals. Such kind of access needs to have more coordination to maintain parallelism, which can be provided through leader election algorithms. In leader election algorithms, users elect one of them as a leader to coordinate the work and to prevent conflicts. This paper introduces an adoptive leader election algorithm (ALEA) that considers medical and healthcare specifications, since it uses leader election algorithm for E-PHR in the cloud environment. The use of ALEA improves performance by allowing more parallelism and reducing the number of coordinating messages within the system, as well as facilitating the medical specifications such as having a primary doctor or handling emergency situations. Moreover, the paper highlights the strengths and weaknesses of using Blockchain technology in the field of healthcare. In fact, the paper investigates the implementation challenges of ALEA concepts using Blockchain technology.

1. Introduction

Within the last decades, the development of technologies, the Internet and digital applications makes them essential components in some other fields such as education and healthcare. This paper focuses on the development of e-health systems using some supportive algorithms and modern technologies. Actually, this paper is an extension of work originally presented in the 2ndInternationalConference on New Trends in Computing Sciences [1].

The competition among healthcare organizations encourages them to involve modern and advanced technologies to increase stakeholders' satisfaction. These technologies help to improve the provided services. For examples, a patient can schedule an appointment online; doctors (physicians) can access, maintain, and transfer E-PHRs electronically anytime and from anywhere; doctors would be able to diagnose, complete the required treatment and even participate in surgery remotely; prescriptions are sent to the corresponding pharmacy electronically. In addition, these technologies enable costs and managerial efforts to be reduced.

The services costs can be time, physical space, energy and infrastructure. Besides that, it gives deep and clear insight for better administration and decision making. The use of E-PHR as a digital version of PHR allows information to be accessed, updated and transferred in an electronic manner [2], which enhances information accessibility, availability, security, privacy, completeness and consistency. It also helps to avoid the risk of having traditional PHR in case of natural disasters such as Hurricane Katrina [2]-[6].

Moreover, an efficient pattern of storage such as cloud storage is required to support the use of E-PHR in distributed systems. Actually the E-PHRs are stored in servers and can be accessed securely on the Internet [7], [8].

The presence of E-PHR, servers, cloud storage and many connected devices creates a parallel and distributed system. In distributed systems, devices are connected through networks to perform specific tasks. Thus, it helps to improve efficiency and throughput of the process of sharing resources but it also requires more control and coordination. Therefore, leader election algorithms can be used to coordinate the parallel tasks and to preserve the consistency of E-PHRs [9]. Indeed, having

*Basem Assiri, 45142 - Jazan University - Jazan City – Saudi Arabia,
+966599933185 & babumusmar@jazanu.edu.sa, bas0911@hotmail.com.

parallelism could result in conflicts, especially when a process tries to update an E-PHR, while another one is reading or updating it at the same time. In fact, having a leader allows for an exclusive access (token) to an E-PHR, which keeps it consistent.

In addition, Blockchain technology allows decentralized ledger to be managed within a distributed system. The ledger contains a chain of blocks (records). In Blockchain many nodes share and process distributed copies of the same ledger. In fact, when a node proposes a new block of transactions, the other nodes process it and vote to commit it if its valid or to abort to if it is not. Based of the votes of the majority (consensus), the block commits and every node updates its copy of the ledger or the block is ignored [10].

Nakamoto uses Blockchain to produce the Bitcoin as a first cryptocurrency, in which users execute electronic financial transactions without banks [10]-[12]. After that, Blockchain technology has been involved in many other areas such as judiciary, notary, copyrights, education and healthcare [12].

This paper introduces an adoptive leader election algorithm(ALEA) that is suitable for E-PHRs and healthcare systems. The paper proposes the principles of a primary and a secondary leader as well as having multiple tokens. ALEA allows the number of communication messages to be reduced in case of failures. Moreover, this work highlights the strengths and weaknesses of using Blockchain technology in the field of healthcare to implement the concepts of ALEA.

The rest of our paper is organized as follows: in Section 2, some related work is discussed. Sections 3 and 4, introduce our proposed system model and algorithm. Section 5 discusses many important issues and techniques such as algorithm correctness, consistency, synchronization, file sharing, traffic flow and replication. Finally, Sections 6 and 7 focus on the advantages and disadvantages of using Blockchain technology, while Section 8 provides conclusion.

2. Related Work

There are many techniques to preserve data consistency in cloud storage [13], [14]. Some research proposes strict consistency while the others relax this concept and accept weaker levels of consistency. Coppieters et al. provide an algorithm with strict consistency, where they order all concurrent processes on all replicas. Actually, the concurrent execution of processes should be matched with a correct sequential execution [15]. Zellag and Kemme introduce an efficient relaxed consistency model for cloud storage [16].

Some others use leader election algorithms for consistency, whereby a leader grants exclusive access to some memory objects to prevent conflict [9], [17]. For leader electing, a bully algorithm [9, 18], enables every user to have a unique identifier (Id) and every user sends its Id to all other users. So, they select the node with the maximum Id as a leader. The complexity of this approach reaches $O(n^2)$ messages, which is considered expensive. With a token ring algorithm [19], the users are structured in a linked-circle and every user sends its Id to the next one. After receiving the message, the user compares its own Id with the received one and sends the greater one to the next user. The complexity of this approach costs $O(n)$ messages. Numan et al. provide an algorithm that uses a centralized linked-list queue of all users. The leader is

the head of the queue; and when it fails, another user acquires a lock and dequeues the old head. The complexity of this approach is $O(I)$ [20].

At the same time, many countries and institutions have started using E-PHRs. For example, at the beginning of 2014, the American Recovery and Reinvestment Act enforced all healthcare agencies (public and private) to use E-PHR. This facilitates accessibility, utilization and management. However, such a change requires specific technical and infrastructural support to be adopted [21].

In addition, Blockchain technology helps to allow decentralized management of E-PHRs, where there is no need for a third party such as hospitals or healthcare agencies. However, Blockchain has been enhanced with fairness and freedom [22]. Therefore, Blockchain provides many advantages for many areas such as health-care systems. First, it supports the availability, robustness and security of E-PHR. In fact, many projects and companies, such as Data Gateways, Gem Health Network, Deloitte and Guardtime have started using Blockchain to manage their E-PHR [11], [23]. Second, it supports all related financial operations such as funding, donations and insurance payments through cryptocurrencies. Third, Blockchain supports scientific clinical and biomedical research such as in the MedRec Health bank. Indeed, companies and organizations use Blockchain for data sharing and verification, ownership proofs and privacy of patients and organizations. In addition, it could apply the principle of "gain as you contribute" in scientific clinical and biomedical research. For example, while Bitcoin (which is the first cryptocurrency) is earned through solving puzzles, Gridcoin, is another cryptocurrency that is earned based on the contribution to scientific research [10], [11] and [24]. Fourth, Blockchain can be used to manage and process any kind of data such as the records of employees, healthcare facilities, medical instruments, medicines and pharmaceutical supply chains.

Laure and Martha use Blockchain as a control manager to manage the access of E-PHR. Actually, this work focuses on Blockchain's advantages such as scalability, security and privacy. Indeed, the work proposes using a Blockchain system as an access control manager, so it only has indexes of records, while the real records are stored in separate storage (out of the Blockchain). This would help to avoid the negative impact of data redundancy, where every node in the Blockchain has a copy of all indexes instead of having a copy of all records [21].

Kevin et al. suggest the use of Blockchain in healthcare to solve the issues of hardware and software heterogeneity. The work focuses on the quality of sharing E-PHR in an understandable and meaningful manner. So, they use the concept of Fast Healthcare Interoperability Resources with specific Application Programming Interfaces as standard for data formatting and presentation [25].

3. The System Model

Before you ALEA is designed according the concepts of in well-organized bully algorithm for leader election [20]. The well-organized bully leader election algorithm is implemented using a linked-list queue to minimize the cost of leader election to $O(I)$. ALEA is a modified algorithm that is efficiently applicable for medical and healthcare systems.

ALEA creates a linked-list queue *Queue* of size *Z*, where *Z* is the number of nodes in the *Queue*. The *Queue* is stored in shared memory, so it is always visible to all nodes. A node (processor/doctor) is denoted as *Nod* and is represented with a unique identifier (*doctorId*), a pointer that is pointing to the next node, and a *token* flag. The new inserted node (doctor) can read the E-PHR but it has to get exclusive access to update it. The exclusive access is given by changing the *token* value from 0 to 1. The queue *head* is the leader (*PL*). To change the leader, dequeue the current *head* node and move the leadership (*PL pointer*) to the next node.

For emergency cases where the *PL* is not accessible, a temporary queue *TQ* is created with a secondary leader called *SL*.

Figure 1 shows *Queue* where the list has three doctors, and *PL* points to the head of the queue. It also shows the emergency block which has another list of *TQ* and *SL*. In reality, *TQ* is not usually there.

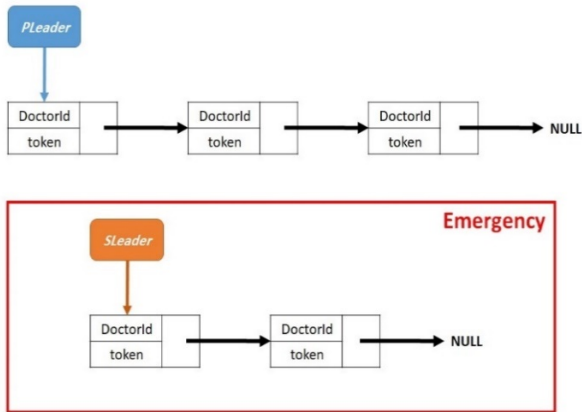


Figure 1: Leader election linked-list *Queue* with a *PL* pointing to the queue head; while the emergency linked-list *TQ* and a *SL* appear in the emergency block.

It is obvious that many doctors can read files in parallel, but conflict arises at the time of the update. For instance, the situation arises when one doctor is reading a file and another updating it. In this case, the data updated last should be visible to the reader. Also, when two doctors are writing to (updating) the same files, the two updating processes conflict with each others. To overcome this situation, it incorporates the idea of *TokenPtr*, which shows who holds the *token* flag to update the files.

4. Proposed Algorithm

- A hospital creates the patient’s E-PHR with a primary doctor (leader) *PL*, (see Function 1). Actually, the node (doctor) has three attributes as follows: first, data to have the *doctorId*; second a *pointer* pointing to the next node; and third a *token* flag. When a new doctor is inserted into the *Queue*, the doctor can read the E-PHR directly, but for update permission, the *token* has to be changed to 1. The *TokenPtr* is another pointer pointing to the node that has the *token* (*token=1*). Finally, it increases the size of *Queue*.
- If *PL* cannot be accessed for any reasons, except in the case of failure, it creates an emergence or temporary queue *TQ* with a secondary leader *SL*, (see Function 2). With the

creation of a new node, the size of *TQ* is increased and the same procedures as in Function 1 are used.

- Now Function 3 shows the process of adding a new node to the *Queue*. Considering the medical needs, *PL* can add a new doctor to the team, by creating and enqueueing a new node to the *Queue*. Then it increases the *Queue* size. This is also applicable to *TQ*.
- Sometimes because of medical needs, the leader has to reorder doctors in the *Queue*, so it swaps the nodes as shown in Function 4. Based on the doctor’s *Id*, *TempPtr1* starts searching from the head position, until it finds the first doctor. Then, *TempPtr2* starts searching from the head position, until it finds the second doctor. After that, it swaps the doctors by inserting *doctorId1* in the node of *TempPtr2* and *doctorId2* in the node of *TempPtr1*.
- When *PL* retires from leadership of the E-PHR’s team, it follows the procedures for Function 5. If it is the only doctor who handles the E-PHR, which means *Queue* has one node only, then, the retirement is not allowed. Else, it uses *TPtr* to point to the *PL* node; moves the *PL* pointer to the next node; moves the *token* (if it is needed) and finally it dequeues the *TPtr* node and enqueues it again from the other end of the *Queue*. This is also applicable for *SL* and *TQ*.
- In Function 6, a doctor leaves the E-PHR’s team (complete clearness). If it is the only doctor who handles the E-PHR, which means *Queue* has one node only, then, the clearness is not allowed. Else, it is dequeued from the *Queue* (as in Function 5), but there is no need to enqueue it again.
- Function 7, explains how to move the *token* among nodes. First, *PL* finds the targeted node according to its *doctorId*, then it gives the *token* (makes *token=1*), or gets it (makes *token=0*). In addition, *PL* allows parallelism by giving the *token* to many nodes simultaneously, which is explained in details later.
- Function 8, shows the case of a doctor requiring the *token*, so it sends an acquiring message to *PL* and waits for some time (*Timeout*). It should wait until the time becomes equal to *T* where $T = \text{currenttime} + \text{Timeout}$. In fact, it is supposed to receive a reply message (acknowledgment) from *PL*. However, if the timeout finishes without receiving the reply message, then *PL* fails, and it calls *PLFailure()*. Actually many nodes may detect *PL* failure simultaneously, so each node has to copy the *Id* of *PL* (failed leader’s *doctorId*) (more details are given in Function 9).
- Function 9, shows the case of leader *PL* failing. Upon the detection of *PL* failure, the detector node calls *PLFailure(Id)*. It also passes *PL*’s *doctorId*. *PLFailure(Id)* moves the *PL* pointer to the next node and

1. || **Algorithm 1: ALEA1.**

2. **1. Initialization()**
3. //To create the linked-list

```

4. Queue Z=0; //The queue size
5. Nod=newnode();
6. Nod->data=doctorId;
7. Nod->next=NULL;
8. Nod->token=0;
9. PL←node;
10. TokenPtr=PL;
11. TokenPtr->token=1;
12. Z++;
13. Return

```

14. 2. Emergency()

```

15. //Adding a new doctor as SL & creating a temporary queue
16. Z=0; //The queue size
17. Nod=newnode();
18. Nod->data=doctorId;
19. Nod->next=NULL;
20. Nod->token=0;
21. SL←node;
22. TokenPtr1=SL;
23. TokenPtr1->token=1;
24. Z++;
25. Return

```

26. 3. AddDoctor()

```

27. //Adding a new doctor to Queue
28. Nod=newnode();
29. Nod->data=doctorId;
30. Nod->next=NULL;
31. Nod->token=0;
32. Queue←enqueue();
33. Z++;
34. Return

```

36. 4. SwapNodes(doctorId1, doctorId2)

```

37. //Swapping the doctors in Queue
38. TempPtr1=PL;
39. TempPtr2=PL;
40. i=1;
41. while(i ≤ Z) do
42. {
43.   if (TempPtr1->data, doctorId1) then
44.   {
45.     TempPtr1=TempPtr1->next;
46.     i++;
47.   }
48.   else
49.   {
50.     //First doctor is found, now find the other one
51.     Break;
52.   }
53. }
54. i=1;
55. while(i ≤ Z) do
56. {
57.   if (TempPtr2->data, doctorId2) then
58.   TempPtr2=TempPtr2->next;
59.   i++;
60.   }
61. else

```

```

62. {
63.   //Now Second doctor is found, so swap them
64.   TempPtr1->data=doctorId2;
65.   TempPtr1->token=0;
66.   TempPtr->data=doctorId1;
67.   TempPtr2->token=0;
68.   Break;
69. }
70. }
71. Return

```

72. 5. Rretirement()

```

73. //Retiring from the leadership
74. if (PL->next=NU LL) then
75. Return False;
76. else
77. {
78. TPtr=PL;
79. PL=PL->next;
80.   if (TokenPtr=TPtr) then
81.   {
82.     TokenPtr->token=0;
83.     TokenPtr=TokenPtr->next;
84.     TokenPtr->token=1;
85.   }
86. TPtr.dequeue();
87. T Ptr.enqueue();
88. }
89. Return

```

90. 6. Clearness()

```

91. //Clearness
92. if (PL->next=NU LL) then
93. return False;
94. else
95. {
96. TPtr=PL;
97. PL=PL->next;
98.   if (TokenPtr=TPtr) then
99.   {
100. TokenPtr->token=0;
101. TokenPtr=TokenPtr->next;
102. TokenPtr->token=1;
103. }
104. TPtr.dequeue();
105. }
106. Return

```

dequeues the failed leader. If many nodes detect *PL* failure simultaneously, all of them invoke *PLFailure(Id)*, which may cause multiple unnecessary dequeues. Thus, it is mandatory to use a Compare-and-Swap atomic operation (CAS statement), by which only one node changes the leader [26]. Using a CAS statement, one node checks if the *PL* is still in a failure (*PL's doctorId=Id*), and it invokes *Clearness()*. In *Clearness()*, *PL* (failed leader) is dequeued and another leader is elected. Therefore, the other nodes that detected the failure of *PL* also use CAS, but find *PL's doctorId≠Id*, since they find *doctorId* of the new *PL* that is not equal to the value of *Id*, and do nothing.

107. **7. Leadership(doctorId1)**

```

108. //Moving the token among nodes
109. //Getting the token
110. TokenPtr→token=0;
111. //Find the targeted node based on Id
112. i=1;
113. while(i ≤ Z) do
114. {
115.   if (TokenPtr→data, doctorId1) then
116.   {
117.     TokenpPtr=TokenPtr→next;
118.     i+ +;
119.   }
120.   else
121.   {
122.     //Give the token
123.     TokenPtr→token=1;
124.     Break;
125.   }
126. }
127. Return
    
```

128. **8. Reminder()**

```

129. //Node sends a message to acquire the token
130. Sendmsg(PL, "Acquire token");
131. //Waiting (Timeout)
132. T=CurTime()+Timeout;
133. while(receiveack() = false && CurTime() < T) do
134. wait();
135. //Timeout finishes and no response (PL fails)
136. if (receiveack()=false) then
137. {
138. Id=PL→data;
139. PLFailure(Id);
140. }
141. Return
    
```

142. **9. PLFailure(Id)**

```

143. //Leader still in failure
144. CAS (PL→data, Id, Clearness());
145.Return
    
```

Figure 2 is an example of a failed leader. Node 3 and 4 discover that the leader has failed. In Figure 2 (a), the two detector nodes elect a new leader in parallel, so both dequeue and move PL pointer. This causes one unnecessary dequeue. In Figure 2 (b) the two nodes elect a new leader in parallel using CAS (the lock can also be used), so both of them copy the doctorId of the PL in Id (Id=1). Both of them apply CAS such that one node will have successful CAS, and it dequeues the failed leader. However, since a new leader is already there, the doctorId of the new leader does not equal to Id anymore. Obviously, the doctorId is 2, while Id=1. As a result of this, the other node gets a failed CAS, so it does nothing.

5. Analysis

5.1. Correctness

For ALEA correctness, the correctness of concurrent operations, that are made by doctors, are proved by satisfying www.astesj.com

Linearizability [26]. This requires the concurrent operations to be ordered to match a correct (valid) sequential execution. Indeed, ALEA is considered an event-based model [9], where a doctor performs the operations(a read/update) on the E-PHR, and each operation is represented by two instantaneous events (*begin()* and *end()*). A complete execution is a sequence of operations where there is no pending operation and every operation has the two events. For the correctness and legality of all operations, it is not difficult to argue about the correctness and legality of a sequential execution where all operations are running in one processor and one after another. This helps to prove the memory consistency and to predict the situation of the file, before and after each operation. Since ALEA has two kinds of operations which are update and read, an update operation is legal if it appears instantaneously and all later reads read it, until another update takes place. A read operation is legal if it reads the last written data, and all later reads read the same data until another update takes place. A sequential execution is legal if all its operations are legal. Then, the concurrent execution is correct and the memory is consistent if the order of concurrent execution (including events of all concurrent operations) matches the order of a legal sequential one; this is known as Linearizability [26].

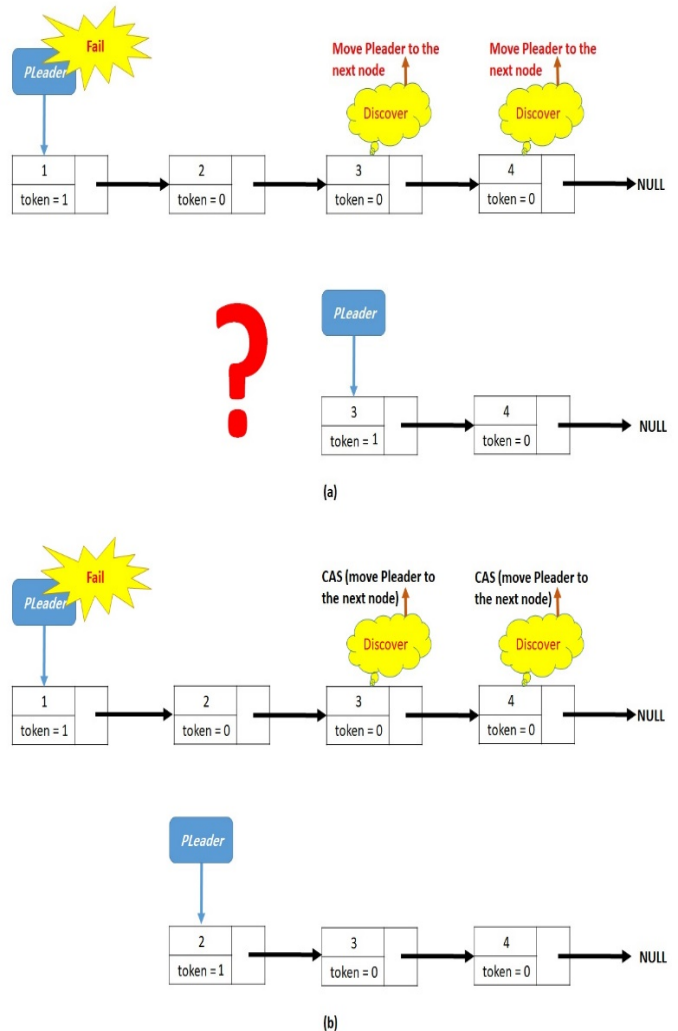


Figure 2: (a): As result of the failure of the leader, two nodes are electing a new leader in parallel, which causes one more unnecessary dequeue: (b) As result of the failure of the leader, two nodes use CAS to elect a new leader in parallel, so one node dequeues the failed leader and the other does nothing

In ALEA, there are two kinds of operations, some operations process the linked-list queue and the other operations process the E-PHR. The first kind is to modify the linked-list queue of doctors. To update the queue, there are operations such as enqueue, dequeue (even electing new leader is conducted through dequeuing) or move the token. In fact, ALEA maintains a queue of linked-list, so it follows the concepts of Michael and Scot [27], that is considered as one of the best lock-free algorithms in the field. Actually the enqueue and dequeue operations are conducted based on Michael and Scot's algorithm [27], which is linearizable. For the operation of moving the token, it is executed only by the leader, so it is serialized. The read operation simply tries to find the current leader which is always able to see the last change. In fact, the linked-queue is stored on a shared memory which makes all updates visible instantaneously. The second kind of operations is to update the E-PHR. Actually, some procedures that prove the correctness of the bully leader election algorithm [9, 18], and the well-organized bully leader election algorithm [20], are applicable in ALEA. To update the E-PHR it has to use the token which gives exclusive access, so it does not conflict with other operations and the operations can be ordered based on the token movement. On the other hand, all read operations are executed concurrently and are ordered with respect to the real-time order. Indeed, ALEA satisfies Linearizability.

5.2. Synchronization

Processes synchronization allows resources to be shared without causing any conflicts (it preserves consistency). Clearly, ALEA uses application level synchronization, but at some points it also relies on operating system primitives such as using locks and CAS atomic operation. Focusing on how to order events, Linearizability follows the real-time order of the concurrent operations (execution). In fact, when there is one thread, events are not interleaved on the same object. However, when there are many threads the synchronization satisfies a happens-before relationship. Therefore, the synchronization of events respects a well-formed clock. Actually, the order depends on the exact time in terms of where and when the operation takes effect. However, having multiple time zones of doctors and patients due to remote access or travel; challenges the use of physical clock. Therefore, a logical clock such as Lamport's logical clock [9, 23], is preferred. Lamport's logical clock is one of the famous approaches that uses the happens-before (\rightarrow) relationship such that, for any two operations a and b , it is said a happens-before b ($a \rightarrow b$) if and only if the $end()$ of operation a occurred before the $begin()$ of b . However, if there is an overlap between operations then the two possibilities are considered ($a \rightarrow b$) or ($b \rightarrow a$). Using ALEA, it orders the operations even if they are executed in multiple processors since all of them are executed on a single version of E-PHR. The read operation is ordered easily, while the update operation gets exclusive access(token), so it is also ordered based on the token movements.

5.3. Parallel Access of E-PHR

In this part, it is suggested the E-PHR to be divided into multiple sections s , and doctors are able to access different sections concurrently. Therefore, ALEA has to use multiple tokens, let say k tokens, where $k=s$ (a token for each section). Now, the leader gives a suitable token to a doctor according to the needed section. Thus, *TokenPtr* of the original ALEA is replaced with a two-dimensional array of pointers (with k rows and 2 columns). Each

row of this array shows the doctor's id and the corresponding section.

5.4. Traffic Flow

Processes It is known that the election of a new leader requires a number of messages to be passed and that may cost $2n$ messages for n nodes [17, 18]. On the other hand, having one leader has severe negative impacts on the system in case of leader failure. To handle this situation, decentralized leader election algorithms enable many replicas for each file and more than one leader. To access any replica, voting is conducted and access is allowed according to the majority (consensus). Such type of permission requires approximately $2n$ messages [17, 18]. However, with ALEA the number of messages is reduced to θ , as it dequeues the head node (leader) and the new head will be the new leader (no traffic). In addition, other messages are sent to acquire the token. In ALEA the leader can hold a token or move it to the targeted node as needed with no messages. Rarely, if a doctor needs to get the token, the doctor sends an acquire message (as shown in Function 8), and it receives a reply message from the leader, which causes no traffic.

5.5. Fairness and Starvation (timeout)

This part focuses on fairness of the leader election process and fairness of token movements. For the fairness of leader election process, the leadership appointment has to follow the queue property first-in-first-leader. However, this is relaxed to approximate-first-in-first-leader to handle medical and healthcare requirements. Indeed, in ALEA, the leader is able to swap the doctors in the queue based on the medical needs, which is completely fair from the medical point of view. On the other hand, ALEA allows the leader to move token among nodes as needed, which is also fair from the medical point of view.

5.6. Replication

This part It would never be advisable to use a single centralized copy of the E-PHR since there is no way for recovery in case anything goes wrong. In this regard, the idea of having multiple replicas of the same file is integrated, which is very important for reliability and performance. Firstly, many replicas allow recovery of files if there is an issue with a server, security problem, file corruption or failed operation (read/update). Secondly, the replicas allow improved performance as they can be distributed based on system capabilities, competences, load balancing or geographical distribution. However, having many replicas requires more effort to keep them consistent. For such an issue, it is suggested that the number of replicas be reduced to three copies. The main (permanent) replica is stored in a suitable place bearing in mind the geographical location of the patient and all doctors accessing this replica. The second replica (backup1) should be stored in the same system or in a very close one, so it can be used easily for recovery. The third replica (backup2) must be stored in a different server that is geographically located far away from the main one; so it can be used in case of natural disasters, for example. In order to maintain consistency in replicas, it is suggested to use the eventual consistency criteria where the consistency of file is relaxed [9, 15]. Eventual consistency is very suitable to E-PHRs and cloud storage. Using eventual consistency, read operations do not cause any harm, as the replicas remain consistent, but this scenario will be different in the update operation because the update operation changes the content of the replica. After

confirmation of an update, the change should be reflected in all other replicas in an asynchronous manner (over the network). In ALEA, the token can be given to more than one doctor working on different parts of the file, which causes a write-write conflict. Therefore, every update is labeled with two timestamps, which are the timestamp when the update takes effect, and the timestamp of the last update before this one took place. This helps all replicas to order the changes and maintain consistency among all of them. Another way is to use Primary-Based Protocol to conduct remote write/update; in this scheme, the user holds a local copy of a file and after performing the update operation locally, it sends a request to the server for final approval [9]. This way, the main server takes responsibility for preventing conflicts. To avoid any type of conflict, it updates the other replicas instantly. This protocol (of maintaining the local replica) is an actual support for mobile applications, where the user works on those selected files from different locations, which may affect the connection with the server. Besides this, a huge number of update operations on the system and the processes of maintaining replica consistency result in very high contention on the network. To solve this difficulty, the techniques below are effective:

- Directly, send the modified part of a file to replicas. This can be used in case of a considerable number of update operations.
- Only send a notification to invalidate the other replicas. The invalidation notification has a smaller size compared to the update messages. Indeed the invalidation can be more specific to tell which part of the replica is not valid anymore. Then, the system denies access to that part until it finds a suitable time to update it.
- For some kind of update, send the computation itself, so the others process it locally and make the update themselves.

6. The Use of Blockchain

Blockchain technology is a distributed ledger that is shared and processed by nodes according to consensus [10, 11]. In fact, Blockchain algorithms have three major phases which are proposing a new block, voting on the new block, committing or aborting it based on the consensus [10, 11, 22]. The implementation of ALEA concepts using Blockchain technology requires some modifications. The creation of a new queue in Functions 1 and 2, will be based on the consensus of users, rather than hospitals or healthcare agencies. Moreover, adding, removing and swapping doctors (in Functions 3, 4, 5 and 6) will not be executed by the leader; instead, they have to make a proposal, vote and then take a decision according to the majority. The same thing is applicable to token movements in Functions 7 and 8, as well as in case of the failure of the leader in Function 9.

6.1. Transaction's Validation

The Blockchain consists of a number of nodes (processors). Each node can access the patients' E-PHRs and has a copy of the ledger. The ledger contains blocks and each block contains a number of transactions. The transaction contains some operations that are executed on E-PHRs. The operations on the E-PHR are either read or update (the update includes creating, editing and deleting E-PHR). After the execution of transactions, the miners

validate the transactions by validating the output of the operations of those transactions. Indeed, the miners also consider the concurrency of transaction and can use some standard property for such issues such as Opacity [28]. According to this correctness property, a valid transaction commits and is placed on the blocks, while an invalid transaction aborts [29]. Note that more than one miner may validate the same transaction and every one places it in a different block. or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

6.2. Block's Validation

When the block is full (based on the size of the block), it is proposed to the validators (who are the team of doctors), so they validate the correctness of the block, considering the concurrent blocks as well. In fact, the content of the block is hashed to secure it. The hashing can be produced in many ways such as Proof-of-Work (PoW) or Proof-of-Stack (PoS) [30]. In addition, validators validate the signature of the proposer node to verify that the block has been created by a legitimate node. Actually, the validators validate the identity of the proposer through its digital signature, the PoW and content of the block. According to this, validators vote to commit the block or abort it.

6.3. Consensus

Blockchain uses consensus to consider the validation processes of all validators. If the majority votes to commit the block, then it is added to the ledger, otherwise it is ignored (and then there is no need for a leader like in ALEA). The majority of votes means more than 50% is needed, and in some systems they increase it to 67% [31, 22]. In addition, the order of the committed blocks in the ledger is decided based on who gets PoW first. Some work uses a unique timestamp for each block, so the one with a smaller timestamp is added first [22]. The blocks are chained in the ledger using the hash number, as every block has the hash of the previous one; therefore the ledger is unchangeable.

7. Discussion on the Use of Blockchain

The implementation of ALEA's concept using Blockchain has some strengths and weaknesses. Thus, this section discusses the issues of security, privacy, immutability, decentralization, robustness, availability, ownership, computational costs and system traffic flow [10, 11, 24, 21].

7.1. Security, Privacy and Immutability

Using Blockchain technology the identities of doctors are hidden. The E-PHRs and the operations on them are hashed and encrypted. Such security and privacy are positive points that encourage everyone to use Blockchain. On the other hand, users will be untraceable and that is an issue for healthcare systems, especially in terms of tracking suspicious and illegal behaviors.

In addition, using Blockchain guarantees immutability, since the ledger (that has the records of all operations on E-PHRs) is unchangeable. Obviously, this is a positive point; however, such a system does not allow to rollback.

7.2. Decentralization

To avoid the presence and the control of the centralized third party such as a hospital, decentralization enables queue to be

created or emergency to be handled using consensus. The negative impact of consensus is the voting delays; as well as having non-trusted nodes or misleading votes.

7.3. Robustness

To The distribution of Blockchain technology helps to prevent a single point of failure, especially in the case of leader failure. Instead, the work will continue as long as the majority of nodes is still working.

7.4. Availability

Blockchain provides high availability, since all nodes have replicas of the files. The large redundancy of replicas requires more space (memory), communication and processing (to preserve the consistency of the replicas). To avoid such an issue, it is proposed to use Blockchain as an access control manager that shares copies of indexes rather than the real records, and the records are stored in separate centralized storage [21], as shown in Figure 3.

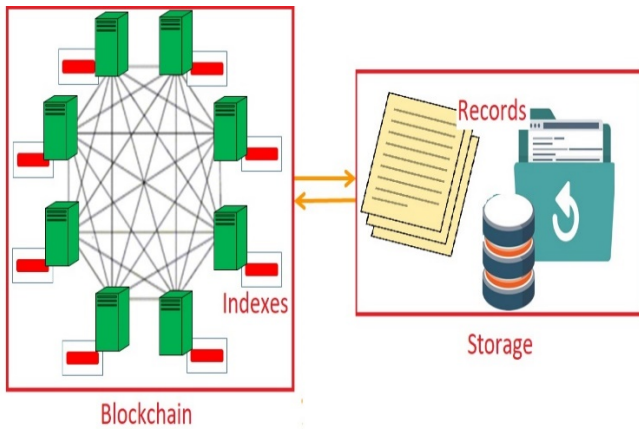


Figure 3: Blockchain with indexes and separated centralized storage of records.

7.5. Performance Cost

The use of Blockchain technology results in huge increases in computational and communication costs. In ALEA, the computation is executed by one node, while in Blockchain all validators execute the computation to confirm its correctness. In addition, when a node proposes a new block, it broadcasts to all validators, so if there are n validators, it broadcasts n messages. After the validators execute the computation, they send votes to all nodes, let's say m nodes, which costs m^2 messages. Actually, the set of the validators is a subset of all nodes in the Blockchain. Then, every node calculates the majority and broadcasts messages of the decision (commit/abort), which also costs m^2 messages. Finally, in case of commit, all nodes update their own copy of the ledger, rather than a limited number of replications. Thus, the use of Blockchain negatively affects the speed of the system and its traffic flow.

7.6. Fault Tolerance

The consensus is a fault tolerant correctness property, where some validators do not confirm the validity of the block. In medical

cases, it is not suitable to ignore the votes of 49% or even 33% of doctors because they are not the majority. This means, the Blockchain may still allow for some errors, which is very critical for healthcare specifications.

8. Conclusion

This paper proposes an adaptive algorithm, for E-PHRs in a cloud environment, so it can be easily used with minimal infrastructure. ALEA enhances parallelism using an alternative and modified leader election technique that suits medical and healthcare systems. This work investigates the performance and the characteristics of the ALEA in comparison to Blockchain technology, which shows that the use of Blockchain technology may result in some negative impacts

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] Basem Assiri. Leader election and blockchain algorithm in cloud environment for e-health. In 2019 2nd International Conference on new Trends in Computing Sciences (ICTCS), pages 1–6. IEEE, 2019.
- [2] Paul C Tang, Joan S Ash, David W Bates, J Marc Overhage, and Daniel Z Sands. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. *Journal of the American Medical Informatics Association*, 13(2):121–126, 2006.
- [3] Selena Davis, A Roudsari, and Karen L Courtney. Designing personal health record technology for shared decision making. *Studies in health technology and informatics*, 234:75–80, 2017.
- [4] Janet Woollen, Jennifer Prey, Lauren Wilcox, Alexander Sackeim, Susan Restaino, Syed T Raza, Suzanne Bakken, Steven Feiner, George Hripcsak, and David Vawdrey. Patient experiences using an inpatient personal health record. *Applied clinical informatics*, 7(02):446–460, 2016.
- [5] Arloc Sherman and Isaac Shapiro. Essential facts about the victims of hurricanekatrina. *Center on Budget and Policy Priorities*, 1:16, 2005.
- [6] Shayne Sebold Taylor and Jesse M Ehrenfeld. Electronic health records and preparedness: lessons from hurricanes katrina and harvey. *Journal of medical systems*, 41(11):173, 2017.
- [7] Mu-Hsing Kuo. Opportunities and challenges of cloud computing to improve health care services. *Journal of medical Internet research*, 13(3):e67, 2011.
- [8] Hoang T Dinh, Chonho Lee, Dusit Niyato, and Ping Wang. A survey of mobile cloud computing: architecture, applications, and approaches. *Wireless communications and mobile computing*, 13(18):1587–1611, 2013.
- [9] Andrew S Tanenbaum and Maarten Van Steen. *Distributed systems: principles and paradigms*. Prentice-Hall, 2007.
- [10] Satoshi Nakamoto. *Bitcoin: A peer-to-peer electronic cash system*. Technical report, Manubot, 2019.
- [11] Tsung-Ting Kuo, Hyeon-Eui Kim, and Lucila Ohno-Machado. Blockchain distributed ledger technologies for biomedical and health care applications. *Journal of the American Medical Informatics Association*, 24(6):1211–1220, 2017.
- [12] Michael Crosby, Pradan Pattanayak, Sanjeev Verma, Vignesh Kalyanaraman, et al. Blockchain technology: Beyond bitcoin. *Applied Innovation*, 2(6-10):71, 2016.
- [13] Ibrahim Abaker Targio Hashem, Ibrar Yaqoob, Nor Badrul Anuar, Salimah Mokhtar, Abdullah Gani, and Samee Ullah Khan. The rise of “big data” on cloud computing: Review and open research issues. *Information systems*, 47:98–115, 2015.
- [14] Divyakant Agrawal, Sudipto Das, and Amr El Abbadi. Big data and cloud computing: current state and future opportunities. In *Proceedings of the 14th International Conference on Extending Database Technology*, pages 530–533, 2011.
- [15] Tim Coppieters, Wolfgang De Meuter, and Sebastian Burckhardt. Serializable eventual consistency: consistency through object method replay. In *Proceedings of the 2nd Workshop on the Principles and Practice of Consistency for Distributed Data*, pages 1–3, 2016.
- [16] Kamal Zellag and Bettina Kemme. How consistent is your cloud application? In *Proceedings of the Third ACM Symposium on Cloud Computing*, pages 1–14, 2012.

- [17] Gerard Tel. Introduction to distributed algorithms. Cambridge university press, 2000.
- [18] George Coulouris, Jean Dollimore, and Tim Kindberg. Distributed systems: Concepts and design edition 3. System, 2(11):15.
- [19] P Beaulah Soundarabai, J Thriveni, HC Manjunatha, KR Venugopal, and LM Patnaik. Message efficient ring leader election in distributed systems. In Computer Networks & Communications (NetCom), pages 835–843. Springer, 2013.
- [20] Muhammad Numan, Fazli Subhan, Wazir Zada Khan, Basem Assiri, and Nasrullah Armi. Well-organized bully leader election algorithm for distributed system. In 2018 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET), pages 5–10. IEEE, 2018.
- [21] Laure A Linn and Martha B Koo. Blockchain for health data and its potential use in health it and health care related research. In ONC/NIST Use of Blockchain for Healthcare and Research Workshop. Gaithersburg, Maryland, United States: ONC/NIST, pages 1–10, 2016.
- [22] Basem Assiri and Wazir Zada Khan. Enhanced and lock-free tendermint blockchain protocol. In 2019 IEEE International Conference on Smart Internet of Things (SmartIoT), pages 220–226. IEEE, 2019.
- [23] Marc Pilkington. Blockchain technology: principles and applications. In Research handbook on digital transformations. Edward Elgar Publishing, 2016.
- [24] Usman W Chohan. Environmentalism in cryptoanarchism: Gridcoin case study. Available at SSRN 3131232, 2018.
- [25] Kevin Peterson, Rammohan Deeduvanu, Pradip Kanjamala, and Kelly Boles. A blockchain-based approach to health information exchange networks. In Proc. NIST Workshop Blockchain Healthcare, volume 1, pages 1–10, 2016.
- [26] Maurice Herlihy and Nir Shavit. The art of multiprocessor programming. Morgan Kaufmann, 2011.
- [27] Maged M Michael and Michael L Scott. Simple, fast, and practical non-blocking and blocking concurrent queue algorithms. In Proceedings of the fifteenth annual ACM symposium on Principles of distributed computing, pages 267–275, 1996.
- [28] Rachid Guerraoui and Michal Kapalka. On the correctness of transactional memory. In Proceedings of the 13th ACM SIGPLAN Symposium on Principles and practice of parallel programming, pages 175–184, 2008.
- [29] Nir Shavit and Dan Touitou. Software transactional memory. Distributed Computing, 10(2):99–116, 1997.
- [30] Luon-Chang Lin and Tzu-Chun Liao. A survey of blockchain security issues and challenges. IJ Network Security, 19(5):653–659, 2017.
- [31] Basem Assiri and Costas Busch. Approximate consistency in transactional memory. International Journal of Networking and Computing, 8(1):93–123, 2018.

Organizational, Social and Individual Aspect on Acceptance of Computerized Audit in Financial Audit Work

Bambang Leo Handoko*, Nada Ayuanda, Ari Tihar Marpaung

Accounting Department, Faculty of Economics and Communication, Bina Nusantara University, Jakarta, Indonesia, 11480

ARTICLE INFO

Article history:

Received: 26 February, 2020

Accepted: 14 April, 2020

Online: 03 May, 2020

Keywords:

Organizational

Social

Individual

Audit

Software

ABSTRACT

Auditors now can no longer rely on the old-fashioned way of auditing manually. More and more jobs, increasingly complex work environment, the demands of the times, accuracy and speed of work require auditors inevitably must adopt technology. This research began with our success as academics in the audit family. Related to Indonesia, a large country and has several hundred public accounting firms and thousands of auditors, but computerized use of audits using software is still very little. The public accounting firm still uses a manual system, using normally typed paperwork. We want to find out what can boost the use of software among auditors. Our results are useful for auditors in Indonesia. From the results of statistical tests we found that auditors use compilation software by individual auditors themselves rather than organizations and individuals

1. Introduction

This paper is an extension of work originally presented in 2019 International Conference on Information Management and Technology (ICIMTech) [1]. The development of the information system used by the client has an impact on the expertise that must be mastered by the auditor who was originally approached manually, so with these changes the auditor is required to master the information system process used by the client and computerized audit technique for adjusting the audit process and the procedures used when carrying out field work such as changes in the manual accounting system environment into a computer-based accounting information system causes the auditor to study a system. Purpose auditing for carried out effectively and efficiently, the auditor should adjust his audit techniques to the client's information system [2].

General Audit Software or commonly abbreviated as GAS is an auditing with computer technique, or better known as Computer Assisted Audit Techniques. In Western countries, this GAS has been widely applied by audit companies. [3] Revealed that "Adoption of General Audit Software (GAS) can improve audit quality, and this is the reason why its use is applied by US Auditing Standards".

In terms of regulation (environment), the Indonesian Audit Standards suggest GAS in conducting works [4], but its use is not

mandated / charged. Therefore, the use of technology in audit companies in Indonesia is not fully regulated. One of the Public Accounting Firm's Big Four states they use Generalized Audit Software because they have the software and own the resources.

In terms technology, information technology proficiency including English as foreign language is very influential for public accounting firms medium and small. In terms of technological skills, the Big Four is not in doubt. Obviously, become problem in domain of public accounting firms that smaller. On the contrary, is nor serious problem; the elite office believe language matters won't hamper their work, it used in operation every day. In addition, auditors must understand compatible software that is compatible to check/audit clients for accuracy in testing data integrity.

Associate terms of audit companies (organizations), IT Capital Budget is very influential on the implementation of General Audit Software. Big companies tend having customized audit tools, while medium and small audit companies usually rely on commercially available and cheaper software. Besides that, the IT skills of auditors in the company are also very influential.

On the other hand, the attitudes and intentions of auditors are also very influential in adopting General Audit Software (GAS). One senior auditor from the intermediate audit commented "Software is to help auditor work, nor to make it more complicated". Difficult to operate software make hamper usage.

*Bambang Leo Handoko, Email: bambang.handoko@binus.edu

As a result, the auditor will need even more effort to be able to use it. Interestingly, the intermediate audit office found young employee pay more interest in Generalized Audit Software (GAS) [5]. These premises also supported by research conducted by [6] which resulted that people willing to adopt, when it simplify audit process times and do their jobs more efficiently. Auditors who are young and have IT insights will usually be more interested if given the opportunity to enhance their skills.

2. Theoretical Framework and Hypothesis

2.1. Technology Acceptance Model

The Technology Acceptance Model (TAM) developed by [7] a successful and highly acceptable model for predicting willingness to adopt current information approach. Until now this theory was a pioneer and the foremost in uncovering the reasons people want to change to keep abreast of technological developments. Many studies have re-examined, expanded, and used TAM.

The TAM model originated from theory of reason action [8], that is grounded from study of reaction perception for something, effect on person's will and behavior. Reaction and perception affect in acceptance of technology. One factor that can influence is the user's will, so that the reason someone sees the actions / behavior as a benchmark in the acceptance of a technology. The TAM model that is grounded with a premise reaction or perception for something which outcome in result behavior. Reaction and perception affect in acceptance of technology. One factor that can influence it is the user's will his or herself.

According to [9] people willing to leave old method and change to new modern one, because they think it was useful and easy. Both of these components when associated with TRA are part of belief.

Basically it will be very strong if someone already has a perception, perception will be a suggestion, it will make someone want to do it [10]. The main perception here is the perception of usefulness, namely that this technology is very useful. But in addition to usability, it must also be supported by perceptions of ease, because we find many useful but difficult, it will also be left by people. Based on this premises and previous study by [11], [12] we formulated our premise as follows:

Hypothesis 1: PU has positive influence on System Usage (SU).

Hypothesis 2: PEU has positive influence on System Usage (SU).

Hypothesis 3: PEU has positive influence on Perceived Usefulness (PU).

2.2. Auditing Software

CAATs according to [13] software designed to enable auditors with less sophisticated computer skills to carry out audits related to data processing functions. These packages can carry out certain analytical calculations, thus detect anomalies. Audit software is also interpreted as a computer program that allows automatic decision. Conventional tools such as system use programs, information reappearance programs, or high-level

programming languages can be used for this audit. Real data processing goes through an audit program. Outputs are simulated and compared with regular outputs for monitoring purposes. Parallel simulation, redundant processing of all input data by conducting a separate program test, allows comprehensive and very precise validation to be carried out on important transactions that require 100% audit [14]. The audit program used in parallel simulations is usually a type of general audit program that processes data and produces output that is identical to the program being audited. Audit software is one of the software that adapted in educational work, along with other software used in college/university [15].

2.3. Social, Individual, and Organizational

Reflecting on the preliminary research, we find that there are recognizable 3 factors that are triggers for the perception of ease and perception of use. these three factors are individual, social and organizational [16]. The organization in our research context is the office where the auditor works. The policies taken by partners, which are the highest leaders in the public accounting firm, determine whether software is used or not. In accordance with [16] organization impact both Perceived Usefulness, Perceived Ease of Use (PEU).

The second factor to be investigated is social. Socials in our study are peers or peers of the auditor. When working of course the auditor works in a team, where in the team consists of several people. This is where social interactions were influenced [16]. [11] said people tend to be the same and not too different from their peers or social community. If one uses software easily, the others don't want to be left behind to use it. this opinion is reinforced by the results of the study [17], [18].

The third factor is individual. These factors originated from within the auditor himself. Where they have their own intentions, have their own perceptions before being influenced by the surrounding environment [16]. Individual factors are suspected to have strong potential, because it involves beliefs from within [13], Based on this premises, we form these several hypotheses related to the social, individual and organizational factors to PU and PEU.

H4: Organizational factor has positive influence on Perceived Usefulness

H5: Social factor has positive influence on Perceived Usefulness

H6: Individual factor has positive influence on Perceived Usefulness

H7: Organizational factor has positive influence on PEU

H8: Social factor has positive influence on PEU

H9: Individual factor has positive influence on PEU

3. Research Methodology

Our research was an associative research using path analysis test. We use a total of six factors. The six variables contain three independents, namely: individual, social, organizational, also two intervening variables namely perceived usefulness and perceived ease of use and one dependent variable, namely system usage. Hypothesis testing performed using statistical software.

Hypothesis testing is done after the precondition namely validity and reliability.

3.1. Population and Samples

Respondents in our study were people who worked as independent auditors or financial auditors in public accounting firms. We do not limit the accounting firm, whether the big four, ten or others. The population size is difficult to know, because the number of auditors is very large and changing, based on the approach taken by Chassan [19]. it was concluded that the respondents were at least 30, so we decided to use 100 respondents similar like [20].

3.2. Measurements of Factors

Factors used within our research are adjectives, which are basically abstract. For this reason, in order to be concrete and can be explored with certainty, then we made the operation of variables to make measurements. Variable measurements are made to make abstract variables (derived from adjectives) more tangible and can be calculated quantitatively. Variable measurement is based on preliminary research and grand theory. Here in table 1 presented the operation of variables.

Table 1: Table Operation of Variables

Operation of Variables		
Variable	Main indicator	Source
Organizational Factor (X1)	1. Support 2. Training 3. Management Support	[21]
Social Factor (X2)	1. Internalization 2. Image	[21]
Individual Factor (X3)	1. Job relevance 2. Output quality 3. Result demonstration	[21]
Perceived Usefulness (Z1)	1. Improve job performance 2. Useful	[9]
Perceived Ease of Use (Z2)	1. Easy to Understand 2. Easy to Use	[9]
System Usage (Y)	1. Frequency of Use 2. Use anytime	[9]

4. Research Result

We use path analysis to examine whether there is influence and how strong is the influence between variables in our study. However, before we conduct hypothesis testing, the data must first be tested for validity, reliability and classic assumptions. This is done to ensure that the data is valid, reliable and worth testing.

4.1. Reliability and Validity Testing

A variable is said to be reliable when the Cronbach alpha value is above 0.7 [22], reliable means that even if tested over and over again, the results will remain consistent.

Table 2: Cronbach's Alpha

No.	Variable	Reliability indicator
	List of variables	Cronbach's Alpha
1.	X1 Organizational factor	0.950
2.	X2 Social factor	0.830
3.	X3 Individual factor	0.879
4.	Z1 Perceived usefulness	0.820
5.	Z2 Perceived ease of use	0.861
6.	Y System Usage	0.947

Different from reliability, validity is intended to know whether if a question is asked to many parties then the answer remains the same. Validity can be seen from the r count greater than r table.

Table 3: Validity Testing

Variable			Variable		
Indicator	r count	r table	Indicator	r count	r table
OF			PU		
OF1	0.842	0.195	PU1	0.697	0.195
OF2	0.932	0.195	PU2	0.697	0.195
OF3	0.917	0.195	PEU		
SF			PEU1	0.756	0.195
SF1	0.712	0.195	PEU2	0.756	0.195
SF2	0.712	0.195	SU		
IF			SU1	0.900	0.195
IF1	0.697	0.195	SU2	0.900	0.195
IF2	0.862	0.195			
IF3	0.751	0.195			

4.2. Classical Assumption

This must be performed as a data quality test in testing the linear regression model. Quality multiple regression analysis in this study are free from deviations of assumptions. The classical assumption test, there are several tests including normality, multicollinearity, heteroscedasticity, and autocorrelation.

4.2.1. Normality Test

In this study normality testing is performed using graph analysis and statistical analysis.

The statistical test used in the study is Kolmogorov-Smirnov (K-S) non-parametric statistical test. Data requirements are

normal if the probability or $p > 0.05$ in the Kolmogorov-Smirnov normality test. Kolmogorov-smirnov test value > 0.05 means that the data is normally distributed [22]. If the Kolmogorov-Smirnov test value < 0.05 then the data is not normally distributed. In Table 4 the K-S test value is 0.200 which means it is greater than 0.05, which means that the data are normally distributed.

Table 4: Normality Test

		Unstandardized Residual
N		100
Normal Parameters ^{ab}	Mean	.0000000
	Std. Deviation	1.95132147
Most extreme differences	Absolute	.088
	Positive	.088
	Negative	-.082
Test Statistic		.088
Asymp. Sig. (2 tailed)		.200 ^c

The criteria used in chart analysis are, if the data spreads around a diagonal line and follows the direction of the diagonal line or the histogram graph shows a normal distribution pattern, then the regression model meets the normality assumption.

Based on Figure 1 it can be concluded that the data is spread around the diagonal line, so it can be concluded that the regression model meets the normality assumption

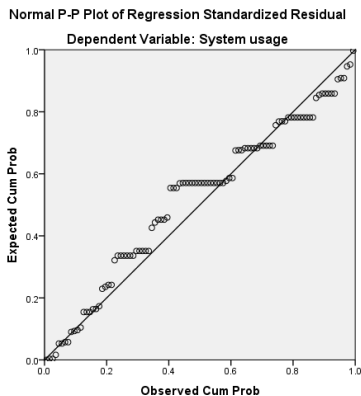


Figure 1: Graphic Normality

4.2.2. Multicollinearity Test

Multicollinearity test aims to test whether the regression model found a correlation between independent variables. A good regression model should not accept a correlation between independent variables. To find out the presence / absence of multicollinearity is to use Variance Inflation Factor (VIF) and tolerance. A low tolerance value is the same as a high VIF value (because $VIF = 1 / \text{Tolerance}$). If the tolerance value ≤ 0.10 or VIF value ≥ 10 , it means that there is multicollinearity, while the tolerance value ≥ 0.10 or VIF value ≤ 10 , it means there is no multicollinearity.

Based on Table 5, the tolerance values for the OF, SF, and IF variables are 0.597, 0.338 and 0.486, also for the PU and PEU variables, each at 0.532, which is greater than 0.100. In addition, the VIF values of the OF, SF, and IF variables of 1,674, 2,956 and

2,059, also for the PU and PEU variables, respectively 1,880, which are all smaller than 10. Based on these results, we can conclude that the research data free from multicollinearity

Table 5: Multicollinearity Test

Model		Collinearity Statistics	
		Tolerance	VIF
1.	(Constant)		
	PU	.532	1.880
	PEU	.532	1.880
	Organizational	.597	1.674
	Social	.338	2.956
	Individual	.486	2.059

4.2.3. Heteroscedasticity Test

Heteroscedasticity test aims to determine whether in the residual regression model there is an imbalance of variance from one observation to another. If the variance from one observation to another is the same, it can be called homoscedasticity and if it is different it can be called heteroscedasticity. Testing heteroscedasticity in this study is using statistical tests and plot graph tests.

The statistics used in this study to determine the presence or absence of heteroscedasticity is the Goldfeld-Quandt Test. The testing steps are as follows:

1. Sort the independent variable X from the smallest largest
2. Then make two separate regressions, first for the smallest X value. Second for large X values and omit some data in the middle.
3. Make the ratio of RSS (Residual Sum of Square = error sum if square) from the second regression to the first regression ($RSS2 / RSS1$) to get the calculated F value.
4. Perform the F test using degrees of freedom of $(n-d-2k) / 2$, where

n = number of observations,

d = the amount of data or observation values lost

k = estimated number of parameters.

F test criteria if:

F arithmetic $>$ F table, then there is heteroskedasticity

F arithmetic $<$ F table, then there is no heteroskedasticity

Results of group I regression with $RSS1 = 381.974$. The results of group II regression with $RSS2 = 814.614$. $F\text{-stat} = RSS2 / RSS1 = 814.614 / 381.974 = 2.132$. F-table is 4.85. It can be concluded that F-statistic $<$ F-table it means that no heteroscedasticity

Heteroscedasticity test plot graph to detect the presence or absence of heteroscedasticity by looking at the plot graph, if there are certain patterns such as points that form certain patterns and orderly, then heteroscedasticity has occurred. If the pattern image spreads above and below the number 0 on the Y axis indicates the absence of heteroscedasticity.

Figure 2 illustrates the spread of scattered data so that it can be concluded that the data in this study is free from heteroscedasticity.

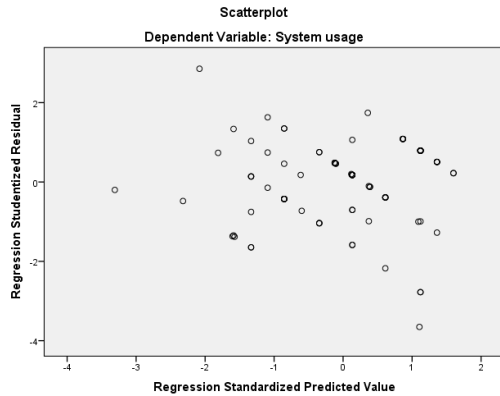


Figure 2: Graphic Heteroscedasticity

4.2.4. Autocorrelation Test

Autocorrelation test was performed using Durbin Watson. If the Durbin Watson value ranges between the upper limit values (dU), then an autocorrelation violation is not expected. The following is a table of autocorrelation test results:

Table 6: Autocorrelation Test

dL	dU	4-dU	4-dL	DW
1.61	1.74	2.26	2.39	1.740

Based on Table 6 it is known that the Durbin Watson (DW) value is 1,740. In the table, we see DW to obtain dL value of 1.61 and dU of 1.74. So, that in the regression equation, the DW value is in the $dU < d < 4-dU$ region. Then H_0 is accepted meaning that the DW value is in the criteria of no autocorrelation. Thus, the assumptions on autocorrelation in the regression equation model have been fulfilled.

We also used Breusch-Godfrey or often called the LM test. In order to detect the presence of autocorrelation, the following are things that can be done:

1. Pay attention to the t-statistic value, R^2 , F test, and Durbin Watson (DW) statistics.
2. Perform LM test (Breusch Godfrey method). This method is based on the values of F and Obs*R-squared, where if the probability value of Obs*R-squared exceeds the level of confidence, then H_0 is accepted. This means that there is no autocorrelation problem.

Testing the autocorrelation hypothesis:

- a. H_0 : autocorrelation does not occur

- b. H_a : autocorrelation occurs
3. If the p-value $Obs^*R\text{-square} < \alpha$, H_0 is rejected.

Following are the results of autocorrelation testing with the Breusch-Godfrey test:

Table 7: Breusch Godfrey Test

F-Statistic	0.316	Probability	0.726
Obs*R-Squared	0.835	Probability	0.657

Based on the results of calculations using the Breusch-Godfrey test, Obtained probability value Obs*R-square is equal to 0.657. This matter means probability $> \alpha = 0.01$, then the conclusion is the level of confidence 99% of the regression models are free from autocorrelation problems.

4.3. Hypothesis Testing

The data has passed the prerequisite test, which means it is time to test the initial assumption / hypothesis. The test is presented in Tables 8 and 9. If the p-value is significant below 0.05 we find that the variable has a significant effect. After having an effect, we also see whether the effect is positive or negative. The direction is seen from the value of the beta coefficient, whether positive or negative. If positive means positive and vice versa. When referring to tables 8 and 9 we can conclude that hypotheses 1, 2 and 3 are all accepted, have a significant effect and are in a positive direction, this result supported previous result in preliminary study by [23].

Table 8: Hypothesis Testing 1

No.	Variable	Unstandar dized B	t	Sig.
1.	PU	-0.319	3.439	0.001
2.	PEU	0.977	10.458	0.000
Dependent Variable: SU				

Table 9: Hypothesis Testing 2

No.	Variable	Unstandar dized B	t	Sig.
1.	PEU	0.689	9.285	0.000
Dependent Variable: PU				

H4 can't be accepted and H5 and H6 can't be rejected, this resulted in line with [24], in other word not supported [16] and [25].

Table 10: Hypothesis Testing 3

No	Variable	Unstandar dized B	t	Sig.
1.	OF	0.009	0.135	0.893
2.	SF	0.297	2.074	0.041

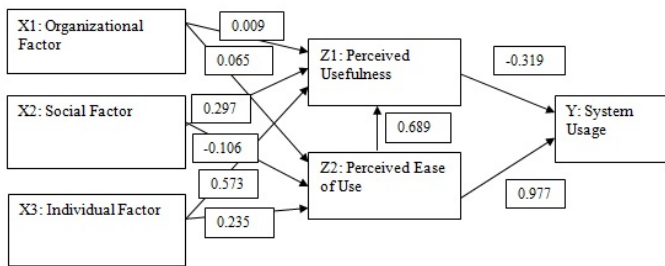
3.	IF	0.573	7.219	0.000
Dependent Variable: PU				

While H7 is can' be accepted, H8 also can't be accepted and H9 positively influenced. We have result that support [26],[25], [27].

Table 11: Hypothesis Testing 4

No.	Variable	Unstandardized B	t	Sig.
1.	OF	0.065	0.830	0.408
2.	SF	-0.106	-0.599	0.551
3.	IF	0.235	2.398	0.018
Dependent Variable: PEU				

The research path coefficient is presented below.



5. Conclusion and Suggestion

After getting the test results statistically, we followed up. The follow up that we do is to conduct interviews and observations of respondents. Our results have found that audit judgment expectancy and expectancy incentives can influence auditors' interest in using GAS. This is because the auditor's performance will increase, when performance increases, then incentives such as salary increases, benefits, bonuses, which are material and also praise/acknowledgement, non-material recognition will accompany their careers. Defers case with age similarity influence; It is true that auditors in their daily life prefer to associate with the same age. However, this association only affects behavior outside of work obligations. For example, it only affects hobbies, games and favorites foods. Meanwhile, to use audit software, it still has to be mandatory from the lead leader.

While other results state that the availability of software and the desire to adopt influence people to actually use it. This is reasonable, because how is it possible for an auditor to use, if the software is not provided by the relevant accounting firm where he works. It is impossible for the auditor to install himself on a personal laptop and use it without the company's approval. For the influence between intention and behavior clearly follows psychological rules, that someone who already wants, will usually continue to use.

Future studies can examine the actual process, on internal auditors and small accounting firms outside the big four and big ten. Future studies can test the extent to which the ability of the www.astesj.com

software. Future studies examine affect partners or owners can increase utilization by providing facilities and infrastructure as well as conducting training and socialization of its use.

References

- [1] B. L. Handoko, Meiryani, S. Sabrina, and N. Ayuanda, "Admission of Information Technology in External Audit Profession: Impact of Organizational, Social and Individual Factors," in *Proceedings of 2019 International Conference on Information Management and Technology, ICIMTech 2019*, 2019, pp. 36–41.
- [2] N. Mahzan and A. Lymer, "Examining the adoption of computer-assisted audit tools and techniques," *Manag. Audit. J.*, vol. 29, no. 4, pp. 327–349, 2014.
- [3] J. J. Schultz, J. L. Bierstaker, and E. O'Donnell, "Integrating business risk into auditor judgment about the risk of material misstatement: The influence of a strategic-systems-audit approach," *Accounting, Organ. Soc.*, 2010.
- [4] S. M. Glover, M. H. Taylor, and C. Western, "Mind the Gap: Why Do Experts Have Differences of Opinion Regarding the Sufficiency of Audit Evidence Supporting Complex Fair Value Measurements?," *Contemp. Account. Res.*, 2019.
- [5] N. K.-G. J. of and U. 2017, "Attitudes and Perceptions Towards Incorporating Computer Assisted Audit Techniques in an Undergraduate Auditing ...," *Theibfr.Com*, vol. 11, no. 3, pp. 55–71, 2017.
- [6] M. Mustapha and S. J. Lai, "Information Technology in Audit Processes: An Empirical Evidence from Malaysian Audit Firms," *Int. Rev. Manag. Mark.*, vol. 7, no. 2, pp. 53–59, 2017.
- [7] F. D. Davis, "Perceived Usefulness , Perceived Ease Of Use , And User Acceptance," *MIS Q.*, vol. 13, no. 3, pp. 319–339, 1989.
- [8] M. Fishbein and I. Ajzen, "Belief, Attitude, Intention and Behaviour: An Introduction to Theory and Research," *Read. MA AddisonWesley*, no. August, p. 480, 1975.
- [9] V. Venkatesh, F. D. Davis, and S. M. W. College, "Theoretical Acceptance Extension Model : Field Four Studies of the Technology Longitudinal," vol. 46, no. 2, pp. 186–204, 2012.
- [10] A. Rogers, "Examining Small Business Adoption of Computerized Accounting Systems Using the Technology Acceptance Model.," *Walden Diss. Dr. Stud.*, p. 126, 2016.
- [11] Y. L. Chen, H. T. Chih, and C. C. Wan, "The relationship between attitude toward using and customer satisfaction with mobile application services: An empirical study from the life insurance industry," *J. Enterp. Inf. Manag.*, vol. 28, no. 5, pp. 680–697, 2015.
- [12] H.-J. Kim, A. Kotb, and M. K. Eldaly, "The use of generalized audit software by Egyptian external auditors," *J. Appl. Account. Res.*, vol. 17, no. 4, pp. 456–478, 2016.
- [13] R. Widuri, B. L. Handoko, and I. E. Riantono, "Perception of Accounting Student on Learning of Generalized Audit Software," in *Proceedings of 2019 International Conference on Information Management and Technology, ICIMTech 2019*, 2019.
- [14] H. J. Kim, A. Kotb, and M. K. Eldaly, "The use of generalized audit software by Egyptian external auditors: The effect of audit software features," *J. Appl. Account. Res.*, 2016.
- [15] 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)*, 2016, pp. 14–19.
- [16] K. Rosli, P. H. P. Yeow, and E.-G. Siew, "Adoption of Audit Technology in Audit Firms," *Proc. 24th Australas. Conf. Inf. Syst.*, pp. 1–12, 2013.
- [17] A. S. Shammari, "An Examination Factors Influencing The Intention To Use Instructional Technologies: An Extension Technology Acceptance Model (TAM)," *Int. J. Inf. Res. Rev.*, vol. 04, no. 02, pp. 3637–3641, 2017.
- [18] A. Ahmi and S. Kent, "The utilisation of generalized audit software (GAS) by external auditors," *Manag. Audit. J.*, vol. 28, no. 2, pp. 88–113, 2013.
- [19] J. B. Chassan, *Research Design in Clinical Psychology and Psychiatry*. New York: Irvington Publishers Inc, 1979.
- [20] J. B. Chassan, "Intensive design in medical research," *Pharmacol. Ther. Part B Gen. Syst.*, vol. 1, no. 1, pp. 139–148, 1975.
- [21] H. J. Kim, A. Kotb, and M. K. Eldaly, "The use of generalized audit software by Egyptian external auditors: The effect of audit software features," *J. Appl. Account. Res.*, vol. 17, no. 4, pp. 456–478, 2016.
- [22] U. Sekaran and R. Bougie, "Research Methods For Business. A Skill Building Approach. 7th Edition," *Book*, 2016.
- [23] A. Tarhini, K. Hone, and X. Liu, "A cross-cultural examination of the impact of social, organisational and individual factors on educational technology

acceptance between British and Lebanese university students,” *Br. J. Educ. Technol.*, vol. 46, no. 4, pp. 739–755, 2015.

- [24] M. Sharif Abbasi, A. Tarhini, M. Hassouna, and F. Shah, “Social, Organizational, Demography and Individuals’ Technology Acceptance Behaviour: a Conceptual Model,” *Eur. Sci. J.*, vol. 11, no. 9, pp. 1857–7881, 2015.
- [25] K. Rosli, P. Yeow, and E.-G. Siew, “Factors Influencing Audit Technology Acceptance by Audit Firms: A New I-TOE Adoption Framework,” *J. Account. Audit. Res. Pract.*, vol. 2012, pp. 1–11, 2012.
- [26] C. Chiu, S. Chen, and C. Chen, “An Integrated Perspective of TOE Framework and Innovation Diffusion in Broadband Mobile Applications Adoption by Enterprises,” *Int. J. Manag. Econ. Soc. Sci.*, vol. 6, no. 1, pp. 14–39, 2017.
- [27] R. Widuri, B. L. Handoko, and I. C. Prabowo, “Adoption of Information Technology in Public Accounting Firm,” 2019.

Performance Effects of Algorithmic Elements in Selected MANETs Routing Protocols

Mutuma Ichaba^{1,*}, Felix Musau², Simon Nyaga Mwendia¹

¹KCA University, Faculty of Computing and Information Management, 00200, Kenya

²Riara University, School of Computing Sciences, 00100, Kenya

ARTICLE INFO

Article history:

Received: 21 February, 2020

Accepted: 20 April, 2020

Online: 03 May, 2020

Keywords:

Mobile Ad Hoc Networks (MANETs)

Routing Protocols

Reactive

Proactive

Hybrid

Location-Aided

ABSTRACT

Over time, several routing protocols have been suggested for use in Mobile Ad Hoc Networks (MANETs). Because of availability of so many MANETs routing protocols, network engineers and administrators face difficulties in identifying an appropriate routing protocol for a particular scenario. This challenge results from the unavailability of adequate technical analytic studies designed to examine the effects of various algorithmic aspects of the available routing protocols. Availability of such studies are critical in routing identification and selection process, thus making the work of network engineers and administrators more manageable. Moreover, such studies can guide future development and implementation of MANETs routing protocols. Although there are studies meant to gauge comparative performance of various routing protocols, very little or no attempts have made to ascertain the effects of nodal topological position-information data on overall performance. This study used purposive sampling to select the routing protocols for study, literature review process to review and critique the available studies and simulation to determine the effects of nodal position-location information. Largely, MANETs routing protocols are either characterized as reactive, proactive, hybrid or location-aided. Through purposive sampling, we selected protocol Zone Routing Protocol (ZRP) to represent hybrid routing whereas Ad hoc On-Demand Distance Vector routing (AODV) to represent reactive routing. Destination-Sequenced Distance Vector routing (DSDV) was selected purposively to represent proactive while Greedy Perimeter Stateless Routing (GPSR) was selected purposively to represent location-assisted routing. Initial elementary scalar variables of data throughput, packet drop rates, average delay and the number of packets received are simulated on NS2—to simulate ZRP and OMNET++—to simulate AODV, DSDV and GPSR. Simulation data from the two simulators was analyzed on RapidMiner. Simulation results indicate that GPSR outperforms other selected routing protocols. This result can possibly be attributed to presence of nodal topological position-location data in GPSR algorithm. However, this study held the number of nodes constant throughout the simulation process. Simulation results suggest that GPSR has better output in packet delivery ratio, delay and overall data packet throughput. The results suggest that inclusion of position-location algorithm in a routing protocol algorithm may enhance its performance. Clearly, the study findings suggest that it is prudent to select and implement a routing protocol that uses nodal position-location in its algorithm.

1. Introduction

In mobile computing, one of the least explored area is Mobile Ad Hoc Networks (MANETs). MANETs is part of wireless networks but does not depend infrastructure to function. Routing is very critical in MANETs because it is very dynamic networks—mobile

nodes can join or leave a MANET autonomously. Indeed, selection of routing protocols to be used in a particular scenario is one of the critical decisions that face network engineers and administrators. There many factors to consider while deciding on a routing protocol to use. Among these factors is overhead (mainly power) utilization by the mobile nodes and efficient propagation of data

* Mutuma Ichaba, Email: icmutuma@gmail.com

packets. Propagation of data needs reduced delay, drop rates and maximized amount of data packets throughput.

Comparative studies on MANETs routing protocols generated from Google Scholar, IEEE depositories, and Springer Publishers reveal that majority of available articles and opinions are very generic in their methodology. For this reason, it is very challenging to attribute performance variations to precise algorithmic characteristics of the examined routing protocols.

Because MANETs are mostly deployed in areas whose infrastructure is non-existent or damaged, power management and reliable transfer of data packets is very vital. For example, in a military operation, the mobile devices must operate on minimum power to avoid catastrophic failures. But this power efficiency is only achievable through keen selection of an appropriate routing protocol that functions on minimum power as an overhead maximized data packets delivery is paramount. Reliable data packets transmissions demand minimization of other aspects of routing such a jitter, delay (latency) and packet's drop rates.

Determining the best combination of these factors or aspects of MANETs routing is paramount. However, this is only possible through a focused comparative study on specific algorithmic characteristics of the routing protocols. Evidently, however, majority of the comparative literature on MANETs routing protocols are very general. That is, performance comparison is conducted without stipulating the possible effecting traits behind the variations in results [1].

In most cases, routing in MANETs is either proactive, reactive or a combination of both (hybrid routing). So far, very little research has been conducted on location-assisted routing protocols such a Location-Aided routing protocol (LAR) or Greedy Perimeter Stateless Routing (GPSR). Studies [2] and [3] are some of the limited studies that do not consider specific aspects of MANETs routing protocols.

Indeed, majority of research papers concentrate too much on either proactive, reactive or hybrid routing protocols—see [4] and [5] as examples. This imbalance in research leaves the location-assisted routing protocols largely unexplored. Because comparative examination of proactive, reactive and hybrid routing protocols require analysis of two counterintuitive routing concepts only, it remains unclear how inclusion of algorithmic position-location information may affect their performance. This is particularly problematic in routing protocols implementation by network engineers and administrators because the current research is not comprehensive and specific enough to allow informed selection and implementation of routing protocols.

Typically, MANETs are deployed in situations occurring in places whose communication infrastructure is either unreliable or lacking altogether. Subsequently, MANETs are habitually installed in disaster and rescue operations, military operations and lately, in the delivery of learning to the marginalized areas [6]. All these situations require reliable routing protocols that guarantee

intelligent utilization of critical overhead such as power and maximized delivery data of packets at steady speeds.

A key trait in MANETs is the ability of mobile nodes to self-configure wirelessly. According to [7], nodes in MANETs use adjacency and immediacy of nodes in signal transmission and data packet delivery. Therefore, nodes that are next to each other convey signal and data packets between them. This chain of events repeats itself throughout the network. Another important characteristic of MANETs is ability node to play a multi-purpose role of sensing, receiving, routing and transmitting. Additionally, MANETs can multi-hop and self-configure. This trait introduces fluidity and flexibility as nodes autonomously come in or out of the network. Mutability of MANETs make them very useful for varied emergency situations such as floods, fire, storm or earthquake evacuations [7].

To achieve appropriate characterization of such emergent scenarios and response, it is possible to introduce and use the concept of clustering in MANETs [8]. Clusters in MANETs are composed of nodes dedicated for communication in such cases—military operations or emergency.

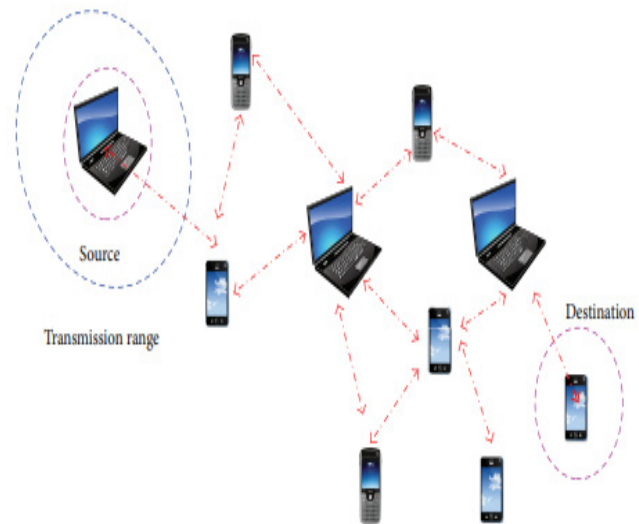


Figure 1: Example of a MANET network. Adapted from [8]

Description of various ways of signal transmission and data packet delivery in MANETs is conducted through examination of routing protocols [9]. Due to the abrupt, endless and unregulated participation in the network by the nodes, MANETs topological information is perpetually changing. This constant change of topological information of nodes in MANETs is considered as one of the hindrances to the possible reliability and security. Different and new nodes are included or excluded from the network. However, the efficacy of packets and the signal broadcast inside a network is primarily reliant on the routing protocol selection [9]. A routing protocol in MANET decides on the utmost proficient route from the source node to the destination node.

Categorization of routing protocols in MANETs is carried out through various methods. Generally, the topological information

of network is used in categorizing a routing protocol as either reactive or proactive. On the other hand, the plan, strategy or the general design of a MANET communication applied during signal and data packet transmission may result to either multicast, broadcast, or unicast routing [10]. Moreover, MANETs routing can be categorized into groups that depend on routing network topological tables or into groups that require route request from a source to initiate routing [11]. However, some studies suggest that this classification is arrived at through application of routing strategy.

The structure of the network is another method applicable in the classification of MANETs routing protocols. Classifying of routing protocols on network structure, generally, three classes are derived. That is, the group of routing that is aided by geographic information, flat routing and hierarchical routing. Network topological routing table dependent and the source node route discovery routing protocols are included in flat routing. On the other hand, routing protocols in MANETs mainly lie in either proactive, reactive and or hybrid categories [11].

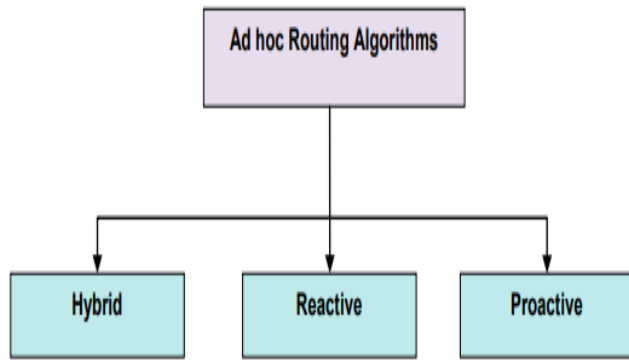


Figure 2: Routing protocols groupings in MANETs based on design philosophy [11].

Network topological routing tables maintain the network information on routing over the connectedness period. However, the topological routing information maintained by proactive routing protocols are not used constantly during the network’s state of connectedness. Because MANETs connect on peer-to-peer basis, every mobile node in a proactive network maintains routing information of the neighboring devices. Proactive protocols rely on network topological routing tables to accumulate and maintain routing data on every neighboring mobile node.

One of the shortcomings of table-driven routing is the need to occasionally update according to the constant and abrupt inclusion and exclusion of mobile nodes. MANETs nodes signal and data packets propagation derive its guidelines from the link state of nodes. Therefore, a signal propagation and routing procedure is established through routing information as stored and maintained in a network topological routing table. Furthermore, proactive routing protocols can be made up of more than one type of data in a routing table [12]. For example, a routing table may contain the nodal-zonal information or the link-state of a mobile device/node.

Because table-driven routing protocols preserve routing statistics of every mobile node, they are wasteful and unfitting for use in

highly interconnected ad hoc networks. But, if table-dependent protocol is used in a highly interconnected network, the routing information available for storage and maintenance may exceed the capacity of routing tables. Consequently, this may lead into additional overhead of more power consumption. Oppositely, reactive routing protocols avoid holding onto routing information—topological, of nodes in routing tables. Similarly, reactive protocols do not maintain network events in tables. Reactive protocols initiate routing process on request by the source nodes. During the on-demand event, the concerned nodes articulate the most efficient routes in a network. Reactive routing transmits data and the related signals as requested by other nodes. However, lack of routing network topological tables in reactive routing may require route flooding during route discovery process [12].

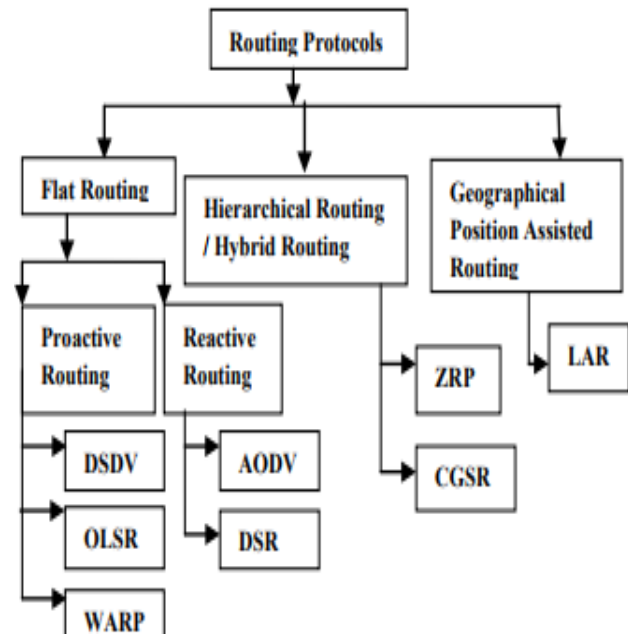


Figure 3: Classification of MANET Routing [12]

In hybrid routing protocols, the operational design and capabilities of proactive and reactive routing are unified to maximize effectiveness. Because hybrid routing is designed to minimize routing overhead—such as power consumption, they are mostly deployed in large networks. Routing overhead results from the systematic topological data apprisers and maintenance in the routing tables. Besides, hybrid routing minimizes data packet delivery delay in table-driven (reactive) routing protocols [13]. Latency results from route discovery procedure initiated in the reactive part of hybrid routing. Over period of time, latency reduces the overall number of data packets delivery per unit.

In Zone Routing Protocol (ZRP), zones—determined through hops, are composed of nodes whose location is defined by the maximum size of the radius. Zones in ZRP can be conceptualized as small networks within larger networks. The radius of a subnetwork is determined by the number of hops. For example, a radius of two is equivalent to the zone whose maximum hops are 2. Because ZRP uses hybrid routing, it utilizes operational technicalities of both proactive and reactive routing. A local zone uses table

driven (proactive routing), while nodes outside local zones use reactive routing. Local zones use proactive routing protocols for speedy transmission of signals and data packets among nodes. To minimize possible overhead, ZRP uses reactive routing for inter-zonal transmissions [14].

Since DSDV is a table-drive (proactive) routing protocol, it relies on network topological routing tables to store and preserve the network information [15]. Mobile nodes in DSDV uphold routing tables containing topological information and data of the network. These routing tables maintain information on addresses, the number and subsequent hops to destination node, and the sequence assigned to the various nodes—particularly, the destination node. DSDV uses the concept of Routing Information Protocol (RIP). In RIP, each node carries routing tables whose routing information includes probable end points of communications (messages) and the hop quantity vital in enabling transmission of the messages in the ad hoc network. Another trait of DSDV is distance vector routing. That is, it utilizes bidirectional links. A key short coming of DSDV is its limitation to utilize one/single route in internodal communications.

Greedy Perimeter Stateless Routing (GPSR) protocol relies on location information of nodes during route discovery. In GPSR, forwarding a packet is built on the known position of the routers/source node and the destination node. It a “novel routing protocol for wireless datagram networks that uses the positions of routers and a packet’s destination to make packet forwarding decisions” [15, p.5]. The information stored in topological routing tables is utilized in greedy propagation of data throughout the nodes in ad hoc network. Upon reaching the limit of greedy transmission, GPSR uses a set of rules (algorithm) to traverse the limits of the network.

Notably, with the increase in the quantity of destination nodes, GPSR “scales better in per-router state than shortest-path and ad-hoc routing protocols” [15, p.1) because it only maintains the topological link state the local topology in a network. Since Manets are dynamic—mobile nodes are autonomous to either join or leave the network, their topologies are subject to variations. But, GPSR handles this topological dynamism and its effects through exploitation of the local network topological data for route discovery. Mostly, GPSR is appropriate require highly connected/dense networks.

Apparently, all MANETs routing protocols are unique in their algorithmic composition. However, virtually all performance comparative studies evaluate these routing protocols on general terms without ascribing the variations of their performance to any exact algorithmic feature. With such generality, it becomes problematic to use them for routing selection by the network engineers and administrators. Moreover, it becomes difficult to use such studies for future MANETs routing designs. For this reason, it is critical that comparative performance analysis grounded on particularity of these unique algorithmic foundations. The research question of this study is:

1. What algorithmic elements attributable to variations in performances of selected MANETs routing protocols?

2. Theoretical Background

Performance comparison studies of MANETs routing protocols are not new. However, majority of these studies conduct performance analysis in general, without pinpointing to specific possible algorithmic aspects that affects variations in performance. For instance, study [16] compares the general performance of Dynamic Source Routing (DSR), Ad hoc on demand distance Vector Routing (AODV), and Destination Sequenced Distance Vector (DSDV) to establish the effectiveness of each of the protocols. Study [16] measures packet delivery ratio, average throughput, average end-to-end delay, and packet loss ratio based on nodes as the variable.

Likewise, paper [17] studies ad hoc on demand distance vector (AODV), dynamic source routing (DSR), and destination-sequenced distance vector (DSDV) routing protocols by means of various parameters of Quality of Service (QoS). These metrics include packet delivery ratio (PDR), normalize routing overhead, throughput, and jitter. The study aimed to establish the difference in performance of the listed protocols. Characterization of the simulation scenario in this study was based on parameters depicting large-area MANET with high-speed mobile nodes. Simulation results are thereafter analyzed through AWK and Xgraph.

Meanwhile, study [18] compares the performance of three routing protocols—DSDV, DSR and AODV. Finally, the paper presents the overall properties and performance of these protocol. Markedly, however, the study is not clear on the significance and types of parameters simulated or analyzed. Other works such as [19], uses literature review to compare various protocols on the basis of some selected performance metrics such as scalability, overhead, reliability. In this study, however, the tools of research or study are not clear. Additionally, the criteria for selecting routing protocols for analysis is not clearly indicated.

Paper [20] comprehensively compares AODV, DSR and OLSR on a real-word scenario—office. The study compares the routing protocols in three probable situations (based on mobility) via UDP, Ping and TCP traffic. The study also tries to determine their comparative performance variation among simulation, emulation and the real world. Matching protocol employments applied in each of the environments. A key shortcoming of this paper is its failure to highlight the reasons behind the selection of study and comparison routing metrics. As such, this disadvantage diminishes the contribution of this study. Other performance based comparative studies include [21] that presents comparison of the prevalent broadcast-based MANETs routing protocols. The study examines the performance of Dynamic Source Routing (DSR)—IEEE 802.11 standard based protocol, on performance metric parameters of throughput, end-to-end packet delay, jitter and salvaged packets using. Performance evaluation was done on QualNet 5.0.2 network simulator.

Studies such as [22] evaluates and compares the performance DSDV, AODV and DSR. It bases the examination on metrics of throughput, packet delivery ratio and average end-to-end delay simulated NS-2. However, the paper fails to present a detailed case where either a single trait or a group of routing characteristics are pinpointed based on their effects on the presented performance. Similarly, studies [23]-[30] conducted various analysis on basic and common metrics such as data packet delivery ratio, general packet outputs, delays and latency. A key limitation of these studies is the lack of precision on routing protocol selection methodology and the lack of specificity on effects of various characteristics of the compared routing protocols. Consequently, it is challenging to attribute the differences in performance to any particular characteristic or principle of the routing protocols.

3. Study Methodology

To identify and select the MANETs routing protocols to be evaluated, this study used purposive sampling method. A purposive sample is also called judgmental or expert sample [31]. It is non-probabilistic in its approach to sampling. The core goal of a purposive sampling is to obtain a sample that is a reasonable representative of the population. To select the routing protocols to study and compare, this study used the following criteria;

- Zone Routing Protocol (ZRP) represents hybrid routing protocols
- Ad hoc On-Demand Distance Vector routing (AODV) represents reactive routing protocols.
- Destination-Sequenced Distance Vector routing (DSDV) represents proactive routing
- while Greedy Perimeter Stateless Routing (GPSR) represents location-based routing protocols

Review of the selected routing protocols followed the steps outline in [32]. Before simulation, we formulated the following research questions from pre-liminary review of the existing studies on MANETs routing protocols.

- How are current comparative performance studies in MANETs limited?
- What is the general approach in MANETs routing protocols comparative performance studies?

Identification of answers to the above questions was followed by critical analysis and evaluation of the current studies on MANETs routing protocol comparative performance. Literature review and analysis suggests that the current studies do not try to associate certain traits of routing protocols to differences in performance. The following research questions cannot be answered by the current studies:

- What algorithmic aspects of the routing protocols that could possibly cause the differences in performances?
- What is attributable to the differences in performance in MANETs routing protocols?

Without an answer to the above research questions, it introduces generality that is not helpful in future designs of MANETs routing protocols. Additionally, without answers to above questions, it becomes challenging for network engineers and administrators to select the appropriate routing for given scenarios.

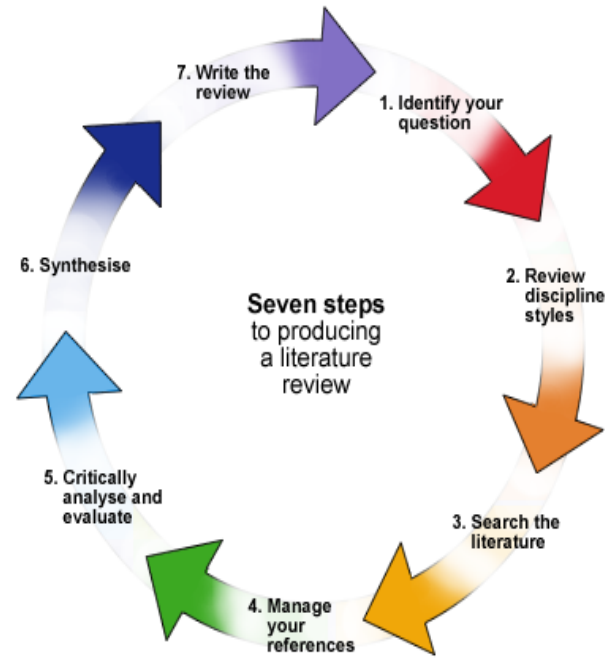


Figure 4: Literature Review and Analysis Steps [32]

After selection of the above listed routing protocols, both NS-2 and OMNET++ was used for simulation. NS2 was used to simulate ZRP while AODV, DSDV and GPSR were simulated on OMNET++. Simulation data from the two simulators was analyzed on RapidMiner. General review of literature followed the steps outlined in [32].

4. Results and Discussion

Metrics on data packets received, latency, drop rates and throughput of packets relating to ZRP were obtained from NS2 simulator while OMNET++ was used to generated simulation data for AODV, DSDV and GPSR. It should be clear that this study seeks only to analyze the effects of algorithmic nodal position information in selected MANETs routing protocols and not to a comprehensive study of performances under combinations of various parameters. Indeed, this study seeks to examine the inclusion of algorithmic location information at basic level. Notable from the parameter table 2 is the lack of various common parameters necessary for comprehensive performance analysis in MANETs routing protocols.

This has been done through design because the purpose of this study is to uses basic parameters to determine the algorithmic elements of initiating performance variations in selected routing protocols. For example, this study leaves out jitter, power and simulation node density.

Table 1: Preliminary parameters

Parameter	Metric
Quantity of simulation nodes (x)	100
Simulation Area (a)	500m*500m
Simulation Packet Size (p)	512 bytes
Simulation time (t)	1000s
Simulations(n)	50

Data packets received refers to the packets difference between the quantity sent and the quantity lost. Data packets successfully transmitted is entire amount of data packets that effectively reaches the intended destination node over the period of simulation. Data packets transmitted from the source node comprise of the packets sent. This fractional metric is created on the count of Constant Bit Rate (CBR) from the source node to the destination. To calculate the metric for the packets delivery ratio (PDR), the following formula is used:

$$PDR = \frac{\text{Packets Received}}{\text{Packets Sent}}$$

However, the aim of this study was not to conduct a general performance analysis, but rather, basic examination of algorithmic code on the selected MANETs routing protocols. Packet Delivery Ratio, however, shows the relationship between the data packets received successfully and the packets sent. A low count of packets received results to a lower PDR and indication of a big loss across the network. A big packet loss may suggest a possible technical problem in the network layer. Low PDR may also suggest a reduced normalized routing load (NRL), which measures the ratio between data packets delivered successfully and the original quantity of data transmitted.

Nonetheless, this study is not designed to conduct a substantive analysis of comparative performance but rather to answer the following research question:

- What algorithmic aspect of the selected routing protocols is attributable to difference in performance?

Table 1 and figure 1 show the variation in the quantity of data packets successfully transmitted from a source node to the destination. From the table and the figure, it is evident enough that, however small in variation, GPSR steadily outperforms ZRP, DSDV, and AODV. Algorithmically, DSDV and AODV use hops to determine a route from the source node to the destination node. Despite DSDV depending on network topological tables to initiate routing process, it steadily outperforms AODV which depends on queries from source nodes to initiate routing. This could be attributable to the algorithmic element to constant update network topological tables, hence faster access to route information. Easily accessible route information speeds up the routing process. The constant updates of routing tables by may utilize more power in

the form of overhead, but this pays off in form of augmented routing efficiency.

Moreover, ZRP—a hybrid protocol, seems to outperform both DSDV and AODV. This variation in routing performance can be credited to its utilization of algorithmic operational qualities in both AODV and DSDV. That is, its algorithmic properties to combine proactive and reactive routing features. Conspicuous in the table 1 and figure 1 is the abrupt improved performance of GPSR at around 700th second of simulation. Consequently, the usage of nodal location as key determinant of routes in GPSR seems to outperform other selected routing protocols. But this phenomenon appears to exemplify with simulation time. One implication of such a routing behavior is dependability. Evidently, figure 1 indicates that inclusion of nodal location information can be dependable network stabilizer, thus enhancing its dependability. In this case, for example, it is plausible for a network engineer or administrator to suggest deployment of GSPR in a military or rescue operations because they all demand stable and dependable networks.

Because distance vector routing protocols determine best routes on relative distance—hops, AODV and DSDV appear to be efficient but relatively unreliable for networks that require stability. ZRP on the other hand, seems relatively less reliable to GPSR. Although distance vector routing algorithm decides on best routing paths on basis of fewest hops, they are unstable and unreliable compared to location-assisted and hybrid routing.

Table 2: A Summary of Data Packets Successfully in Selected Routing Protocols.

GPSR	ZRP	DSDV	AODV	Time in Seconds
9254	9100	9000	8500	100
8951	8793	8695	8650	200
10861	10733	10620	10531	300
11064	10937	10848	10788	400
12584	13637	12193	11926	500
13502	13111	13099	14000	600
14302	14150	13973	13600	700
13064	11941	9491	10267	800
12584	12391	10321	11927	900
12987	11926	11871	11876	1000

Visible in figure 5 is the trend that packets delivery is growing steadily with time. Particularly, on average, GPSR has more packets delivered from the source to the destination while experiencing least packet loss. Exhibiting better performance in comparison to the other two protocols—DSDV and AODV, is ZRP. One of the possible causes of this better performance can be attributed to presence of location-position data in the routing tables of a nodes in GPSR. With such information, possible flooding of nodes is

avoided during route discovery process. Flooding avoidance in IERP could also be associated to its second-best data delivery.

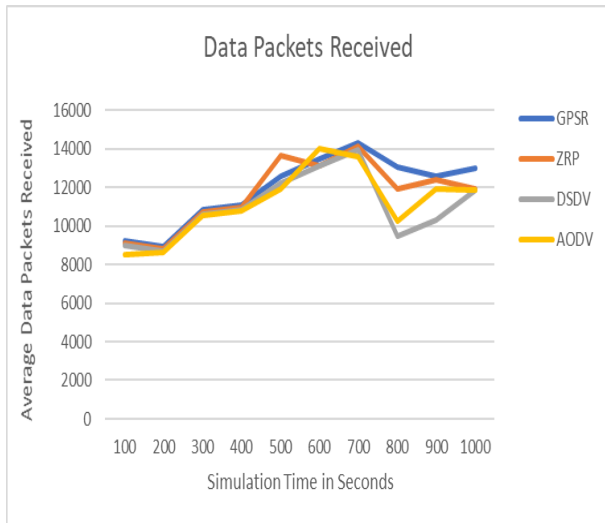


Figure 5: Data Packets Received Successfully in Selected Routing Protocols

Average delay determines the average time take by a single data packet from the source node to the destination node. Data packet delivery delay can be caused by route discovery process, broadcasting/transmission buffer, retransmission requests, or general signal latency. High delay time is an indication of inefficiency in the network layer—as outlined in OSI model. Highly delay can be caused by factors such as network congestions or low network transmission capacity among other reasons. Average delay is derived from subtracting Packet Send Time (PST) from Packet Receive Time (PRT), then divide it by the successfully received packets. Packet Transmission Delay (PTD) metric is calculable by using the following formula:

$$PTD = \frac{\text{Packet Receive Time} - \text{Packet Send Time}}{\text{Successfully Received Packets}} \quad (1)$$

Table 3 and figure 3 represent average delay results as recorded from the simulation of the selected routing protocols. Delay in GPSR seems to be lowest among the selected routing protocols. Perhaps, inclusion of nodal location information element in its routing algorithm speeds up route discovery process. Delay in GPSR seems to start comparatively high but reduces with time. This is explaining why GPSR, comparatively, takes longer to outpace other selected routing protocols. Similarly, ZRP begins on a relatively higher delay but makes up the ground with simulation time. Delay dormancy in both of these routing protocols—GPSR and ZRP, is attributable to their uniquely algorithmic routing features compared to the other selected routing protocols.

As GPSR and ZRP minimize delay with simulation time, AODV and DSDV begin to increase their delay. Since AODV and DSDV use hops—however differently, to determine the best paths, their initial low delay can be attributed to average distances of adjacent/neighbor mobile nodes. Initial low delay suggests high reachability of the neighboring nodes, while high delay indicates

low reachability of nodes further away from the source node. High delay at the beginning simulation in ZRP points to an existence of complexity as a result of combining two mutually exclusive algorithmic features—reactive and proactive routing elements. Similarly, GPSR introduces complexity in the initial stages due to location calculation. The following formulas determine the location—distance and angular location, information in GPSR respectively.

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (2)$$

$$\angle = \tan^{-1} \left(\frac{y_2 - y_1}{x_2 - x_1} \right) \quad (3)$$

GPSR	ZRP	DSDV	AODV	Time in Seconds
0.098	0.099	0.0119	0.0125	100
0.087	0.092	0.0108	0.0121	200
0.068	0.073	0.081	0.095	300
0.045	0.051	0.062	0.072	400
0.019	0.021	0.031	0.043	500
0.06	0.067	0.073	0.089	600
0.055	0.059	0.064	0.071	700
0.068	0.074	0.074	0.078	800
0.045	0.059	0.069	0.072	900
0.019	0.041	0.035	0.069	1000

In figure 4, GPSR consistently outperforms ZRP, DSDV and OADV in average delay of packet delivery. Although, generally the packets delay in all four routing protocols decrease overtime, delay in GPSR declines the most. A key deduction from this metric—packet delivery delay, is that MANETs routing protocols utilizing the position-location information reduces flooding and overlapping of zones, thus achieving overall better performance over a period of time. This holds true for packets drop rate and packets throughput as manifested in figures 7 and 8, respectively.

Packets dropped during transmission represent the lost packets during routing. Table 4 and figure 7 show that comparatively, GPSR has a slight edge in the number of packets dropped during routing. Generally, however, all other three selected routing protocols significantly experience reduced packets drop rates over the course of simulation period. Lower packets drop rates in GPSR can be ascribed to availability of location information of destination nodes. Location information increases precision of location of network nodes—despite their movements within a network. This means that GPSR, unlike the other three selected routing protocols, constantly calculates the locations of its network nodes. Equipped with nodal location data, GPSR is able to propagate data packets with more precision than other selected routing protocols.



Figure 6: Packets Delay

Dropped or lost data packets can be calculated by subtracting the number of packets successfully transmitted from the number of data packets successfully received, then divided by the transmission time. The following formula is applicable:

$$= \frac{\text{Dropped or Lost Packets}}{\text{Successfully Received Packets} - \text{Successfully Sent Packets}} \times \text{Transmission Time} \quad (4)$$

Table 4: Packets Drop Rates in Selected Routing Protocols

GPSR	ZRP	DSDV	AODV	
5326	5339	5438	5531	100
6784	6869	6912	6954	200
4586	4625	4735	4891	300
3400	3981	3999	4103	400
1105	1230	1421	1531	500
627	936	999	1023	600
521	950	962	964	700
625	1020	1025	1037	800
500	925	947	981	900
607	890	903	955	1000

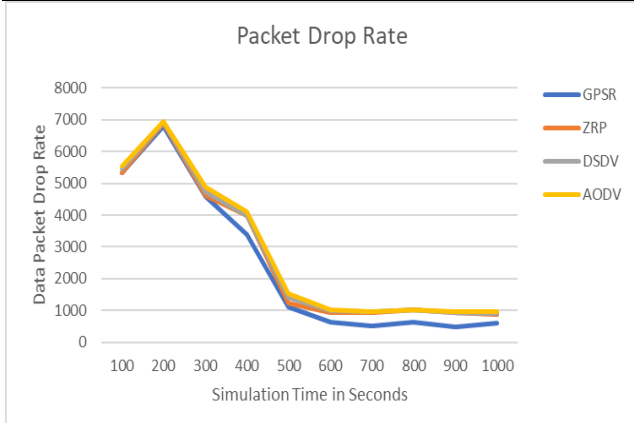


Figure 7: Packets Drop Rate

Table 5: Packets Throughput Across Selected Routing Protocols

GPSR	ZRP	DSDV	AODV	
13624.01	13523	13267	12348	100
14215.32	13181	13080	12670	200
14832.97	14735	14621	14210	300
15462.76	15371	15213	15104	400
18952.85	17731	17512	17111	500
18108	18000	17836	17000	600
18321	17900	17520	17390	700
19001	18273	18040	17900	800
15000	17671	9439	9100	900
19000	17943	17391	16890	1000

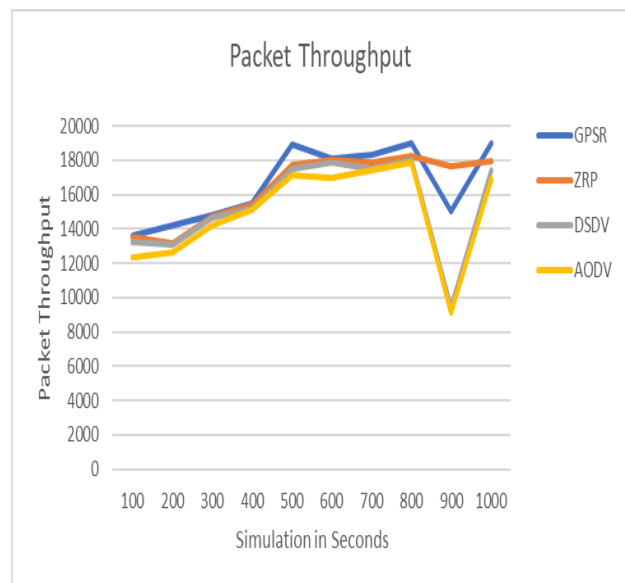


Figure 8: Packets Throughput

5. Conclusion and Further Recommendations

Comparative performance studies on MANETs routing is one of the highly examines areas of networking. A keen enquiry, however, reveals that majority of performance analysis research works are very general. Absence of provenance to exclusive algorithmic elements in MANETs routing protocols make hard for network engineers and administrators to develop, select and implement routing. Moreover, such studies fail to make significant contributions vital in future design, development and implementation of MANETs routing.

This study attempts to investigate the effects of various unique algorithmic elements in selected MANETs routing protocols. Through purposive sampling, systematic literature review and simulation tools and analysis (NS2, OMENT++ and RapidMiner), we strive to associate variations in MANETs performances to identified algorithmic elements such position-location nodal topological information.

However, we must note that this study is basic in its approach. For instance, we left out critical parameters during simulation. This was done purposely to delimit the research to its topic variables. By simulating these purposively selected MANETs routing protocols—GPSR, ZRP, DSDV, AODV, the results suggest that GPSR, over time, outperforms the selected MANETs routing protocols on: data packets delivery, delay and overall throughput. Better performance of GPSR could be, generally, attributed to its use of algorithmic nodal location-position information elements.

However, this study used time (fixed at 1000 seconds) as variable metric for simulation. Consequently, it is suggested that further simulations be conducted on composite variables of number of nodes. Using number of nodes as the basis (variable factor) for investigation, can reveal the most effective MANETs routing protocols for small, medium and large networks.

Conflict of interest

The authors declare that there is no conflict of interests in relation to the publication of this paper.

References

- [1] P. Manickam, T. Guru Baskar, M. Girija and D. Manimegalai, "Performance Comparisons of Routing Protocols in Mobile Ad Hoc Networks", *International Journal of Wireless & Mobile Networks*, vol. 3, no.1, pp.98-106, 2011. Available: 10.5121/ijwmn.2011.3109 [Accessed 14 April 2020].
- [2] M. Kassim, R. Ab. Rahman and R. Mustapha, "Mobile ad hoc network (MANET) routing protocols comparison for wireless sensor network", *2011 IEEE International Conference on System Engineering and Technology*, 2011. Available: 10.1109/icsengt.2011.5993439 [Accessed 14 April 2020].
- [3] R. Ferdous and V. Muthukumarasamy, "A Comparative Performance Analysis of MANETs Routing Protocols in Trust-Based Models", *2016 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2016. Available: 10.1109/csci.2016.0171 [Accessed 14 April 2020].
- [4] K. Zaini, A. Monzer Habbal, F. Azzali and M. Abdul Rejab, "Comparative Study on the Performance of TFRC over AODV and DSDV Routing Protocols", *Software Engineering and Computer Systems*, pp.398-407, 2011. Available: 10.1007/978-3-642-22203-0_35 [Accessed 14 April 2020].
- [5] P. Nand and S. Sharma, "Comparison of Routing Protocols for MANET and Performance Analysis of DSR Protocol", *Communications in Computer and Information Science*, pp. 406-412, 2011. Available: 10.1007/978-3-642-18440-6_52 [Accessed 14 April 2020].
- [6] D. Ismail and M. Ja'afar, "Mobile ad hoc network overview", *2007 Asia-Pacific Conference on Applied Electromagnetics*, 2007. Available: 10.1109/apace.2007.4603864 [Accessed 16 February 2020].
- [7] S. Meshram and P. Dorge, "Design and performance analysis of mobile Ad hoc network with reactive routing protocols", *2017 International Conference on Communication and Signal Processing (ICCS)*, 2017. Available: 10.1109/iccsp.2017.8286396 [Accessed 16 February 2020].
- [8] S. Omari and P. Sumari, "An Overview of Mobile Ad Hoc Networks for the Existing Protocols and Applications", *International Journal on Applications of Graph Theory in Wireless Ad Hoc Networks and Sensor Networks*, vol. 2, no. 1, pp. 87-110, 2010. Available: 10.5121/jgraphoc.2010.2107 [Accessed 16 February 2020].
- [9] L. Cai, J. Pan, X. Shen and J. Mark, "Peer Collaboration in Wireless Ad Hoc Networks", *NETWORKING 2005. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems*, pp. 840-852, 2005. Available: 10.1007/11422778_68 [Accessed 16 February 2020].
- [10] A. Veerasamy, S. Madane, K. Sivakumar and A. Sivaraman, "Angle and Context Free Grammar Based Precarious Node Detection and Secure Data Transmission in MANETs", *The Scientific World Journal*, vol. 2016, pp. 1-12, 2016. Available: 10.1155/2016/3596345 [Accessed 16 February 2020].
- [11] S. Omari and P. Sumari, "An Overview of Mobile Ad Hoc Networks for the Existing Protocols and Applications", *International Journal on Applications of Graph Theory In Wireless Ad Hoc Networks And sensor Networks*, vol.2, no.1, pp. 87 110, 2010. Available: 10.5121/jgraphoc.2010.2107 [Accessed 16 February 2020].
- [12] R. Thiagarajan and M. Moorthi, "Efficient routing protocols for mobile ad hoc network", *2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*, 2017. Available: 10.1109/aeecib.2017.7972346 [Accessed 16 February 2020].
- [13] H. Moudni, M. Er-rouidi, H. Mounicif and B. Hadadi, "Secure routing protocols for mobile ad hoc networks", *2016 International Conference on Information Technology for Organizations Development (IT4OD)*, 2016. Available: 10.1109/it4od.2016.7479295 [Accessed 16 February 2020].
- [14] N. Saudi, M. Arshad, A. Buja, A. Fadzil and R. Saidi, "Mobile Ad-Hoc Network (MANET) Routing Protocols: A Performance Assessment", *Proceedings of the Third International Conference on Computing, Mathematics and Statistics (iCMS2017)*, pp. 53-59, 2019. Available: 10.1007/978-981-13-72797_7 [Accessed 16 February 2020].
- [15] M. Ichaba, "Examining Possible Supplementary Nature of Routing Protocols in Mobile Ad-hoc Networks (MANETs): A Discussion", *OMICS Online*, vol. 09, no. 03, 2018. Available: 10.4172/2229-8711.1000229 [Accessed 16 February 2020].
- [16] T. Feiroz Khan and D. Sivakumar, "Performance of AODV, DSDV and DSR protocols in mobile wireless mesh networks", *Second International Conference on Current Trends In Engineering and Technology - ICCTET 2014*, 2014. Available: 10.1109/icctet.2014.6966324 [Accessed 16 February 2020].
- [17] S. Ali and A. Ali, "Performance Analysis of AODV, DSR and OLSR in MANET", Masters, Blekinge Institute of Technology, 2020.
- [18] E. Cana, "Comparative Performance Simulation of DSDV AODV and DSR MANET Protocols in NS2", *International Journal of Business & Technology*, vol. 2, no. 1, pp. 24-31, 2013. Available: 10.33107/ijbte.2013.2.1.4.
- [19] M. Alslaim, H. Alaqel and S. Zaghoul, "A comparative study of MANET routing protocols", *The Third International Conference on e-Technologies and Networks for Development (ICeND2014)*, 2014. Available: 10.1109/icend.2014.6991375 [Accessed 16 February 2020].
- [20] S. Mohapatra and P. Kanungo, "Performance analysis of AODV, DSR, OLSR and DSDV Routing Protocols using NS2 Simulator", *Procedia Engineering*, vol. 30, pp. 69-76, 2012. Available: 10.1016/j.pro-eng.2012.01.835 [Accessed 16 February 2020].
- [21] T. Devi, "IMPLEMENTATION OF DYNAMIC SOURCE ROUTING (DSR) IN MOBILE AD HOC NETWORK (MANET)", *International Journal of Research in Engineering and Technology*, vol. 02, no. 11, pp. 339-345, 2013. Available: 10.15623/ijret.2013.0211053 [Accessed 16 February 2020].
- [22] T. Devi, "IMPLEMENTATION OF DYNAMIC SOURCE ROUTING (DSR) IN MOBILE AD HOC NETWORK (MANET)", *International Journal of Research in Engineering and Technology*, vol. 02, no. 11, pp. 339-345, 2013. Available: 10.15623/ijret.2013.0211053 [Accessed 16 February 2020].
- [23] E. Mahdipour, A. Rahmani and E. Aminian, "Performance Evaluation of Destination-Sequenced Distance-Vector (DSDV) Routing Protocol", *2009 International Conference on Future Networks*, 2009. Available: 10.1109/icfn.2009.51 [Accessed 16 February 2020].
- [24] M. Alslaim, H. Alaqel and S. Zaghoul, "A comparative study of MANET routing protocols", *The Third International Conference on e-Technologies and Networks for Development (ICeND2014)*, 2014. Available: 10.1109/icend.2014.6991375 [Accessed 16 February 2020].
- [25] S. Goswami, S. Joardar, C. Das, S. Kar and D. Pal, "Performance Analysis of Three Routing Protocols in MANET Using the NS-2 and ANOVA Test with Varying Speed of Nodes", *Ad Hoc Networks*, 2017. Available: 10.5772/66521 [Accessed 16 February 2020].
- [26] A. Alfa, S. Misra, A. Adewumi, F. Salami, R. Maskeliūnas and R. Damaševičius, "Implementation of MANETs Routing Protocols in WLANs Environment: Issues and Prospects", *Advances in Intelligent Systems and Computing*, pp. 252-260, 2018. Available: 10.1007/978-3-319-74980-8_24 [Accessed 16 February 2020].

- [27] D. Sudarsan, P. Mahalingam and G. Jisha, "Distance Aware Zone Routing Protocol for Less Delay Transmission and Efficient Bandwidth Utilization", *Advances in Intelligent Systems and Computing*, pp. 63-71, 2012. Available: 10.1007/978-3-642-30111-7_7 [Accessed 16 February 2020].
- [28] S. Shelly and A. Babu, "Link Reliability Based Greedy erimeter Stateless Routing for Vehicular Ad Hoc Networks", *International Journal of Vehicular Technology*, vol. 2015, pp. 1-16, 2015. Available: 10.1155/2015/921414 [Accessed 16 February 2020].
- [29] B. Karp and H. Kung, "GPSR", *Proceedings of the 6th annual international conference on Mobile computing and networking - MobiCom '00*, 2000. Available: 10.1145/345910.345953
- [30] S. Adam and R. Hassan, "Delay aware Reactive Routing Protocols for QoS in MANETs: a eview", *Journal of Applied Research and Technology*, vol. 11, no. 6, pp. 844-850, 2013. Available: 10.1016/s1665-6423(13)71590-6 [Accessed 16 February 2020].
- [31] H. Ames, C. Glenton and S. Lewin, "Purposive sampling in a qualitative evidence synthesis: a worked example from a synthesis on parental perceptions of vaccination communication", *BMC Medical Research Methodology*, vol. 19, no. 1, 2019. Available: 10.1186/s12874-019-0665-4 [Accessed 17 April 2020].
- [32] L. Machi and B. McEvoy, *The literature review*, 2nd ed. Corwin, 2012.

Intrusion Detection in Cyber Security: Role of Machine Learning and Data Mining in Cyber Security

Gillala Rekha¹, Shaveta Malik², Amit Kumar Tyagi^{3,*}, Meghna Manoj Nair³

¹Koneru Lakshmaiah Education Foundation, Department of Computer Science and Engineering, Hyderabad, India – 522502

²Terna Engineering College, Department of CSE, Navi Mumbai, Maharashtra, India.

³Vellore Institute of Technology, School of Computer Science and Engineering, Chennai Campus, Chennai, 600127, Tamilnadu, India.

ARTICLE INFO

Article history:

Received: 10 August, 2019

Accepted: 04 March, 2020

Online: 03 May, 2020

Keywords:

Cyber Security

Intrusion Detection System

Machine Learning

Data Mining

ABSTRACT

In recent years, cyber security has been received interest from several research communities with respect to Intrusion Detection System (IDS). Cyber security is “a fast-growing field demanding a great deal of attention because of remarkable progresses in social networks, cloud and web technologies, online banking, mobile environment, smart grid, etc.” An IDS is a software that monitors a single or a network of computers from malicious activities (attacks). Detecting an intrusion or prevention (due to increase the usage of internet), is becoming a critical issue. In past, several techniques have been proposed to overcome or detect intrusion in a network. But most of the techniques (used now days in detecting IDS) are not able to overcome this problem (in efficient manner). Together this, Machine Learning (ML) also has been adopted in various applications (due to providing good accuracy results (in respective domain)). Hence, this work discusses “How machine learning and data mining can be used to detect IDS in a network” in near future. ML use efficient methods like classification, regression, etc., with efficient results like high detection rates, lower false alarm rates and less communication costs. This work also provides a detail comparison with metrics in table 1-3 (with their performance/ algorithms/ dataset or metrics used).

1. Introduction

Cyber security involves the practice of preventing the exposure of computers, programs, etc. from attacks, unauthorized usage, modifications, destructions, etc. It's a common practice to find every Cyber Security system to have a firewall, antivirus techniques and Intrusion Detection System (IDS). IDS are a crucial component as they help in spotting any undesirable and unwanted changes in the system [1]. Intruders are mainly categorized as External Intrusions/Intruders (i.e., attack by the people who don't belong to the organization) and Internal Intrusions/Intruders (i.e., attack by the people from within the same establishment). However, cyber analytics can be separated on the following bases: i) on the basis of misuse or signatures ii) on the basis of anomalous encryptions iii) on the basis of hybrid nature.

The first form of classification is created to represent attacks following an ordered pattern to spot and prevent a similar attack in the further years along with the detection of famous attacks

(though they become hard to use in the case of naïve outbreaks). It is to be pointed out that this method can't be used for the identification of novel (or zero day) catastrophes. The second classification (i.e., based on anomaly) replicates the behavioristic approach by developing an activity profile, hence differentiating the ambiguity from the normal attitude. This method can be used for the detection of novel-attacks and hence are deeply encouraged. Furthermore, it customizes the normal activity routine for every instance, ensuring that the intruders are unable to comprehend which of the activities can be performed incognito. But just like how every coin has two sides, this technique too has its own disadvantage – it is likely for False Alarm Rates (FARs). The last categorization involves the combination of the first two methods – misuse and anomaly detection. They are mainly implemented to raise the rate of detection of common attacks and reduce the False Positive (FP) rate for the minor attacks. IDS's can also be divided based on network or host. An IDS which depends on the network identifies attacks by keeping an eye on the traffic through the network devices. A host-based IDS screens all processes and file activities related to the software with a host.

*Corresponding Author: Amit Kumar Tyagi, amitkryagi025@gmail.com

- a) Host-based IDS (HIDS): It mainly focuses on analyzing the internal functioning of a computing system. It might detect activities like which program is trying to access which particular resource and are there any attempts on illegitimate access. For example, a word processor which spontaneously alters the system password database.
- b) Network-based IDS (NIDS): It focuses on analyzing and filtering the traffic among network device. It's commonly found that intrusions occur as ambiguous patterns. These are mainly caused by the attacks launched by external intruders who wish to access the network to gamble the network and destroy it.

Hence, the article is organized into a number of sections. Section 2 discusses several classifications (like signature and anomaly) with respect to cyber security. Further, section 3 discusses several cyber data sets available for making a comparison and later on the significance of machine and data mining in detection of intrusion detection in cyber security/ applications (in near future) has been discussed in section 4. Further Section 5 discusses "how machine learning and artificial intelligence can be more useful for cyber security professionals for detecting vulnerabilities or preventing attacks". Finally, this work is concluded with some future enhancements in brief in section 6.

2. Cyber Security's Classifications

The three types of Intrusion detection in support of cyber security are [2]-[4]: *Misuse-based or Signature based, Anomaly-based, and Hybrid*. Here, each one can be discussed in detail as:

2.1. Misuse-Based or Signature Based

There are multiple ways to replicate an attack. The attack can be a pattern, or a signature used to identify the deviation. They are bound to detect a majority or most of the common attack techniques. However, they come to be of little use in the case of minor or unidentified attack patterns. These systems try to spot and differentiate on the principle of "bad" behavior. The prime obstacle to overcome is on how to create a signature that combines all the varieties of a consistent attack. A plethora of Machine Learning methodologies have been put into use for the detection of misuse in these systems. These detections prove to be useful to identify the outbreaks on networks by associating the routine activities with that of the expected actions of an intruder.

In [5], the author proposed a framework to identify and classify network activities based on Artificial Neural Network (ANN). The data sources are based on various formats, i.e., limited, incomplete, and nonlinear in nature. They implemented data detection that utilizes the analytical strengths of neural networks. A multi-layer classification prototype using MLP is used to detect the misuse by developing the architecture containing four fully connected layers. The neural architecture consists of 9 nodes as input and 2 output nodes. The data pre-processing were conducted at three different levels includes a) Protocol Identifier (PID) – the rules and regulations pertaining to an event (TCP = 0, UDP = 1, ICMP = 2, and Unknown = 3) b) Source Port c) Destination d) Source Address (IP address corresponding to a source) e)

Destination Address (IP address of a destination) f) ICMP type (like echo requests, null, etc.) g) ICMP Code h) Raw Data Length (length of the data packets) i) Raw Data.

The neural network model was trained using a back-propagation algorithm for 10,000 iterations of the selected training data. Out of 9,462 records, 1000 were randomly selected for testing and the remaining was used to train the system. The neural network model required 26.13 hours to complete. The results reveal that on training data the root mean square error is 0.058298 and on Test data root mean square error is 0.069929. Finally, an accuracy of 93% can be considered based on RMS, where each data packet was classified as either a normal or an attack set.

In [6], the authors proposed Online Analytical Processing (OLAP) Mining and Classification based IDS, (OMC-IDS). OMC_IDS handle any intrusion detection data using historical data analysis from heterogeneous sources and summarize them by filtering the data by removing the irrelevant data. Apart, a data cube is constructed and integrate OLAP techniques. They applied association rule mining to extract the interesting patterns and classify each connection as normal or any attack. They proposed association rules to find the correlation between TCP/IP parameters and the types of attack on DARPA 1998 data set. They generated rules and less constraint is retained. After the rules are generated, a C4.5 classifier is applied for new connection records. The experiments were carried out on DARPA19985 dataset. The training data and test data are generated in the first seven weeks and in the next two weeks respectively. The results show that total of detection rates as 99%, 97%, 86% and 74%, respectively. The main drawback of association rule mining is that the generated rules may express correlation, but the approach is promising for attack signature building.

Further in [7], authors proposed an algorithm to use the existing signature data and find the signature of the related attack in less time. They compared their approach with algorithm based on Apriori called Signature Apriori (SA) and found that it takes less processing time. Such algorithms can be used to generate new signatures, i.e., used into misuse detection systems such as Snort. The proposed method finds newly attack signature based on the known signature. Scan Reduction method is also used for the reduction of time consumed for scanning of databases. This method involves the determination of a new attacking signature in an efficient way when compared to the Signature Apriori algorithm. Authors have implemented the data mining approach to complement the signature discovery in IDS based on network [8]. This not only generates signatures for the detection of misuses dependent on transfer protocols, but also for those based on content of traffic. The Signature Apriori (SA) is based on the typical association rules algorithm – Apriori algorithm [8]. The experiments have two parts to it: a) Speed testing of SA algorithm b) Accuracy testing of the signatures being mined. This evolves 70% support and the time consumed is extremely less (one is less than 50111s the other is 330 ms). On the whole, the techniques which are applied to tackle the cyber-attacks have been active

predominantly as they emphasize on screening the traffic in the network, identification of anomalies and traffic sequences of cyber-attack. Apart from this, the misuse detection can be enforced for the detection of these outbreaks prior to them actually being a part of the attack. Some authors have spotted the command and control traffic (C2C) in Internet Relay Chat using the technique of machine learning to adhere to the botnet existence, for which TCP level data sets have been put into use. Wireless traffic sniffers were used extensively to gather complete TCP/IP headers from around 18 locations around the campus. This was divided into two major stages: (i) The initial stage involved the distinction between IRC and non-IRC traffic, (ii) after which, there was a distinction between botnet and real IRC traffic. For the initial stage, the comparison of performance is done between J48, naive Bayes, and Bayesian network classifiers to identify IRC and non-IRC traffic damages by attaining an excellent overall classification accuracy. Only the naïve Bayes classifiers were capable of achieving reduced false negative rate.

The naive Bayes classifiers accurately classified 35 out of the 38 botnet IRC (which flows correctly and achieves False Negative Rate (FNR) of 7.89%) [9]. In Stage (ii), by applying classification they accurately labelling IRC traffic as botnet and non-botnet were more challenging. In [10], author proposed an adaptive intrusion detection system which is considered as a framework for detecting intrusion detection using Naïve Bayesian network. The DARPA KDD99 dataset with 38 attacks are used to find the new intrusion signature like DoS, r21,u2r and probe. The dataset consists of 9 features in the inference network such as Protocol type, Service, Land, Wrong fragment, Numerous failed login, Logged in, Root shell, Is guest login. In the first stage, a junction tree inference technique is used to identify the normal or attack data with performance detection rate 87.68% on normal and 88.64% on intrusion. In the second stage, the dataset is classified into 4 classes: DoS, Probing, R2L and U2R. The performance determine a detection rate of 88.64% for DoS, 99.15% for Probing, 20.88% for R2L, 6.66% for U2R and 66.51% for other classes.

In [11], authors used reliable signatures generated based on supervised clustering algorithm and updating them in real-time using unsupervised clustering technique. The signature updating is done to change attack methods while retaining the signatures useful information. They used a simple density-based clustering algorithm, called Simple Logfile Clustering Tool (SLCT) to create clusters of regular and anomaly traffic. The study made use of a new user stricture, M , in SLCT which mentions the percentage of fixed attributes to be spotted out of all the attributes that a potential cluster is expected to have. If the value of M equates to 0, it then allows the formation of clusters irrespective of the number of fixed attributes. By equating the value of M to greater values they recapitulate the intruder ones, thus classifying the original data. This is inferred to with the help of parameter M as SLCT attack. Both the clustering techniques are implemented for the detection of normal or attack traffic and for identification of the usual traffic in a supervised manner accordingly. In [11], the author treated anomalous centroid of cluster as a signature. The

experiments are carried out using KDD data sets using different attack percentages (0%, 1%, 5%, 10%, 25%, 50%, and 80%) and the author reported impressive results without prior knowledge of any attacks in the KDD datasets. Further, Kruegel et al. [12] installed an intrusion detection signature using clustering algorithms to derive decision tree for intrusion detection. It was a placement with Snort. With the help of a decision tree, we are able to choose the features which highly distinguish the characteristics of the rule set, permitting parallel evaluation for every unique feature. It provides a better performance with respect to Snort. In [12], the author make use of the tcpdump files as the necessary dataset for the ten days of test data when considering the evaluation of 1999 DARPA intrusion detection. On comparing and contrasting the rate of processing of Snort and the decision tree for the above data, it was observed that real performance gain vary drastically depending on the basis of the comprehended traffic. 103% was found to be the maximum speed, while 5% turned out to be the minimum. The decision trees performed better as they result in an average speed of 40.3%. The second task was also conducted with increased number of protocols right from 150 up to 1581. The results proved that the approach of the decision tree works efficiently, especially with respect to large rule sets. This approach notifies that the clustering action based on decision trees will definitely reduce the operating time, thus enhancing the processing speed. Furthermore, it portrays a generic solution to many of the other IDSs like host and network-based, and firewall and packet filters.

Zhang et al. [13] study proposed a complete intrusion detection framework containing a detector used for signature-based attack prediction and a database to identify outlier. All the anomaly patterns identified by the system or user either manually or automatically are stored in the database. Because of the extremely quick nature of its implementation, it's often used as an online solution. Gharibian et al. [14] has put forth a comparative study with the help of probabilistic and futuristic ML methods and processes for detection of intruders and their malicious acts namely, Naïve Bayes and Gaussian along with those of Decision Tree and Random Forests. A lot of the training data sets which have been constructed from KDD99 are being deployed for effective functioning today and each of the methods have been used for categories of attack like DoS, Probe, R2L and U2R with a proper analytical study of their results. Normalization used in the formation of these datasets, complementing the argument that the features in KDD are not similar to those of the others and they possess high variance scales. The executional capability of Decision Trees (DT) and Random Forests (RF) portray valid results and operations in the identification of DoS. On the contrary, Gaussian and Naïve Bayes results shows much better in few of the varied attack domains like Probe, R2L and U2R. Based on the results, the author stated that the probabilistic techniques are more robustness in nature than predictive techniques for intrusion detection.

Mukkamala et al. [15] considered the performance of ANN, SVM and Multivariate Adaptive Regression Splines (MARS) and

proved that ensembles of ANNs, SVMs and MARS is of top priority for individualized perspectives for the detection of these attacks with respect to precision of division. The five class classification experiments were performed on 11,982 records. They applied 3 classification algorithms like SVMs, MARS and ANNs. The ensemble of SVMs, MARS and ANNs approach outperforms with accuracies of 99.71% for Normal, 99.85% for Probe or Scan, 99.97% for DoS, 76% for U2R, and 100% for R2L are reported respectively. The accuracy of four classes are 99% using SVM, RP, SCG, OSS algorithms and the accuracy on the U2R class is much less with 76%. In this paper [16] the author used genetic algorithms to generate simple rules for network traffic.

These rules are used to differentiate normal network connections from anomalous connections and these anomalous connections refer to events with probability of intrusions. Abraham et al. [17] applied genetic programming algorithms such as Linear Genetic Programming (LGP), Multi Expression Programming (MEP) and Gene Expression Programming (GEP) in attack classification. In [18], Hansen et al. used GP with homologous crossover for performing intrusion detection. Arnes et al. [19] proposed a novel approach to network risk assessment. The approach considers the risk level of a network as the composition of the risks of individual hosts. It is probabilistic and uses Hidden Markov models (HMMs) to represent the likelihood of transitions between security states. They tightly integrate the risk assessment tool with an existing framework for distributed, large scale intrusion detection, and apply the results of the risk analysis to prioritize the alerts generated by the intrusion detection sensors.

An HMM is denoted by (P, Q, Π) . Lee et al. [20] developed a systematic framework using data mining techniques for automated IDS. In [21] the author trained Naïve Bayes classifier on KDD 1999. The data is partitioned into training set and test set and the data was grouped into four attacks (1. probe or scan, 2. DoS, 3. U2R, and 4. R2L). The author stated an accuracy of 96%, 99%, 90% and 90% for the respective attacks. Hu et al. [22] proposed a framework for malicious transactions. A cyber-attack detection model is needed as a prerequisite for fast damage recovery. The framework employed a sequential mining algorithm for finding the dependencies in database and presented as classification rules. The data captured from database logs including (Tname) transaction name, (TID) transaction ID, begin and end time, etc. They applied the framework for identifying U2R attacks as part of cyber security. The result presented 91% of TP (True Positive) rate and 29% of FP (False Positive) rate.

In [23], the author presented an IDS model with high accuracy and efficiency using machine learning algorithms including K-means, Support Vector Machine (SVM). They also employed feature reduction methods to eliminate the unwanted features. Table 1 shows the algorithm, data set, metric used for misuse-based intrusion detection.

2.2. Anomaly and Hybrid Detection

Lippmann et al. [24] proposed an IDS system on transcripts of telnet sessions. The combination of training data and new keywords were used to find the common attacks using neural network model. The system achieves 80% of high detection rate. Palagiri et al. [25] proposed a model for learning the normal traffic patterns from TCP/IP port. They applied preprocessing techniques then perform clustering on normal traffic and final trained using Artificial Neural Network (ANN). The study reported a 100% normal behavior.

Apiletti et al. [26] proposed NETMINE framework which classifies the traffic data using data mining/ machine learning techniques. The framework performs data stream processing, refinement analysis by using general association rule extraction for profile data, anomaly detection, and identifying recurrent patterns.

Intrusion Detection Systems (IDS) mainly intend towards protection of computerized systems and helps in spotting vulnerabilities and other attack exposures. A novice structural outline which has its' roots based on data mining methods have been put forth [27] for the creation of an IDS. This framework proposes Association Based Classification (BC) which is dependent on rules linked to fuzzy logic for the development of classifiers and this helps in categorization of normal and abnormal records. Compatibility threshold is the central parameter in this application. The approximate value for this depends on the ROC curve of the system which is produced by carrying out lots of tests on datasets, with varied threshold values. Therefore, 0.06 becomes the compatibility threshold which is to be dealt with in the detection of anomalous behavior. The FP error produced can be reduced to the level of that of misuse detection situation and there's a huge decrease in the detection rate of existing attacks. In the case of unforeseen intrusions, the ambiguous case outshines the misuse perspective, and this is the key advantage of anomaly-based approaches.

Luo et al. [28] has combined the association logic along with the frequency episodes with that of fuzzy logic to determine the sequence in the data. This produces short and flexible variations for intrusion detection as a lot of quantifying features come into play. To ensure that data instances don't outshine the contribution of that of the others, normalization is carried out before retrieving the fuzzy association rules. The required simulations have been conducted by customized programs and the results have proved the necessity of fuzzy rules and its' frequency occurrences in intrusion detection. Kruege et al. [29] implemented an intrusion detection system for identifying attacks against Operating System (OS), they analyzed OS calls to detect attacks against daemon applications and set uid programs. Also implemented on machines running with Linux or Solaris with individual system calls. A feature vector is represented which captures information specific to each system call such as the system call number, its return code, and its arguments. They applied Bayesian network to classify events during open and executive OS calls.

Table 1: The algorithm, data set, metric used for misuse-based intrusion detection.

Paper Citation	Algorithm Used	Data Set Used	Metric Used
[5]	Artificial Neural Network	RealSecure network monitor (Internet Security Systems)	Accuracy
[6]	OMC-IDS (OLAP and Association rule mining)	DARPA 1998	Accuracy
[7]	Signature Apriori (SA)	Signature based data	Accuracy
[8]	Apriori algorithm	SigSniffer architecture	Accuracy
[9]	J48, Naïve Bayes and Bayesian network	Dartmouth's wireless campus network (TCP level)	Accuracy
[10]	Bayesian network	DARPA KDD	Accuracy
[11]	Density-based clustering algorithm (SLCT)	KDD	Accuracy
[12]	Decision Tree	DARPA	Accuracy
[13]	Random Forest	KDD	Accuracy
[14]	Random Forest (Predictive techniques)	KDD	Accuracy
[15]	ANN, SVM and MARS	DARPA	Accuracy
[16]	Genetic algorithms	DARPA	Accuracy
[17]	Genetic algorithms	DARPA	Accuracy
[18]	Genetic algorithms	KDD	Accuracy
[19]	Hidden Markov Network	KDD	Accuracy
[20]	RIPPER	DARPA	Accuracy
[21]	Naïve Bayes	KDD	Accuracy
[22]	Apriori algorithm	Sequence patterns of log files from database are examined to find database intrusions.	Performance
[23]	Ant Colony Optimization (ACO)	KDD	Accuracy

Table 2: The algorithm, data set, metric used for anomaly and hybrid-based intrusion detection.

Paper Citation	Algorithm Used	Data Set Used	Metric Used
[24]	Artificial Neural Network (ANN)	Transcripts of telnet sessions	Accuracy and False alarm
[25]	Artificial Neural Network	DARPA	-----
[26]	NETMINE framework	Network capture tools are used to capture the network traffic packets and it was developed at Politecnico di Torino	Support

[27]	Fuzzy Association Based Classification (ABC)	KDD	Accuracy and FP rate
[28]	Fuzzy Logic	Tcpdump	Accuracy
[29]	Bayesian network	DARPA	Accuracy and False Alarm Rate (FAR)
[30]	Naïve Bayes algorithm	DARPA	-----
[31]	sequence matching algorithms	User command level (shell commands)	Accuracy and False Alarm Rate (FAR)
[32]	EXPOSURE (C4.5 Decision Tree algorithm)	DSN	Accuracy and False Alarm Rate (FAR)
[33]	EXPOSURE (C4.5 Decision Tree algorithm)	Real-World Network	Accuracy and False Alarm Rate (FAR)
[34]	Genetic algorithms	KDD	Accuracy and False Alarm Rate (FAR)
[35]	Genetic Programming	DARPA	ROC (Receiver's Operating Curve) and False Alarm Rate (FAR)
[36]	Hidden Markov Network	KDD	(False Positive) FP rate and (False Negative) FN rate
[37]	RIPPER	DARPA	(False Alarm Rate) FAR
[38]	Bayesian network	KDD	Accuracy and False Alarm Rate (FAR)
[39]	Apriori algorithm	DARPA	Support
[40]	Robust Support Vector Machines	DARPA	Accuracy and False Alarm Rate (FAR)
[41]	Support Vector Machine	NetFlow data (Flame tool)	Accuracy and False Positive (FP) rate
[42]	Self-Organizing Feature Map (SOFM), Genetic Algorithms (GA), and Support Vector Machine (SVM)	DARPA 1999	Accuracy, (False Positive) FP rate and (False Negative) FN rate

The DARPA 1999 data set is used to excite the OS kernel by TCP/IP packets. These features are fed to Bayesian network model and if the output is close to zero it indicates normal or anomaly state.

In [30], the author proposed alert correlation method based on naïve bayes algorithm. 2000 DARPA dataset with their intrusion objective are used to train Bayesian network. In [31], the author proposed a model for differentiating masquerader’s users from

real users. The study stated a detection rate as high as 80.3% and a false positive rate as low as 15.3%. Table 2 shows the algorithm, data set, metric used for anomaly and hybrid-based intrusion detection. Now, next section will discuss availability of cyber security dataset (in current) globally.

Bilge et al. [32] introduced EXPOSURE, a system that employs large-scale, passive DNS analysis techniques to detect domains that are involved in malicious activity. Bilge et al. [33] presented

DISCLOSURE, a large-scale, wide-area botnet detection system that incorporates a combination of novel techniques to overcome the challenges imposed by the use of NetFlow data. In [34] the author broadly demonstrates how information of the network connection can be replicated as genes and how the parameters in GA can be define in this respect. Lu et al. [35] presented a rule evolution approach based on Genetic Programming (GP) for detecting novel attacks on networks. Joshi et al. [36] classify the TCP network traffic as an attack or normal using HMM and to build an anomaly detection system. Fan et al. [37] proposed an algorithm to generate artificial anomalies to coerce the inductive learner into discovering an accurate boundary between known classes of normal connections and known intrusions, and anomalies. Amor et al. [38] uses a simple form of a Bayesian network that can be considered a Nave Bayes classifier in intrusion detection. Li et al. [39] applied AprioriAll, an algorithm for mining frequent sequential pattern in Data mining field, to discovery multistage attack behavior patterns. Hu et al. [40] presented a new approach, based on Robust Support Vector Machines (RSVMs) for anomaly detection. Wanger et al. [41] proposed an approach for evaluating Netflow records by referring to a method of temporal aggregation applied to Machine Learning techniques. In paper [42], they proposed a new SVM approach, named Enhanced SVM, which combines soft-margin SVM and one class SVM methods

3. Cyber-Security Datasets

Data plays an important role for ML and DM models. Today data is new oil for digital world (or for industries), i.e., based on collecting data, competitors can launch affordable services in market. For example, based on collecting requirements/ demands of particular things in an area, companies can shift towards to sell their product in that specified area/ region. The necessary elements for the efficient conduction of research related to cyber security includes the right choice of data and its' proper utilization. To comprehend the ML and DM algorithms, put forth by a number of authors, requires a better understanding of data sets. We can achieve cyber security of data with the help of different gatherings like Win Dump or Wireshark tool to acquire the network data packets. It can also be done using the current public datasets.

a. **DARPA:** DARPA (Defense Advanced Research Projects Agency) intrusion detection datasets was collected and published by the Cyber Systems and Technology Group MIT/LL (Massachusetts Institute of Technology Lincoln Laboratory). The data was generated using network simulation and compiled based on TCP/IP network data. The datasets can be downloaded from the website and it primarily includes: DARPA 1998, 1999, 2000. DARPA 1998 consists of data collected for 9 weeks, which includes training data (seven weeks) and of test data (two weeks). Similarly, DARPA 1999 consists of data collection for five weeks wherein training data is for three weeks and the last two weeks is test data. DARPA 2000 includes scenario-specific datasets. Table 3 lists the complete basic features of TCP connection.

b. **KDD 1999 cup datasets:** The most popular and widely used datasets for intrusion detection are KDD 1999 datasets created by KDD cup challenge. This dataset is based on DARPA 1998 dataset with 4 million records. The KDD 1999 datasets consist of normal and 22 attacks categorized into 5 main components. Dos (Denial of Service attacks), R2L (Root to Local attacks), Probe (Probing attacks), U2R (User to Root attack) and normal. There exist 41 number of attributes containing features related to basic, content and traffic.

Table 3: List of the Complete Basic Features of TCPconnection

Basic Features	Type	Represented	Description
Duration	Continuous	Integer	Time duration of connection
Protocol, type	Symbolic	Nominal	Type of the protocol (TCP, UDP and ICMP)
Service	Symbolic	Nominal	HTTP, Telnet, FTP, SMTP and others
Flag	Symbolic	Nominal	Connection status
Src bytes	Continuous	Integer	Number of bytes sent per connection
Dst bytes	Continuous	Integer	Number of bytes received per connection
Land	Symbolic	Binary	Value=1 if port numbers and src/ dst IP address are same
Wrong fragment	Continuous	Integer	Total of bad checksum packets
Urgent	Continuous	Integer	Sum of urgent packets

Hence, this section discusses current cyber security datasets in detail. Now next section will discuss a brief introduction of data mining and machine learning and necessary uses in detecting vulnerabilities or intrusion over cyber – network (cyber space).

4. Introduction to Data Mining (DL) and Machine Learning (ML) for Cyber Security

The terms Machine Learning (ML), Data Mining (DM), and Knowledge Discovery in Databases (KDD) are often used interchangeably. As per research, KDD process is represented as whole and deals with extracting valuable, earlier unknown knowledge/information from data. Fayyad et al. [43], has clearly mentioned and explained the process of DM as a specific step in KDD which handles the implementation of algorithms for retrieval of sequences from data. It can hence, be observed that

they possess common characteristics between ML and DM. The steps involved in KDD process are as follows: data selection, data cleaning and pre-processing, data transformation, application of DM algorithms, result interpretation/ evaluation. DM is one step among all and used for extracting patterns from data by applying algorithms. It's to be pointed out that there is a plethora of publications [e.g., Cross Industry Standard Process for Data Mining (CRISP-DM) [44] along with industry participants who consider the process DM.

These two terms are commonly discussed together and are applied interchangeably. According to Arthur Samuel Creator of Machine Learning (ML) defined "ML as a field of study that makes the computers to learn by itself without being explicitly programmed". The machine learning algorithms mainly focus on classification and prediction techniques. The ML algorithms learn from the training/ past data and finds the insights for future/unknown conditions. The various classification algorithms in general applied to cyber security are discussed as below.

- Decision Trees

Decision trees are the important and popular techniques used for classification. A decision tree is nothing but a simple flowchart similar to that of the structure of a tree which has every internal node denoting a test with respect to an attribute such that each branch indicates the outcome of the test and each leaf node acquired a class label. ID3 (Iterative Dichotomiser) is a decision tree algorithm which was developed by Ross Quinlan. He then represented the successor of ID3 – C4.5 which has turned out to be a benchmark for comprehending algorithms

- C4.5 Algorithm

This model forms its basis from ID3 algorithm along with additional characteristics to acknowledge the issues faced by that of ID3. It's considered to have a greedy approach and it is said to possess a top-down recursive divide and conquer method. Given a data samples S, C4.5 applies divide and conquer algorithm for tree generation and the process is stated as follows:

- a) If S is small or all the data samples in S belong to the same class, then the leaf node is labeled with the most frequent class in S.
- b) Or else, the process of selecting attributed is made use of to control the criterion of the splitting process. The criterion for the process of splitting indicates which attribute is to be tested at node S by identifying the most efficient way to distinguish the tuples into separate classes.

The process continues recursively to form a decision tree.

- Naive Bayes Algorithm:

The Naive Bayes algorithm (NB) employs a simplified version of Bayesian learning method. It involves statistical classifiers. The probabilities of membership can be determined with the help of these classifiers and it has its foundation on Bayes theorem. It's

assumed that the effect of a feature value of a given class doesn't depend on the values from other features and is called conditional independence. One of the most efficient, robust and best methods to prevent noisy data is by making use of Naive – Bayes classifiers. The highlight feature being that it calls for only a small amount of training data to approximate the strictures needed for categorization.

- K-Nearest-Neighbor

K-Nearest-Neighbor (k-NN) is a classification which is one of the simplest and fundamental ones, working well even in the presence of little or absolutely no prior knowledge regarding the data distribution and it's based on the process of learning by equivalence. 'm' dimensional numerical attributes are used for describing the training samples with each sample replicating a certain point in the m-dimensional space. Hence, we can see that all the points are stored in an m-dimensional pattern space. In the case of an unknown data sample, a k-nearest neighbor classifier checks out the pattern space for the k training data modules which are quite close to that of an unknown sample. 'Closeness' refers to Euclidean distance. The new and unknown sample is designated with the most common class from it is nearest k neighbors.

- Support Vector Machine

It mainly plots the input vector into a space of very high dimensions and helps in the construction of a hyper plane. The hyper plane has the capacity to separate the data points into different classes. A great level of distinction is obtained by hyper planes which has the greatest distance to the closest training data point of any class which is called as the functional margin. It's observed that with increase in margin, there's a lower generalization error for the classifier. The hyperplane is a decision boundary for the two classes. In reality, the persistence of a decision boundary ensures the detection of a misclassification which is created by a particular method. Classification, regression, and other jobs are implemented with SVM.

- Repeated Incremental Pruning to Produce Error Reduction (RIPPER)

RIPPER, is a generic methodology used for effectively applying separate-and -conquer rule learning. It helps in increasing the precision of protocols by replacing or re-enforcing the individual norms. Reduce Error Pruning was implemented to create the rule and the created rules are often restricted to a smaller number. It ensures the pruning of each rule right after the creation and removal of data samples. Reduced error pruning facilitates the handling of huge training sets, thus improving the precision. The below mentioned steps are carried out: Spot the characters/features from the training data and identifies the split of all attributes essential for categorization (i.e., feature/dimensionality reduction). Comprehend models using the training data and use the trained model to segregate the unknown data. In the initial

stage of training, each feature with a corresponding class is acquired by using suitable algorithms from the training set. The perspectives of ML/DM are mainly categorized into three classes supervised, unsupervised and semi-supervised. The different machine learning and data mining methods applied for cyber security is mentioned in Table 1-2.

5. Role of Machine Learning and Artificial Intelligence towards Cyber Security

Today cyber security has put everything on risk, due to attracting billions of online users over internet and storage of data over internet (at cloud side). Everyday every country is facing critical attacks by enemy nations on their computer labs, systems or network, which can create a situation of third world war. Till today, we are detecting cyber attackers or hackers through human work-force, for that we require a huge number of skilled workforce to look over or prevent against any cyber threats. But in near future, there is a possibility that intrusion or vulnerabilities detection can be done by using machine learning and artificial intelligence. Also, it will provide several benefits to society and avoid the problem of weaker security, lower efficiency, leaking of personal information by Internet of Things, increasing vulnerabilities on cyber and physical space or cyber physical systems. Note that recently many critical attacks have been measured by several countries on their nuclear programs/ sites [45]-[48] by their enemy nations. On other side, Artificial Intelligence (AI) will reduce required workforce (requirement of cyber security professionals), speed of detection of intrusion, etc. AI can help in living life longer and better through its emerging innovations. Such benefits of AI are listed in following ways.

- Handling huge volumes of security data
- Picking out threat needles in cyber haystacks
- Acceleration of detection and response times
- Keeping up in the Artificial Intelligence arms race
- Breathing space for human cyber security teams.

Hence, data mining, machine learning and artificial intelligence are necessary components for 21st century generation. So, we will see the tremendous uses of Machine learning, Artificial intelligence in next 20-30 years, which will do many/ everyday task and will serve humanity better and better.

6. Conclusion and Future Enhancements

In the recent/ several decades, several attacks have been measured/ noticed. Due to this reason, cyber security and intrusion detection has been coined in this smart era. Due to enormous internet usage (in the past decade), the vulnerabilities of network security (in a network) need to be overcome. Overcoming such issue has become an important issue today. In general terms, Intrusion detection system is used to identify the flaws in the system such as unauthorized access and unusual attacks over the secured networks. Hence, to solve this issue, several authors had discussed many studies. In that, we found that (from literature, refer section 2 and 3) machine learning can be

more useful in solving these issues/such problems using regression, prediction, and classification techniques. In this smart era, we have large amount of data (generated from internet/ web browsing) and shortage of talented employees in cyber-security domain/area. So, Machine Learning is the only solution to provide efficient results in minimum time. Hence, in order to understand importance of ML techniques for solving the IDS problems, which focus on the design of the single, hybrid and ensemble classifier models (with discussing several algorithms, used datasets). This work also discussed “How Machine learning, and data mining can be useful in identifying/ detecting intrusion, in section 4”?

Hence, we found that uses of different classifier/ ML techniques in IDS a promising study in cyber security and artificial intelligence. It will make attraction of young scientists from research communities for a long time. For future work, this work has identified some valid points which are: removal of data redundancy and irrelevant features for the training phase (have important role in system performance), i.e., consideration of best feature selection algorithm will play an important role in the classification techniques in near future. Also, multiple or different selection of algorithms for featured selection will provide best possible solutions in various scenarios/ intrusion detection in a network. Last, but not the least, cyber security and intrusion detection systems works well and shows a better performance with ensemble classification algorithms when compared to single classification algorithms.

Authors' Contributions

Gillala Rekha drafted this manuscript, whereas Shaveta Malik and Meghna Manoj Nair have put this article's content in correct order. In last, Amit Kumar Tyagi has approved this manuscript.

Acknowledgement

This research is funded by the Anumit Academy's Research and Innovation Network (AARIN), India. The author would like to thank AARIN India, an education foundation body and a research network for supporting the project through its financial assistance.

Conflict of interest

The authors declare that they do not have any conflict of interest with respect to publication of this research work.

Scope of the Work

This work has been written through collecting articles from several international journals like ACM, IEEE, Springer, Wiley, etc. This work will be useful for future researchers who are working towards computer vision/ the use of machine learning or artificial intelligence towards cyber security.

References

- [1] S. Mulkamala, A. Sung, A. Abraham, Cyber security challenges: Designing efficient intrusion detection systems and antivirus tools, Vemuri, V. Rao, Enhancing Computer Security with Smart Technology. (Auerbach, 2006) 125-163.

- [2] A. Sundaram, An introduction to intrusion detection *Crossroads* 2 (4) (1996) 3–7.
- [3] V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: A survey, *ACM computing surveys (CSUR)* 41 (3) (2009) 15.
- [4] B.-C. Park, Y. J. Won, M.-S. Kim, J. W. Hong, Towards automated application signature generation for traffic identification, in: *Network Operations and Management Symposium, 2008. NOMS 2008. IEEE, IEEE, 2008*, pp. 160–167.
- [5] J. Cannady, Artificial neural networks for misuse detection, in: *National information systems security conference*, Vol. 26, Baltimore, 1998.
- [6] H. Brahmi, I. Brahmi, S. B. Yahia, Omc-ids: at the cross-roads of OLAP mining and intrusion detection, in: *Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer, 2012*, pp. 13–24.
- [7] H. Zhengbing, L. Zhitang, W. Junqi, A Novel Network Intrusion Detection System (NIDS) based on signatures search of data mining, in: *Proceedings of the 1st international Conference on Forensic Applications and Techniques in Telecommunications, information, and Multimedia and Workshop, ICST, 2008*, p. 45.
- [8] H. Han, X.-L. Lu, L.-Y. Ren, Using data mining to discover signatures in network-based intrusion detection, in: *Machine Learning and Cybernetics, 2002. Proceedings. 2002 International Conference on*, Vol. 1, IEEE, 2002, pp. 13–17.
- [9] L. Carl, et al., Using machine learning techniques to identify botnet traffic, in: *Local Computer Networks, Proceedings 2006 31st IEEE Conference on*. IEEE, 2006.
- [10] F. Jemili, M. Zaghdoud, M. B. Ahmed, A framework for an adaptive intrusion detection system using bayesian network, in: *Intelligence and Security Informatics, 2007 IEEE, IEEE, 2007*, pp. 66–70.
- [11] G. R. Hendry, S. J. Yang, Intrusion signature creation via clustering anomalies, in: *Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security 2008*, Vol. 6973, International Society for Optics and Photonics, 2008, p. 69730C.
- [12] C. Kruegel, T. Toth, Using decision trees to improve signature-based intrusion detection, in: *International Workshop on Recent Advances in Intrusion Detection, Springer, 2003*, pp. 173–191.
- [13] J. Zhang, M. Zulkernine, A. Haque, Random-forests-based network intrusion detection systems, *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 38 (5) (2008) 649–659.
- [14] F. Gharibian, A. A. Ghorbani, Comparative study of supervised machine learning techniques for intrusion detection, in: *Communication Networks and Services Research, 2007. CNSR '07. Fifth Annual Conference on, IEEE, 2007*, pp. 350–358.
- [15] S. Mukkamala, A. H. Sung, A. Abraham, Intrusion detection using an ensemble of intelligent paradigms, *Journal of network and computer applications* 28 (2) (2005) 167–182.
- [16] W. Li, Using genetic algorithm for network intrusion detection, *Proceedings of the United States Department of Energy Cyber Security Group 1* (2004) 1–8.
- [17] A. Abraham, C. Grosan, C. Martin-Vide, Evolutionary design of intrusion detection programs., *IJ Network Security* 4 (3) (2007) 328–339.
- [18] J. V. Hansen, P. B. Lowry, R. D. Meservy, D. M. McDonald, Genetic programming for prevention of cyber-terrorism through dynamic and evolving intrusion detection, *Decision Support Systems* 43 (4) (2007) 1362–1374.
- [19] A. Arnes, F. Valeur, G. Vigna, R. A. Kemmerer, Using hidden markov models to evaluate the risks of intrusions, in: *International Workshop on Recent Advances in Intrusion Detection, Springer, 2006*, pp. 145–164.
- [20] W. Lee, S. J. Stolfo, K. W. Mok, A data mining framework for building intrusion detection models, in: *Security and Privacy, 1999. Proceedings of the 1999 IEEE Symposium on, IEEE, 1999*, pp. 120–132.
- [21] M. Panda, M. R. Patra, Network intrusion detection using naive bayes, *International journal of computer science and network security* 7 (12) (2007) 258–263.
- [22] Y. Hu, B. Panda, A data mining approach for database intrusion detection, in: *Proceedings of the 2004 ACM symposium on Applied computing, ACM, 2004*, pp. 711–716.
- [23] Y. Li, J. Xia, S. Zhang, J. Yan, X. Ai, K. Dai, An efficient intrusion detection system based on support vector machines and gradually feature removal method, *Expert Systems with Applications* 39 (1) (2012) 424–430.
- [24] R. P. Lippmann, R. K. Cunningham, Improving intrusion detection performance using keyword selection and neural networks, *Computer networks* 34 (4) (2000) 597–603.
- [25] C. Palagiri, Network-based intrusion detection using neural networks, department of Computer Science Rensselaer Polytechnic Institute Troy, New York (2002) 12180–3590.
- [26] D. Apiletti, E. Baralis, T. Cerquitelli, V. DElia, Characterizing network traffic by means of the netmine framework, *Computer Networks* 53 (6) (2009) 774–789.
- [27] A. Tajbakhsh, M. Rahmati, A. Mirzaei, Intrusion detection using fuzzy association rules, *Applied Soft Computing* 9 (2) (2009) 462–469.
- [28] J. Luo, S. M. Bridges, Mining fuzzy association rules and fuzzy frequency episodes for intrusion detection, *International Journal of Intelligent Systems* 15 (8) (2000) 687–703.
- [29] C. Kruegel, D. Mutz, W. Robertson, F. Valeur, Bayesian event classification for intrusion detection, in: *Computer Security Applications Conference, 2003. Proceedings. 19th Annual, IEEE, 2003*, pp. 14–23.
- [30] S. Benferhat, T. Kenaza, A. Mokhtari, A naive bayes approach for detecting coordinated attacks, in: *Computer Software and Applications, 2008. COMPSAC'08. 32nd Annual IEEE International, IEEE, 2008*, pp. 704–709.
- [31] K. Sequeira, M. Zaki, Admit: anomaly-based data mining for intrusions, in: *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2002*, pp. 386–395.
- [32] L. Bilge, E. Kirda, C. Kruegel, M. Balduzzi, Exposure: Finding malicious domains using passive dnsanalysis., in: *Ndss, 2011*.
- [33] L. Bilge, D. Balzarotti, W. Robertson, E. Kirda, C. Kruegel, Disclosure: detecting botnet command and control servers through large-scale netflow analysis, in: *Proceedings of the 28th Annual Computer Security Applications Conference, ACM, 2012*, pp. 129–138.
- [34] M. S. A. Khan, Rule based network intrusion detection using genetic algorithm, *International Journal of Computer Applications* 18 (8) (2011) 26–29.
- [35] W. Lu, I. Traore, Detecting new forms of network intrusion using genetic programming, *Computational intelligence* 20 (3) (2004) 475–494.
- [36] S. S. Joshi, V. V. Phoha, Investigating hidden markov models capabilities in anomaly detection, in: *Proceedings of the 43rd annual Southeast regional conference-Volume 1, ACM, 2005*, pp. 98–103.
- [37] W. Fan, M. Miller, S. Stolfo, W. Lee, P. Chan, Using artificial anomalies to detect unknown and known network intrusions, *Knowledge and Information Systems* 6 (5) (2004) 507–527.
- [38] N. B. Amor, S. Benferhat, Z. Elouedi, Naive bayes vs decision trees in intrusion detection systems, in: *Proceedings of the 2004 ACM symposium on Applied computing, ACM, 2004*, pp. 420–424.
- [39] Z. Li, A. Zhang, J. Lei, L. Wang, Real-time correlation of network security alerts, in: *e-Business Engineering, 2007. ICEBE 2007. IEEE International Conference on, IEEE, 2007*, pp. 73–80.
- [40] W. Hu, Y. Liao, V. R. Vemuri, Robust support vector machines for anomaly detection in computer security., in: *ICMLA, 2003*, pp. 168–174.
- [41] C. Wagner, J. Francois, T. Engel, et al., Machine learning approach for ip-flow record anomaly detection, in: *International Conference on Research in Networking, Springer, 2011*, pp. 28–39.
- [42] T. Shon, J. Moon, A hybrid machine learning approach to network anomaly detection, *Information Sciences* 177 (18) (2007) 3799–3821.
- [43] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, The KDD process for extracting useful knowledge from volumes of data, *Communications of the ACM* 39 (11) (1996) 27–34.
- [44] C. Shearer, The crisp-dm model: the new blueprint for data mining, *Journal of data warehousing* 5 (4) (2000) 13–22.
- [45] Tyagi, Amit Kumar, Building a Smart and Sustainable Environment using Internet of Things (February 22, 2019). *Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur - India, February 26-28, 2019*. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.3356500>
- [46] Tyagi, Amit Kumar, Cyber Physical Systems (CPS)- Opportunities and challenges for improving cyber security, *International Journal of Computer Applications*, 2016, 137 (14).
- [47] Sravanthi Reddy, M. Shamila, Amit Kumar Tyagi, Cyber Physical Systems: The Role of Machine Learning and Cyber Security in Present and Future, *Computer Reviews Journal, PURKH, Vol. 5* (2019).
- [48] Meghna Manoj Nair, Amit Kumar Tyagi, Richa Goyal, Medical Cyber Physical Systems and Its Issues, *Procedia Computer Science* Volume 165, 2019, Pages 647-65.

Evolution of Privacy Preservation Models in Location-Based Services

A B Manju¹, Sumathy Subramanian^{2,*}

¹*School of Computer Science and Engineering, Vellore Institute of Technology, Vellore.*

²*School of Information Technology and Engineering, Vellore Institute of Technology, Vellore.*

ARTICLE INFO

Article history:

Received: 17 February, 2020

Accepted: 10 April, 2020

Online: 03 May, 2020

Keywords:

Location-based services

Privacy preservation

Fog computing

challenges

benefits

ABSTRACT

Location-based services have become increasingly prevalent with the advancement in the positioning capabilities of smart devices and their emergence in social networking. In order to acquire a service, users must submit their identity, query interest and location details to service providers. Such information shared by users are accumulated continuously, stored and analyzed in order to extract the knowledge base from it. Generally, this extracted information is used by service providers to provide users with personalized services. The accumulated data have enormous market value which is found to be used for many lucrative purposes. This work presents a detailed study on the evolution of existing privacy preservation models need to preserve privacy, and the opportunities to integrate fog computing services into privacy architectures. The study proposes a fog integrated privacy preservation model exploring the benefits and open research issues in traditional models and recent integrated fog models. Future directions of fog incorporated privacy preservation models are presented.

1. Introduction

Though location-based services (LBS) originated in the early 1990s, they became significant only after 2000. Since then, massive improvements have been made in facilitating technologies (e.g. telecommunications services), expanding applications (e.g. from outdoor to indoor environments), delivering interfaces (e.g. Smartphone, smart devices) and increasing technological innovations that have made the ambient environment more user-friendly (e.g. an increasing number of devices connected to the Internet and access to 5G). Meteoric development of the functionality of mobile devices play a vital role in bringing comfort to people's everyday lives [1]. Low-cost positioning devices with acceptable power consumption have made location-based services accessible to the common man and, in addition to providing profitable business opportunities [2]. Although it comforts end-users with on-demand and recommendation based services, significant concerns about privacy [3] have become a dominant issue. In order to make use of location-based services (LBSs), service users must disclose their private data, such as their identity, location and query information, to third-party service providers who cannot be trusted. The exposed data is accessed through snapshot queries

(single query) and continuous queries (continuous follow-up queries). When user data is collected over a period of time, a short user profile is created on the basis of the data accumulated. User profile data [4] is used profitably at the discretion of the service provider and moreover most location based-services are typically offered free of charge. When a service user is at a particular point on Earth, LBS providers infer users' interest on the basis of the user's time, location and query data. The point on earth is therefore considered to be significant data in the LBS, represented in latitude and longitude data. The amount of data accumulated by service providers infer the user's private data, which leads to user tracking, gathering user's daily activities, finding the user's home and office address, and children's school or college. Remarkable real-world case studies represent the unauthorized use of users' private data for monetary profits, cyber-stalking of victims, intrusion of thieves, and many such activities. Current location-based service policies need to be revised with stronger security standards to support hesitant location-based service users.

Developing cloud computing technology has facilitated many location service providers to outsource their data in order to use the cloud storage service efficiently [5]. Security issues occur as location data is outsourced to cloud service providers, because cloud providers may benefit from location data. The plain text is therefore encrypted before being outsourced to the cloud.

*Sumathy Subramanian, Email: ssumathy@vit.ac.in

Cryptographically signed data cannot be transferred directly to location service users. Users should therefore be assured of the key to the decryption of the data. Users must receive encrypted data from the cloud and keys from service providers. But, this track of users-cloud and, users-LSP (Location Service Provider), have privacy issues. The cloud service provider operates as a user and collect decryption keys (dual identity attacks [6]). Integrating cloud servers into location privacy models have increased complexities at user and LSP end. Simultaneously, fog computing systems have developed to provide distributed services at the edge of the networks [7]. Contemporary development in cloud computing technology has introduced fog computing, with features such as distributed architecture, location awareness, enhanced security, local storage, processing, increased latency, and connectivity support. The fog integrated design models for location-based services have become significant [8]. Generally, location based privacy schemes support peer-to-peer and trusted third party (TTP) models. Users need to undertake privacy and security policies in peer-to-peer models, as they do not implement intermediate servers, while in TTP-based models, intermediate servers manage privacy and security protocols. Recent work, such as K-anonymity [9] dummy based [10] and mix zone [11] models, have adapted TTP servers to ensure privacy and security. Adapting TTP servers have some drawbacks, such as an intruder hacking TTP to access confidential user data. This has prompted many design models to incorporate fog services and enhance protection and privacy in location-based services. Fog servers can replace conventional TTP servers by preventing TTP vulnerabilities such as single point failures and security issues. Although recent studies have introduced fog servers [12] as an intermediate server, the dynamics of fog servers have not been used exhaustively in privacy preservation models. This work explores privacy preservation in location-based services, as well as the feasibility and benefits of integrating fog servers as middleware instead of TTP servers.

The survey explores traditional privacy preservation models and recent fog integrated models to understand the benefits of integrating fog into privacy architectures. In addition, two different types of privacy preservation models are proposed, such as the integration of fog in the user-collaborative approach and the trusted third-party approach. Overall, the survey presents opportunities for future directions for the preservation of privacy in location-based services and benefits in integrating fog into the privacy preservation architecture.

The paper is structured as follows. Section 2 presents the evolution of privacy concern in location-based services, followed by need for privacy preservation in location-based services in section 3. Traditional privacy models for location-based services are detailed in section 4; section 5 uncovers the location privacy preservation attributes for location-based services. Motivation of integrating fog computing in privacy models is presented in Section 6, followed by Section 7 covering existing fog integrated privacy preservation models. Section 8 details the proposed fog incorporated approaches and Section 9 presents the conclusion of the work.

2. Evolution of privacy concern in location-based services

The scientists at MIT initiated the concept of GPS for the first time on October 4, 1957 and observed that the frequency of the

radio signals from the Russian satellite increased as it approached closer and decreased as it moved away. They were able to track the position of the satellite and the speed of movement using the frequency of the signals. Using the distance from the satellite, the position of the receiver in the ground can be tracked. The theory has grown, creating a huge impact in the field of GPS systems. Currently (2019) there are 74 GPS satellites operating in space where 31 are operational, 9 are being assessed for failure replacement, 2 are being tested, 2 lost during launch and 30 have expired. At the early stage of development, location-based services were segmented into location-based tracking applications and position-based applications. In tracking applications, push services, such as local fast food commercials, are pushed into users' smart devices, and in the positioning applications, the device location is used to update the timing of the mobile phone. The lightweight dynamic pseudonym approach [13] was developed as part of the service agreement and the active pseudonym is chosen by the user and submitted to the service provider. In order to provide the service, the service provider logs into the dynamic pseudonym. The pseudonym is dropped by the user at the end of the service. The service provider logs into a complex alias to provide the service. However, in agreement with the service provider, the pseudonyms are created by taking into account the service providers as a fully trusted party, whereas the trust agreement is not defined.

The need for trustworthy and intelligent middleware telematics (location-based telematics) is addressed by the authors in [14], who pioneered the idea of middleware servers to forward the user's request to telematics servers. But the principle of confidence for smart middleware is not obvious. The work proposed by [15] elaborates the privacy concerns of mobile users and the importance of designing an option that allows users to turn off the location of their devices. The authors survey a community of location-based service users in another distinct study [16] and conclude that "service users are not very concerned about their privacy when the services are helpful in emergencies. A comprehensive risk prediction analysis of LBS adoption is presented in [17]. This analysis reveals different ways in which consumers can adapt LBS, such as the revision of device policies towards consumers, the social contract between service providers and consumers, the integration of third parties for privacy services and privacy preferred services. The research in [18] states that most LBS providers are mobile communication providers, and hence privacy risks are higher than individual LBS providers. Mobile communication providers can easily track the mobile users' through cell tower information. The user cooperative method has been suggested in [19], where an agent is randomly chosen from the user group to forward group communications to service providers. However, collaborative user selection policies have not been established. The proposed work in [20] implements Casper server to respond to requests of, especially anonymous queries. The incorporation of Casper increases the complexity of the service providers' architecture in the LBS. Region-aware privacy protection technique is proposed in [21]. Two types of dummy selection strategies, such as circular area-based and grid area-based, have been developed. Compact processing is designed on the server-side for the processing cost reduction of dummy users. The practical feasibility of modifying the architecture of the server-side is challenging, as multiple users have different adoptions of privacy protection. The proposed hybrid approach

[22] allow users to switch between the peer-to-peer and collaborative approach based on the number of neighboring peers present. The aim of this work is to provide users with privacy in either case. An anonymous server is used in another query based privacy protection system [23] to forward queries from users to location service providers. The anonymous server shields the identity of the users, sending users request as anonymous request to achieve privacy. Nevertheless, the trust between the user and anonymous server has not been discussed in this study. The proposed work in [24] aims to protect the privacy of users without a trusted third party. Cloud server support is used to evaluate the user density present in each region in order to achieve user-side spatial cloaking. Distributed anonymous servers are deployed as a proxy between the LSP and users in order to forward the spatial cloaking area of group of users to LSP [25]. Yet it is burdensome to deploy and manage the distributed network of servers. The query privacy scheme proposed in [26] has a trusted agent to maintain network parameters such as key management and data management between a service provider and a cloud server. To access the device parameters, the user registers with service providers and requests the query from cloud service providers. The system must, however, maintain a completely trusted agent. Moreover adversaries target trusted agents. In [27] the dummy-based approach enhances the dummy features. The dummies are placed at the level as of the speed of the real users. The dummies and the real users are crossed to recover from the accidental reveal of a real user. There is no emphasis on the consistency of the number of dummies to choose. In [9], the collaborative scheme, user device memory is utilized to cache the query request. The trusted agents are eliminated, and the user's collaborative cache enhances the system by sharing the query among the collaborative users. The trust between the users are not elaborated. The dual protection model in [28], provide data privacy to service providers and query privacy to users'. The proposed model outsources the database of service providers to cloud servers by encrypting it. Users availing the service, register with service providers and obtain the secret key. The encrypted data from the cloud is decrypted at the users' side with keys. This model considers that cloud service providers do not collude with any other entities; however, possibilities are not focused. The system model proposed in [29] has a convertor and anonymizer in between user and the LBS provider. The convertor defines the user-defined grid to a uniform grid and is sent to the user. The encrypted request is forwarded to the anonymizer and the encrypted response from a service provider is forwarded from anonymizer to user. The anonymizer also maintains cache of the data for future queries. The maintenance of more than one middle agent increases maintenance complexities and have practical feasibility concerns. The R-constrained dummy based scheme proposed in [30], constructs virtual circles throughout the trajectory of the users for trajectory protection. The cost of processing the dummies is burdensome for the system. The semantic information of the location is utilized to generate fake queries [31]. In this approach, the queries are generated by the system based on the time and the location semantic information. However, the users' queries are not always related to the semantic location information of the users. To enhance the caching based design, a trusted agent in the middle is utilized to cache the efficient data that is frequently requested by the majority of users [8]. The trusted agent combines the K-spatial request from many users and eliminates the

duplicates, to enhance the processing time at the server and to reduce the network traffic. The agent may collect sensitive users' information and use them profitably. The ongoing research evolution in location-based services is described widely in [32].

At the initial stages (2000-2004) users had less concentration on their privacy as they are helpful in emergency services. Moreover, awareness of nefarious activities was less. The awareness of users' information collection at the service providers' side was increasing (2002-2005), hence users' hesitation towards the usage of LBS was increasing. As a result, the service providers started revising their privacy policies in making them transparent to the service users' (2005). In the period (2000-2006), most of the location-based services are provided by mobile communication providers; hence providing privacy protection becomes complex. As mobile communication providers monitor the users' location based on cell tower information. In (2005-2006) simple pseudonym exchange models were proposed to hide the identity of the users. Random user collaboration approaches were initiated to eliminate agent in the middle. As the users' devices are not much capable of storing the queries for the future, the caching was not feasible. In (2007-2009), the random selection models were proposed to select the dummy users' and the behavioral pattern of the dummies was not much concentrated. In (2010-2012), many trusted third parties in the middle were proposed. They were deployed as a single agent or multiple agents as per the requirements. Various levels of user side caching have been proposed during (2012-2016), as the storage capabilities of smart devices have been enhanced. In addition, the number of location-based services increased dramatically, with third-party service providers starting to use cloud storage services. During (2016-2019), dummy-based strategies have been provided at the level of real user activity by improving dummies' behaviors and concentrating on the locations where dummies are chosen. In addition, distributed computing, such as fog computing, was incorporated into privacy preservation models instead of trusted agents.

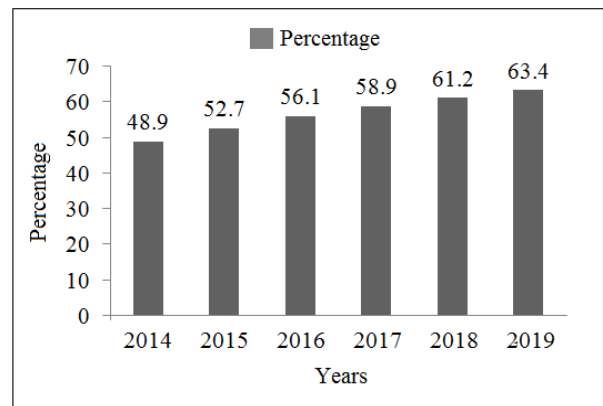


Figure 1: Percentage of mobile internet users from 2014 to 2019 [34]

3. Need for privacy preservation in location-based services

According to Allied Market Research [33], location-based service market is expected to grow from 23.74 billion in 2018 to 157.34 billion in 2026. The enormous increase in the number of mobile Internet users has also triggered an increasing number of location-based applications. Figure 1 shows the rate of increase in the number of mobile phone internet users [34].

The need for privacy and awareness among service users is increasing as a result of the increasing number of online cybercrime cases. The trendy and dominant online social networking websites, such as Facebook and Twitter, provide registered customers with a free-of-cost subscription. These giant organizations have users from all over the world, where activities such as user's personal information, official follow-ups are deliberately uploaded to users. However, there are other unknown sources of information service providers can avail, such as locations from where the users' login, monitoring users' online activities to extract the users' interest for personalised advertisements and recommendations that have huge business profits. Service providers have no right to use the information of users for unauthorized purposes. There is no law controlling the information distributed from the service providers end to other third parties. The information includes privacy data that contain user's personal habits, regular timeline visits, business profits, banking details, and family member information. When this private information reaches the hands of targeted attackers, it can lead to unwanted stalking, theft, abuse of women, and kidnapping.

Increasing numbers of cyber-crime cases enable privacy breaches from small organizations to large service providers. In 2011, iPhone's hidden location synchronization was uncovered, and user locations were sent to Apple without user's knowledge. Similarly, angry bird game collects the age, gender, and location of the user [35]. The main concern of users is that the private information is trapped in the hands of adversaries that result in vulnerabilities.

- Privacy threats

Location service users are exposed to threats in many ways, such as tracking service users (tracking threats), mapping online identity to real-world identity (identification threats) and uncovering online behavioral patterns (user profiling threats).

- Tracking threat

Service users need to use location-based services in many situations to know the location information. Timing information related to the service request is the significant data that links the day-to-day activities of users in accordance with time [36]. When these private data are analyzed, the adversary may be able to track the location of the user throughout the day [37]. With accurate data analytics, past, present and future locations of users are easily exposed to attackers.

- Identification threat

The online identification used by users can be linked to the real-world identity of users with the help of quasi-identification attributes such as geographical tags in uploaded photos, home and office addresses from personal websites [38]. Adversaries may be able to identify the real identity of the user and map the data of the user.

- User threat profiling

The location information associated with the time exposed by the user reveals the user's private information [39]. When the online activities of users are documented for a period of time, the data analyzed reveals user profiles containing health conditions,

religious beliefs, marital status, political interest, business details, and the home branch of the bank [40].

4. Traditional system models of privacy preservation for location-based services

The generic framework of location-based service is the Peer-to-Peer and trusted third party model.

4.1. Peer-to-Peer model

The basic structure of the peer-to-peer model is illustrated in Figure 2, which consists of three entities, such as location-based service users, location providers and location-based service providers. With the help of GPS technology, the user acquires the current location from the location provider via his smart device. The current location of the user is then sent to location-based service provider along with the user identity and query interest to avail the service. The location-based service provider responds with the location of the user based on the query request.

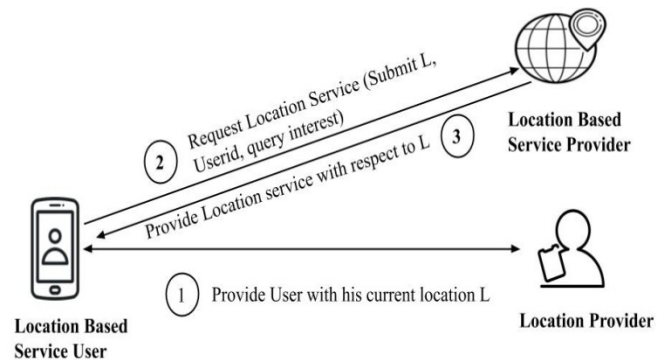


Figure 2: Peer-to-Peer Architecture

In the peer-to-peer model, data is communicated between the service user and the service provider directly. The privacy of users are defined by a user-trusted collaboration model [9] or with the help of online friends circles such as social networking friends [41]. In general, users' do not send the unprocessed data to the service providers, hence in [42], the users get collaborated locally with n-hop distance and the data grouped is sent to the service providers. The advancement of social networking sites has increased social friendships and their bonding. Trust between social friends is used to hide private data sent to the (Location Service Provider) LSP. Location obfuscation models have been implemented as dummy-based approaches [27] and location perturbation approaches [43]. In obfuscation models, users independently outsource their data as anonymous data such as enlarging their position in an area and adding dummy users. The main drawback of the model is the service user, who becomes solely responsible for the outsourcing of the data. However, these models do not require additional systems to support. In feelings-location privacy [3], the authors have presented the depersonalization of location-based on user-desired location The k-locations chosen for the protection of privacy are based on the priority of the user and also on the popularity of the region. In this approach, finding k-locations is a complicated process, as all locations chosen must be equally popular as users' location.

The dummy-based approach formulated in [31] considered the correlation between the location subject and the query type to

generate more robust dummies. Dummies are created in locations where the semantic subject information and the query type are the most appropriate. Finding dummy locations to match semantic information is complex as it depends on the spatial distribution of a region. The authors of the user-centric location privacy architecture [44] proposed the user-desired level of privacy in which the service user decides which data to be sent and which should not. Although service automation has enhanced the process, finding dummy locations with service similarities in all geographic regions is still complex. The asymmetric encryption technique involved in [45] preserves the sharing of private locations with social friends. The location can only be known to friends by decrypting the location information. Encryption and decryption increase the number of messages being exchanged between friends. The dummy-based approach proposed in [27] creates dummies with a replica of the actual user. Dummies are limited to travel and are managed to keep similar to the actual user. User membership benefits could save the cost of dummy user processing in location-based services. However, in the current research, the conditions for the number of dummies to be created have not yet been defined.

4.2. Trusted third party-based architecture

The architecture of the trusted third party server is shown in Figure 3. Unlike the peer-to-peer model, third-party servers have been integrated between the location-based service user and the location-based service provider. The user obtains the current location from the location provider, and then sends the user's identity, current location, and query interest to the trusted third party server. A trusted third-party server also receives a request from other service users. The user-identity is hidden and the query is sent to a location-based service user as an anonymous query using a third party server. Service response from location-based service providers is sent to trusted third party servers, and third party servers finally segregate user response and forward it to service users.

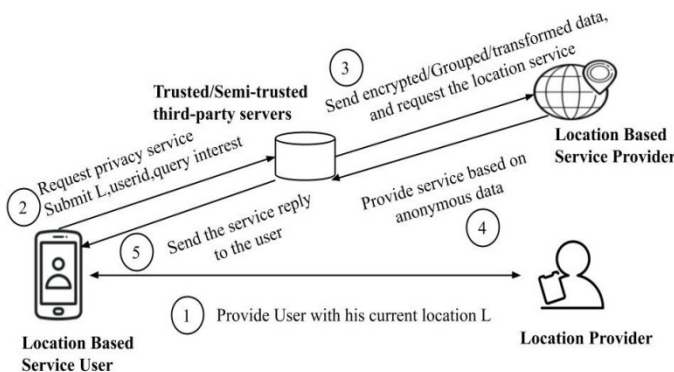


Figure 3: Trusted third party based architecture

The TTP servers are responsible for the data submitted by users. The cost of incorporating TTP servers is an additional burden for service users. Attackers may target TTP servers to access private data. Trust maintenance between the service user and the TTP server defines the robustness of the system. TTP servers are implemented as fully trusted [46] or, in some cases, as semi-trusted [6], to overcome privacy threats. The location transformation approach [47] replaces a fully trusted third-party

server with a semi-trusted server and a function generator as an intermediary. Semi-trusted servers and function generators work independently. The function generator transforms the location coordinates, and without the knowledge of the transformation parameter, the semi-trusted server does not have a chance to learn the true location. This approach has an additional burden on the execution of a function generator and a semi-trusted server, making the model expensive. The trusted third party intermediary servers involvement in [48] forms a user group under the intermediary server. The authors argue that there is no need to exchange pseudonyms (in order to hide the real identity of the users) as they do not enhance security; instead, the members of the group are qualified on the basis of positive, negative and no change in membership to continue with the group on the basis of their activities. However, trusted third-party servers are expensive and maintaining a group becomes more complex.

The k-anonymity approach [49] involves the location perturbation server in the middle. The intermediate server maintains the private data of the k-users and sends the group request to the LBS to protect the identity, location and query of the k-users. However, trust issues arise from the trusted location perturbation server implementation. Incorporating more than one TTP server has also been experimented to prevent TTP from learning private data of users [50].

5. Attributes of location privacy

Figure 4 shows the attributes of location privacy. The service user identity can be an email ID, phone number, unique login ID and device ID. There are many LBS applications that require verification of email ID or phone number before acquiring the service. Few LBS applications do not ask for any user identities, such as finding "My location" in Google. However, the service is used based on the continuous monitoring of the type of service obtained with a many context-based link information about the user obtained at service providers end. The user's identity is protected by the use of pseudonyms acquired from TTP servers. The real identity of the user is replaced by a pseudonym (fake ID). Whenever different pseudonyms are used, the adversaries find it hard to track the user. However, when TTP servers work with LSPs, real users can be easily tracked. As a result, user collaboration approaches have evolved [51]. In collaborative user approaches, users exchange their user ID in a temporary collaboration that eliminates TTP servers.

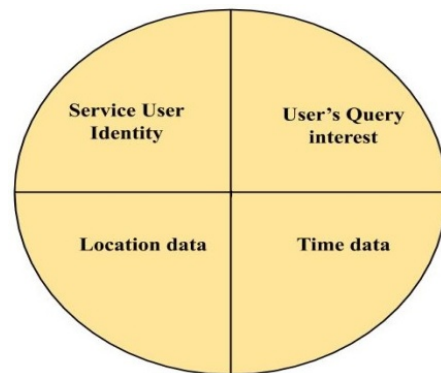


Figure 4: Attributes of location privacy

The location data of the user is the primary information used as the basic attribute that is required for the LBS service. Based on the location data, service users home location, the office address and the school address of the children can be obtained. The security attribute "location" is obtained from service users based on their service usages. When users obtain location-based services on a continuous basis, user trajectory data can be easily mapped with location updates. The location of the user is protected by

- Add noisy locations to the real location of the user [52]
- Location transformation approaches: real location of the user is transformed to other neighboring locations [53]
- Group-based service procurements: Exchange the location of the users with group members [54]

The time data, acquired by the adversaries is linked with other attributes to coordinate the users' activities depending on the time of the day. The exact time of the service request is delayed by the user to hide the timing of the service request from the user. Time transformation will help third party intruders to record the delay in the user's request for service instead of actual time.

Information on the users' **query** reveal their personal interest, which becomes the key information for personalized online advertisements. To hide users' query interest, there are models such as crypto-algorithms that encrypt the query at the users' end and decrypt it in the service providers end. There are different types of models that add noise data to the exact query information. The addition of noise data to the exact query must not delay the quality of the service provided to the user. In this regard, the user collaboration approaches have helped by adding a collaborative user query, and the intermediate server approach has helped users to hide their identity by adding other users from the same region.

The PIR (Private Information Retrieval) technique was used to reduce computational latency [55]. The specific query of the user is anonymized with the help of expanding the size of the region from the location of the user. The query response for all the locations in the region is stored in the local database. Users retrieve an accurate query response from the local database. The requirement and trust of the local database is not defined. Protecting the privacy in continuous queries creates more complexity [56]. Fake queries are added to the actual queries in order to anonymize the query of the user. Fake queries inserted into the actual queries must be contextually linked to the location of the user in order to avoid attackers to remove the fake queries. In [56], a query pool is built for each location, which provides queries from the historical query request provided by different users. However, the processing of large queries for single users creates additional burden for service providers.

In general, the privacy preservation models degrade the precision of information submitted to the service providers and target to acquire the services accurately without compromising the service quality.

5.1. Location Privacy

Location privacy threat is the leakage or misuse of service users' location information by the service providers or other adversaries [57]. The most popularly used protection scheme is the dummy-based models. Semantic location-related information is used to generate realistic dummies [31]. However, the number

of dummies to be generated is not defined. Dummy locations are generated within a circular area where the actual location is centered [30]. However, there is a high risk of exposure to the centered real location. Most dummy-based models are designed to generate realistic dummies, and similarly, the attribute-conscious scheme [58] uses location attributes to generate dummies in the context of location query probabilities. Vehicular location privacy protection based on the vehicles in their proximity is implemented in [51], but the dynamic collaboration has the risk of management of the dynamic group. Another vehicular privacy protection model [59] enhanced the dynamic group formation technique by introducing local hotspots and global hotspots. The positive activities of the group members act as a credit to join the group each time. The collaborative approach in [40] suggests that TTP servers to be replaced by service users device resources, and user collaborative groups need a central controller. The multidimensional privacy protection model [60] with both location and query protection is targeted and based on the model, the semi-anonymous server is incorporated to direct the request to a service provider. The accuracy of the results of the query may be degraded in this model.

5.2. Query Privacy

Query privacy protection prevents adversaries from accessing accurate query information. Basically, query protection schemes fall into two broad categories, such as query obfuscation techniques [61] and dummy query insertion techniques [62]. Query protection scheme "Dummy-Q" [56] eliminates TTP servers by using mobile resources to store the query pool system in order to optimally store the queries used by the quad tree system. Cloud servers [63] prevent location and query directly from being submitted to service providers. The user submits the enlarged region where the service is needed, and the encrypted data related to the region is sent to the cloud server where the cloud server assists the user in the requested service. In [64], the TTP server is used to collaborate with users and to send a collaborative query request to the service provider where the TTP servers trust is not defined. The work proposed in [5] presents on how cloud servers can effectively replace TTP servers, and how users can gain greater privacy on the basis of homomorphic encryption.

5.3. User Identity Privacy

The online user identity associated with online user activities defines user behavior patterns and serves as the perfect information for personalized recommendations and advertisements. Identity protection models of users are comparatively less focused than the location and query protection, as the achievement of location privacy and query privacy completely undermine the identity of the user. In general, the pseudonyms replace the identity of the user. Dynamic pseudonyms are used to protect the identity of users. Pseudonyms are used to hide the identity of service users [65]. The game-theoretic approach has been implemented in [66] in order to protect the identity and location of the service user's privacy.

5.4. Privacy metrics

Wide range of privacy metrics are used to measure the protection achieved. Privacy in location-based service is achieved by using a fake identity, encrypted or anonymized query and

location instead of actual name, query and location. The purpose of the privacy measure is to evaluate any breach attempted by an attacker. The most common privacy metrics used are

- Entropy (H)

The most widely used metric inspired by the Shannon Information Theory is entropy [67]. Entropy is a logarithmic measure of the number of states with a significant probability, explicitly and the states with a substantial probability of being occupied. Entropy is used by modifying it according to their considered parameters. In general entropy is defined in Eq.(1)

$$H = -\sum_{i=1}^K P_i \log P_i \quad (1)$$

Where 'P_i' is defined in Eq.(2)

$$P_i = Q_i / (\sum_{i=1}^K Q_i) \quad (2)$$

Maximum entropy is achieved when all K locations have the same query probability Q, as shown in Eq.(3)

$$H_{max} = \log_2 K \quad (3)$$

The higher the value of entropy, greater the privacy achieved.

- Single location exposure risk (SE)

The probability of exposing the actual location of the service user from the group of locations chosen for anonymity is single location exposure risk [30]. The exposure probability of single location from the group of location points D_i is shown in Eq.(4)

$$\frac{1}{D_i} \quad (4)$$

The probability of exposing the actual location of the user from the group of locations termed as 'set n' is defined in Eq.(5)

$$SE = \frac{1}{n} \sum_{i=1}^n \frac{1}{|D_i|} \quad (5)$$

The lesser the value of SE, the greater the privacy.

- Trajectory Exposure Risk (TE)

Number of dummy trajectories m, in which s defines the number of trajectories that overlap, and (m-s) trajectories that do not overlap [30]. The trajectory exposure risk is determined as in Eq.(6)

$$TE = \frac{1}{(m-s) + T_s} \quad (6)$$

where T_s is the total overlapping trajectories present in the group of trajectories formed by the user. TE value is aimed at a minimum to achieve higher privacy.

- Distance deviation (dd)

The mean value of offset distance between the location position of real trajectory and the dummy trajectory [30] is defined in Eq.(7)

$$dd = \frac{1}{n} \sum_{t=1}^n \left(\frac{1}{m} \sum_{j=1}^m L_{dist}(RL^t, DL_j^t) \right) \quad (7)$$

where m is the number of dummy trajectories, t is time instance and (RL^t, DL^t) is the distance between each location position in the real and dummy trajectory. Minimum distance deviation defines maximum privacy.

- Distance deviation degree (D_{degree})

The distance deviation degree is the mean value of distance deviation (dd) and radius (|R|) of the circular area defined for dummy locations generation [30]. When n is the number of dummy locations the D_{degree} is defined as in Eq. (8)

$$D_{degree} = \left(\frac{1}{n} \sum_{i=1}^n dd_i \right) / |R| \times 100 \quad (8)$$

Lesser the D_{degree} value, maximum privacy is achieved.

- Temporal similarity between real and the dummy trajectory (Sim_t)

The temporal difference between the real and the dummy trajectory should be minimum in order to increase privacy [30]. The temporal similarity is defined in Eq.(9)

$$Sim_t = \frac{\|(\tau' - \tau)\|}{\Theta} \quad (9)$$

where τ' is the query request time of real trajectory and τ is the query request time of dummy trajectory. 'Θ' is maximum time threshold defined by the user and '|| · ||' is the normalization. The higher the value of Sim_t, the maximum privacy is achieved.

- Spatial similarity between the real and the dummy trajectory (Sim_s)

The spatial similarity between the real and the dummy trajectory is measured using Eq.(10)

$$Sim_s = \frac{\| \langle x, y \rangle, \langle x', y' \rangle \|}{\delta} \quad (10)$$

in which <x,y> is the spatial position of real location and <x',y'> is the spatial position of the dummy location. δ is the maximum spatial threshold set by the users [30]. The higher value of Sim_s achieves higher privacy value.

- Anonymous area requirement (A_{Area})

The anonymous area requirement is 100% when the anonymous area determined satisfies the anonymous area defined by the user [12]. If (A_{min} = A), then A_{Area} is 100%, where A_{min} is the minimum area defined by the user and A is the anonymous area determined. The higher the value of A_{Area}, the maximum the privacy. However, it increases the processing cost as well.

- Position protection (PP)

The position protection [12] is defined in Eq.(11)

$$PP = \left(\frac{(x', y')^p - (x, y)}{(x', y')^p} \right) \quad (11)$$

in which (x',y')^p is the number of all dummy positions and (x,y) is the actual position. The maximum value of PP leads to maximum privacy.

- Trajectory protection (TP)

The trajectory protection [12] is determined based on the number of valid trajectories. When S_T represents the total valid trajectories, the trajectory protection is determined using Eq.(12)

$$TP = (S_T - 1) S_T \quad (12)$$

The maximum the value of TP, the maximum is the privacy achieved.

6. Research motivation of integrating fog computing in privacy models

The analysis of the privacy protection models presents the existing downsides in the current protection models. For example, in the context of peer-to-peer architecture, users are solely responsible for forwarding the request to service providers. Users need to take care of protection techniques such as data encryption, user collaboration, local storage, and user mobility. When users are in emergency situations such as the abduction of attackers, road accidents, trapped in the forest, and other such activities, they will not be able to take complex steps to use the location services and could increase their risk of being trapped in the hands of strangers.

Considering TTP as a solution for peer-to-peer approaches, they are also of concern to users. TTP servers are designed to forward user service requests in a privacy-friendly manner, but the risk of single-point failure attacks [68] is unsolvable. The risk of trusting anonymous servers is always a matter of concern. Instead of defining TTP servers as fully trusted, research focused on semi-trusting servers. The semi-trusted servers [6] were implemented to store encryption keys, encrypted data of the users' collaboration details and other details related to the users. The problem with semi-trusted servers is that they might get collaborated with third party-service providers to map the users' private data.

Traditional location-based services are provided to customers through centralized cloud-based approaches. Cloud computing policies have been satisfying customers as they evolved, but the huge increase in data movements from and to cloud computing has degraded the quality of the service it provides to customers. Centralized cloud computing has therefore evolved to serve customers in a promising way, without compromising the quality of services in a decentralized manner. The promising solution for decentralized computing is termed as fog computing by Cisco [69]. Fog computing serves customers at the edge of the network (at their local ends) and the data is being processed with the help of networking resources [68]. Further, if necessary, the data will be sent to the cloud for processing. The promising fog computing solution has the following advantages: proximity to end-users, geographical distribution, optimum resource utilization, low latency, reduced network traffic, improved service quality and a superior computing environment for users. By acquiring the benefits of fog in research, fog computing can fit into privacy-preserving LBS models instead of traditional TTP servers.

7. Existing fog integrated privacy preservation models

The characteristics of fog computing, such as improved security, decentralized control, improved latency, local computing, stimulate many researchers to incorporate fog. The fog computing technique is used in many recent works as local computing. The benefits of fog computing are used by the incorporation of fog

nodes [70]. IoT devices are used to implement the location privacy protection algorithms by incorporating surveillance cameras in the location of users, and to forward the user request to service providers [71]. Fog-based privacy preservation technique [72], implemented fog servers to store encrypted data, and users are provided with decryption keys from the location service provider. Keys are generated based on the region of division; therefore, vehicles entering the region will only be able to access the region key. In [73], TTP servers are replaced by fog servers to eliminate single point failure attacks and to store cache data.

8. Proposed Fog incorporated approaches

Based on previous deliberations, it is clear that there is a strong need for privacy preservation techniques that makes existing privacy policies more user-friendly. In addition, the integration of fog servers will bring enormous benefits to service users, service providers and global green computing benefits [74]. Resources between the source (end users) and the destination (cloud servers) are referred to as fog resources.

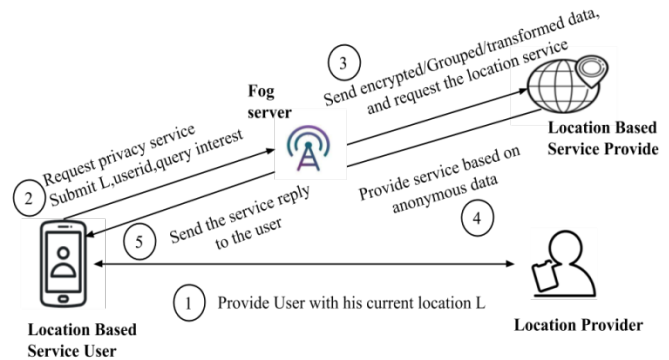


Figure 5: Trusted third party as fog server

8.1. Fog Server as TTP

The architecture of fog server as intermediate server is shown in Figure 5. The user obtains the current location from the location provider and then sends the user's identity, current location, and query interest to the fog server. A fog server also receives a request from other service users. The user-identity is hidden and the query is sent to a location-based service user as an anonymous query using a fog server. Service response from the location-based service provider is sent to fog server, and fog server finally segregates user response and forwards it to service users. Fog servers are intermediate servers set up by fog service providers at the edge of the network with the help of edge resources. The fog servers are proposed to establish at the locations where the fog services are required the most (based on the number of tasks forwarded to the cloud from that location). The fog servers established in such locations can act as the TTP servers for privacy protection models. The local map information can be stored in fog servers for easy updates and retrievals.

The research gaps identified are

- Frequency of the cache data update

Fog servers are deployed with available idle resources from end-users [75] and therefore have fewer resources than the cloud. Cache memory in fog is used for faster access to location-based data stored in fog [73]. Cache memory must be updated on the

basis of the newest data accessed by users. It is therefore necessary to focus on updating the cache memory frequency based on the availability of the cache memory and to use the cache efficiently.

- Optimal utilization of fog resources

The fog resources are deployed at each location based on the requirements and the focus of utilizing the resources optimally plays a major role in fog services. Fog resource optimization focused on recent works [76] [77] emphasizes the importance of optimized resource utilization in fog services.

- Security issues in fog storage

Fog computing services are expected to face many security issues other than those inherited from cloud [78] Cloud servers are deployed and maintained by a single party, while fog servers take on a variety of deployment options, such as end-users, cloud providers, and internet providers [79]. Trust issues while using e-commerce services, insiders attack fog providers, secure data storage and authentication issues prevail in fog services [80].

8.2. User collaborative approach

The architecture of user collaborative fog incorporation is shown in Figure 6. Initially, the users from a proximity collaborate into a group and a group representative is chosen among them. Then each user obtains the current location from the location provider, then sends the user's identity, current location, and query interest to the fog server as a group request. A fog server also receives a request from other service users groups. The user-identity is hidden and the query is sent to a location-based service user as an anonymous query using a fog server. Service response from location-based service provider is sent to fog server, and fog server finally segregates user response and forwards it to service users. Further, the group representative segregates the users request and sends it to each user. In this approach two level anonymization is achieved, one at user level and the other at fog server level.

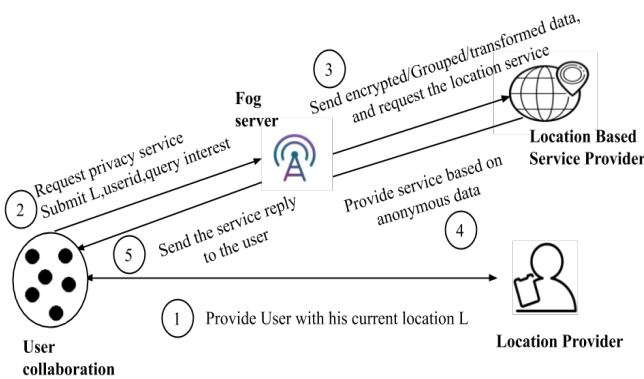


Figure 6: User collaborative approach incorporating fog

User collaborative techniques reduce the risk of third party dependence; however, they increase the burden on the user side. Peer-to-peer computing [81] is gaining popularity due to the increasing number of smart devices and their computational capabilities. Users are reluctant to establish peer-to-peer computation as there is a need to establish trust between peers. The key challenge is to establish cooperation between devices owned by different individuals [82]. Influenced by online social

networking sites, individual social relations are on the rise every day [83]. Consider, for example, a device user from home interacts with a friend in the office or a family member in the neighborhood. These interactions will set up a device-to-device relationship to work together for peer computational task. Relationships are established in the framework of mutual cooperation. However, for the benefit of others, no device voluntarily establishes communication. In such situations, the previous history of device assistance helps the devices to help each other. Incentive mechanisms will bring satisfactory benefits for users of devices in order to build a fair relationship between devices.

The ultimate aim of collaborative computing is to establish a user group that is physically or socially connected. These group members exchange their privacy attributes, such as their identity, location and query in order to acquire privacy-friendly location-based services.

The research gaps identified are

- Defining trust between users

Trust between users during collaboration is critical in collaborative approaches. In collaboration with a group, no single user must leave the system until all the users in the group are prompted to benefit equally. Trust models are developed based on user history, and online or offline user relationships.

- Central authority to manage the user group

The group's central representative leads the group in a positive direction in the models of user collaboration. The selection of a central group representative is an open issue in collaborative strategies.

- Automated dynamic group formation

The pervasive nature of mobile devices always forms a dynamic group, as proximity users are not always the same. The policy focusing on the benefit of the group members must push LBS users to join the secure LBS group.

- Incentives for the collaborative members

Collaborative user work is being developed to eliminate third-party servers [77]. Voluntary involvement of users in collaboration is difficult because users become greedy and are unable to use their resources for the benefit of others. Incentive mechanisms for such collaborative models are needed to encourage users to participate in collaboration, where a user acting as a representative will receive incentives from all other users [84]. The implementation of an effective reward system will ensure productive cooperation between users.

9. Conclusion

The location-based services are increasingly gaining significance along with the increasing utilization of mobile devices. The survey presents an overview of the evolution of the privacy preservation models of location-based services. The work describes the current research attainment of fog-integrated models of privacy preservation for location-based services. The research benefits and issues are described in detail and the opportunities to integrate the fog into the current user-collaboration approach and trusted third party approach are proposed. The research outcomes

provide a better understanding of the current research scenarios of privacy preservation techniques and future directions in integrating advanced computing paradigms such as fog computing in the privacy preservation approaches. The survey provides directions for many fog integrated robust privacy approaches in order to gain market adoptions soon.

References

- [1] J. Raper, G. Gartner, H. Karimi, and C. Rizos, "A critical evaluation of location based services and their potential," *J. Locat. Based Serv.*, vol. 1, no. 1, pp. 5–45, 2007.
- [2] Ryan Goodrich, "Location-Based Services: Examples and Uses," <https://www.businessnewsdaily.com>, 2013. [Online]. Available: <https://www.businessnewsdaily.com/5386-location-based-services.html>. [Accessed: 06-Dec-2019].
- [3] T. Xu and Y. Cai, "Feeling-based location privacy protection for location-based services," in *Proceedings of the ACM Conference on Computer and Communications Security*, 2009, pp. 348–357.
- [4] W. Luo, Y. Lu, D. Zhao, and H. Jiang, "On Location and Trace Privacy of the Moving Object Using the Negative Survey," *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 1, no. 2, pp. 125–134, 2017.
- [5] H. Zhu, R. Lu, C. Huang, L. Chen, and H. Li, "An efficient privacy-preserving location-based services query scheme in outsourced cloud," in *IEEE Transactions on Vehicular Technology*, 2016, vol. 65, no. 9, pp. 7729–7739.
- [6] J. Chen, K. He, Q. Yuan, M. Chen, R. Du, and Y. Xiang, "Blind Filtering at Third Parties: An Efficient Privacy-Preserving Framework for Location-Based Services," *IEEE Trans. Mob. Comput.*, vol. 17, no. 11, pp. 2524–2535, 2018.
- [7] "Cisco Fog Data Services," 2015. [Online]. Available: <https://www.cisco.com/c/en/us/products/cloud-systems-management/fog-data-services/index.html>. [Accessed: 17-Apr-2019].
- [8] S. Zhang, X. Li, Z. Tan, T. Peng, and G. Wang, "A caching and spatial-anonymity driven privacy enhancement scheme in continuous location-based services," *Futur. Gener. Comput. Syst.*, vol. 94, pp. 40–50, 2019.
- [9] T. Peng, Q. Liu, D. Meng, and G. Wang, "Collaborative trajectory privacy preserving scheme in location-based services," *Inf. Sci. (Nj.)*, vol. 387, pp. 165–179, 2017.
- [10] D. Wu, Y. Zhang, and Y. Liu, "Dummy location selection scheme for k-anonymity in location based services," in *Proceedings - 16th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, 11th IEEE International Conference on Big Data Science and Engineering and 14th IEEE International Conference on Embedded Software and Systems*, 2017, pp. 441–448.
- [11] Q. A. Arain et al., "Clustering Based Energy Efficient and Communication Protocol for Multiple Mix-Zones Over Road Networks," *Wirel. Pers. Commun.*, vol. 95, no. 2, pp. 411–428, Jul. 2017.
- [12] G. Li, Y. Yin, J. Wu, S. Zhao, and D. Lin, "Trajectory Privacy Protection Method Based on Location Service in Fog Computing," in *Procedia Computer Science*, 2019, vol. 147, pp. 463–467.
- [13] T. Rodden, A. Friday, H. Muller, and A. Dix, "A lightweight approach to managing privacy in location-based services," in *2nd International workshop on mobile commerce*, 2002, no. Equator-02-058, pp. 15–24.
- [14] C. Bisdikian et al., "Intelligent pervasive middleware for context-based and localized telematics services," in *Proceedings of the ACM International Workshop on Mobile Commerce*, 2002, pp. 15–24.
- [15] A. Dey and L. Barkuus, "Location-Based Services for Mobile Telephony: a Study of Users' Privacy Concerns," in *IFIP Conference on Human-Computer Interaction*, 2013, pp. 702–712.
- [16] L. Barkhuus, "Privacy in Location-Based Services, Concern vs. Coolness," 2004.
- [17] H. Xu, H. H. Teo, and B. C. Y. Tan, "Predicting the adoption of location-based services: The role of trust and perceived privacy risk," in *Association for Information Systems - 26th International Conference on Information Systems, ICIS 2005: Forever New Frontiers*, 2005, pp. 897–910.
- [18] V. Johnson, R. Torres, B. Phillips, and A. Rahnamae, "Continued Usage of Location-Based Services: Privacy Risk Impact on Motivation and Adoption," *Inf. Q.*, vol. 1, no. 1, pp. 1–17, 2014.
- [19] C. Y. Chow, M. F. Mokbel, and X. Liu, "A peer-to-peer spatial cloaking algorithm for anonymous location-based service," in *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems*, 2006, pp. 171–178.
- [20] C. Y. Chow, M. F. Mokbel, and W. G. Aref, "The new casper: Query processing for location services without compromising privacy," *ACM Trans. Database Syst.*, vol. 34, no. 4, pp. 763–774, 2009.
- [21] B. Niu, Z. Zhang, X. Li, and H. Li, "Privacy-area aware dummy generation algorithms for location-based services," *2014 IEEE Int. Conf. Commun. ICC 2014*, pp. 957–962, 2014.
- [22] C. Zhang and Y. Huang, "Cloaking locations for anonymous location based services: A hybrid approach," *Geoinformatica*, vol. 13, no. 2, pp. 159–182, 2009.
- [23] F. Liu, K. A. Hua, and Y. Cai, "Query l-diversity in Location-Based Services," in *Tenth International Conference on Mobile Data Management: Systems, Services and Middleware*, IEEE., 2009, pp. 436–442.
- [24] S. Wang and X. Sean Wang, "In-device spatial cloaking for mobile user privacy assisted by the cloud," in *Proceedings - IEEE International Conference on Mobile Data Management*, 2010, pp. 381–386.
- [25] Y. Wang, D. Xu, X. He, C. Zhang, F. Li, and B. Xu, "L2P2: Location-aware location privacy protection for location-based services," in *Proceedings - IEEE INFOCOM*, 2012, pp. 1996–2004.
- [26] H. Zhu, F. Liu, and H. Li, "Efficient and Privacy-Preserving Polygons Spatial Query Framework for Location-Based Services," *IEEE Internet Things J.*, vol. 4, no. 2, pp. 536–545, 2017.
- [27] T. Hara, A. Suzuki, M. Iwata, Y. Arase, and X. Xie, "Dummy-Based User Location Anonymization under Real-World Constraints," *IEEE Access*, vol. 4, pp. 673–687, 2016.
- [28] M. Zeng, K. Zhang, J. Chen, and H. Qian, "P3GQ: A practical privacy-preserving generic location-based services query scheme," *Pervasive Mob. Comput.*, vol. 51, pp. 56–72, 2018.
- [29] S. Zhang, K. K. R. Choo, Q. Liu, and G. Wang, "Enhancing privacy through uniform grid and caching in location-based services," in *Future Generation Computer Systems*, 2018, vol. 86, pp. 881–892.
- [30] J. Zhang, X. Wang, Y. Yuan, and L. Ni, "RcDT: Privacy Preservation Based on R-Constrained Dummy Trajectory in Mobile Social Networks," *IEEE Access*, vol. 7, pp. 90476–90486, 2019.
- [31] G. Sun et al., "Location privacy preservation for mobile users in location-based services," in *IEEE Access*, 2019, vol. 7, pp. 87425–87438.
- [32] H. Huang, G. Gartner, J. M. Krisp, M. Raubal, and N. Van de Weghe, "Location based services: ongoing evolution and research agenda," *J. Locat. Based Serv.*, vol. 12, no. 2, pp. 63–93, 2018.
- [33] "Location-Based Services Market Predicted to Hit \$157.34 billion by 2026." [Online]. Available: <https://www.alliedmarketresearch.com/press-release/location-based-services-market.html>. [Accessed: 06-Dec-2019].
- [34] "Global mobile phone internet user penetration 2019 | Statista," 2018. [Online]. Available: <https://www.statista.com/statistics/284202/mobile-phone-internet-user-penetration-worldwide/>. [Accessed: 06-Dec-2019].
- [35] J. Bell, "Angry Birds and 'leaky' phone apps targeted by NSA and GCHQ for user data," *The Guardian*, 2014. [Online]. Available: <http://www.theguardian.com/world/2014/jan/27/nsa-gchq-smartphone-app-anxious-birds-personal-data>. [Accessed: 05-Dec-2019].
- [36] O. Jan, A. J. Horowitz, and Z. R. Peng, "Using global positioning system data to understand variations in path choice," *Transp. Res. Rec.*, no. 1725, pp. 37–44, 2000.
- [37] A. Y. Xue, R. Zhang, Y. Zheng, X. Xie, J. Huang, and Z. Xu, "Destination prediction by sub-trajectory synthesis and privacy protection against such prediction," *Proc. - Int. Conf. Data Eng.*, pp. 254–265, 2013.
- [38] C. Bettini, X. S. Wang, and S. Jajodia, "Protecting privacy against location-based personal identification," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2005, vol. 3674 LNCS, pp. 185–199.
- [39] K. Fawaz, H. Feng, and K. G. Shin, "Anatomization and Protection of Mobile Apps Location Privacy Threats," in *24th USENIX Security Symposium (USENIX Security 15)*, 2015, pp. 753–768.
- [40] R. Gupta and U. P. Rao, "Achieving location privacy through CAST in location based services," *J. Commun. Networks*, vol. 19, no. 3, pp. 239–249, 2017.
- [41] M. Han et al., "Cognitive Approach for Location Privacy Protection," *IEEE Access*, vol. 6, pp. 13466–13477, 2018.
- [42] W. Sheng, H. Jiafeng, Z. Hui, W. Hanyi, and L. Fenghua, "A collaboration-based scheme for location-based services with incentive mechanism," *Chinese J. Electron.*, vol. 27, no. 2, pp. 310–317, 2018.
- [43] T. Dimitriou and N. Al Ibrahim, "'I wasn't there'—Deniable, privacy-aware scheme for decentralized Location-based Services," *Futur. Gener. Comput. Syst.*, vol. 86, pp. 253–265, 2018.
- [44] R. Dewri and R. Thurimella, "Exploiting service similarity for privacy in location-based search queries," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 2, pp. 374–383, 2014.
- [45] R. Schlegel, C. Y. Chow, Q. Huang, and D. S. Wong, "Privacy-Preserving

- Location Sharing Services for Social Networks,” *IEEE Trans. Serv. Comput.*, vol. 10, no. 5, pp. 811–825, 2017.
- [46] S. Wang, Q. Hu, Y. Sun, and J. Huang, “Privacy Preservation in Location-Based Services,” *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 134–140, 2018.
- [47] T. Peng, Q. Liu, and G. Wang, “Enhanced Location Privacy Preserving Scheme in Location-Based Services,” *IEEE Syst. J.*, vol. 11, no. 1, pp. 219–230, 2017.
- [48] X. Pan, J. Xu, and X. Meng, “Protecting location privacy against location-dependent attacks in mobile services,” *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 8, pp. 1506–1519, 2012.
- [49] B. Gedik and L. Liu, “Protecting location privacy with personalized k-anonymity: Architecture and algorithms,” *IEEE Trans. Mob. Comput.*, vol. 7, no. 1, pp. 1–18, 2008.
- [50] S. Zhang, G. Wang, M. Z. A. Bhuiyan, and Q. Liu, “A Dual Privacy Preserving Scheme in Continuous Location-Based Services,” *IEEE Internet Things J.*, vol. 5, no. 5, pp. 4191–4200, 2018.
- [51] J. Cui, J. Wen, S. Han, and H. Zhong, “Efficient Privacy-Preserving Scheme for Real-Time Location Data in Vehicular Ad-Hoc Network,” *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3491–3498, 2018.
- [52] H. Lu, C. S. Jensen, and M. L. Yiu, “PAD: Privacy-area aware, dummy-based location privacy in mobile services,” *MobiDE 2008 - Proc. 7th ACM Int. Work. Data Eng. Wirel. Mob. Access*, pp. 16–23, 2008.
- [53] T. Peng, Q. Liu, and G. Wang, “Privacy preserving for location-based services using location transformation,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8300 LNCS, 2013, pp. 14–28.
- [54] A. Solanas and A. Martínez-Ballesté, “A TTP-free protocol for location privacy in location-based services,” *Comput. Commun.*, vol. 31, no. 6, pp. 1181–1191, 2008.
- [55] F. Olumofin, P. K. Tysowski, I. Goldberg, and U. Hengartner, “Achieving efficient query privacy for location based services,” 2010.
- [56] A. Pingley, N. Zhang, X. Fu, H. A. Choi, S. Subramaniam, and W. Zhao, “Protection of query privacy for continuous location based services,” in *Proceedings - IEEE INFOCOM*, 2011, pp. 1710–1718.
- [57] B. Niu, Q. Li, X. Zhu, G. Cao, and H. Li, “Achieving k-anonymity in privacy-aware location-based services,” in *Proceedings - IEEE INFOCOM*, 2014, pp. 754–762.
- [58] Y. Li, W. Li, C., & Geng, “APS: Attribute-aware privacy-preserving scheme in location-based services,” *Information Sci.*, 2019.
- [59] D. Liao, H. Li, G. Sun, M. Zhang, and V. Chang, “Location and trajectory privacy preservation in 5G-Enabled vehicle social network services,” *Journal of Network and Computer Applications*, vol. 110, pp. 108–118, 2018.
- [60] T. Peng, Q. Liu, G. Wang, Y. Xiang, and S. Chen, “Multidimensional privacy preservation in location-based services,” *Futur. Gener. Comput. Syst.*, vol. 93, pp. 312–326, 2019.
- [61] J. Shao, R. Lu, and X. Lin, “FINE: A fine-grained privacy-preserving location-based service framework for mobile devices,” in *Proceedings - IEEE INFOCOM*, 2014, pp. 244–252.
- [62] H. Niu, B. Zhu, X., Lei, X., Zhang, W., & Li, “Eps: Encounter-based privacy-preserving scheme for location-based services,” in *IEEE global communication conference*, 2013.
- [63] Z. Liu, L. Wu, J. Ke, W. Qu, W. Wang, and H. Wang, “Accountable Outsourcing Location-Based Services With Privacy Preservation,” *IEEE Access*, vol. 7, pp. 117258–117273, 2019.
- [64] P. Zhao, J. Li, F. Zeng, F. Xiao, C. Wang, and H. Jiang, “ILLIA: Enabling k-Anonymity-Based Privacy Preserving Against Location Injection Attacks in Continuous LBS Queries,” *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1033–1042, 2018.
- [65] R. Yu, J. Kang, X. Huang, S. Xie, Y. Zhang, and S. Gjessing, “MixGroup: Accumulative Pseudonym Exchanging for Location Privacy Enhancement in Vehicular Social Networks,” *IEEE Trans. Dependable Secur. Comput.*, vol. 13, no. 1, pp. 93–105, 2016.
- [66] Y. Qu, S. Yu, L. Gao, W. Zhou, and S. Peng, “A hybrid privacy protection scheme in cyber-physical social networks,” *IEEE Trans. Comput. Soc. Syst.*, vol. 5, no. 3, pp. 773–784, 2018.
- [67] Jianhua Lin, “Divergence Measures Based on the Shannon Entropy,” *IEEE Trans. Inf. Theory*, vol. 37, no. 1, pp. 145–151, 1991.
- [68] M. Aazam and E. N. Huh, “Fog computing micro datacenter based dynamic resource estimation and pricing model for IoT,” in *Proceedings - International Conference on Advanced Information Networking and Applications, AINA*, 2015, vol. 2015-April, pp. 687–694.
- [69] Cisco, “Cisco Fog Computing Solutions : Unleash the Power of the Internet of Things,” *White Pap.*, pp. 1–6, 2015.
- [70] Y. He, J. Ni, B. Niu, F. Li, and X. (Sherman) Shen, “Privbus: A privacy-enhanced crowdsourced bus service via fog computing,” *J. Parallel Distrib. Comput.*, vol. 135, pp. 156–168, 2020.
- [71] Y. Tian, M. M. Kaleemullah, M. A. Rodhaan, B. Song, A. Al-Dhelaan, and T. Ma, “A privacy preserving location service for cloud-of-things system,” *J. Parallel Distrib. Comput.*, vol. 123, pp. 215–222, 2019.
- [72] S. Liu, A. Liu, Z. Yan, and W. Feng, “Efficient LBS queries with mutual privacy preservation in IoV,” *Veh. Commun.*, vol. 16, pp. 62–71, 2019.
- [73] T. Wang et al., “Trajectory Privacy Preservation Based on a Fog Structure for Cloud Location Services,” *IEEE Access*, vol. 5, pp. 7692–7701, 2017.
- [74] S. Sarkar and S. Misra, “Theoretical modelling of fog computing: A green computing paradigm to support IoT applications,” *IET Networks*, vol. 5, no. 2, pp. 23–29, 2016.
- [75] Y. Sun and N. Zhang, “A resource-sharing model based on a repeated game in fog computing,” *Saudi J. Biol. Sci.*, vol. 24, no. 3, pp. 687–694, 2017.
- [76] V. Mushunuri, A. Kattapur, H. K. Rath, and A. Simha, *Resource optimization in fog enabled IoT deployments*. 2017.
- [77] O. Skarlat, M. Nardelli, S. Schulte, M. Borkowski, and P. Leitner, “Optimized IoT service placement in the fog,” *Serv. Oriented Comput. Appl.*, vol. 11, no. 4, pp. 427–443, Dec. 2017.
- [78] Y. Zhi, Z. Fu, X. Sun, and J. Yu, “Security and Privacy Issues of UAV: A Survey,” in *Mobile Networks and Applications*, 2019, pp. 1–10.
- [79] O. Skarlat, S. Schulte, M. Borkowski, and P. Leitner, “Resource provisioning for IoT services in the fog,” in *Proceedings - 2016 IEEE 9th International Conference on Service-Oriented Computing and Applications, SOCA 2016*, 2016, pp. 32–39.
- [80] F. Alghamdi, S. Mahfoudh, and A. Barnawi, “A Novel Fog Computing Based Architecture to Improve the Performance in Content Delivery Networks,” *Wirel. Commun. Mob. Comput.*, pp. 1–13, Jan. 2019.
- [81] D. S. Milojevic et al., “Peer-to-Peer Computing,” 2003.
- [82] S. Ye, F. Makedon, and J. Ford, “Collaborative automated trust negotiation in peer-to-peer systems,” *Proc. - 4th Int. Conf. Peer-to-Peer Comput. P2P2004*, pp. 108–115, 2004.
- [83] M. Nitti, G. A. Stelea, V. Popescu, and M. Fadda, “When social networks meet D2D communications: A survey,” *Sensors (Switzerland)*, vol. 19, no. 2, 2019.
- [84] X. Li, M. Miao, H. Liu, J. Ma, and K. C. Li, “An incentive mechanism for K-anonymity in LBS privacy protection based on credit mechanism,” *Soft Comput.*, vol. 21, no. 14, pp. 3907–3917, Jul. 2017.

A Harmonized European Drone Market? – New EU Rules on Unmanned Aircraft Systems

Anna Konert*, Tadeusz Dunin

¹Faculty of Law and Administration, Łazarski University, Warsaw 02-662, Poland

ARTICLE INFO

Article history:

Received: 17 February, 2020

Accepted: 10 April, 2020

Online: 03 May, 2020

Keywords:

Air law

BVLOS operations

Drones

EU law

European Union drone market

Polish law

UAS

UAV

Unmanned aircraft systems

VLOS operations

ABSTRACT

The European drone market has been showing steady growth year after year. New EU drone rules will come into force as of July 1st, 2020, with the European Union setting itself the target to replace national rules with a common regulation with the ultimate goal of creating a harmonized European drone market. This study will clarify that the EU's regulatory framework covers all types of existing and future drone operations, creating an international market for unmanned aircraft services. This move will facilitate the enforcement of citizen's privacy rights and address security issues and environmental concerns to the benefit of the EU citizens. Moreover, this study will show that national legislators are now faced with the difficult task of replacing their national regulations with EU rules, however these were drafted so fast that they still leave a number of issues to be decided on by national legislators. The method of this study comprised of content analysis of existing legislation. The current doctrine were confronted with existing regulations, documents, and materials.

1. Introduction

The UAV and UAS markets have been showing steady, increasing growth year after year in Europe. Some of the world's top drone service providers are based here, and their results are proof of continuous and considerable market growth [1]. Some forecasts predict that the European drone market will more than double in size between the years 2018 and 2024 [2]. Others predict that revenues generated from commercial drone use in Europe will grow from USD 251 million in 2020 to USD 3 billion in 2025 [3]. Different estimates see the region's drone market bringing in EUR 10 billion annually by 2035, and over EUR 15 billion annually by 2050 [4], with countries like Germany, France and the UK representing the lion's share of this total mainly due to their regulatory structure and the sheer size of their economies.

This predicted growth will most likely stem from an increase in popularity in commercial drone usage in many areas. According to SESAR (Single European Sky ATM Research), there are presently around 10,000 commercial UAV's in use in Europe – with forecasts for this number to rise to 200,000 in 2025, and up to 395,000 in 2035.

The non-military use of drones is already significant and extensive, spreading to areas never before though possible just a decade or so ago. UAVs are presently being used by the media (TV), in law enforcement and policing activities, in border patrols, for agricultural and environmental monitoring, by fire services, for power line surveying, aerial photography, and in security management systems (mostly in autonomous and semi-autonomous locomotion). For practical applications of UAVs in security management systems, please see “A Cyber-Vigilance System for Anti-Terrorist Drives Based on an Unmanned Aerial Vehicular Networking Signal Jammer for Specific Territorial Security” [5], “Design and implementation of a cyber-vigilance system for anti-terrorist drives based on an unmanned aerial vehicular networking signal jammer for specific territorial security” [6], “An Aerial Landmine Detection System with Dynamic Path and Explosion Mode Identification Features” [7]. Moreover, SESAR predicts that the application of UAVs will be more diverse as we move forward, with rising numbers of drones employed in the energy sector, public safety and security, mobility and transport, the media, insurance, e-commerce and delivery, real estate, academic research, telecommunications, and mining and construction [8].

*Corresponding Author: Anna Konert, a.konert@lazarski.edu.pl

It is no surprise that authorities across the Old Continent are scrambling to make Europe a powerhouse in terms of unmanned aerial business and a global first with regards to the mainstream and commercial use of UAVs. Take, for example, the U-Space initiative and its goal to implement a future air traffic management system that is open to and even capable of monitoring and directing drones (UAS Traffic Management, UTM [9]). Ambitious as this may sound, such initiatives get the ball rolling for a future that is integrated with unmanned vehicles.

As with every new technology, unmanned aircraft vehicles or systems possess their own regulatory and legal constraints. However, the main legal issue is of a regulatory nature, as the central aim of aviation regulations is to provide an adequate level of safety. In the case of UAVs, regulations aim to protect third parties on the ground and in the air, as there are no crew or passengers on board.

In order to enable the safe operation of remotely-controlled aircraft, or their operation by automated flight control systems, it is essential to develop detailed solutions for a number of technical and operational aspects.

This includes but is not limited to the following issues:

- airworthiness standards for UAV and UAS elements,
- data transmission standards for remote pilot stations, including protection against unauthorized interference etc.,
- collision avoidance systems, including motion and obstacle detection,
- emergency systems (e.g. loss of communications),
- unmanned traffic management systems,
- operator competences and training [10].

Drafting regulations that would enable the safe operation and development of unmanned aviation is no easy task. These need to be developed with a thorough understanding of the technology involved and with detailed knowledge of the risks that come with UAV operations. The scope of integration of UAVs with manned aviation is especially significant, as both operate within the same airspace. This is a long-term challenge for UAS, however it is co-dependent on improvements made in the field of manned aviation in the future [11].

From a regulatory point of view, the first step on the road to the integration of manned and unmanned flight is to allow for Visual Line of Sight Operations (VLOS). Most countries have already taken the appropriate steps for this by adopting or developing their national legislation to include drones, usually with weight limits up to 25 kg, in VLOS operations.

The next step forwards is for a regulator to allow Beyond Visual Line of Sight Operations (BVLOS) outside of segregated airspace or autonomous operations – which is now the case in Poland and some other EU countries.

When thinking about the European Union, it is all too convenient to think of its constituents as a single uniform marketplace. In reality, the EU is a group of drone technology markets each with their own regulations, and all of these have distinct specifications characteristic to their respective countries. Over the years, these regulations have been a barrier for the adoption of commercial drone technology across the region, with EASA (European Union Aviation Safety Agency) even claiming that they stifle business development and innovation [12]. But this is about to change with

the agency's publication of a set of common, Pan-European regulations for UAV operations [13].

This paper will analyze current regulations in Poland and new EU regulations, showing how the region's drone market will change in the nearest future.

2. EU Regulations

“Europe will be the first region in the world to have a comprehensive set of rules ensuring safe, secure and sustainable operations of drones both, for commercial and leisure activities. Common rules will help foster investment, innovation and growth in this promising sector” – Patrick Ky, Executive Director of EASA [14].

In 2014, the European Commission published a communication entitled “A new era for aviation. Opening the aviation market to the civil use of remotely piloted aircraft systems in a safe and sustainable manner” [15].

Next, a legislative process aimed at extending the EU's competences in terms of safety regulations for drones and their operations was launched in 2015. In the opinion of EU institutions, the differences that exist between national regulations of Member States hamper market growth for the entire UAS sector.

By adopting Regulation (EC) No 216/2008 of the European Parliament and of the Council on common rules in the field of civil aviation, and by establishing the European Aviation Safety Agency [16] (through the so-called Basic Regulation on EASA), the EU set out to regulate matters related to the operation of civil unmanned aircraft with an operating mass of less than 150 kg. The establishment of regulations that concern unmanned aircraft used by the military, governmental services, average citizens, those designed for experimental purposes, and all other unmanned aircraft with an operating mass of no more than 150 kg fall within the competences of EU Member States. Such a state of affairs resulted from the fact that EASA's competences were extended in 2008 with the adoption of Regulation 216/2008, however the EU did not take into consideration the then unpredictable and rapid development of the unmanned aircraft sector [17].

On September 11th, 2018, Regulation (EC) No 216/2008 was repealed with the purpose of the new regulation establishing and maintaining a high and uniform level of civil aviation safety in all Member States, including the harmonization of regulations on the use of UAVs in all EU countries. Some of the main changes that were introduced included the obligation for operators to register their UAVs online (those weighing over 250 g) and common standards for UAV certification. However, these EU regulations serve as guidelines and do not contain specific provisions on the implementation of UAV operations, which were introduced by the Commission's delegated and implementing regulations.

As of July 2020, national rules in EU Member States will be replaced by a common EU regulation. Since neither the EU Parliament nor the EU Council had any objections, both Implementing and Delegated Acts (Commission Delegated Regulation (EU) 2019/945 [18] and Commission Implementing Regulation (EU) 2019/947 [19]) were published in mid-June of last year and entered into force 20 days later. These regulations will become gradually applicable one year from their date of publication, giving Member States and UAV operators time to prepare with the transitional period being fully completed by 2022. However, UAV operators will be obliged to register in the Member

State of their residence or there where they have registered their main place of business by as early as July 2020.

The new EU regulatory framework will cover all types of existing and future drone operations, enabling operators – once authorized in their state of registration – to freely circulate between Member States.

The purpose of introducing these new regulations is to ensure the safety of UAV operations, as well as protect the privacy of EU citizens, with respect to personal data protection, and the environment while allowing free access to airspace. The new regulations establish technical and operational requirements, provisions for UAS operations and personnel (minimum requirements and operator training), including both pilots and any organizations.

While maintaining its primary aim to ensure safe UAV operations, the new pan-EU regulatory framework will facilitate the enforcement of citizens’ privacy rights and address security issues and environmental concerns. As touched on earlier, these new rules include both technical and operational requirements for UAVs and their operators. They define UAV capabilities, types of operations, and label these into three broad risk-based categories. These three categories of operations are based on the levels of risk involved per UAV flight and each adopts a varied regulatory approach, with UAV flight operational limitations decreasing with the requirement for greater authorization from a Member State’s national aviation authority [20].

Regulation 2019/947 presents a comprehensive system of unified legal regulations which classifies UAVs into three categories based on the risks involved in their operations, their mass, and their application:

- Open. Operations in this category do not require official permission but are subject to a number of restrictions including those on UAV weight (up to 25 kg), types of operations (only VLOS, except when an unmanned aerial vehicle observer is involved), safety requirements (through the use of product safety regulations), obligations for operator registration, and geo-fencing systems. During Open operations, an operator must maintain a safe distance from other people and cannot fly over crowds, at a height of over 120 m, and cannot transport or drop hazardous materials. Open operations are further divided into three subcategories. (Article 4 of Regulation 2019/947);
- Specific. Operations in this category require authorization from aviation authorities based on a risk analysis of the proposed operations. This authorization can either concerns a single operation or a series of operations specified by time or place. Operation authorization must contain a relevant detailed list of risk mitigation measures. (Article 5 of Regulation 2019/947).
- Certified. Operations in this carry the highest risk. This category requires drone certification under Regulation (EU) 2019/945 as well as operator certification, which may, in some cases, include obtaining a drone pilot license. Certified operations include those performed under any of the following conditions: over crowds or gatherings, or involving the transport of persons or dangerous goods. In addition, the competent authority may, based on a risk assessment (Article 6 of Regulation 2019/947).

This new set of common rules (summarized in Table 1 and presented as a visual representation in Figure 1) will provide both

professional and recreational drone operators with a much clearer understanding of what they are and are not allowed to do. Continent-wide and uniform regulations, especially the requirement that UAVs be registered and individually identifiable, will surely prevent events similar to the ones that took place in the UK at Gatwick and Heathrow airports back in 2018 – which grounded aircraft for up to 36 hours, affected over 140,000 passengers and cost airports an estimated GBP 50 million [21] – from occurring again.

Table 1: Authorization and regulations applicable to each of the adopted three categories of UAV operations [22]

Category of operations	Open <i>low risk</i>	Specific <i>medium risk</i>	Certified <i>high risk</i>
Authorisation needed	None	Authorisation from NAA based on operational risk assessment or specific scenario	Authorisation from NAA/EASA
UAS	Compliant with Commission Delegated Regulation on UAS	Compliant with requirements included in the authorisation	Certified UAS
Operations allowed	Restricted to: <ul style="list-style-type: none"> ▪ VLOS ▪ Altitude < 120 m ▪ Other limitations defined by: <ul style="list-style-type: none"> - Commission Regulation on UAS operations - National airspace zones 	Restricted to: <ul style="list-style-type: none"> ▪ Operations specified in the authorisation ▪ Limitations defined by national airspace zones 	Controlled airspace U-Space
Regulations	Commission Regulation on UAS operations in open and specific		Revision of existing aviation regulation
	Commission Delegated Regulation on UAS	No regulatory requirement (UAS requirements included in the authorisation)	

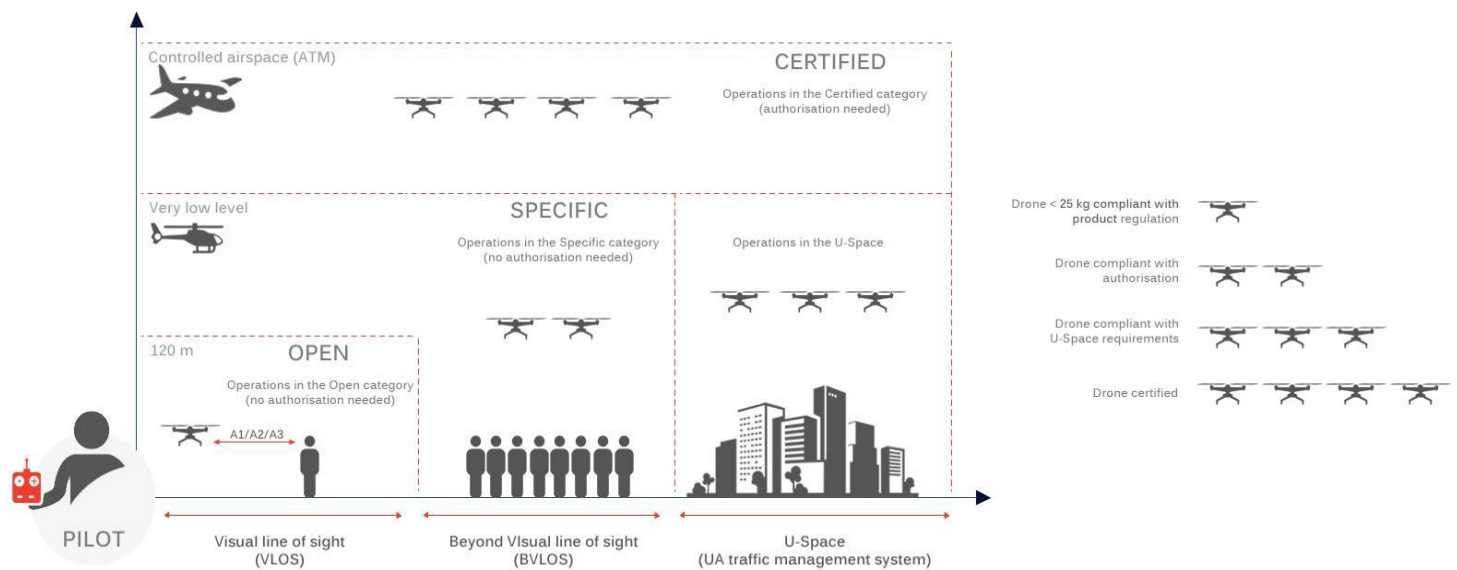
The benefits of facilitating uniform UAV operation requirements across all Member States seem clear. This is an unprecedented step forward, as even the United States are facing problems with UAV regulation unification due to the fact that individual states prefer to enforce local regulations [23].

Also, Member States will be able to define zones, most probably through satellite geo-location, where UAVs will not be permitted to enter or where they will be allowed more freedoms such as BVLOS flights [24], allowing countries to cater to their individual national specificities.

3. Discussion

This section starts with an analysis of Poland’s national regulations. The first national regulations for UAVs were adopted in the country in July of 2011, in the Act of 3rd July 2002 – Aviation Law (Aviation Law Act) [25].

However, the very first implementing rules only concerned VLOS operations. The Regulation of the Minister of Transport, Construction and Maritime Economy of March 26th, 2013, on the exclusion of certain provisions of the Aviation Law Act – for certain types of aircraft and the specification of conditions and requirements for the use of these aircraft [27], was the Polish government’s very first attempt to regulate general requirements for UAV operations, becoming the first national regulation for UAVs in Europe. This regulation was widely amended in 2016 [28], with the amendments being based on the practical aspects of unmanned operations.



A new regulation was adopted on December 20th, 2018 and entered into force on February 1st, 2019. This regulation amended previous rules from 2013 on the exclusion of certain provisions of the Aviation Law Act and included a new set of rules for BVLOS operations (MTOW of up to 25 kg, a maximum height of 120 m, and BVLOS operations outside of segregated airspace) [29].

This created new possibilities for the country's UAV sector, but also led to a series of safety challenges for manned aviation. It is still too early to make an assessment of the new regulation, as the first three months in which it has been in force saw the performance of just 20 UAV BVLOS flights under the new rules. However, this period has led to serious doubts on the risk that such UAV operations pose to manned aircraft operating at lower altitudes.

It must be underlined that the new regulation allows for BVLOS UAV operations outside of segregated airspace. Similar regulations were adopted in other European countries, but these include much stricter limitations. For instance, in France and Spain, such operations are only allowed for UAVs that weigh less than 2 kg [30]. In Spain, however, UAVs with a mass up to 25 kg can operate in BVLOS operation outside of segregated airspace when equipped with an approved detection and avoidance systems [31]. Similar regulations apply in the UK, where operators are required to implement an approved method of aerial separation and collision avoidance as a prerequisite for BVLOS operations [32]. Under Polish law, such systems are not considered as mandatory for BVLOS UAV operations outside of segregated airspace. And this is the main concern regarding the new regulation, even though it also creates the legal framework for the introduction of a UTM system [33].

While operating under national regulations (the details of which may vary between countries), UAV operators are required by law to be prepared for the new EU regulations that will apply across all Member States. The approach, terminology and the requirements used at an EU-level are quite different to national regulations as they are drafted for all EU Member States.

EU countries having their own rules are now obliged to make the transition from old regulations to new as smooth and seamless for drone operators as possible. One example is that the Civil Aviation Authorities will be obliged to transform the current UAV VLOS and BVLOS "licenses" into European equivalents, but it is currently unknown whether operators will be required to undergo supplementary training. The transformation of operator qualification certificates is not the only problem. Poland needs to decide whether new BVLOS regulations will stay in force for operations other than civil.

Due to the fact that EU law applies uniformly in all Member States, it is advisable to discuss the regulations that apply in non-EU countries, and to compile EU countries' own regulations to the extent that EU law allows the state's competence to regulate.

Below is a table that outlines how individual national regulations shape UAV flight and ownership in selected European Union countries.

As shown in the few examples included in Table 2, national regulations concerning UAV operations vary widely from country to country. These differences only hamper the growth of this sector and should be unified in accordance with the EUs Pan-European regulations for UAV operations explained earlier.

4. Results

The EU's new set of UAV rules for the European market which will facilitate the development of drone technologies and create plenty of job and business opportunities. Besides addressing vital issues such as citizen's privacy, security and environmental topics, a far-sighted vision sees this harmonization lead to the creation of a functioning UTM system through U-Space, enabling the large-scale performance of UAV operations in low-level airspace, beyond visual line of sight and in congested areas.

It is therefore critical that these new rules were created by EASA, as this organization is responsible for European air safety is capable of holding extensive stakeholder consultations.

Although it is too early to assess the new EU regulations in practice, since they will only enter into force later this year, this

Table 2: National regulations in selected EU countries

UAV Classification	UAV Requirements	Pilot Requirements	Permits
Austria – Austro Control [34]			
Austria classifies drones by joules of kinetic energy	<ol style="list-style-type: none"> 1. No flying over crowds (e.g. events, sports events, concerts, etc.) without special permission 2. No flying in the immediate vicinity of airports without special permission 	<ol style="list-style-type: none"> 1. Drone pilots must be 16 years of age or older 2. Liability insurance is mandatory for all drone operators 3. Only Austro Control may grant an aviation permit to an operator 	<ol style="list-style-type: none"> 1. A permit is required if 79 joules are exceeded or if a drone is flown above 30 m 2. A permit is also required to fly a drone for commercial purposes or to take photos or record video regardless of operating mass
Belgium – Belgian Civil Aviation Authority (BCAA) [35]			
Belgium has three different classes of operations, with licensing requirements for each class	<p>Class 1a and 1b Operations Definition:</p> <ol style="list-style-type: none"> 1. Up to 300 ft (around 90 m) above ground outside of controlled airspace. 2. Operations can only occur in daylight conditions. 3. Drone must weigh less than 330 lbs. (150 kg). <p>Class 2 Operations Definition:</p> <ol style="list-style-type: none"> 1. Flying height limit of 150 ft (around 45 m) outside of controlled airspace and outside of cities or communities. 2. Operations can only occur in daylight. 3. The drone must weigh less than 11 lbs. (5 kg). 	<ol style="list-style-type: none"> 1. The pilot must maintain a visual line of sight with the drone at all times. 	<p>Licensing Requirements:</p> <ol style="list-style-type: none"> 1. Class 1a and 1b Operations – require a remote pilot license issued by the BCAA. To obtain this license, UAV pilots need to pass both a theoretical examination and a practical skill test. 2. Class 2 Operations – UAV pilots need a certificate of competence issued by the BCAA. To obtain this certificate, UAV pilots need to follow a theoretical course and pass a practical skill test.
France – French Civil Aviation Authority (DGAC) [36]			
Commercial and Recreational	<ol style="list-style-type: none"> 1. Drones may not be flown at night (unless with special authorization from the local prefect). 2. Drones may not be flown over people; over airports or airfields; over private property (unless with owner’s authorization); over military installations, prisons, nuclear power plants, historical monuments, or national parks. Drones may also not be flown over ongoing fires, accident zones, or around emergency services. 3. Drones may not be flown above 150 m (492 ft), or higher than 50 m (164 ft) above any object or building that is 100 m (328 ft) or more in height. 	<ol style="list-style-type: none"> 1. Drone pilots must maintain a visual line of sight with their drones at all times. If a visual observer is tracking the drone, the pilot may fly out of his or her own range of sight. 	<ol style="list-style-type: none"> 1. Commercial – commercial drone pilots must pass a theoretical exam to receive a theoretical telepilot certificate. The pilot must have this printed and with them during all flights. Commercial drone pilots must also undergo basic practical training. Recreational – recreational drone pilots do not need a training certificate when UAV mass is less than 800 g. They must undergo training if UAV mass exceeds this. 2. All drones of 800 g or more must be registered by their owner on AlphaTango, a public portal for users of remotely piloted aircraft. The drone then receives a registration number that must be affixed on the drone.
Italy – Italian Civil Aviation Authority (ENAC) [37]			
Commercial and Recreational	<ol style="list-style-type: none"> 1. Drones being flown for recreational purposes may not fly more than 70 m (230 ft) above the ground. For commercial purposes it is 150 m (492 ft). 2. Drones must be identified by a plate showing the identification of the system and of the operator. An identical plate shall be installed also on the remote ground pilot station. 3. UAVs must be equipped with an Electronic Identification Device. 4. Drones may not be flown within 5 km (3 miles) of any airport or airfield. 5. Drones may not be flown at night. 6. Drones are not allowed to fly over people or crowds, including sports events, concerts, and other large events. 	<ol style="list-style-type: none"> 1. Drone pilots must maintain a direct line of sight with their drone during operations. 	<ol style="list-style-type: none"> 1. Commercial drone pilots conducting low-risk operations must submit a statement of compliance with specific requirements to ENAC along with a 94 EUR processing fee. For higher-risk operations, commercial drone pilots must obtain training and an operating certificate as well as a health certificate.
Netherlands – Netherlands Directorate General of Civil Aviation (DGCA) [38]			
Commercial and Recreational	<ol style="list-style-type: none"> 1. Flying a drone is only allowed in category G airspace. 2. Drones may not fly more than 120 m (394 ft) above the ground. 3. Drones must fly at a safe distance from people and buildings and may not be flown at night. 5. The maximum weight for private drones is 25 kg. 6. UAV pilots may use drones to make aerial films and photographs for their own personal use. This is regarded as recreational use. 	<ol style="list-style-type: none"> 1. Drone pilots must give priority to all other aircraft, such as airplanes, helicopters, gliders, etc. 2. Pilots must maintain a visual line of sight with their drone during operations. 	<ol style="list-style-type: none"> 1. Commercial drone operations in the Netherlands require the drone pilot to hold a pilot’s license, and the company/organization overseeing the operation to hold an ROC permit to fly. 2. If UAV pilots operate a drone weighing no more than 4 kg, then they may apply for a light permit. 3. Drone insurance is required for commercial drone operations in the Netherlands.
Spain – Spanish Aviation Safety and Security Agency (AESA) [39]			
Recreational, Model Aircraft, Professional	<ol style="list-style-type: none"> 1. Drones may be flown up to 120 m (394 ft) above the ground. 2. Drones may only be flown during the day. 3. Flights may be carried out at night for drones with a take-off weight of less than 2 kg (4.4 pounds) as long as a flight altitude of 50 m (164 ft) above the ground is not exceeded. 4. Drone pilots must maintain a distance of 150 m (492 ft) from buildings, and a distance of 50 m (164 ft) or more from people not involved in the flight. 5. Drone pilots must maintain a distance of at least 8 km (5 miles) to airports in uncontrolled airspace, or 15 km (9.3 miles) on approved BVLOS flights. 	<ol style="list-style-type: none"> 1. Drone pilots must maintain a visual line of sight with their drone at all times. 2. BVLOS allowed if a second visual observer monitors the drone and is in direct eye contact with the pilot 	<ol style="list-style-type: none"> 1. A permit is required for commercial drone flights. 2. Liability insurance is required for commercial drone pilots. 3. For flights in national parks, UAV pilots need permission from AESA. 4. The use of drones in no-fly zones must be approved by the Spanish Ministry of Defense.

study shows that certain EU countries (such as Poland) tried to stay ahead of the curve and adapt their national policies to developing and growing markets and technologies. Also, one of the greatest challenges lying ahead for all Member States is making the transition and replacing their current national rules – those that have been widely accepted by UAV operators in them – with an “unknown” set of regulations put forward by a supranational body. Adapting these to the specificity of each Member State’s legal framework will certainly be a difficult task. It must not be forgotten that despite having a unifying nature, the EU’s regulations were created so quickly that they could not possibly cover all UAV-related issues, leaving many blank spaces to be filled in and questions to be answered by national legislators themselves.

5. Conclusion

As we enter the two-year transition period, EASA if aware of the difficulties that lie ahead. The agency even summed up at the recent High Level Conference on Drones, which was held during Amsterdam Drone Week this past December, that a common European vision in which UAVs can be used for emergency services, transportation or even parcel delivery will only become a reality if safety standards are met and solutions are found to pressing issues such as noise and privacy [40]. However, we can be quite certain that this unprecedented continental initiative to unify regulations will foster the growth of the UAV technology sector, allowing private businesses to drive the commercial market forward. May this example of European cooperation serve as a model for other nations around the world as we stand at the brink of a development that could revolutionize how drone technology is used across the region.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Research financed by National Science Center, Poland. Project No 2017/27/B/HS5/0008 Unmanned Aerial Vehicle. A New Era in Aviation Law.

References

- [1] <https://uavcoach.com/droneii-2019-reports/> [Accessed: 13-Jan-2020]
- [2] <https://www.droneii.com/the-drone-market-2019-2024-5-things-you-need-to-know> [Accessed: 09-Jan-2020]
- [3] <https://www.tractica.com/newsroom/press-releases/commercial-drone-hardware-and-services-revenue-to-reach-12-6-billion-by-2025/> [Accessed: 09-Jan-2020]
- [4] European Drones Outlook Study Unlocking the value for Europe. [Online]. Available: https://www.sesarju.eu/sites/default/files/documents/reports/European_Drones_Outlook_Study_2016.pdf, pp.3
- [5] D. Chowdhury, M. Sarkar, M. Z. Haider "A Cyber-Vigilance System for Anti-Terrorist Drives Based on an Unmanned Aerial Vehicular Networking Signal Jammer for Specific Territorial Security" *ASTES Journal* Vol. 3, No. 3, 43-50 (2018)
- [6] D. Chowdhury, M. Sarkar, M. Z. Haider, S. A. Fattah and C. Shahnaz, "Design and implementation of a cyber-vigilance system for anti-terrorist drives based on an unmanned aerial vehicular networking signal jammer for specific territorial security," 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka, 2017, pp. 444-448.
- [7] S. A. Fattah et al., "An aerial landmine detection system with dynamic path and explosion mode identification features," 2016 IEEE Global Humanitarian Technology Conference (GHTC), Seattle, WA, 2016, pp. 745-752.

- [8] European Drones Outlook Study Unlocking the value for Europe. [Online]. Available: https://www.sesarju.eu/sites/default/files/documents/reports/European_Drones_Outlook_Study_2016.pdf, pp.14
- [9] <https://www.sesarju.eu/U-space> [Accessed: 02-Jan-2020]
- [10] P. Kasprzyk and A. Konert, Unmanned aircraft. A new era for aviation. A new era for aviation law?, *Coventry Law Review* no 12 / 2018.
- [11] See more in ICAO RPAS CONOPS (March 2017) and EUROCONTROL RPAS ATM CONOPS (Ed. 4.0, 2017)
- [12] European Aviation Safety Agency, Regulatory framework to accommodate unmanned aircraft systems in the European aviation system, ISSUE 2 — 4.6.2018. [Online]. Available: <https://www.easa.europa.eu/sites/default/files/dfu/ToR%20RMT.0230%20%E2%80%93%20Issue%202.pdf>
- [13] <https://www.easa.europa.eu/newsroom-and-events/news/eu-wide-rules-drones-published> [Accessed: 13-Jan-2020]
- [14] <https://www.easa.europa.eu/newsroom-and-events/press-releases/eu-wide-rules-drones-published> [Accessed: 13-Jan-2020]
- [15] COM/2014/0207 final
- [16] Regulation (EC) No 216/2008 of the European Parliament and of the Council of 20 February 2008 on common rules in the field of civil aviation and establishing a European Aviation Safety Agency, and repealing Council Directive 91/670/EEC, Regulation (EC) No 1592/2002 and Directive 2004/36/EC (Official Journal of the European Union, L 79, 19.3.2008).
- [17] P. Kasprzyk and A. Konert, Unmanned aircraft. A new era for aviation. A new era for aviation law?, *Coventry Law Review* no 12 / 2018.
- [18] Commission Delegated Regulation (EU) 2019/945 of 12 March 2019 on unmanned aircraft systems and on third-country operators of unmanned aircraft systems https://eur-lex.europa.eu/eli/reg_del/2019/945/oj
- [19] Commission Implementing Regulation (EU) 2019/947 of 24 May 2019 on the rules and procedures for the operation of unmanned aircraft https://eur-lex.europa.eu/eli/reg_irimpl/2019/947/oj
- [20] <https://www.easa.europa.eu/document-library/opinions/opinion-012018> [Accessed: 13-Jan-2020]
- [21] <https://www.airport-technology.com/features/new-eu-drone-rules/> [Accessed: 09-Jan-2020]
- [22] <https://dronerules.eu/assets/covers/Table-1.png> [Accessed: 09-Jan-2020]
- [23] <https://www.commercialuavnews.com/europe/easa-publishes-common-drone-rules-for-all-of-europe> [Accessed: 13-Jan-2020]
- [24] https://dronerules.eu/en/professional/eu_regulations_updates [Accessed: 09-Jan-2020]
- [25] Journal of Laws from 2002 No 130, Item 1112 with further amendments. See also M. Żylicz (ed.), *Prawo lotnicze. Komentarz* (Air Law Commentary). Warsaw 2016
- [26] https://dronerules.eu/assets/images/DroneRules_v2.jpg [Accessed: 09-Jan-2020]
- [27] Journal of Laws from 2013, Item 440
- [28] Journal of Laws from 2016, Item 1317
- [29] See more details in P. Kasprzyk & A. Konert, Drones are flying outside of segregated airspace in Poland – new rules for BVLOS UAVs operations, *IEEE proceedings*: <https://ieeexplore.ieee.org/document/8798085>
- [30] LOI n° 2016-1428 du 24 octobre 2016 relative au renforcement de la sécurité de l'usage des drones civils (JORF n°0249 du 25 octobre 2016) (Law n° 2016-1428 of October 24th, 2016 on the reinforcement of the safety of the use of the civil drones)
- [31] Real Decreto 1036/2017, de 15 de diciembre, por el que se regula la utilización civil de las aeronaves pilotadas por control remoto, y se modifican el Real Decreto 552/2014, de 27 de junio, por el que se desarrolla el Reglamento del aire y disposiciones operativas comunes para los servicios y procedimientos de navegación aérea y el Real Decreto 57/2002, de 18 de enero, por el que se aprueba el Reglamento de Circulación Aérea. (BOE Núm. 316 Viernes 29 de diciembre de 2017) (Royal Decree 1036/2017, of December 15, which regulates the civil use of drones and amending Decree 552/2014)
- [32] CAP 722: Unmanned Aircraft System Operations in the UK Airspace – Guidance. [Online]. Available: <https://publicapps.caa.co.uk/docs/33/CAP%20722%20Sixth%20Edition%20March%202015.pdf>
- [33] Such as “Drone Radar” DAMS (Drone Awareness and Monitoring System), see more on <https://droneradar.eu/blog/droneradar-for-ansp/>
- [34] <https://www.austrocontrol.at/jart/prj3/ac/main.jart?reserve-mode=active&rel=en>. <https://uavcoach.com/drone-laws-in-austria/> [Accessed: 14-Feb-2020]
- [35] https://mobilit.belgium.be/fr/transport_aerien/drones, Drone operation type table: <https://149355317.v2.pressablecdn.com/wp-content/uploads/2018/05/bcaa-drone-operations.png> [Accessed: 14-Feb-2020]
- [36] DGAC website: <https://www.ecologique-solidaire.gouv.fr/direction-generale-laviation-civile-dgac>, Vast legal report:

<https://www.loc.gov/law/help/regulation-of-drones/france.php>, AlphaTango website: <https://fox-alphatango.aviation-civile.gouv.fr/en/>, Drone law overview:

https://dronerules.eu/assets/regulationspdfdownloads/NRP_france.pdf [Accessed: 14-Feb-2020]

[37] <https://www.enac.gov.it> [Accessed: 14-Feb-2020]

[38] Recreational drone use: <https://www.government.nl/topics/drone/rules-pertaining-to-recreational-use-of-drones>, Commercial drone use: <https://www.government.nl/topics/drone/rules-pertaining-to-the-commercial-use-of-drones>, DGCA website: <https://www.ilent.nl/zoeken?trefwoord=drone> [Accessed: 14-Feb-2020]

[39] <https://www.boe.es/buscar/doc.php?id=BOE-A-2017-15721>, https://www.seguridadaerea.gob.es/lang_en/cias_empresas/trabajos/rpas/default.aspx, https://dronerules.eu/assets/regulationspdfdownloads/NatinalRegulatoryProfile_Spain.pdf [Accessed: 14-Feb-2020]

[40] <https://www.easa.europa.eu/newsroom-and-events/press-releases/regulators-and-industry-unite-need-address-societal-concerns> [Accessed: 13-Jan-2020]

Personality Measurement Design for Ontology Based Platform using Social Media Text

Andry Alamsyah*, Sri Widiyanesti, Rizqy Dwi Putra, Puspita Kencana Sari

School of Economics and Business, Telkom University, Bandung 40257, Indonesia

ARTICLE INFO

Article history:

Received: 03 March, 2020

Accepted: 22 April, 2020

Online: 03 May, 2020

Keywords:

Big Five Personality

Ontology Model

Personality Measurement

Personality Platform

Social Media

Radix Tree

ABSTRACT

Human behavior quantification is an essential part of psychological science. One of the cases is measuring human personality. Social media provide rich text, which can be beneficial as a data source to get valuable insight. Previous researches show that social media offered favorable circumstances for psychological researchers by tracking, analyzing, and predicting human character. In this research, we propose a personality measurement design to help to assess human character through linguistic usage from human digital traces. We construct our model by classifying social media text to the pre-determined personality facet from Big Five personality traits, mapping the knowledge to the ontology model, and implementing the model as a platform dictionary. Our model is based on the Indonesian language, which to the best of our knowledge is the first in the subject area. The platform is running effectively by using a well-established sorting algorithm, called the radix tree. Our objective is to support psychological science in adapting to a new technological era.

1. Introduction

The presence of advanced technologies, such as social media platforms and mobile devices, are shifting the way on how people communicate. Social media users have the freedom to upload daily routine information, exchanging messages, or even basic conversations [1]. The content created by the users is formed into digital traces. The personality of users revealed through their writing or textual content [2]. Each personality has its own charm that deserves attention over its complex arrangement. There are thoughts, hearts, and feelings that can change over time [3]. Hence, measuring human personality is a hard problem regarding the dynamic feeling's alteration.

Personality measurement has become the most extensive research in the field of Psychology [4]. The legacy methodology of personality assessment performed by interview and written examination [5]. Those methods are integrated method, where an interview is conducted to validate the test result. The characteristic of legacy methodology requires fulfillment instruments, such as the physical tools and psychologists. In order to adapt to the new technological era, some alternative methodologies are developed for bringing a faster process and result. One of them is the measurement through its natural environment or using digital traces on social media [6]. In this research, we utilize digital traces to put forward an automation personality measurement for minimizing fulfill instruments in the legacy methodology. Thus, it

would significantly reduce the cost and reduce the time process of getting the results. This approach needs to be developed, considering that automation is the most demanded characteristic of the Industry 4.0 era.

According to Madden et al. [7], digital traces, which consist of user information, e.g., personal information, shared texts, pictures, and videos, are a proof dataset that cannot be ignored, which expressed online human activity. These footprints are offered valuable opportunities for psychology research in understanding human characteristics [8]. Previous researches have assessed human personality through social media, such as Bhardwaj et al. [9], who assess personality through Facebook and LinkedIn and Park et al. [6] who applying the regression model to predict human character based on social media language.

Most of the research generally uses a machine learning approach, in contrast to our study, which utilizing the ontology approach. Machine learning provides us some leverage, such as the speed of analysis regarding large-scale data [10]. It is also able to predict personality on various forms of data like text, speech, and image [11]. However, the machine learning approach has some weaknesses in processing each meaning and intention of words due to language uniqueness [12]. Ontology afford us a better understanding of contextual knowledge [13]. There is an opportunity to use ontology as a basis for measuring human personality through the words on linguistic usage. Thus, in terms of measuring human nature through social media textual data, the

*Andry Alamsyah, Email : andrya@telkomuniversity.ac.id

ontology model provides a more accurate result than a machine learning approach as long as the collection word type in the ontology model's corpus is prosperous. The following approach scheme is shown in Fig. 1.

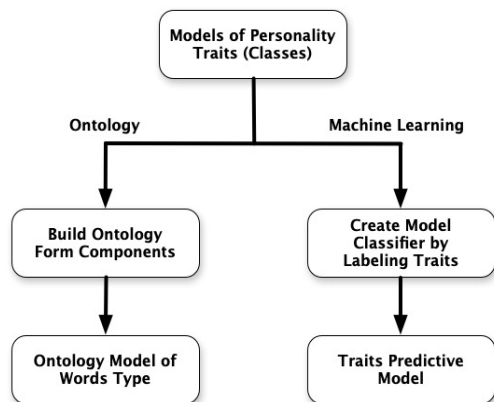


Figure 1: Personality Traits Model Scheme between Ontology and Machine Learning Approach

Fig. 1 shows the difference between ontology and machine learning model schemes. On the ontology model scheme, the personality trait model established by mapping words that reflect personality as a model's forming component. Those collections of words then verified by experts and utilized as a dictionary for measuring personality. This model is a representation of an expert's knowledge and adequate to map more than one personality trait in the complex sentence. Meanwhile, the personality trait model on the machine learning model scheme is a depiction of a machine's algorithm according to the labeling process. The machine interprets the label on the training dataset as a rule for predicting personality. Therefore, the approach of the machine learning model scheme is entirely different than ontology. The ontology model scheme is only mapping the words with reference to an expert's knowledge, while machine learning is predicting based on the learning result of the machine.

To the best of our knowledge, research of personality measurement has put forward in several personality traits taxonomies based on specified purposes, such as for assessing job placement and natural human emotion. In this research, we want to use the most general taxonomy of human's essential persona. The Big Five Personality theory gets a consensus of a general taxonomy regarding human personality traits [14]. It also able to represent and simplify the diverse characteristics of a human's personality [15]. The Big Five Personality Traits is built by examining several unique human attributes in their linguistic usage. This theory also offers prized terms called Revised Neuroticism-Extraversion-Openness Personality Inventory (NEO-PI-R) metric, which facilitate the exposure psychological characteristic of broad trait [16]. In this research, we use the Big Five Personality Traits Theory with the NEO-PI-R metric as a domain knowledge of our ontology model. The personality divided into five domains and further divided to thirty facet scales. Our research also applies the lexical hypothesis as a basis for analyzing textual content in social media. Based on research of Raad and Mlacic [17], the lexical hypothesis is a process of understanding the meaning of textual data. The lexical hypothesis is qualified to map the personality traits through words in a language [18].

In terms of model development, the needs of an open-source platform are essential. The advantage of creating the platform are model crowdsourcing, public corpus enrichment, correction, and verification. Hence, the platform is significantly enhancing the model's value over time. Nowadays, a platform with an ontology-based model for measuring human personality is still rare. In favor of getting a better model for measuring personality, we implement the model into a platform. This research aim is to show the way on how we develop a design in mapping human character by utilizing the ontology model. The model constructed by a collection of words that refer to personality in Bahasa. Our study mapped 2,331 instances in different facets and traits — those instances used as the corpus in our platform.

2. Literature Review

This chapter provides theoretical foundations related to the personality measurement proposed model based on the ontology approach. Literature review is sorted according to the data flow from social media data collection, personality definition, the personality traits, the ontology as the knowledge representation literature, and at last, is the radix tree algorithm used to parse and sort social media texts.

2.1. User-Generated Content

According to Naab and Sehl [19], user-generated content (UGC) has three criteria, there are 1). UGC is characterized by a rate of personal contribution, 2). UGC must be disclosed, and 3). UGC is built outside the sphere of occupation and professional routines. Besides, Wyrwoll stated that user-generated content is a content that is published online via various platforms by its user [20]. The users are not only the person but also the organizations. The substance classified as a UGC has similar characteristics, such as publicly accessible to other users, need a creative effort to create, and not a result of expertise routines and practices. UGC also defined in many forms, like blogs, posts, chat, podcasts, images, videos, tweets, and many other ways. In this modern era, UGC is a substantial information source for discovering knowledge from human digital activities [21]. Existing studies reveal that there is a high correlation between personality and personal inclination [22]. UGC indirectly shows beneficial information, like the user's demand, lifestyle, and personality. Hence, we utilize UGC in social media to measure human personality.

2.2. Personality

Personality is defined as a set of a person's characteristics, including acting, thinking, and feeling. Personality also correlates with emotions, values, attitudes, and talents [23, 24]. These attributes are establishing a unique persona of an individual and differentiate one person from another [25]. According to Stemmler, a person's personality closely related to language usage in speaking or writing [2]. Language is the most prevalent and reliable tool for people to convey their internal thoughts and emotions in giving comprehension to others [26, 27]. Hence, the linguistic usage of a person is an essential subject in the field of psychology and communication.

Previous research has examined human personality through linguistic usage. Howlader et al. predict Facebook user's personality through status and linguistic features by applying

regression models [28]. Boyd and Pennebaker propose a complementary model that provides big data solutions to measure human personality based on words people use [29]. Pietro et al. and Bogolyubova et al. [30, 31] who discovered the dark triad personality of social media users their online communication. Flekova and Gurevych [32] predict the personality of fictional characters in the novel using lexical-semantic features. Another research is Wei et al. [33], who predict human personality via information in digital traces and conversation logs.

2.3. Big Five personality Traits

Big Five Personality Traits is a model that identifies five characteristics of personality. It is also recognized as the OCEAN model, which stands for Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism (OCEAN) [34, 35]. This model is related to the lexical hypothesis and stated that language in daily interactions is a reflection of the most personality characteristics [36]. Thus, the Big Five Personality Traits predict and describe essential personality differences. In favor of getting a clearer explanation, Rossberger in [37] describes five traits below:

Table 1: Big Five Personality Traits

Big Five Personality Traits	Definition
Openness	or usually called openness to experience: extent to which individuals exhibit intellectual curiosity, self-awareness, and individualism/nonconformance.
Conscientiousness	extent to which individuals value planning, acquire the tenacity quality, and achievement oriented.
Extraversion	extent to which individuals involved with the external world, encounter enthusiasm and other positive emotions.
Agreeableness	extent to which individuals' value mutual effort and social harmony, modesty, dignity, and trustworthiness.
Neuroticism	extent to which individuals deal with negative feelings and their propensity to emotionally overreact.

According to Costa and McCrae's study [16], NEO-PI-R was advanced in samples of middle-age and older adults. The NEO-PI-R included scales to measure six conceptually derived facets in each OCEAN. The scales show generous internal consistency, temporal cohesion, convergent and discriminant validity against partners and peer ratings. Table. 2 shows the six facets which explaining each of the factors in personality traits.

2.4. Ontology and Knowledge Representation

Ontology is a collection of concepts which able to model terms of vocabulary into a domain knowledge [38]. From the perspective of computational science, ontology is explained as a concept to model the system structure. For example, the relevant entities and relationships that exist from observations are useful for specific purposes [39].

Table 2: NEO-PI-R Metric

Personality Traits	Facet	Description
Openness	Aesthetics	Artistic interests; believe in the value of art/do not love poetry
	Fantasy	Imagination; have an expressive imagination/seldom daydream
	Actions	Adventurousness; prefer variety /dislike changes
	Ideas	Intellect; admire convoluted problems/avert discussion
	Feelings	Emotionality; encounter emotions intensely/rarely get emotional
	Values	Liberalism; tend to elect for liberals/believe in one true divinity
Conscientiousness	Competence	Self-efficacy; accomplish tasks successfully/misunderstand the situation
	Order	Orderliness; prefer to order/abandon a mess
	Dutifulness	Follow the rules/break the rules
	Achievement-Striving	Work hard/just do the task
	Self-Discipline	Get task done right away/waste the time
	Deliberation	Cautiousness; avoid error carefully/charge into things directly
Extraversion	Warmth	Friendliness; make friends effortlessly/hard to make friends
	Gregariousness	Love large parties/prefer to be alone
	Assertiveness	Take action/wait for others to get the way
	Activity-Level	Regularly busy/love to take it easy
	Excitement-seeking	Love excitement/dislike loud music
	Positive Emotions	Cheerfulness; radiate joy/am seldom amused
Agreeableness	Trust	Trust others/distrust people
	Compliance	Morality; never deceive/use false praise
	Altruism	Make people feel pleasant/despise on others

Personality Traits	Facet	Description
	Straightforwardness	Cooperation; easy to fascinate/have a sharp tongue
	Modesty	Dislike being center of attention/think highly of myself
	Tender-mindedness	Sympathy; sympathize with the destitute/believe in people
Neuroticism	Anxiety	Worry about things/relaxed most of the time
	Angry-Hostility	Get enraged easily/rarely get annoyed
	Depression	Often feel miserable/feel satisfying
	Self-Consciousness	Easily frightened/not being intimidated easily
	Impulsiveness	Immoderation; often be affected by other/easily resist temptations
	Vulnerability	Panic easily/remain calm under pressure

Ontology associated with discovering and modeling reality under particular perspectives [40]. It focused on the structure and nature of an object [41]. Ontology also pointed out to a representational knowledge which indicates the type of class or entity associated with the relationship of the subtype [42].

Ontology generally has fundamental form components, i.e., class, instance, and relation [43]. Class is referring to a set of multiple instances, like words and phrases. An instance is a scope that considered in the ontology domain knowledge. A relation is defined as a relation among classes or instances [44]. This method is flexible, easy to modify, understood by humans and machines, and able to integrate with machine learning. Ontology has at least four evaluation methods [45]: 1). The Golden Standard; 2). The Application-based; 3). The Human Assessment; 4). The Data-Driven. Our study needs validation and evaluation for each instance in the model before implementing it on the platform.

Recent research has shown that ontology is able to represent knowledge for measuring personality. Some of the studies are that applying linguistic feature analysis and ontology model to measure human nature through social media data [46]. Another research computes human personality derivation in the modern physiognomy domain [47]. [48] also shows that ontology is a representative way to measure contextual knowledge, by building a complex domain of music and see how it relates to the domain of personality. A few research examples were clarifying how ontology might contribute to the analysis, comprehension, and research about human behavior and psychological research. Most of these existed researches are mainly focused in the English language, in contrast to our research which concerns in Bahasa.

2.5. Radix Tree

The dictionary which consists of strings data type is very-time consuming. A sorting algorithm is needed in order to get a more efficient process. Radix Tree is one of several algorithms that able to sort data in a database. Radix tree also beneficial for constructing associative arrays that expressed through the keyword in the form of strings. According to Mauro, radix tree is worked by labeling edges with a sequence of strings rather than characters, and constricting chains of nodes into a single element. Hence, radix tree is running more efficiently compared to a regular tree [49]. The instance of radix tree displayed in Fig. 2.

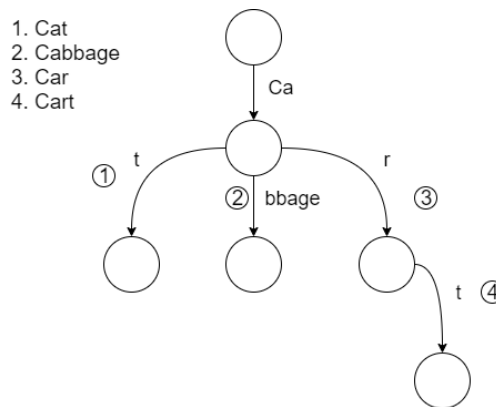


Figure 2: The Instance of Radix Tree Algorithm

3. Methodology

Our idea is to collect social media posts of people who considered influential in society. Their posts most likely to be responded by the public, which will generate conversations, and often set standard for informal language in Bahasa. In the model construction, we map words and phrases to a certain personality, which will validate by experts. There are several steps to construct the proposed model. Those steps are data collection, data preparation, model construction, model validation, platform construction, and conclusion. The research workflow is shown in Fig. 3. In contemplation of earning comprehensive understanding, we also illustrate the conception of our research methodology in Fig. 4.

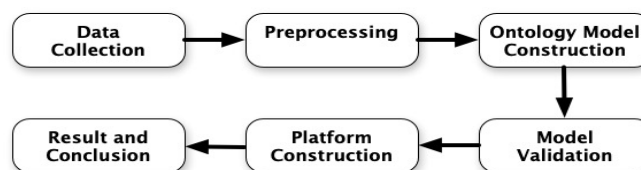


Figure 3: The Proposed Model Workflow

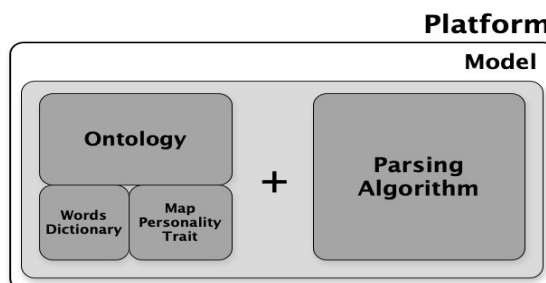


Figure 4: The Conception of Research Methodology

3.1. Data Collection

Our research utilizes real-world conversations on *Twitter* social media. A recent study shows that *Twitter* provides a valuable chance to study human behavior in a natural environment [1]. In the data retrieval process, we use three samples i.e., the famous user in *Twitter* social media with specific criteria. Our samples criteria are:

1. Verified accounts or having tweets with more than 1000 tweets or 500,000 followers.
2. Shows the latest activity with different tweets.
3. Shows many interactions with other accounts.
4. Not a protected account.

Based on those criteria, we have collected 13,047 tweets from three selected users. We comprehend that despite having the same principles in data collection, this collection of famous people who have different character tendencies, and this is the reason why we conduct this research.

3.2. Preprocessing

The data preprocessing is required to clear irrelevant data, such as URLs, symbols, and other terms, which is not beneficial for this research [50]. This process objective is to get substantial information over the data [51]. According to Khadim, preprocessing is a process of improving data quality while reducing barriers that will occur in the classification process [52]. The preprocessing is divided into several steps shown in Table. 3.

Table 3: Preprocessing Steps

Steps	Name of Preprocessing Steps	Definition
1	Case Folding	Transforming all capital letters to lowercase letters.
2	URLs Removal	Eliminating web page links
3	Symbol and Number Removal	Delete the symbol and number in the document
4	Tokenizing	Splitting the sentences into a token
5	Phrases Lookup	Check if there are any phrases in the token
6	Synonym Recognition	Determine the synonym of a word and replace it based on a dictionary
7.	Word Generalization	Replacing a word into a more general word in order to reduce data redundancy

3.3. Ontology Model Construction

This model is established by classifying words in each tweet into thirty classes in the NEO-PI-R metrics and afterward generalize it into Big Five Personality traits. Those collections of words are defined as a corpus or dictionary of our model. In this study, we state the personality class in Big Five Personality Theory as a class, facet in NEO-PI-R metric as a sub-class, and each composed word into an instance. For better understanding, we show our ontology domain model hierarchy in Fig. 5.

The hierarchy of our ontology model is settled in the bottom-up paradigm since this model starting from specific to general or instances to class [53]. There are no properties needed in personality measurement ontology model. As an instance, from Fig. 5., the words *terluka*, *melelahkan*, *penderitaan*, which means hurting, exhausting, and suffering are classified into vulnerability in Neuroticism class. These words arranged to vulnerability facets because it indicates the feelings of flimsy due to the risk of harm from some experiences.

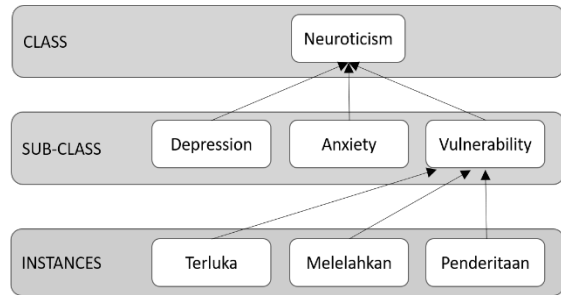


Figure 5: The Hierarchy of Personality Measurement Ontology Model

3.4. Model Validation

The validation process is essential to prevent and minimize any bias from the previous classification process. The human assessment approach is an assessment method with the help of domain experts [45]. For this reason, we assign psychologists to validate our classification result before deployment. In this process, the expert ensures that the classified word has entered the correct personality trait according to the Big Five theory and NEO-PI-R metric.

We conduct a human assessment because a human can represent language in terms of circumstances. The words in Bahasa ordinarily have various meanings depending on contextual purposes. For example, the words *bisa* can be represented to “able” or “can” or “poison from the snake.” Hence, we need the psychologist to check and validate the ambiguity of the words that frequently appeared in people's linguistic usage, especially in social media.

3.5. Platform Construction

Our platform is designed as an open-source platform to encourage the public or crowds to get involved in model development. People can openly provide some enrichment, correction, and verification of the model. Web-based platform is suitable for our research purposes due to its characteristics compared to other forms such as a mobile app or desktop app. The features of this platform are to measure personality based on the input. The common format of the collection of short text sentences and large-scale text is in comma-separated value (csv); thus, we use this .csv format file as one of our input files. At last, we visualize the result in a radar chart as the output. The platform’s workflow displayed in Fig. 6.

In this research, we utilize Personality Measurement Ontology (PMO) Platform which has been built in our prior research [51]. The platform is running by employing a radix tree algorithm as an effort to sort and parse the sentence, then find the semantic similarities between the input text and the corpus. The pseudocode presented in Fig. 7. The trait calculation is conducted

by finding similar keywords between processed textual data and the existing corpus in the ontology model. If matched, the data then considered representing one or more personality traits depending on the number of keywords that paired.

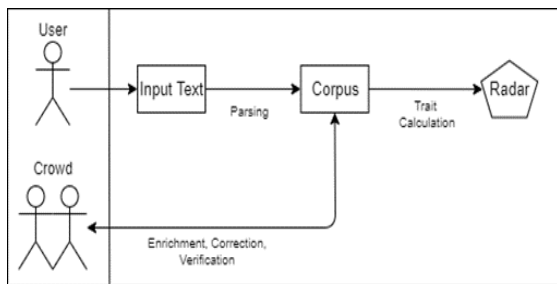


Figure 6: The Platform Workflow

The algorithm of our platform consists of several stages, i.e., convert the model to radix tree, saving database, input processing (sentence and csv format), and trait calculation. The following explanation of each stage are:

1. Algorithm 1: convert the model to radix tree. The established model is sorted using the radix tree algorithm. This algorithm composed of a) csv input file, which a text file contains of collections of keywords, b) a radix tree algorithm named 'tree'; c) an algorithm that is occupied by the data in csv called 'trie'.
2. Algorithm 2: saving database (keywords, traits, and facets). Generate a row for each keyword, facet, and traits into a 'database'.
3. Algorithm 3: input processing (sentence). Parsing sentence by utilizing 'trie'. The result of this process is called 'set_of_words,' which contains detected keywords.
4. Algorithm 4: input processing (csv). Parsing csv data into the database. The result is 'set_of_word' which loaded by detected keywords.
5. Algorithm 5: trait calculation. Calculate ('set_of_word') to get the desired form in the personality measurement results.

```

algorithm 1   convert model to radix tree
input        csv (keywords)
output       radix tree (keywords)
1             csv = input (csv file)
2             tree = new radix_tree
3             for each row in csv do:
4               trie <- row.keyword
5             output (trie)

algorithm 2   saving database (keyword, trait, facet)
input        csv = [keyword, trait, facet]
output       database(keyword, trait, facet)
1             csv = input (csv file)
2             data = [keyword, trait, facet]
3             for each row in csv do:
4               data.keyword = row.keyword
5               data.trait = row.trait
6               data.facet = row.facet
7             database <- import(data)
8             output (database)

algorithm 3   input processing (sentence)
input        sentence
output       result(set of words)
1             array = input(sentence)
2             trie <- load(database)
3             for each word in array do:
4               if word is in trie:
5                 set_of_words <- word
6             output(set of words)
    
```

```

algorithm 4   input processing(csv)
input        csv(set_of_sentence)
output       result(set of words)
1             csv = input(csv file)
2             load radix_tree
3             trie <- load (database)
4             for each row in csv do:
5               for each word in row do:
6                 if word is in trie:
7                   set_of_words <- word
8             output(set_of_words)

algorithm 5   trait calculation
input        set_of_words
output       result(trait composition)
1             array = input (set_of_words)
2             for each word in array do:
3               trait = trait(where facet =
4                 = word.facet)
5               trait += 1
6             output(trait)
    
```

Figure 7: The Pseudocode of PMO

To facilitate user's convenience, the platform provides a personality measurement process in two ways: 1) by entering textual data such as opinions, statements, or conversations; 2) by uploading the csv file, which consists of numerous textual data. The interface of our platform is shown in Fig. 8.



Figure 8: The PMO Platform Interface

4. Result and Analysis

In order to examine our platform, we measure the personality of the three samples through the platform. We calculate a set of sample tweets in the form of a csv file. The result visualized into a radar chart for getting a better comparison of each personality trait. The measurement results displayed in Table 4.

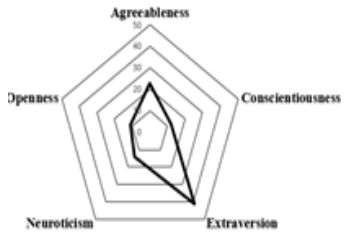
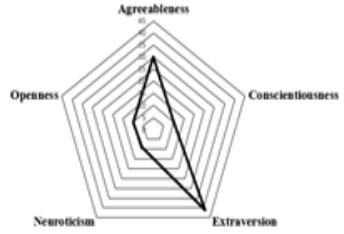

From our measurement result, each sample tends to have one trait as a dominant personality trait. It is a good result because human always has at least one dominant persona over all of the existed character. It also proves that our platform is able to measure all of the traits in the Big Five Personality theory.

As shown in Table 4., our samples have a similar result but not entirely identical. The result also indicates that our samples have a high score in Extraversion personality traits. It means that our samples often show positive feelings, friendliness, and activity-level on their social media activity. For example, @shillicious, who regularly displays his activities, both significant and trivial matters in his daily routine.

The second-highest score of our sample's measurement result is Agreeableness personality trait. This character is representing individuals who have value cooperation and social harmony with another person. @bepe20 is repeatedly showing his compliance and altruism. As a famous professional football player in Indonesia, @bepe20 has a tremendous number of fans. He

generally answers the message of his supporters on social media. He also does not hesitate to praise and welcome others who even he does not know. Thus, this persona makes @bepe20 also has a high score in Agreeableness besides the Extraversion personality trait.

Table 4: The Result of PMO Platform

Account	Result
@jojosuherman	
@bepe20	
@shitlicious	

Even though the platform has proven capacity in measuring personality through linguistic usage, it still has some limitations. Our platform cannot measure complex phrases with different contextual meanings. For example, the sentence *keren gila*, which means he or she expresses a fantastic feeling, is measured into Extraversion and Neuroticism personality traits. The words *keren*, which means impressive, is reflecting Extraversion and *gila*, which means crazy or stupid is representing Neuroticism. That instance shows our platform’s limitation, which only running by mapping word by word. Hence, our platform still needs enrichment and development for measuring sentences with phrases that consist of a different word with different personalities.

5. Conclusion

In this research, we explore human’s linguistic usage in social media to build an ontology model for measuring human character. We have successfully implemented the ontology model to a web-based platform for measuring personality automatically. In this study, the ontology form component defined by assigning personality as a class, facet as a sub-class, and words that are referring personality as instances. The platform is generally running well but cannot handle the complex phrases.

Our model helps us to measure human personality in social media. This research significantly contributed to a psychology study, especially in Indonesia. Our model is adequate to support a psychologist to speed-up the personality measurement process. It also can be improved for reading complex phrases by generating another algorithm. Thus, our suggestion for future research is to employ another parsing algorithm such as n-gram for receiving better results. We also suggest enriching the platform’s corpus by engaging the words that are reflecting personality from different social media. Another recommendation is adding weight to each classified word. Since we only measure the frequency of words, and not considering the context of the tweets, adding weight to the indexed words is required to detect context on a sentence.

References

- [1] A. Java, X. Song, T. Finin, and B. Tseng, "Why We Twitter: Understanding Microblogging Usage and Communities," 9th WEBKDD and 1st SNA-KDD Workshop, San Jose, California, USA, 2007.
- [2] G. Stemmler, and J. Wacker, "Personality, Emotions and Individual Differences in Physiological Responses," Journal of Biological Psychology, vol. 84, July 2010, pp. 541-551.
- [3] S. I. Wahjono, "Perilaku Organisasi," Graha Ilmu, 2010.
- [4] L. Li, A. Li, B. Hao, Z. Guan, and T. Zhu, "Predicting Active Users' Personality Based on Micro-Blogging Behaviors," PLoS ONE 9(1): e84997, January 2014.
- [5] J. L. Farr and N. T. Tippins, "Handbook of Employee Selection 2nd Edition," Taylor & Francis Group, 2017.
- [6] G. Park, H. A. Schwartz, J. C. Eichstaedt, M. L. Kern, M. Kosinski, D. J. Stillwell, L. H. Ungar, and M. E. P. Seligman, "Automatic Personality Assessment Through Social Media Language," Journal of Personality and Social Psychology, November 2014.
- [7] M. Madden, S. S. Fox, A. Smith, and J. Vitak, "Digital Footprints", Internet: <https://www.pewresearch.org/internet/2007/12/16/digital-footprints/>, 2007 [October 18, 2019].
- [8] W. Bleidorn, C. J. Hopwood, A. G. C. Wright, "Using Big Data to Advance Personality Theory," Current Opinion in Behavioral Sciences, vol. 18, December 2017, pp. 79-82.
- [9] S. Bhardwaj, P. K. Atrey, M. K. Saini, A. E. Saddik, "Personality Assessment using Multiple Online Social Networks," Multimedia Tools and Applications, vol: 75 November 2016, pp. 13237-13269.
- [10] B. Y. Pratama and R. Sarno, "Personality Classification Based on Twitter Text Using Naive Bayes, KNN Aand SVM," 2015 International Conference on Data and Software Engineering (ICoDSE), November 2015.
- [11] G. Farnadi, G. Sitaraman, S. Sushmita, F. Celli, M. Kosinski, D. Stillwell, S. Davalos, M. F. Moens, and M. D. Cock, "Computational Personality Recognition in Social Media," User Modeling and User-Adapted Interaction, vol:26, June 2016, pp. 109-142.
- [12] B. Liu, L. Yao, and D. Han, "Harnessing Ontology and Machine Learning for RSO Classification," Springer, 2016.
- [13] N.F. Noy, and D.L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," Stanford Knowledge Systems Laboratory, Technical Report KSL-01-05, 2001.
- [14] O. P. John, R. Robins, and L. Pervin, "Handbook of Personality Theory and Research," New York: The Guilford Press, 2008.
- [15] P. T. Costa and R. R. McCrae, "Normal Personality Assessment in Clinical Practice: The NEO Personality Inventory," Psychological Assessment, 4(1): 5-13, March 1992.
- [16] P.T. Costa Jr and R.R. McCrae, "Domains And Facets: Hierarchical Personality Assessments using The Revised Neo Personality Inventory," Journal of Personality Assessment, vol. 64, February 1995, pp. 21-50.
- [17] B. D. Raad and B. Mlacic, "The Lexical Foundation of The Big Five Factor Model,," in Oxford handbook of the Five Factor Model, US: Oxford University Press, 2017, pp. 191-216.
- [18] B. De Raad, "The Big Five Personality Factors: The Psycholexical Approach to Personality," Hogrefe & Huber Publishers, Göttingen, 2000.
- [19] T. K. Naab and A. Sehl, "Studies of User-Generated Content: A Systematic Review," Journalism 18(10), Sagepub, October 2016.
- [20] C. Wyrwoll, "Social Media Fundamentals, Models, and Ranking of User-Generated Content," Wiesbaden: Springer Fachmedien Wiesbaden, 2014

- [21] M. F. Moens, J. Li, and T. S. Chua, "Mining User Generated Content," Boca Raton: CRC Press, Taylor dan Francis Group, 2014.
- [22] R. Gao, B. Hao, S. Bai, L. Li, A. Li, and T. Zhu, "Improving User Profile with Personality Traits Predicted from Social Media Content," Proceedings of the 7th ACM conference on Recommender systems, October 2013.
- [23] H. Woo and H. J. Ahn, "Big Five Personality and Different Meanings of Happiness of Consumers," *Economics and Sociology*, vol 8(3), November 2015, pp.145-154.
- [24] L. R. Goldberg, "An Alternative "Description of Personality": The Big-Five Factor Structure," *Journal of Personality and Social Psychology*, 59(6) December 1990, pp. 1216–1229.
- [25] B. B. Lahey, "Psychology: An Introduction (Ed.11)", New York: McGraw-Hill, May 2011.
- [26] H. A. Schwartz, J. C. Eichstaedt, M. L. Kern, L. Dziurzynski, S. M. Ramones, M. Agrawal, A. Shah, M. Kosinski, D. Stillwell, M. E. P. Seligman, and L. H. Ungar, "Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach," *PLoS ONE* 8(9): e73791, September 2013.
- [27] Y. R. Tausczik and J. W. Pennebaker, "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods," *Journal of Language and Social Psychology*, vol 29(1), March 2010, pp. 24-54.
- [28] P. Howlader, K. K. Pal, A. Cuzzocrea, and S. M. Kumar, "Predicting Facebook-Users' Personality Based on Status and Linguistic Features via Flexible Regression Analysis Techniques," the 33rd Annual ACM Symposium, April 2018.
- [29] R. L. Boyd and J. W. Pennebaker, "Language-based personality: a new approach to personality in a digital world," *Current Opinion in Behavioral Sciences*, vol:18, December 2017, pp. 63-68.
- [30] D. P. Pietro, J. Carpenter, S. Giorgi, L. H. Ungar, "Studying the Dark Triad of Personality through Twitter Behavior," the 25th ACM International, October 2016.
- [31] O. Bogolyubova, P. Panicheva, R. Tikhonov, V. Ivanov, and Y. Ledovaya, "Dark Personalities on Facebook: Harmful Online Behaviors and Language," *Computers in Human Behavior*, vol: 78, January 2018, pp. 151-159.
- [32] L. Flekova and I. Gurevych, "Personality Profiling of Fictional Characters using Sense-Level Links between Lexical Resources," Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, January 2015.
- [33] H. Wei, F. Zhang, N. J. Yuan, C. Chao, H. Fu, X. Xie, Y. Rui, and W. Y. Ma, "Beyond the Words: Predicting User Personality from Heterogeneous Information," the Tenth ACM International Conference, February 2017.
- [34] R. A. Power and M. Pluess, "Heritability Estimates of The Big Five Personality Traits based on Common Genetic Variants," *Translational Psychiatry* 5(7): e604, July 2015.
- [35] M. Ziegler, K. T. Horstmann, and J. Ziegler, "Personality in Situations: Going Beyond the OCEAN and Introducing the Situation Five," *Psychological Assessment* 31(4), March 2019.
- [36] N. Ramdhani, "Adaptasi Bahasa dan Budaya dari Skala Kepribadian Big Five," *Jurnal Psikologi*, vol: 39, December 2012, pp. 189-207.
- [37] I. Ali, "Personality Traits, Individual, Innovativeness and Satisfaction with Life," *Journal of Innovative & Knowledge*, February 2018.
- [38] W. Wu, H. Li, H. Wang, and K. Q. Zhu, "Probase: A Probabilistic Taxonomy for Text Understanding," Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data, May 2012.
- [39] S. Staab and R. Studer, "Handbook on Ontologies 2nd Edition," Springer Publishing Company, 2009.
- [40] G. A. Silver, J. A. Miller, M. Hybinette, G. Baramidze, and W. S. York, "DeMO: An Ontology for Discrete-event Modeling and Simulation," *Simulation*, vol. 87(9), September 2011, pp. 747–773.
- [41] V. Devedzic, "Understanding Ontological Engineering," *Communications of the ACM - Supporting Community and Building Social Capital*, vol: 45, April 2002, pp. 136-144.
- [42] R. Arp, B. Smith, and A. D. Spear, "Building Ontologies with Basic Formal Ontology". The MIT Press, 2015.
- [43] C. Roussey, F. Pinet, M. A. Kang, O. Corcho, "An Introduction to Ontologies and Ontology Engineering," *Ontologies in Urban Development Projects*, Springer, pp. 9-38.
- [44] A. Alamsyah, M. R. D. Putra and D. D. Fadhilah, "Ontology Modelling Approach for Personality Measurement based on Social Media Activity," in International Conference on Information and Communication Technology, Bandung, 2018.
- [45] K. Sekandar, A Quality Measure for Automatic Ontology Evaluation and Improvement, Utrecht: Utrecht University, 2018.
- [46] D. Sewwandhi, K. Perera, S. Sandaruwan, O. Lakchani, A. Nugaliyadde, and S. Thelijagoda, "Linguistic Features based Personality Recognition Using Social Media Data," 6th NCTM, Malabe, Sri Lanka, January 2017.
- [47] A. K. Awaja, "Human Personality Derivation Using Ontology Based Modern Physiognomy," PhD and MSc Theses- Faculty of Information Technology, Islam University of Gaza, 2017.
- [48] C. M. Marques, J. V. Zuben, and I. Z. Guilherme, "FTMOntology: An Ontology to Fill the Semantic Gap between Music, Mood, Personality, and Human Physiology," *Model-Driven Approach to XML Schema Evolution*, October 2011, pp. 27-28.
- [49] C. Blochwitz, J. Wolff, J. M. Joseph, S. Werner, D. Heinrich, S. Groppe, and T. Pionteck, "Hardware-Accelerated Radix-Tree Based String Sorting for Big Data Applications," *Architecture of Computing Systems - ARCS 2017: 30th International Conference*, March 2017, pp. 47-58.
- [50] M.D.N. Arusada, N.A.S. Putri, and A. Alamsyah, "Training Data Optimization Strategy for Multiclass Text Classification," 5th International Conference on Information and Communication Technology, 2017.
- [51] A. Alamsyah, M. F. Rachman, C. S. Hudaya, and R. R. Putra, "A Progress on the Personality Measurement Model using Ontology based on Social Media Text," 2019 International Conference on Information Management and Technology (ICIMTech), August 2019.
- [52] A.I. Khadim, Y. N. Cheah, and N.H. Ahamed, "Text Document Preprocessing and Dimension Reduction Techniques for Text Document Clustering," 4th International Conference on Artificial Intelligence with Applications in Engineering and Technology, 2014.
- [53] A. Herdiani, N. Selviandro, and M.F.A. Azka, "Pemanfaatan Ontologi dengan Paradigma Pembangunan Combined Hierarchy Dalam Pengukuran Indeks Kebahagiaan Masyarakat Kota Bandung," *e-Proceeding of Engineering*, vol.3, no.2, 2016.

Experimental Studies of the Silicon Photomultiplier Readout Electronics Based on the Array Chip MH2XA030

Oleg Dvornikov¹, Vladimir Tchekhovski², Yaroslav Galkin³, Alexei Kunz³, Nikolay Prokopenko^{4,5,*}

¹Department No. 1, JSC "Minsk Research Instrument-Making Institute", MNIPI JSC, Kolasa St., 73, Minsk, 220113, Belarus

²Laboratory of electronics methods and tools of experiments, Institute for Nuclear Problems of Belarusian State University, Bobruyskaya St, 11, Minsk, 220030, Belarus

³Micro- and Nanoelectronics Department, Belarusian State University of Informatics and Radioelectronics, Brovka St., 6, Minsk, 220013, Belarus

⁴Information System and Radio Engineering Department, Don State Technical University, Gagarin Sq., 1, Rostov-on-Don, 344000, Russia

⁵Institute for Design Problems in Microelectronics of Russian Academy of Sciences, IPPM RAS, Sovetskaya St., 3, Zelenograd, 124681, Russia

ARTICLE INFO

Article history:

Received: 26 March, 2020

Accepted: 22 April, 2020

Online: 03 May, 2020

Keywords:

Silicon Photomultiplier

Readout Electronics

Array Chip

Charge-Sensitive Amplifier

Comparator

ABSTRACT

The experimental findings of the main units of readout electronics of silicon photomultipliers (SiPMs) based on array chip (AC) MH2XA030: a charge-sensitive amplifier (CSA) with an adjustable conversion factor and a base line restorer (BLR) circuit and two types of voltage comparators are considered. The electrical circuits of the units, the measurement results of static and dynamic parameters are given.

1. Introduction

Silicon photomultipliers are successfully used in a number of fields of science and technology for recording various types of electromagnetic radiation [1-10], because they have smaller dimensions, supply voltage and noise [11-13] in comparison with vacuum PM tubes. To process SiPM signals are usually used specialized analog integrated circuits (ICs), optimized according to the parameters of the signal source – internal capacitance, amplitude and duration of the current pulse.

SiPM readout electronics are relatively not required and their development costs too much. This explains the importance of creating specialized ICs on the ACs. So, we developed the main units of SiPM readout electronics at AC MH2XA030:

- the CSA with an adjustable conversion factor and a BLR circuit called ADPreamp13 [11];
- analog interface with a large dynamic range IBUF [14];
- different types of voltage comparators [15].

The purpose of this article is to consider the experimental findings of amplifier ADPreamp13 and comparators ADComp1, ADComp3 made on AC MH2XA030.

2. Circuitry Features of the Analog Units under Study

A detailed description of the operation of amplifier ADPreamp13 and comparators ADComp1, ADComp3 is given in [11,15]. When they were implemented at the AC, minimal corrections were made, caused by a slight change in the resistances of the resistors and the formation of diodes at the p-n junctions of the field effect transistor.

*Corresponding Author: Nikolay Prokopenko, Don State Technical University, Russia | Email: prokopenko@sssu.ru

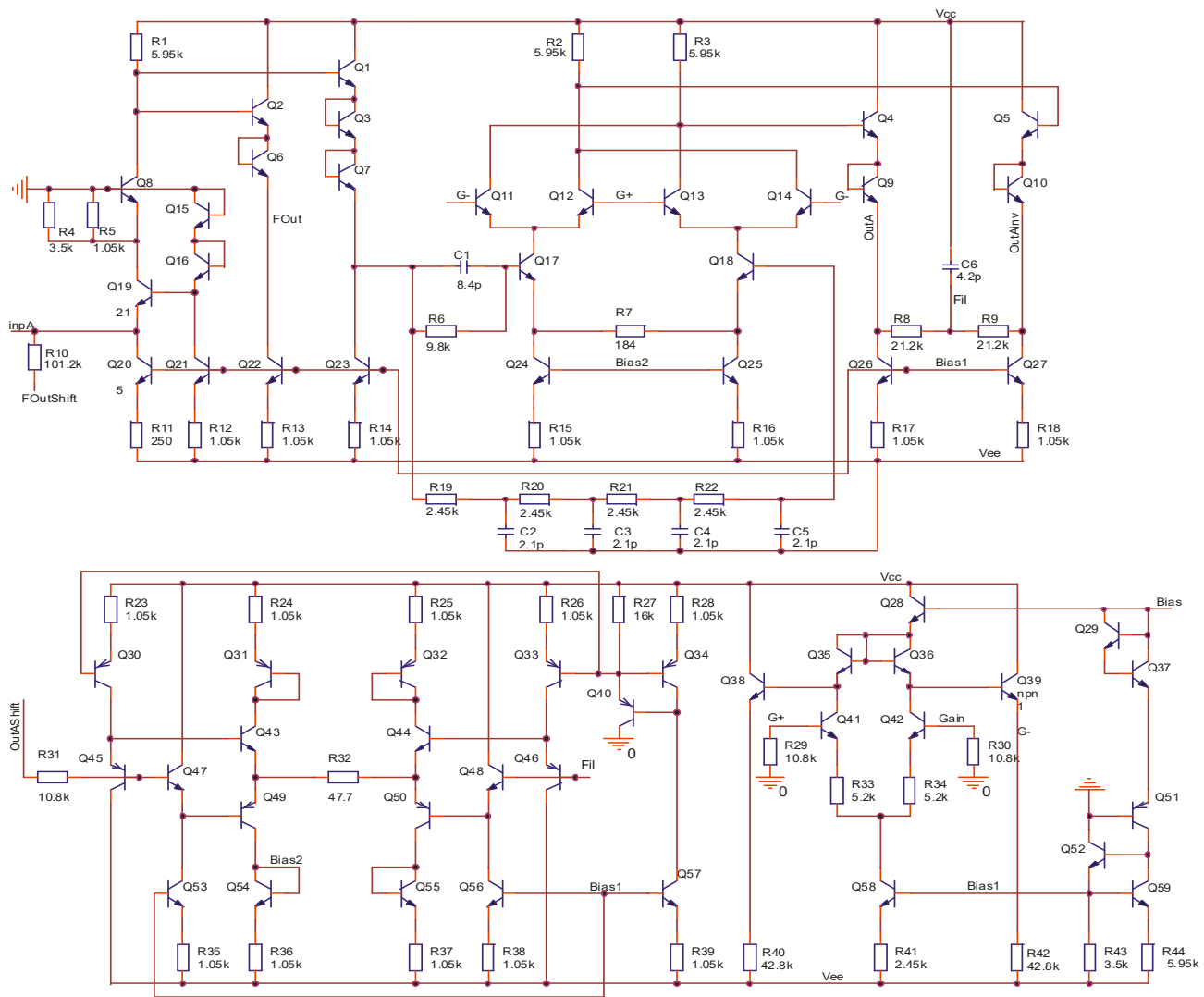


Figure 1: The electrical circuit of amplifier ADPreamp13

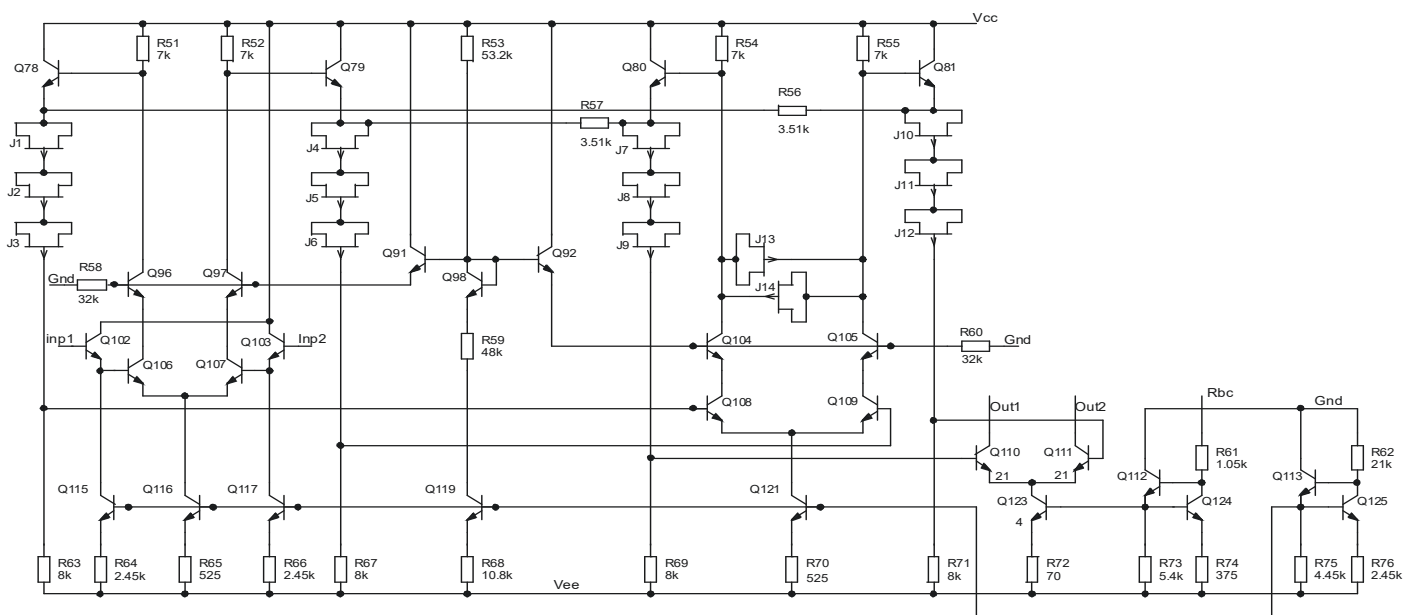


Figure 2: The electrical circuit of comparator ADComp1

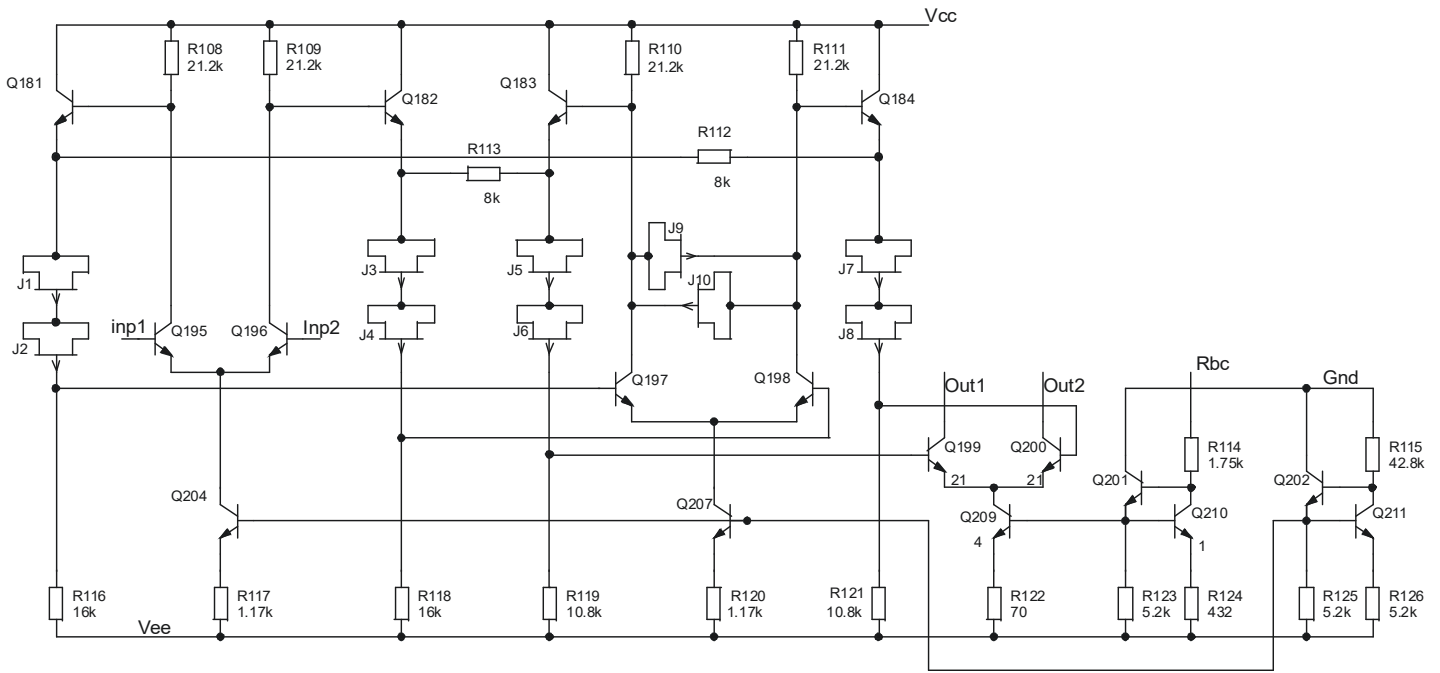


Figure 3: The electrical circuit of comparator ADComp3

Electrical circuits of ADPreamp3, ADComp1, ADComp3 are shown in Figures 1-3. Note that all nodes with the same name, for example, V_{CC} , V_{EE} , G-, G+, Bias1, Bias2, Fil in Figure 1 are interconnected, and, if necessary, the scaling factors are given in the diagrams, showing the number of the AC transistors connected in parallel. So, the emitter junction area of transistor Q19 in Figure 1 corresponds to the area of the emitter junctions of the 21st parallel-connected n-p-n-transistor of the AC.

The values of the resistances $R7 = 184$ Ohms, $R74 = 375$ Ohms, $R124 = 432$ Ohms are shown in Figure 1, Figure 2 and Figure 3 with great accuracy, but these are not necessary conditions. In the circuit, these resistor values are obtained with parallel and serial connection of the 1.05 kOhms and 2.45 kOhms resistors available on the AC.

It is worth noting that there is no need for special use of the p-JFET. These transistors are formed on the AC and are used only in diode connections.

Assignment of the nodes in the circuits of Figures 1-3 are the following.

For ADPreamp3: V_{CC} , V_{EE} – positive and negative supply voltage; InpA – amplifier input; FOut – output of the fast signal circuit (SC), OutA, OutAinv – direct and inverse output of the slow SC, FOutShift – voltage that sets the base line (dc output voltage in the absence of an input signal) by the output of the fast SC, OutAShift – voltage that sets the base line by the outputs of the slow SC, Gain – voltage that sets the value of the charge-voltage conversion factor K_{QV} by the outputs of the slow SC, Bias – node for connecting a current source that sets the operation mode of the amplifier.

For ADComp1, ADComp3: V_{CC} , V_{EE} - positive and negative supply voltage; Gnd - bus of null potential (“ground”); Inp1, Inp2 – inputs; Out1, Out2 – outputs; Rbc – node for connecting a voltage source that sets the maximum output current of the comparator, usually this node is connected to null potential.

Comparator ADComp3 is a simplified version of ADComp1 and is designed to reduce the AC area occupied by the comparator circuit. So, in total 8 units of ADComp1 or 16 units of ADComp3 can be manufactured at AC MH2XA030. On the manufactured experimental samples, comparator ADComp3 was made in two-channel design.

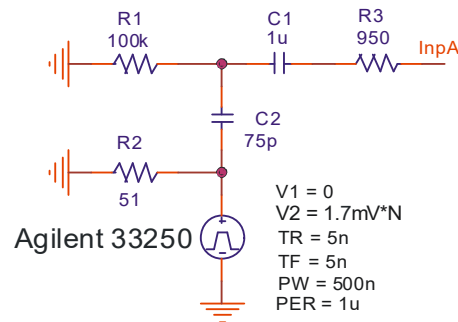


Figure 4: The simplified electric equivalent circuit of SiPM Photonique with 516 microcells, used in measurements

3. Measurement Results

Experimental samples of AC MH2XA030 with analog units were manufactured at JSC “Integral” (<http://www.integral.by/>). Measurements were performed using a set of equipment, including signal generators Agilent 33250 and Tektronix AFG3152C, a set of fixed attenuators, digital oscilloscope GDS 73354, dual-channel voltage supply source Agilent 3646A, data acquisition system 34970A with a set of modules, and digital multimeter Agilent 34410A. The signal came at the input of ADPreamp3 from the electric equivalent circuit SiPM Photonique with 516 microcells [11] (Figure 4), where TR, TF, PW, PER, V1, V2 are parameters of the rectangular voltage source used, namely, the rise and fall times, pulse duration, period, initial and final voltage values, correspondingly. By setting different voltages of V2, it is possible to continuously change the input charge of the amplifier in

accordance with the relation $Q_{INP}=C_2 \cdot V_2$, or when the voltage of V_2 is a multiple of 1.7 mV, which corresponds to one switched on cell, set the value of input charge N of the switched-on microcells of SiPM Photonique. The pulse duration of the generator is chosen equal to 500 ns in order to register the output pulses of ADPreamp13 of different polarity after the output signal reaches the base line.

In the comparator measurements, the same common-mode signal V_{CM} came simultaneously to both inputs, and a rectangular voltage pulse symmetrical with respect to the common-mode signal with a rise and fall front duration of 3 ns and a peak-to-peak value equal to V_{INP} was additionally received at one of the inputs. Both outputs of the comparators were connected through resistors with a resistance of 50.5 Ohms to the bus of null potential.

In total 20 units of amplifiers ADPreamp13, 10 units of dual-channel comparators ADComp3 and single-channel comparators ADComp1 were measured. Figures 5-14 and the table show the main measurement results, and the figures show the dependencies closest to the average.

Measurements made it possible to establish the following.

1) The spread of the base line by output FOut ranged from minus 24 mV to 276 mV with an average value of 85.6 mV. In this case, a voltage change in the node FOutShift from -3 V to 3 V is sufficient to establish a base line value of FOut output close to zero.

2) When the BLR circuit was switched off (the node OutAShift was not connected to any voltage), the spread of the base line at the output OutA was from 300 mV to 800 mV. When connecting the node OutAShift to the bus of null potential, the average base line value for the output OutA was 3.72 mV, and for the output OutAinv it was minus 2.42 mV.

3) The BLR circuit enables to smoothly change the base line at the outputs OutA, OutAinv in the range of ± 0.9 V.

4) When developing analog units, it was assumed that the output signal of amplifier ADPreamp13 would come to the input of the comparator with a small input capacitance. In connection with the above, the output stages of the amplifier are simple emitter followers on n-p-n-transistors and the voltage pulse shape at the outputs FOut, OutA, OutAinv depends on the input capacitance of the oscilloscope, which leads to different output pulses in the positive and negative half-waves. If necessary, the influence of the load capacitance on the shape of the output pulse can be reduced by using an emitter follower on complementary bipolar transistors, as in operational amplifiers previously developed on AC MH2XA030 [16]. Such an output stage increases the current consumption of the amplifier and therefore its use is carried out only in justified cases. To exclude the influence of the capacitance of the oscilloscope on the shape of the output pulses, a non-inverting voltage follower on chip AD8132 connected to the pins FOut, OutA, OutAinv was used in the measurements.

As can be seen from Figure 5, Figure 6, the inclusion of the follower provides almost the same shapes of the output pulses of both polarities.

5) At the maximum gain required to register the signal of 2 microcells of SiPM Photonique switched-on, the dynamic range of

ADPreamp13 exceeds 20 dB, however, in this case, the conversion factor K_{QV} depends on the input charge Q_{INP} (Figure 7, 8).

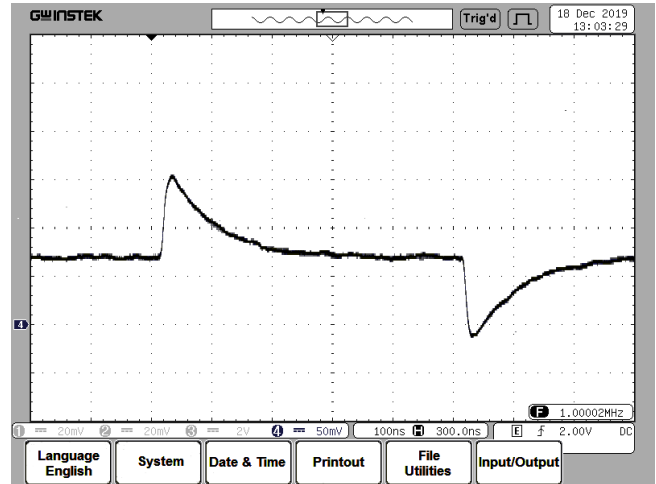
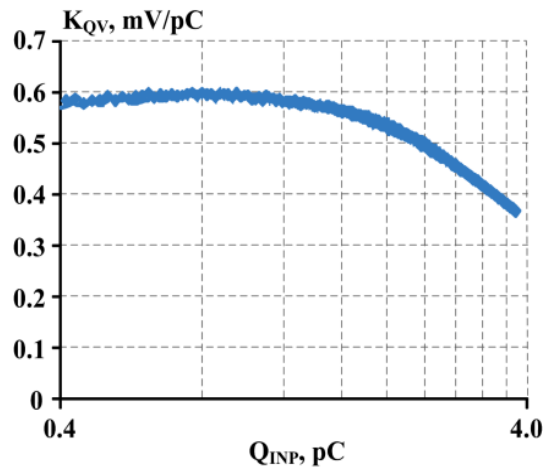
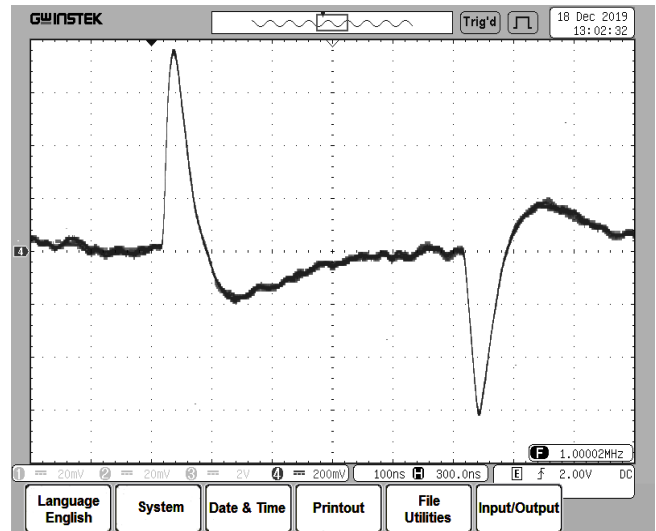


Figure 5: The voltage pulses at the output FOut for 10 switched-on microcells of SiPM Photonique when the voltage follower is used



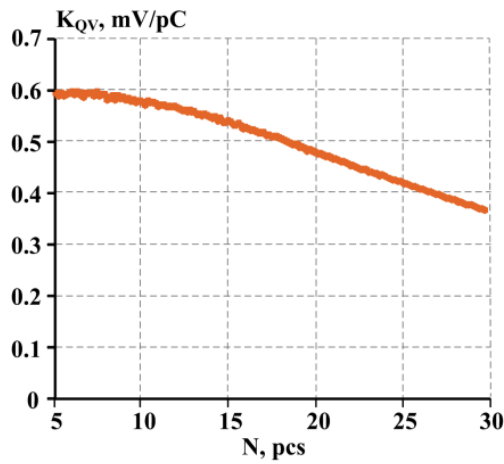


Figure 8: The dependence of the maximum value of the conversion factor K_{QV} at the pin OutA on the value of N switched-on microcells of SiPM Photonique

In the case when it is necessary to register the input charge from more than 30 switched-on microcells of SiPM Photonique, it is recommended to reduce the output pulse by reducing the voltage at the Gain pin (Figure 9) or use the signal from the pin FOut (Figure 10). To significantly increase the dynamic range, it is advisable to use new circuitry solutions, for example, considered in [14].

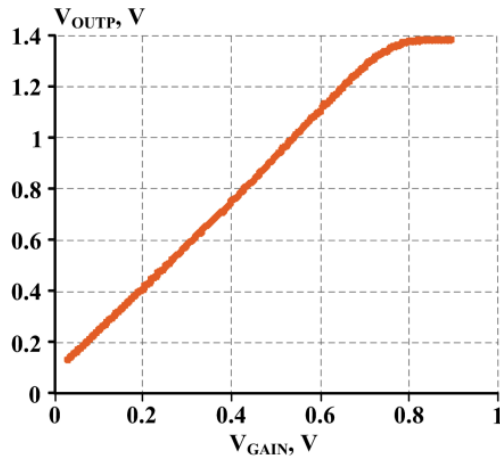
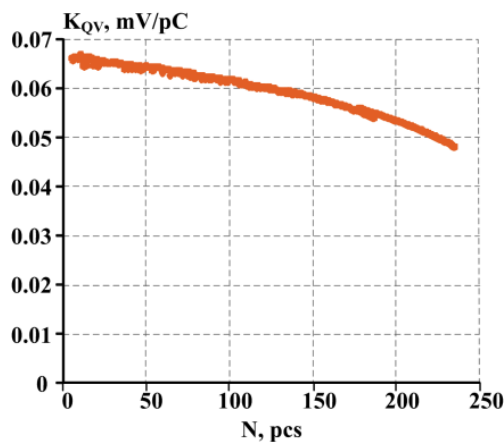


Figure 9: The dependence of the value of the pulse at the output OutA (V_{OUTP}) on the voltage at the pin of the Gain (V_{GAIN})



6) As follows from the table, the measured values of the static parameters of the comparators are close to the simulation results.

Parameter Name	Parameter value			
	ADComp1		ADComp3	
	measurements	simulation	measurements	simulation
Input current I_{INP} , μA	0.7-0.8	0.58	0.5-0.7	0.5
Maximum output current I_{OUT} , mA	3.2-3.4	4.03	2.8-3.0	4.0
Sensitivity, mV	2.0	2.0	2.0	2.0

The output current of the comparators is less according to the measurement results than in the simulation, as shown in Table 1. The dynamic parameters of the comparators characterize the dependencies shown in Figure 11 - Figure 14. It is possible to connect the Rbc node to an external voltage source by adjusting the output current. However, this requires an additional highly stable voltage source. We plan to reduce the resistance of the current-setting resistors with the closest adjustment of the interconnections of the elements of the AC. In this case, the output current of the comparators will increase even when the Rbc node is connected to zero voltage.

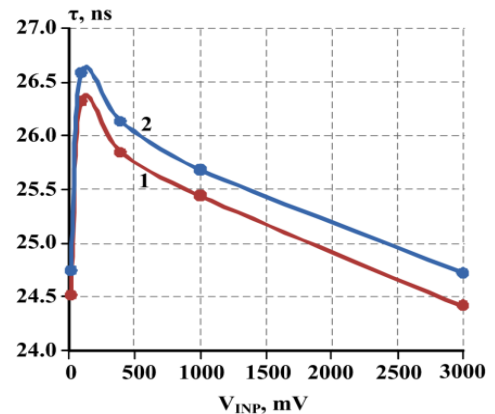


Figure 11: The dependence of the average signal propagation delay in two-channel comparator ADComp3 on the amplitude of the input signal for the first (1) and second (2) channels

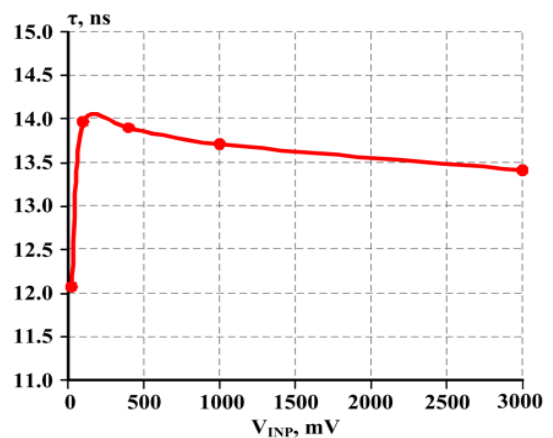


Figure 12: The dependence of the average signal propagation delay in comparator ADComp1 on the amplitude of the input signal

7) The measurements revealed a lower value of the maximum output current of the comparators and a significant excess of the measured values of the peak time of ADPreamp13 and the signal propagation delay of the comparators over the values obtained during the simulation.

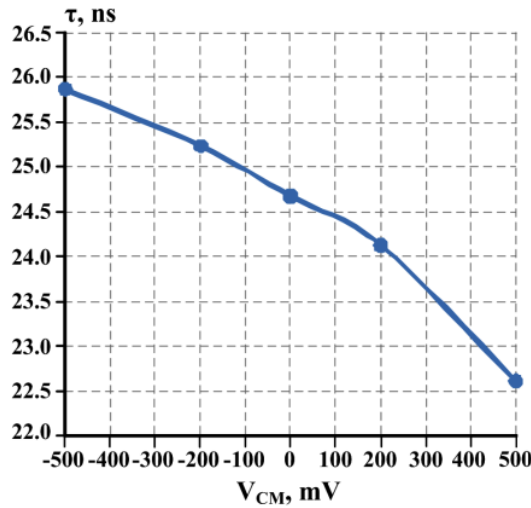


Figure 13: The dependence of the average signal propagation delay in two-channel comparator ADComp3 on the value of the common-mode signal at $V_{INP} = 20$ mV

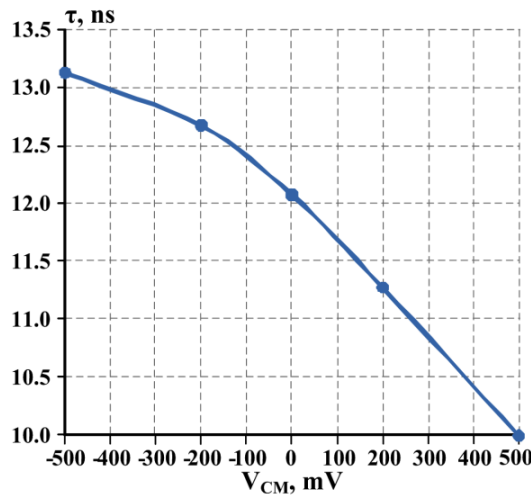


Figure 14: The dependence of the average signal propagation delay in comparator ADComp1 on the value of the common mode signal at $V_{INP} = 20$ mV

From our point of view, the insufficient value of the maximum output current of the comparators is due to the difference in the current-setting resistors of the output stages (R72 in Figure 2, R122 in Figure 3) from the required value and can be eliminated when adjusting the interconnects of the AC elements.

The Spice parameters CJE, CJC, CJS of the AC MH2XA030 transistors used in the simulation were obtained from the identified parameters for transistors with a large layout area by decreasing in proportion to the area of the emitter junction. For the transistors of AC MH2XA030 made according to the minimum design rules, such scaling of the Spice parameters turned out to be incorrect and identification of the Spice parameters CJE, CJC, CJS by the results of measurements of experimental samples is required.

4. Conclusion

It was experimentally established that the created base line restorer circuit reduces the spread of the base line of amplifier ADPreamp13 by almost 100 times and enables to smoothly change the base line at the outputs OutA, OutAinv in the range of ± 0.9 V.

At maximum gain, the dynamic range of ADPreamp13 exceeds 20 dB, however, in this case, the conversion factor depends on the value of the input charge. To register large input charges, it is recommended to reduce the output pulse by reducing the voltage at the Gain pin or process the signal from the pin FOut.

The static and dynamic parameters of created comparators ADComp1, ADComp3 satisfy the requirements for readout electronics of silicon photomultipliers.

The measurements revealed a mismatch between the results of simulation and measurements of the maximum output current of the comparators, the peak time of the amplifier and the signal propagation delays of the comparators, therefore, in the near future, it is planned to adjust the interconnects of the AC elements to increase the maximum output current and refine the Spice parameters of transistors according to the measurement results of experimental samples.

The developed units can be used to create multichannel signal processing circuits of SiPM based on AC MH2XA030.

Conflict of Interest

The authors declare that there is no conflict of interests regarding publication of this paper.

Acknowledgment

The study has been carried out at the expense of the grant from the Russian Science Foundation (Project No. 16-19-00122-P).

References

- [1] G. Paternoster, L. Ferrario, F. Acerbi, A. G. Gola and P. Bellutti, "Silicon Photomultipliers Technology at Fondazione Bruno Kessler and 3D Integration Perspectives," ESSDERC 2019 - 49th European Solid-State Device Research Conference (ESSDERC), Cracow, Poland, 2019, pp. 50-53. DOI: 10.1109/ESSDERC.2019.8901738
- [2] W. Jiang, Y. Chalich, M. J. Deen, "Sensors for Positron Emission Tomography Applications," J. Sensors 2019, 19, 5019; DOI: 10.3390/s19225019.
- [3] R.R. Raylman, A. Stolin, S. Majewski, and J. Proffitt, "A large area, silicon photomultiplier-based PET detector module," Nucl. Instrum. and Methods A, vol. 735, pp. 602-609, 2014. DOI: 10.1016/j.nima.2013.10.008
- [4] A.L. Goertzen, et al., "Design and performance of a resistor multiplexing readout circuit for a SiPM detector," IEEE Trans. Nucl. Sci., vol. 60, no 3, part 1, pp. 1541-1549, 2013. DOI: 10.1109/TNS.2013.2251661
- [5] S. Dey, E. Myers, T.K. Lewellen, R.S. Miyaoka, and J.C. Rudell, "A row-column summing readout architecture for SiPM based PET imaging systems," IEEE Nucl. Sci. Symp. and Med. Imag. Conf. (NSS/MIC), pp. 1-5, Oct. 27 2013-Nov. 2 2013. DOI: 10.1109/NSSMIC.2013.6829062
- [6] S. Gundacker, et al., "A systematic study to optimize SiPM photodetectors for highest time resolution in PET," IEEE Trans. on Nucl. Sci., vol. 59, no 5, part 1, pp. 1798-1804, 2012. DOI: 10.1109/TNS.2012.2202918
- [7] A. Del Guerra, et al., "Silicon Photomultipliers (SiPM) as novel photodetectors for PET," Nucl. Instrum. and Methods A, vol. 648, pp. S232 - S235, 2011. DOI: 10.1016/j.nima.2010.11.128
- [8] Xiaoli Li, C. Lockhart, T.K. Lewellen, and R.S. Miyaoka, "Study of PET detector performance with varying SiPM parameters and readout schemes," IEEE Trans. on Nucl. Sci., vol. 58, no 3, part 1, pp. 590-596, 2011. DOI: 10.1109/TNS.2011.2119378

- [9] B. Seitz, A. G. Stewart, K. O'Neill, L. Wall, and C. Jackson, "Performance Evaluation of Novel SiPM for Medical Imaging Applications," *J. Photonics*, 2014, no. 1 (43), Pp.104-113. DOI: 10.1109/NSSMIC.2013.6829685
- [10] H. Sabet, et al., "High-performance and cost-effective detector using microcolumnar CsI:Tl and SiPM," *IEEE Trans. on Nucl. Sci.*, vol. 59, no 5, part 1, pp. 1841-1849, 2012. DOI: 10.1109/TNS.2012.2202248
- [11] O.V. Dvornikov, V. A. Tchekhovski, N. N. Prokopenko, Ya. D. Galkin, A. V. Kunts, A. V. Bugakova, "Implementation of Reading Electronics of Silicone Photomultiplier Tubes on the Array Chip MH2XA030," *Visnyk NTUUKPI Seriya-Radiotekhnika Radioapara to buduvannia*, 2019, Iss.78, pp.60-66. DOI: 10.20535/RADAP.2019.78.60-66.
- [12] O. V. Dvornikov, V. A. Tchekhovski, V. L. Diatlov, "Equipments to single photon registration. Part 1. Features and possibilities of multi-channel photodetectors with intrinsic amplification," *Devices and methods of measurements : Scientific and Engineering J.*, 2012, no. 2(5), Pp. 5-13. (In Russian)
- [13] O. V. Dvornikov, V. A. Tchekhovski, V. L. Diatlov, "Equipments to single photon registration. Part 2. Silicon photomultiplier signal preliminary processing. (Review) ," *Devices and methods of measurements : Scientific and Engineering J.*, 2013, no. 1 (6), Pp. 5-13. (In Russian)
- [14] O. V. Dvornikov, N. N. Prokopenko, V. A. Tchekhovski, Y. D. Galkin, A. E. Titov and A. V. Bugakova, "Silicon Photomultipliers' Analog Interface with Wide Dynamic Range," *Proceedings of 17th IEEE East-West Design & Test Symposium (EWDTS-2019)*, September 13-16, 2019, Batumi, Georgia, pp. 270-273. DOI: 10.1109/EWDTS.2019.8884430.
- [15] O.V. Dvornikov, N. N. Prokopenko, V. A. Tchekhovski , Ya. D. Galkin, A. V. Kunz, A. E. Titov, "Radiation-Hardened Voltage Comparators for the Synthesis of Microcircuits on the Array Chip MH2XA030," *Proceedings of IEEE International Conference on Computation, Automation and Knowledge Management (ICCAKM - 2020)*, 9-11 January 2020, Dubai, Pp. 1-4.
- [16] O. V. Dvornikov, V. L. Dzatlau, V. A. Tchekhovski, N. N. Prokopenko and A. V. Bugakova, "BiJFet Array Chip MH2XA030 — a Design Tool for Radiation-Hardened and Cryogenic Analog Integrated Circuits," *2018 IEEE International Conference on Electrical Engineering and Photonics (EExPolytech)*, St. Petersburg, 2018, pp. 13-17. DOI: 10.1109/EExPolytech.2018.8564415

University Students Result Analysis and Prediction System by Decision Tree Algorithm

Md. Imdadul Hoque¹, Abul kalam Azad^{*1}, Mohammad Abu Hurayra Tuhin¹, Zayed Us Salehin²

¹Computer Science and Telecommunication Engineering, Noakhali Science and Technology University, Noakhali-3814, Bangladesh

²Information and Communication Engineering, Noakhali Science and Technology University, Noakhali-3814, Bangladesh

ARTICLE INFO

Article history:

Received: 13 January, 2020

Accepted: 24 March, 2020

Online: 03 May, 2020

Keywords:

Data mining

Decision tree

J48 algorithm

WEKA

Predicting Performance

ABSTRACT

The main assets of universities are students. The performance of students plays a vital role in producing excellent graduate students who will be the future viable leader and manpower in charge of a country's financial and societal progress. The purpose of this research is to develop a "University Students Result Analysis and Prediction System" that can help the students to predict their results and to identify their lacking so that they can put concentration to overcome these lacking and get better outcomes in the upcoming semesters. The prediction system can help not only the current students but also the upcoming students to find out exactly what they should do so that students can avoid poor achievement that will help to increase their academic results and other skills. To train the system, we collected data from the university student's database and directly from students by survey using Google form containing information, such as gender, extracurricular activities, no of tuition, programming skills, class test mark, assignment mark, attendance, and previous semester Grade Point Average (GPA), where the main aim is to relate to student performances and Cumulative GPA (CGPA). We use Weka tools to train the system and to develop the decision tree. In decision tree, the acquired knowledge can be expressed in a readable form and produced classification rules that are easy to understand than other classification techniques. These rules used to develop a web-based system that can predict the grade points of students from their previous records. Moreover, the system notifies students' lack and gives suggestions to improve their results. Finally, we compared the performance of three (J48, REPTree, and Hoeffding Tree) different decision tree algorithms, and comparative analysis shows that for our system, the J48 algorithm achieves the highest accuracy.

1. Introduction

University students' academic acquirement is one of the main factors thought by employers in recruiting employees especially new graduates. So, students have to engage in their lessons to obtain good results for fulfilling the employer's demand [1]. In the twenty-first century, it may not be a new message that the importance of science and technological education is growing across the whole world. Newer and newer inventions are happening day by day. So, persons with science and technological knowledge and skills are highly demandable for our country's economic and social development.

In Bangladesh, advancement in technology is also remarkable. Science and technology education are getting a new dimension. A

*Corresponding Author: Abul Kalam Azad, Department of CSTE, NSTU, Noakhali-3814, Bangladesh, Email: ak_azad@nstu.edu.bd

large number of students are interested to admit in different science and engineering universities. It is a very good sign, but we need to ensure the quality of the graduates that we cannot ensure in many cases. So, we need to find out the factors with which student's academic performance depends and addressing them in a paradigm so that academic performance can increase.

Although some academic performance prediction approaches proposed for various countries, in Bangladesh, enough works have not done yet in this area. We find only works in this area. The authors of [2] have introduced an approach to predict the student's results in the form of CGPA using neural networks technique where they collected data from a university of Bangladesh. They showed that student's yearly performance greatly depends on both academic and non-academic activities. However, they work on a comparatively small dataset and do not provide suggestions for improvement. So, this paper aims to the

development of an effective university student result analysis and prediction system that can predict semester results, will show their lacking, and give valuable suggestions to overcome the lacking and hence to improve their performance.

The remainder of the paper formed by covering the following: Section 2 briefly explain the related research works, Section 3 shows the major contributions of the work, Section 4 describes the proposed methodologies of this research, Section 5 covers the implementation and output of this research and Section 6 concludes the paper with the outline of future works.

2. Literature Review

For the last few years, researchers are working to address the issue of student result analysis and prediction. In [3] a decision tree-based classification technique to predict students' final exam results has presented. The authors stated that educational databases' hidden information could play a vital role in students' performance development.

Surjeet Kumar Yadav and Saurab Pal [4] proposed a data mining approach to predict good students to enroll in the Master of Computer Application (MCA) course in India using their past academic records. They conclude that Bachelor of Computer Application (BCA) and B.Sc. students with mathematics performed better in the MCA course, and B.A. without mathematics did not perform well for the course.

Cristobal Romero et. al [5] collected real data of seven Moodle courses from Cordoba University students to develop a specific Moodle data mining tool. The authors compared different data mining techniques to classify students based on their Moodle usage data and the final marks obtained in their respective courses. The authors concluded that a classifier for educational decision making should be both comprehensive and accurate.

Lewis Adam Whitley's [6] research to predict the most affluent learning environment. Within this research project, the author would attempt to use data from the University of North Carolina at Pembroke and process the data into environmental factors that may or may not influence a student's learning ability. The author determined the best method in order to seek a learning environment and try to discover the factors that could impact on a student's academic performance.

Authors of [7] analyzed a system that will predict student's grades using the ID3 decision tree algorithm, where data gathered from the academic department of Redeemer's University, Nigeria.

Gorikhan [8] presents the implementation of different classification techniques for vocational institutional analysis that help teachers to work on weak students to improve their performance and claimed that decision tree is the accurate prediction model for institutions students' analysis.

Bhardwaj and Pal [9] conducted another study on predicting the students' performance by choosing 300 students from 5-degree college conducting BCA (Bachelor of Computer Application) course in Dr. R. M. L. Awadh University, Faizabad, India. Using the Bayesian classification technique on 17 attributes, they showed that students' academic performance relates to both the academic and non-academic attributes like family annual income and students' family status, etc. This study would help the

students to improve their performance. This study would also help to identify those students who needed special attention and by taking appropriate action at the right time the fail ratio could also be reduced. However, specific suggestions to overcome the lacking were absent in the study. So, more works need to be done to find more effective solutions for university students' result analysis and prediction systems, especially in Bangladesh.

3. Major Contribution of the Work

This paper differs from previously proposed result prediction approaches. The major contribution of the work is as follow:

- The attributes considered for this paper consist of a mixture of students' general and academic data. Moreover, the work aims to identify the important attributes for predicting the results.
- The predicted results were verified to make sure its accuracy.
- In this work, we designed a web-based environment where a student can input the required (e.g. gender, programming skill, CT marks, etc) data, and the system shows them the predicted results based on their entry data.
- The developed system will provide the reasons behind the results fall, and also provide valuable suggestions to improve their performance.
- Finally, the paper incorporates the comparison among some prediction models to find out the best one with high accuracy.

4. Proposed Study of Students Result Analysis and Prediction System

The proposed methodology used in this work for predicting students' performance using decision tree algorithms belongs to the data mining technique and make the prediction model that classifies the students' records.

Said differently, using this decision tree algorithm, we wish to guide the students towards the acquirement of better results that we feel they could achieve. To classify the instances, Tree-based methods classify instances from the root node to leaf nodes where each branch downward from a node represents one of the possible values for that attribute [10]. The stages in the process include the following:

4.1. Dataset Collection

The dataset collected from a university that has introduced different attributes. The data collected in two steps. The first step is to collect a partial part of the data from the university student database. Then the rest of the data was taken from the students through the Google doc survey. This dataset contains 850 instances, each of which has 14 different attributes.

4.2. Data Preprocessing

To apply data mining techniques, the dataset needs to be prepared. This stage involves dataset preparation before applying data mining techniques. At first, data collected in an excel sheet. Then, fill the missing data by standard values, and removed another

variance manually to hold classifier quality. The original data includes student details such as (id, name, age, date of birth, address, TGPA, gender, extracurricular activities, no of tuition, CT mark, assignment mark, programming skills, attendance, no of backlog). During analysis, attributes like student Id, age, date of birth, address, no of backlog did not help to predict the class. So it excluded from the training data set. To rank the attributes, we use gain ration measures. Then we select 8 attributes based on their rank to use for the study. Here, Table 1 shows the description of the attributes and their possible values.

Table 1: Used attributes in the experiment

Attribute	Description	Possible Values
Gender	Students Gender	Male / Female
E. activities	Extra curriculum activities	Yes / No
No of Tuition	Number of tuition	Zero, One, Two, Three, Greater Than Three.
Programming Skills	Programming skills	Poor, Average, Good.
CT Mark	Class test marks	1 st class, 2 nd class, 3 rd class.
Assignment Mark	Assignment marks	First-class, Second class, Third class.
Attendance	Percentage of attendance	60%-69%, 70%-79%, 80%,-90%, 90%-100%.
TGPA	Term grade point average	POOR (less than 3.10), AVERAGE (3.10-3.29), GOOD(3.30-3.49), VERY GOOD (3.50-3.69), EXCELLENT (upto 3.69).

The data in the excel sheet exists in CSV (.csv) format. Then, we convert it to ARFF(.arff) format by using the WEKA tool.

4.3. The Data Mining Tools

In this work, WEKA (Waikato Environment for Knowledge Analysis) used as an experimental tool for classifying data. Weka is one of the popular machines learning software that contains a variety of algorithms and visualization tools for data analysis and prediction system design. Moreover, Weka offers a user-friendly graphical user interface to access its functionality [11].

Weka facilitates a variety of data mining tasks like data preprocessing, clustering, classification, regression, with an assumption that all data is available as a single file or relation where each data should contain a fixed number of attributes. However, Weka supports other types of attributes as well.

4.4. Technologies Used

Some other technologies we used for the system design and implementation are HTML, CSS, PHP, and Laravel framework, etc.

4.4.1 HTML and CSS

HTML (Hypertext Markup Language) is a type of language used to develop web pages and applications to access over the Internet. It allows us to include objects and images to form interactive documents using several structural semantics.

CSS (Cascading Style Sheets) is used to enable the separation between the HTML or other makeup language contents and its presentation. It used to ameliorate the accessibility of the contents, improving the flexibility and control the contents' specification and presentation characteristics. CSS also facilitates multiple pages to share the formatting to scale down the redundancy [12].

4.4.2 PHP and the Laravel Framework

PHP (Hypertext Preprocessor) is a server-side scripting language that can embed with HTML code. It is an open-source language widely used for web development [12].

Laravel is one of the popular PHP frameworks generally used to build small to big size web applications. It follows the MVC (Model View Controller) structure which gets in easy to learn and rapid development of web application prototypes. It provides many built-in features like routing, mail, sessions, and authentication, etc.

4.5. Design of the Prediction System

The system comprises a knowledge basis that has compiled experience and a set of rules for utilizing the knowledge base for each case. It is the knowledge-based system where knowledge extracted in the form of IF-THEN rules from the decision tree that uses different attributes, including the semester examination records denoted by TGPA. Through the user interface, the system provides advice and instructions to the user to improve the students' performance. The prediction system architecture [7] given in Figure 1.

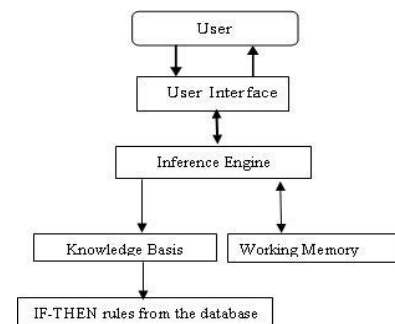


Figure 1: System flowchart of the prediction system

Knowledge Base: It's the knowledge domain in the form of IF-THEN rules that will be returned from the WEKA tools.

Working Memory: The user enters the required information into the working memory. The prediction system compares this information with the knowledge carried in the knowledge base to

deduce the new facts. The matching procedure continues by inserting these new facts into memory. Finally, the system reached some completion and enters into the working process.

Inference Engine: The inference engine interacts with the information in the knowledge base and the working memory. It tried to find out the rules to make a match between its premises and the information there have in the working memory. After finding a match, it reached a decision in memory and looks for a new match.

End-User: End users are those individuals who intend to consult with the system to learn their predicted outcome and to get advice from the system that they could follow to improve their performance.

5. Implementation and Result

This section describes the detailed process of the J48 decision tree algorithm in machine learning software WEKA for the implementation of university students' result analysis and prediction system.

5.1. Selected WEKA Algorithm

In a decision tree algorithm, one of the main complications is the optimal size of the final tree. All theoreticians and specialists are still now searching for techniques to make this algorithm more accurate, cost-effective, and efficient. A tree if it is too big, then there is a risk to overfit the training data and poor generalization to a new sample. If a tree is too small, then it might not get significant structural data about the sample space. Sometimes, an error could drastically reduce after adding one more extra node. This phenomenon named the horizon effect. A common technique is to use tree pruning. Here, the concept is to eliminate the nodes that do not have any additional information. It reduced the overfitting and can remove a section of a classifier that may base on noisy or erroneous data to reduce the classifier complexity and improve the prediction accuracy. Tree pruning could do using the following approaches.

- Pre-pruning
- Post-pruning

In the pre-pruning approach, the tree construction halts early by deciding not to partition the instance in a node that then becomes a leaf. In the pre-pruning approach, the choice of threshold to halt the tree construction is very crucial because a low threshold could result in slight simplification and a high threshold could result in oversimplification [13].

On the contrary, post-pruning is a popular and reliable tree pruning technique that eliminates subtree from a fully developed tree. In this case, the removed branches replace with a leaf marked with the most regular class among the subtrees that have to replace [14]. However, this approach requires more computation than the pre-pruning approach.

In our decision tree algorithm, we use the J48 algorithm to classify our instances.

About J48 Pruned Tree:

J48 algorithm follows the divide and conquers strategy to create a small decision tree. It is a prolongation of the ID3

algorithm. In the J48 algorithm, each node of the tree effectively selects the best attribute of the data that can divide its samples into subsets prolonged in one class or the others. The dividing measure is the normalized information gain (variety in entropy). The decision takes based on the attribute with the highest normalized information gain. Eventually, the J48 algorithm recurs on the smaller sub-lists.

A. Construction

The following steps have to follow to construct the tree:

1. Make sure whether every case fits in the same class.
2. Then, calculate the information and information gain for each attribute.
3. Finally, have to discover the best splitting attribute (based on the present selection principle) [15].

B. Calculating Information Gain

Shannon's theory is at the base of the J48 algorithm. Shannon entropy is the best known and most popular technique applied to calculate the information gain. It could define in this way, the measure of information supplied by a consequence, the higher the chance of happening the occurrence is low (it is rare), the more information it supplies [16]. In the following, all logarithms are base 2.

Shannon Entropy

For a sample S, if the given probability distribution $P = (p_1, p_2, \dots, p_n)$, then the Information carried by this distribution, also called the entropy of P is given by:

$$Entropy(p) = -\sum_{i=1}^n P_i \times \log(P_i)$$

The information gain G (p,T)

The information gain function allows us to assess the degree of mixing of classes for all samples to measure the attributes position during tree construction. It allows defining a function to choose the test that must label the current node. For a test **T** and a position **p**, the gain equation is as follows:

$$Gain(p,T) = Entropie(p) - \sum_{j=1}^n (p_j \times Entropie(p_j))$$

Where, the value of (p_j) is the set of all possible values for attribute T. To rank the attributes, this measure is necessary and hence to build the decision tree. In this process, the attributes with the highest information gain placed top positions in the path from the root of the tree [17].

Based on the highest information gain, splitting should do to form a small and efficient tree. Suppose, there are 11 male(m) and 7 female(f) in a class instance. Based on their calculated entropy and information gain, it could be divided further into two different groups of instances. So 4m and 5f as left instance and 7m and 2f as the right instance. After inserting the values to calculate the entropy and information gain, the formula will be as follows [18]:

$$Entropy_before = -7/18 * \log(7/18) - 11/18 * \log(11/18)$$

$$Entropy_left = -4/9 * \log(4/9) - 5/9 * \log(5/9)$$

$$Entropy_right = -7/9 * \log(7/9) - 2/9 * \log(2/9)$$

$$\text{Entropy}_{\text{after}} = 9/18 * \text{Entropy}_{\text{left}} + 9/18 * \text{Entropy}_{\text{right}}$$

$$\text{Information Gain} = \text{Entropy}_{\text{before}} - \text{Entropy}_{\text{after}}$$

In the case of our research, now we calculate entropy and information gain:

Imagine, the entropy of the full training set, S is:

$$\begin{aligned} \text{Entropy}(S) &= \sum_{i=1}^n P_i * \log(P_i) \\ &= -P_{\text{poor}} * \log_2(P_{\text{poor}}) - P_{\text{average}} * \log_2(P_{\text{average}}) - \\ &P_{\text{Good}} * \log_2(P_{\text{Good}}) - P_{\text{Very Good}} * \log_2(P_{\text{Very Good}}) - \\ &P_{\text{Excellent}} * \log_2(P_{\text{Excellent}}). \end{aligned}$$

Then, the information gain calculation for the first attribute is:

$$\text{Information Gain}(S, \text{Gender}) = \text{Entropy}(S) - P_{\text{male, Gender}} * \text{Entropy}(S_{\text{male}}) - P_{\text{female, Gender}} * \text{Entropy}(S_{\text{female}})$$

Now, the calculations of entropies are:

$$\begin{aligned} \text{Entropy}(S_{\text{male}}) &= \sum_{i=1}^n P_i * \log(P_i) \\ &= -P_{\text{poor, male}} * \log_2(P_{\text{poor, male}}) - P_{\text{average, male}} * \\ &\log_2(P_{\text{average, male}}) - P_{\text{Good, male}} * \log_2(P_{\text{Good, male}}) - P_{\text{Very Good, male}} * \\ &\log_2(P_{\text{Very Good, male}}) - P_{\text{Excellent, male}} * \log_2(P_{\text{Excellent, male}}). \end{aligned}$$

$$\begin{aligned} \text{Entropy}(S_{\text{female}}) &= \sum_{i=1}^n P_i * \log(P_i) \\ &= -P_{\text{poor, female}} * \log_2(P_{\text{poor, female}}) - P_{\text{average, female}} * \\ &\log_2(P_{\text{average, female}}) - P_{\text{Good, female}} * \log_2(P_{\text{Good, female}}) - \\ &P_{\text{Very Good, female}} * \log_2(P_{\text{Very Good, female}}) - \\ &P_{\text{Excellent, female}} * \log_2(P_{\text{Excellent, female}}). \end{aligned}$$

In this way, we can calculate information gain for all attributes (No. of tuition, Programming skills, CT mark, assignment mark, etc.). The attribute having the highest information gain is the root node of the decision tree. The process has been continuing for selecting the whole tree root node and prioritization of those attributes frequently.

5.2. Results from WEKA Analysis

In an academic session, generally, each student has to go through four years of study, and each year consists of 2 terms. The ‘Years’ and ‘Terms’ selected as the possible nodes of the tree in WEKA. Here in our research, we put a total of five class attributes, which are: Poor, Average, Good, Very Good, and Excellent. To build the model, we applied a variety of algorithms using WEKA explorer. The classification algorithms used for this work are REPTree, Hoefding Tree, and J48 tree. The classify panel of WEKA offers to apply a variety of classification and regression algorithms to the resulting dataset to measure the accuracy of the resulting prediction model, and to look out the flawed predictions or the model itself.

At first, load the data file- NSTU.arff into WEKA. The file contains the information of students. We need to divide up the records, so some data instances can use to create the model, and some data can use to test the model to ensure that we didn’t overfit it. This is necessary to evaluate the generated models’ accuracy as there is no separate data to test the model. Then, we got the model as a decision tree.

In step by step, now we will depict the whole process of WEKA simulation software.

Preprocessed the data by converting CSV to ARFF format:

The raw data which contains CSV(comma-separated values) format on the excel sheet was selected and convert it to ARFF(.arff) format by the WEKA tool (Figure 2). The decision tree-based system model was then generated after the processing of the NSTU.arff file.

Figure 2: ARFF conversions by Weka tools

The preprocessed data was then used to train the system using the WEKA implementation tool (Figure 3).

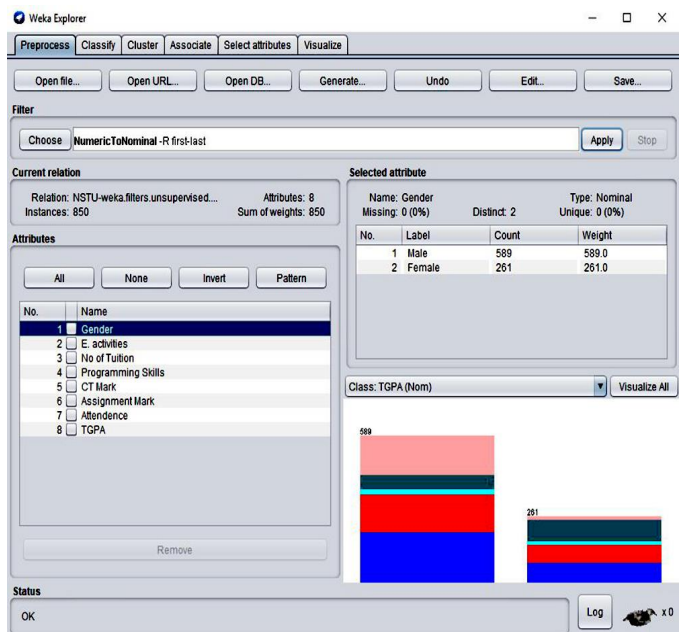


Figure 3: WEKA interface for the pre-processed data

Then the data is used to classify using WEKA implemented J48 algorithm. The visualization of the attributes presented in Figure 4.

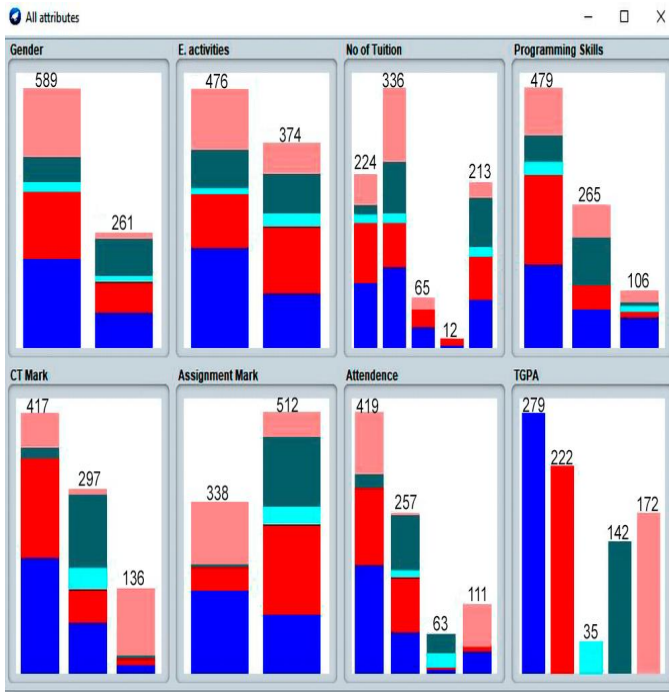


Figure 4: Attributes visualization

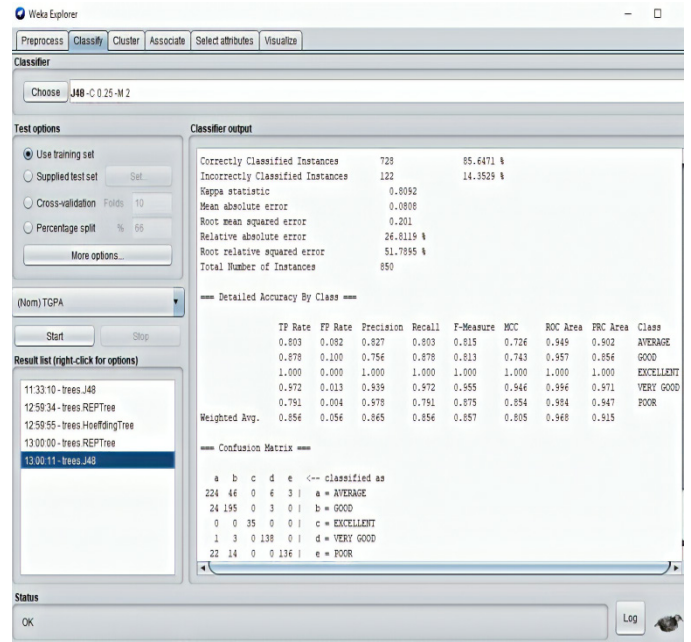


Figure 6: Result of the classifier

Figures 5 show the set of generated rules, and Figures 6 shows the corresponding result of the classifier using WEKA tools.

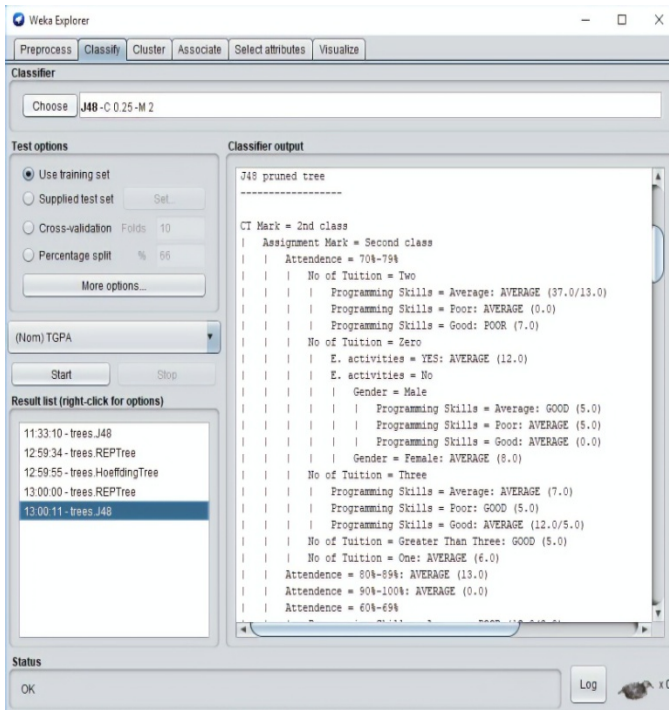


Figure 5: J48 rules generated using WEKA

There are 850 instances and 8 attributes. In the tree shown in Figure 7 has the number of leaves: 89, tree size: 130. Here, the percentage of correctly classified instances is 85.6471%, and the percentage of incorrectly classified instances is 14.3529%. The accuracy model ascertains that the proposed model could be a very good choice to predict student's results and for further development.

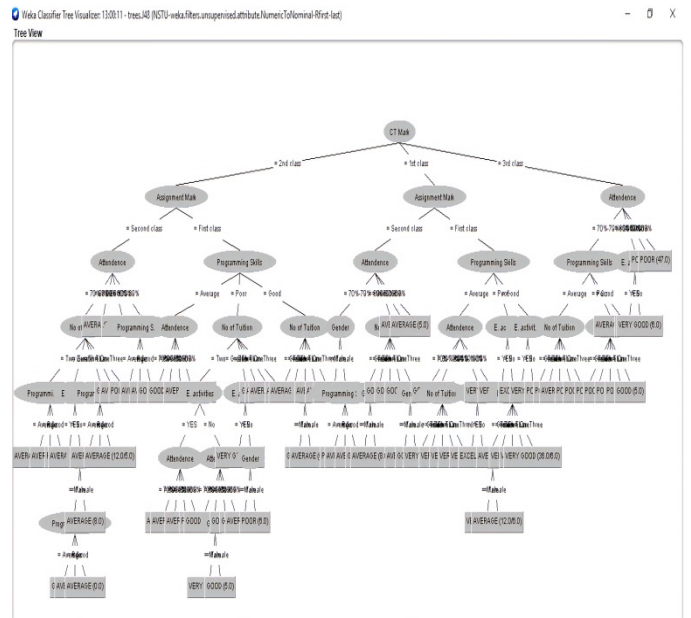


Figure 7: Decision tree produced using WEKA

5.3. Final Result Exploration

The decision tree knowledge was deduced as IF-THEN format rules and demonstrated in Figure 8.

In this way, we got 89 possible IF-THEN rules. Then we use these IF-THEN rules on the backend program for system development.

5.4. Login Page

It is the initial page that the user can see through which a user can get access to the proposed web-based student result analysis and prediction system. Use must put their credentials (username, password) to get access (Figure 9).

```

if then rules - Notepad
File Edit Format View Help
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Two,Programming Skills = Average Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Two,Programming Skills = Poor Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Two,Programming Skills = Good Then TGPA = POOR
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Two,Programming Skills = Average Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Zero,E. activities = YES Then TGPA = AVERAGE

IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Zero,E. activities = No,Gender = Male,
Programming Skills = Average Then TGPA = GOOD.
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Zero,E. activities = No,Gender = Male,
Programming Skills = Poor Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Zero,E. activities = No,Gender = Male,
Programming Skills = Good Then TGPA = AVERAGE

IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Zero,Gender = Female Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Three,Programming Skills = Average Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Three,Programming Skills = Poor Then TGPA = GOOD
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Three,Programming Skills = Good Then TGPA = AVERAGE

IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = Greater Than Three Then TGPA = GOOD
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 70%-75%,No of Tuition = One Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 80%-85% Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 90%-100% Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 60%-65%,Programming Skills = Average Then TGPA = POOR
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 60%-65%,Programming Skills = Poor Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = Second class,Attendance = 60%-65%,Programming Skills = Good Then TGPA = AVERAGE

IF CT Mark = 2nd class,Assignment Mark = First class,Programming Skills = Average,Attendance = 70%-75% Then TGPA = GOOD
IF CT Mark = 2nd class,Assignment Mark = First class,Programming Skills = Average,Attendance = 80%-85% Then TGPA = GOOD
IF CT Mark = 2nd class,Assignment Mark = First class,Programming Skills = Average,Attendance = 90%-100% Then TGPA = GOOD
IF CT Mark = 2nd class,Assignment Mark = First class,Programming Skills = Average,Attendance = 90%-100% Then TGPA = GOOD
IF CT Mark = 2nd class,Assignment Mark = First class,Programming Skills = Average,Attendance = 80%-85% Then TGPA = AVERAGE

IF CT Mark = 2nd class,Assignment Mark = First class,Programming Skills = Poor,No of Tuition = Two,E. activities = YES,
Attendance = 70%-75% Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = First class,Programming Skills = Poor,No of Tuition = Two,E. activities = YES,
Attendance = 80%-85% Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = First class,Programming Skills = Poor,No of Tuition = Two,E. activities = YES,
Attendance = 90%-100% Then TGPA = AVERAGE
IF CT Mark = 2nd class,Assignment Mark = First class,Programming Skills = Poor,No of Tuition = Two,E. activities = YES,
Attendance = 60%-65% Then TGPA = POOR
    
```

Figure 8: Rule set generated by J48

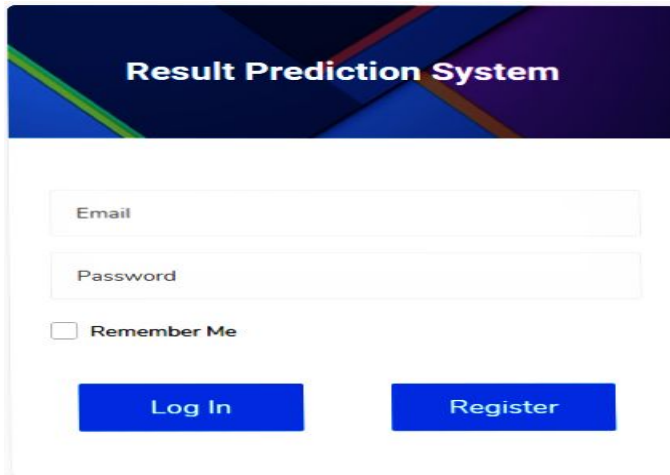


Figure 9: Login page

5.5. Sample Output for the System

The system asks the user questions related to their CT marks, assignment mark, and other related questionnaires of the previous semester, which results have not yet published. Then based on the user response, the system can predict those semester results for the user and give a suggestion that provides the reason behind the corresponding results. The system also able to categories the students' results. This also enables the system to classify the categories (Excellent, Very good, Good, Average, Poor) of student's records in their academic period. A sample output of the proposed prediction system demonstrated in Figure 10.

5.5.1 Class-wise Accuracy and Accuracy model for Class Prediction

For different outcome categories, the true positive, false positive, and correct precision results presented in Table 2.



Figure 10: Sample output for the prediction system

Table 2: Class-wise accuracy for five class prediction

TGPA CLASS	True Positive (TP)	False Positive (FP)	Correct Precision (%)
EXCELLENT	1.000	0.000	100%
VERY GOOD	0.972	0.013	97.2%
GOOD	0.878	0.100	87.8%
AVERAGE	0.803	0.082	80.3%
POOR	0.791	0.004	79.1%

We found the accuracy percentage of the J48 decision tree algorithm is about 85.6471% for the dataset used in this study (Table 1) and shows in Table 3.

Table 3: Accuracy percentages

Algorithm	Correctly Classified Instances(CCI)	Incorrectly Classified Instances(ICI)
J48	85.6471%	14.3529%

5.5.2 Comparison of the performance metric for other classification models

Table 4 shows the accuracy comparison of J48, REPTree, and Hoeffding tree algorithms for the dataset used in this study (Table 1) observed as follows:

Table 4: Comparison of other classification models

Algorithm	Correctly Classified Instances(CCI)	Incorrectly Classified Instances(ICI)
J48	85.6471%	14.3529%
REPTree	82.2353%	17.7647%
Hoeffding Tree	65.1765%	34.8235%

So from the comparison, we see that J48 is the best classifier for the data set.

6. Conclusion and Future Work

As a countries development mostly relies on the educational outcome of the students, especially on the performance of university graduates, the performance improvement of the students is highly desirable. The proposed system can predict the students' results and will help them for their performance improvement by knowing their lacking and following the pieces of advice from the system based on their current conditions. Moreover, the proposed system would help the students to cut down the overall failure rate as they can now be well directed and advised. The experimental results show that the true positive rate for getting the POOR, AVERAGE, GOOD, VERY GOOD, EXCELLENT class is 79.1%, 80.3%, 87.8%, 97.2%, and 100% respectively, and J48 is the best algorithm for classification for the dataset used in this study that achieved 85.6471% accuracy. In conclusion, we hope that university students will hugely benefit from using the system for their performance improvement.

The future work would comprise utilizing the proposed system on an extended data set with more attributes (e.g. student financial condition, No of backlog, etc.) to get more accurate results. We also aim to prolong the work with more experiments using other techniques like neural network and clustering, etc.

References

- [1] A. S. Olaniyi, S. Y. Kayode, H. M. Abiola, S. I. T. Tosin, A. N. Babutunde, "Student's Performance Analysis Using Decision Tree Algorithm" *Annals. Computer Science Series*, 15(1), 55-62, 2017.
- [2] M. F. Sikder, M. J. Uddin, S. Halder, "Predicting Students Yearly Performance using Neural Network: A Case Study of BSMRSTU" *5th International Conference on Informatics, Electronics and Vision (ICIEV)*, 524-529, 2016. <https://doi.org/10.1109/iciev.2016.7760058>
- [3] B. K. Baradwaj, S. Pal, "Mining Educational Data to Analyze Students Performance" *International Journal of Advanced Computer Science and Applications*, 2(6), 2011. <https://doi.org/10.14569/IJACSA.2011.020609>
- [4] S. K. Yadav, S. Pal, "Data Mining Application in Enrolment Management: A Case Study" *International Journal of Computer Application (IJCA)*, 41(5), 1-6, 2012. <https://doi.org/10.5120/5534-7581>
- [5] C. Romero, S. Ventura, P. G. Espejo, C. Hervás, "Data Mining Algorithms to Classify Students" *In EDM*, 8-17, 2008.
- [6] L. A. Whitley, "Educational data mining and its uses to predict the most prosperous learning environment" *M.Sc. Thesis, East Carolina University in USA*, 2018.
- [7] A. O. Ogunde, D. A. Ajibade, "A Data Mining System for Predicting University Students' Graduation Grades Using ID3 Decision Tree Algorithm" *Journal of Computer Science and Information Technology*, 2(1), 21-46, 2014.
- [8] N. A. Gorikhan, "A study on Implementation of classification techniques to predict students' results for Institutional Analysis" *MSc Thesis, Faculty of Engineering & IT, The British University in Dubai*, 2016.
- [9] B. K. Bhardwaj, S. Pal, "Data Mining: A prediction for performance improvement using classification" *International Journal of Computer Science and Information Security (IJCSIS)*, 9(4), 136-140, 2011.
- [10] R. S. J. D Baker, K. Yacef, "The State of Educational Data Mining in 2009: A Review and Future Visions" *Journal of Educational Data Mining*, 1(1), 2009. <https://doi.org/10.5281/zenodo.3554657>
- [11] M. Hall, E. Frank, G. Holmes, B. P. P. Reutemann, I. H. Witten, "The WEKA Data Mining Software: An Update" *SIGKDD Explorations*, 11(1), 10-18, 2009. <https://doi.org/10.1145/1656274.1656278>
- [12] K. Adhatrao, A. Gaykar, A. Dhawan, R. Jha, V. Honrao, "Predicting Student's Performance Using ID3 and C4.5 Classifications Algorithms" *International Journal of Data Mining & Knowledge Management Process (IJDKP)*, 3(5), 39-52, 2013. <https://doi.org/10.5121/ijdkp.2013.3504>
- [13] B. Abdullah, I. Abd-Alghafar, G. I. Salama, A. Abd-Alhafez, "Performance evaluation of a genetic algorithm based approach to network intrusion detection system" *in 13th international conference on aerospace sciences and aviation technology, Military Technical College, Kobry Elkobbah, Cairo, Egypt*, 2009, 1-17. <https://doi.org/10.21608/asat.2009.23490>
- [14] K. M. Osei-Bryson, "Post-pruning in decision tree induction using multiple performance measures" *Computers & operations research*, 34(11), 3331-3345, 2007. <https://doi.org/10.1016/j.cor.2005.12.009>
- [15] W. Y. Loh, "Classification and Regression Tree Methods" *Encyclopedia of Statistics in Quality and Reliability*, 315-323, 2008. <https://doi.org/10.1002/9780470061572.eqr492>
- [16] B. Devèze, M. Fouquin, "Data Mining C4.5 DBSCAN, Promotion" *SCIA Ecole pour l'informatique et techniques avancées*, 2005.
- [17] B. Hssina, A. Merbouha, H. Ezzikouri, M. Erritali, "A comparative study of decision tree ID3 and C4.5" *International Journal of Advanced Computer Science and Applications*, 4(2), 13-19, 2014. <https://doi.org/10.14569/SpecialIssue.2014.040203>
- [18] D. N. Bhargava, G. Sharma, D. R. Bhargava, M. Mathuria, "Decision Tree Analysis on J48 Algorithm for Data Mining" *International Journal of Advanced Research in computer Science and Software Engineering*, 3(6), 1114-1119, 2013.

Alternative Real-time Image-based Smoke Detection Algorithm

Sally Almanasra*, Ali Alshahrani

Faculty of Computer Studies, Arab Open University, Riyadh, Saudi Arabia

ARTICLE INFO

Article history:

Received: 16 February, 2020

Accepted: 26 April, 2020

Online: 03 May, 2020

Keywords:

Image Processing

Smoke Detection

Target Extraction

ABSTRACT

Most buildings are equipped with various types of sensors to detect smoke in the event of a fire, though most are located internally. Practically, smoke has to reach the sensor in order for the sensor to react. The limitations of these sensors are their inability to respond in the early stages of a fire, and their questioned efficiency in accurately detecting the source of the smoke and locations in external environments. Image processing techniques are widely used in different critical applications in the domains of security, recognition, detection, etc. In this paper, we present an alternative image-based algorithm that can detect smoke in both indoor and outdoor environments. The algorithm operates over colored images to detect smoke at the early stages of a fire. The core of the algorithm relies on target extraction, color analysis and block subtraction components. Results shows that our proposed algorithm is capable of detecting smoke accurately at a rate of 95.10%, making it suitable for wide range of applications.

1. Introduction

Accidental fires cause severe economic and ecological damage in addition to threatening people's lives. To avoid the catastrophe of a fire, fire-detection techniques were developed which were mainly based on particle sampling [1], temperature sampling [2], in addition to the traditional ultraviolet [3] and infrared [4] fire detectors and sensors. Sensor-based smoke detection systems work effectively if they are located close to the smoke source. One of the critical limitations of these systems is their inability to identify and locate the source the smoke, especially if deployed in outdoor environments.

As alternative to sensors, image-based detection systems of indoor and outdoor fires have been developed as an effective alternative to detect smoke. The concept relies on the development of an algorithm which is able to automatically detect very early signs of smoke in captured images. One of the most challenging issues in such techniques is the recognition of smoke in outdoor environments. This is due to the difficulty of recognizing the visual characteristics of the smoke.

In this paper we present an alternative image-based smoke detection algorithm which works on recognizing fire smoke using colored digital images captured in outdoor environments. The proposed technique assumes that regular images of the area under surveillance are captured and immediately processed. These

images will be analyzed and through a series of stages, including extracting the target, analyzing image colors and subtracting images blocks through an organized process. The significance of the proposed method lies behind its structure's simplicity, smoke detection accuracy and performance. Compared to other techniques, our alternative method is found promising alternative as it offers high detection rates with non-complex design.

2. Literature Review

The authors of [5] proposed a method which can detect smoke based on the concept of source separation. The method treats the captured fire images as a linear combination of smoke and background image pixels. In other words, the method aims to separate the smoke from the background of the captured images. The results published in this research show that the method is able to effectively detect smoke in outdoor environments. However, the performance of the proposed method is not examined, hence the method's efficiency is questioned.

In [6], a method for image-based smoke detection using an image-based technique is presented. The method relies on considering the static and dynamic features of the fire smoke. The dynamic features are disorder, growth, and frequent flicker, while the static features are self-similarity and wavelet energy. The first step is to detect the moving target in the image. This is done through the use of a median filter algorithm to retrain the noise. Consequently, a background subtraction technique based on

* Sally Almanasra, Email : s.almanasra@arabou.edu.sa

www.astesj.com

<https://dx.doi.org/10.25046/aj050316>

adaptive a background update is applied. Upon detecting the moving target, the method aims to extract the target's contour in every binary image of the target in order to extract a frequent-flicker feature in target boundaries. The disorder that may appear in the captured image can be measured based on the ratio of circumference to area for the segmented target region. The frequent-flicker feature is also extracted using a wavelet domain approach. This can help in determining the temporal high-frequency activity in a pixel. The last step carried out in this method is the process of joining all extracted features through a joint feature vector into an artificial neural network.

The work in [7] is presented according to the fact that the boundary of the smoke region is not clear, which makes it difficult to extract the smoke region using fundamental image processing. To resolve this limitation, authors incorporated K-means algorithms to fix the initial point in the domain region. The image is completely black while all pixels are in the initial point and through an image segmentation process, the image of the smoke shape then develops self-similarity properties. The positional relationships of pixels extracted from a single brightness is used to finally apply the fractal encoding techniques used for detecting smoke images [8].

In [9] a method to improve the image quality of the smoke using fuzzy logic is presented. The smoke area is extracted using a Gaussian mixed model. A hyper parameter of a support vector machine is established for smoke recognition.

Different types of RGB-based models can be used for fire/smoke detection systems, including, YCbCr and YUV color models. The YCbCr color model is effective for fire and smoke detection using image processing techniques [10-12]. However, to minimize the chance of selecting moving objects, clustering algorithms such as the K-means algorithm are used.

A different type of technique is presented in [13]. The authors propose a real-time smoke classification method using texture analysis. This method works by background subtraction for moving objects using the Gaussian Mixture Model (GMM), which divides the image into background and foreground. Pixels that don't match Gaussian distribution are considered as foreground pixels. The image is divided into blocks, where each block is composed of 16×16 pixels as shown in Figure 1.

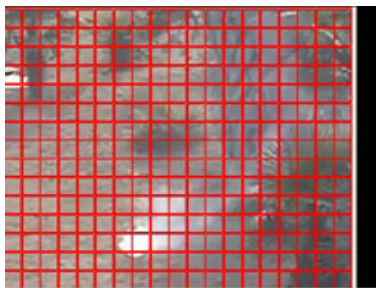


Figure 1: Sub-blocking the captured image [13]

For each block, the number of foreground pixels is counted. The method considers that particular block as a candidate block if and only if it has more than 205 foreground pixels. Consequently, the texture features are calculated for every candidate block using Gray Level Co-occurrence Matrices (GLCM). The texture features

are classified under three categories: energy, contrast and homogeneity. A back propagation Neural Network (NN) algorithm is used to discriminate the smoke features. The three texture features form the input to the NN algorithm. The resulting output is a binary value (either 0 or 1), where the value 1 represents the existence of smoke in the image, and the value 0 represent the absence of smoke, as illustrated in Figure 2.

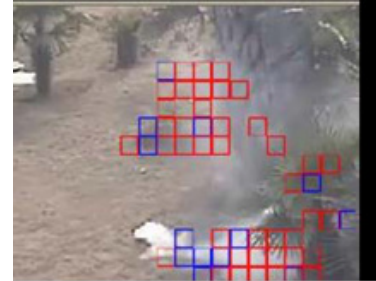


Figure 2: Detected smoke in the captured image [13]

In [14], the proposed method is designed based on deep saliency network. It aims to specify the important objects in an image. The output of this method is a smoke saliency map that is generated through the combination of pixel-level and object-level convolutional neural networks. The results show the method achieves good performance compared to some other method in that particular types of detection.

An improved smoke detection approach based on frame movement is presented in [15]. The approach relies on analyzing the characteristics of smoke at its early stages. The core of the approach is based on converting the captured images (extracted from multiple frames) to their binary representations. The lightness pixels that are not within the scope of the method are then removed. Subsequently, the smoke is detected using gray and transparency features. According to the performance analysis, the approach achieves a detection rate of about 92%.

In [16], the authors present an alternative smoke detection method based on visual smoke characteristics, including movement, color, gray tones, etc. When a region with movement is detected, the pixels within that region are estimated and analyzed to specify possible smoke regions. For higher accuracy, the method uses local binary patterns to characterize each region. Evaluation results shows that this method achieve an average smoke detection accuracy rate of 98.84%.

The authors in [17] proposes an algorithm to detect smoke and flame for video obtained from a camera installed in open area. The algorithm utilizes the adaptive background subtraction, and apply the optical flow-based movement estimation to identify a motion. Wavelet analysis is also applied to achieve moving blobs classification. Experiments results shows that the algorithm achieves a detection rate of 87% for smoke.

Similarly, the method presented in [18] utilizes the appearance and motion information to extract robust information. Machine learning is applied to achieve accurate smoke detection. Experiments results shows that the algorithm achieves a detection rate of 84.08%.

Finally, a video-based smoke-detection method is proposed in [19]. The core of the method relies on analyzing the spatial and

temporal characteristics of video sequences to identify smoke features from possible smoke regions. These features include: edge blurring, gradual energy changes, and gradual chromatic configuration changes. These features are combined using a support vector machine technique and a temporal-based alarm decision unit to enhance the detection rates. Experiments results shows that the algorithm achieves a detection rate of 83.05%.

3. Alternative Smoke Detection Algorithm

In this section we present our proposed real-time smoke detection algorithm. The core of the algorithm relies on two main components: a target extractor and a color analyzer. In addition, the algorithm is provided with a supportive component for controlling the sensitivity of the detection process according to the area of application. The overall structure of the proposed algorithm is presented in Figure 3.

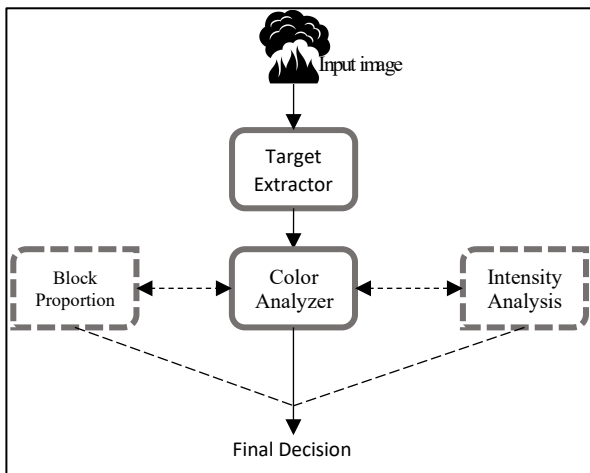


Figure 3: The structure of our proposed smoke detection algorithm

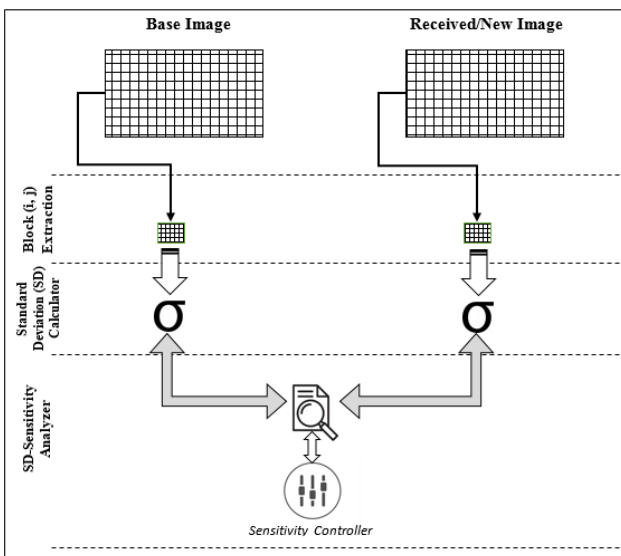


Figure 4: The internal structure of the TE Component

According to the above structure, the received image from the area under surveillance is processed through the Target Extractor (TE). This process assumes that the system is initialized by a base image in order to extract the target blocks which might be a candidate for a smoke block. Note that sensitivity controller is

designed to control the sensitivity of the smoke detection process. Increasing the level of sensitivity results in detecting smoke in earlier stages. However, over-increasing in the sensitivity of the controller might increase the false positive ratios. The internal structure of the TE component is illustrated in Figure 4.

Upon completing the target extraction stage, the received image is then masked to compare the overall intensity against the base image. This operation is carried out as part of the Color Analyzer (CA). The CA component is designed to separate the RGB blocks to examine the standard deviation of the three mean values of the extracted blocks, as shown in Figure 5.

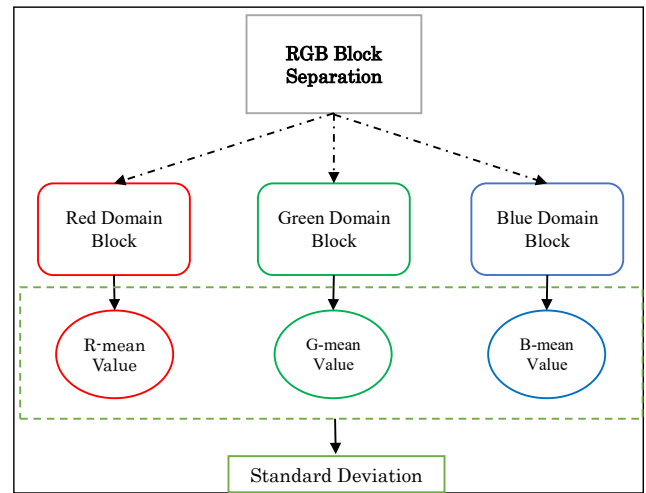


Figure 5: The internal structure of the CA Component

Consequently, a curve-fitting process is applied over the block intensity of both base image and current image obtained from the area under surveillance. The decision on whether smoke is found in the image or not is based on the difference between the fitted curves of both base and current images.



Figure 6: The base image of the area under surveillance



Figure 7: The image of the smoke captured for the area under surveillance

4. Results and Discussion

Our proposed system is tested against a set of images. Initially, the base-image should be stored for future comparison against the received image from the area under surveillance. Accordingly, Figure 6 presents the base image used in this experiment. For testing purposes, a fire has been started and the smoke starts to rise. Figure 7 shows the current image captured by the installed camera in the area under surveillance.

At this stage, a mask is applied over both the base image and the newly captured image to study possible changes in the overall intensity of the image's blocks. As shown in Figure 8, the masking process shows exactly which areas are candidate to represent smoke.



Figure 8: Masking the base-image (A) and the newly captured image (B)

Upon completing both the target extraction and color analysis, our proposed algorithm considers only the blocks which includes the black smoke. In the next stage. The curve fitting process is applied. Figure 9 shows the base image intensity before applying the curve fitting, where the figure shows the mean value of each block in the image.

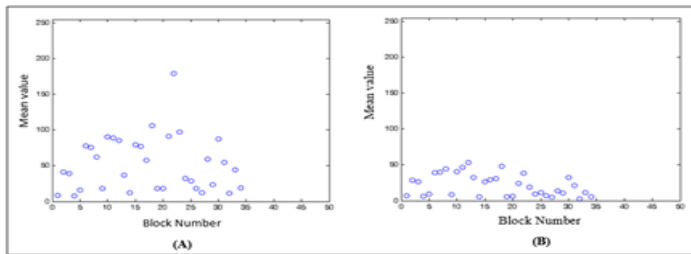


Figure 9: The intensity of the base-image (A) and the current image (B) before applying curve fitting

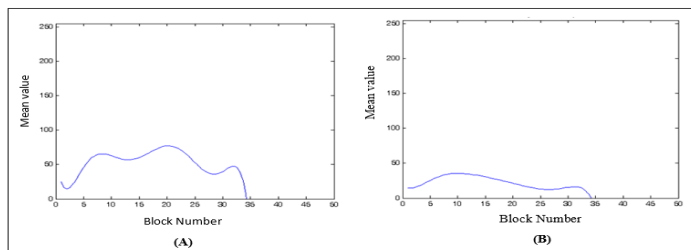


Figure 10: The intensity of the base image (A) and the current image (B) before applying curve fitting

Consequently, the curve-fitting process is applied over the base and current images. This process aims to find the best fit to the series of data points in each image. Figure 10 presents the intensity of the base-image and the current image after curve fitting has taken place.

Based on the two fitted curves of both the base and current images, one can note that the fitted curve of the current image (with smoke blocks) is lower than the fitted curve of the base-image. At this point, the level of sensitivity plays its role in whether to consider a particular block as smoke or not. We have conducted our experiment over a selected set of original smoke and none smoke images found in [20] and illustrated in Figure 11.

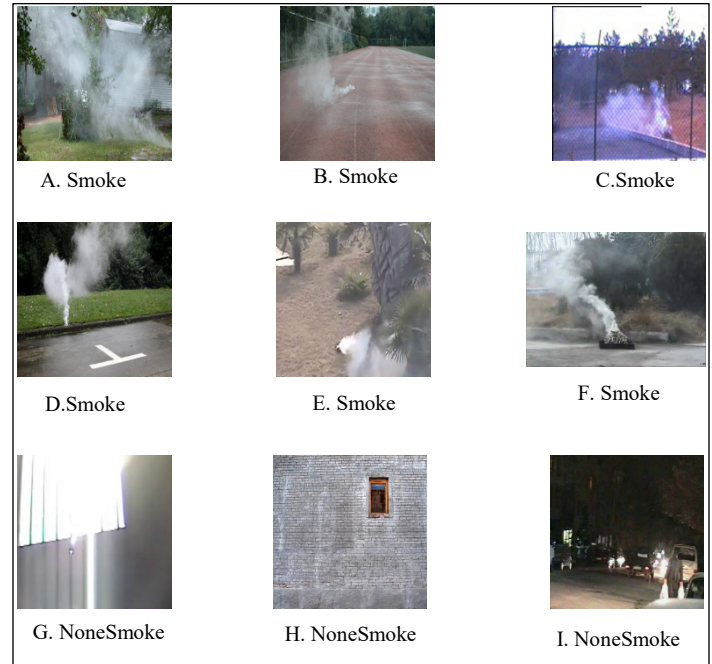


Figure 11: Set of Smoke Images (A-F) and Set of None Smoke Images (G-I)

The experiment results over the set of images shown in Figure 11, reveals that our method is able to detect the smoke-spectrum in most of the images. The results of the detection process applied over smoke and none smoke images are presented in Figure 12 and 13, respectively.

Table 1: Analyzing detection rates over different levels of sensitivity

Image	Level of sensitivity	Decision	False Alarm?
A. Smoke	50%	Smoke Detected	No
A. Smoke	75%	Smoke Detected	No
A. Smoke	95%	Smoke Detected	No
B. Smoke	50%	No Smoke	Yes
B. Smoke	75%	Smoke Detected	No
B. Smoke	95%	Smoke Detected	No
C. Smoke	50%	Smoke Detected	No
C. Smoke	75%	Smoke Detected	No
C. Smoke	95%	Smoke Detected	No
G. NonSmoke	50%	No Smoke	No
G. NonSmoke	75%	No Smoke	No
G. NonSmoke	95%	Smoke Detected	Yes
H. NonSmoke	50%	No Smoke	No
H. NonSmoke	75%	No Smoke	Yes
H. NonSmoke	95%	Smoke Detected	Yes

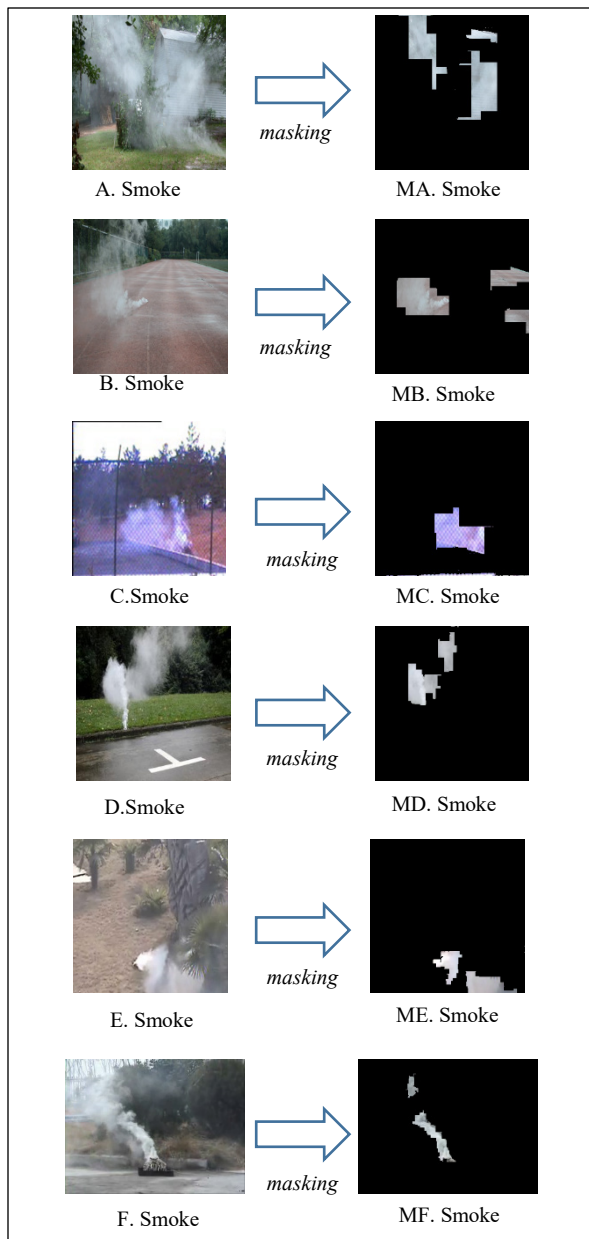


Figure 12: Original Smoke Images before Masking (A-F) and Masked Smoke Images (MA-MF)

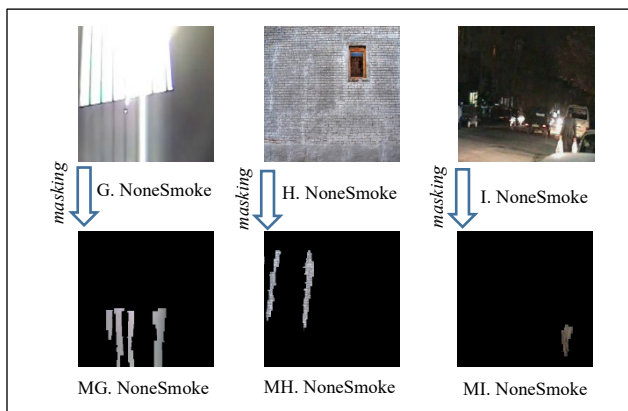


Figure 13: Original None Smoke Images Before Masking (G-I) and Masked None Smoke Images (MG-MI)

Note that our proposed method is able to detect the smoke accurately as illustrated in Figure 12. However, our algorithm has also detected some smoke patterns in non-smoke images. This is due to the value assigned to the sensitivity controller. As highlighted in Section 3, the sensitivity controller plays a pivot role in formulating the final decision on whether a particular block of pixels is recognized as smoke or not. The sensitivity controller works on curve-fitting process. Increasing the sensitivity level to some extent, lead to higher false alarm rates. Table 1 shows how our algorithm react over different levels of sensitivities.

Obviously, the ideal level of sensitivity is found to be at rate 75%. Increasing or decreasing this ratio to some extent, results in higher rates of false alarms. However, it is a trade-off process, as some applications may require higher levels of sensitivity to avoid massive lost.

We run our experiment over 200 images. the results show that the detection rate is 95.1%. For comparison purposes, we compared our results with existing methods published in [15-19]. Results presented in Table 2 show that our method outperforms most of these methods in term of smoke detection rates. This is due to low-cost operations used in our method.

Table 2: Comparing the smoke detection rates of our proposed method against existing methods

Method	Smoke Detection Rate
Method 1 [15]	92.60%
Method 2 [16]	98.84
Method 3 [17]	87.00%
Method 4 [18]	84.08%
Method 5 [19]	83.05%
Our Method	95.10%

In terms of performance, our method was tested on an Intel Core i7 processor of 2.8GHz, a memory RAM of 3 GB, and hard disk of 80 GB. The results show that the average processing time for an image of size 92.5KB takes about 98ms. This performance makes our method suitable for deployment in a wide range of applications and environments.

5. Conclusion

This research presents an early smoke detection method using image processing techniques. The method is based on target extraction, color analysis and curve fitting to delimit the candidate smoke region. The level of sensitivity plays a pivot role in determining whether an alarm of smoke should be activated. The experiments results show that the proposed method provides a good detection rate of smoke against several scenarios. Results show that our method provides outstanding detection results at the sensitivity level of 75%, where the alarm was activated on all images that includes smoke patterns. Similarly, the images without smoke has not activated the alarm in more than 97% of the cases. The main reason behind some false alarm cases is the distance between the recording device and the target area. However, the intensive analysis of our method over a set of 200 images reveals that, the detection rate is 95.10%. In term of processing time, our method is capable of processing an image of

less than 100KB in less than 1 second, making it feasible candidate for many environments.

Acknowledgment

The authors would like to thank the Arab Open University, Saudi Arabia for supporting this research.

References

- [1] Kaiser, L. What Is an Air Sampling Smoke Detection System? Retrieved May 11, 2019, from Mission Critical Fire Protection: <https://www.orrprotection.com/mcftp/blog/air-sampling-smoke-detection-system>, 2015.
- [2] Shixing, L., Xinxin, L., Wei, Y., Changzheng, C., Kun, Y., Maoxiang, Y., Yongming, Z. Aspirating fire detection system with high sensitivity and multi-parameter. 2014 International Conference on Information Science, Electronics and Electrical Engineering. Sapporo, Japan: IEEE, 2014.
- [3] Trumble, T. United States of America Patent No. US3982130A, 1976.
- [4] Bosch, I., Serrano, A., & Vergara, L. Multisensor Network System for Wildfire Detection Using Infrared Image Processing. *TheScientificWorldJournal*, 1-10, 2013.
- [5] Bchir, I. Image-based Smoke Detection Using Source Separation. *Spectroscopy and Spectral Analysis*, 39(3), 982-989, 2019.
- [6] Xu, Z. Automatic Fire Smoke Detection Based on Image Visual Features. 2007 International Conference on Computational Intelligence and Security Workshops, (pp. 316-319). Heilongjiang, 2007.
- [7] Fujiwara, N., Terada, K. Extraction of a smoke region using fractal cording. *IEEE International Symposium on Communications and Information Technology*. Sapporo: IEEE, 2004.
- [8] M., Silva, Abu Baker, M., Petra, M. State of the art of smoke and fire detection using image processing. *Int. J. Signal and Imaging Systems Engineering*, 10(1), 22-30, 2017.
- [9] Yuanbin, W. Smoke recognition based on machine vision. *International Symposium on Computer, Consumer and Control* (pp. 668-671). Xi'an: IEEE, 2016.
- [10] Pu, R., Lee, S. Study of smoke detection by analysis of saturation in video. 2011 International Conference on Fluid Power and Mechatronics (pp. 169-172). Beijing: IEEE, 2011.
- [11] Jinghong, L., Xiaohui, Z., Lu, W. The design and implementation of fire smoke detection system based on FPGA. 24th Chinese Control and Decision Conference (pp. 3919-3922). Taiyuan: IEEE, 2012.
- [12] Hauser, M., Li, Y., Li, J., Ray, A. Real-time combustion state identification via image processing: a dynamic data-driven approach. *American Control Conference* (pp. 3316-3321). Boston: IEEE, 2016.
- [13] Yu, C., Zhang, J., Fang, J. Texture Analysis of Smoke for Real-Time Fire Detection. *Second International Workshop on Computer Science and Engineering* (pp. 511-515). Qingdao: IEEE, 2009.
- [14] Xu, G., Zhang, Y., Zhang, Q., Lin, G., Wang, Z., Jia, Y., Wang, J. Video Smoke Detection Based on Deep Saliency Network. *Fire Safety Journal*, 105, 277-285, 2019.
- [15] Mutar, A., Dway, H. Smoke Detection Based On Image Processing by Using Grey and Transparency Features. *Journal of Theoretical and Applied Information Technology*, 96(21), 6995-7006, 2018.
- [16] Mercado, J., Medina, K., Perez, G., Suarez, A., Meana, H., Orozco, A., Villalba, L. Early Fire Detection on Video Using LBP and Spread Ascending of Smoke. *Sustainability*, 11(3261), 1-16, 2019.
- [17] Ye, S., Bai, Z., Chen, H., Bohush, R., Ablameyko, S. An effective algorithm to detect both smoke and flame using color and wavelet analysis. *Pattern Recognition and Image Analysis*. 27(1): p. 131-138, 2017.
- [18] Avgerinakis, K., A. Briassouli, and I. Kompatsiaris. Smoke detection using temporal HOGHOF descriptors and energy color statistics from video. *International Workshop on Multi-Sensor Systems and Networks for Fire Detection and Management*, 2012.
- [19] Lee, C., Lin, C., Hong, C., Su, M. Smoke detection using spatial and temporal analyses. *International Journal of Innovative Computing, Information and Control*. 8(7): p. 4749-4770, 2012.
- [20] Zhang, Q. Research Webpage about Smoke Detection for Fire Alarm. Retrieved March 10, 2020, from State Key Laboratory of Fire Science: <http://smoke.ustc.edu.cn/datasets.htm>. 2011

Improvement of Desirable Thermophysical Properties of Soybean Oil for Metal Cutting Applications as a Cutting Fluid

Putta Nageswara Rao^{*1}, Suresh Babu Valer², Koka Naga Sai Suman²

¹Department of Mechanical Engineering, VVIT, Guntur, 522508, India.

²Department of Mechanical Engineering, Andhra University, Visakhapatnam, 53000, India.

ARTICLE INFO

Article history:

Received: 09 March, 2020

Accepted: 14 April, 2020

Online: 03 May, 2020

Keywords:

Soybean oil

Reduced graphene oxide

Al_2O_3

Nanoparticle concentration

Properties

Stability

ABSTRACT

Vegetable oils are often proved to be promising for industrial lubrication applications among which soybean oil found to be better due to its attractive and desirable thermo physical properties for machining compared to other vegetable oils. However, already it was established that influenced desirable thermophysical properties necessary for a vegetable oil through which better machining performance can be obtained. Among the various vegetable oils which are practically in use for machining applications as a cutting fluid soybean oil is found to be best and which has the scope for improvement of thermophysical properties nearer to the optimized values obtained. Therefore, the present work aims to improving the influencing thermophysical properties of soybean oil by reinforcing with suitable nanoparticles. Therefore, within the present work two categories of nanoparticles such as metallic-Aluminium oxide (Al_2O_3) and non-metallic reduced graphene oxide (RGO) particles were selected for dispersion in soybean oil with different concentrations for obtaining the required properties. The obtained results reveal that non-metallic nanoparticles i.e RGO with 0.5% concentration in soybean found to be better for imparting nearer to optimum properties required for obtaining better machining performance. Further sedimentation studies were carried out to ascertain the stability of particles. The studies revealed that at 0.5% concentration of RGO assisted with ultrasonication resulted for better stability of suspended nanoparticles for long term usage.

1. Introduction

Vegetable oils are widely used as a cutting fluid for machining applications from last two decades because of their high lubrication performance, less toxicity, eco-friendly and non-hazardous nature. However, during the real time applications especially concerned to machining applications performance of these oils differs one over the other due to the wide variation of fluid properties [1-3]. Therefore Suresh et.al [4] investigated about optimization of vegetable oil properties through CFD and established that thermal conductivity and viscosity play a major role in improving the machining performance apart from another two important properties such as specific heat and density. Within their studies [4] they further established that viscosity (kg/m.s), Density (kg/m^3), Thermal conductivity (w/m-k) and Specific heat (J/kg-k) 0.025, 920, 0.18 and 2000 respectively are the desirable thermo physical properties required for a vegetable oil for

imparting better machining performance. However, within the literature [5-6] it is further observed that among various vegetable oils, soybean oil is found to be better [7] and has the scope for the researchers to improve its thermophysical properties nearer to optimized values established by Suresh et.al [4].

Therefore the present work focuses on improving the thermophysical properties of soybean oil by reinforcement of widely used metallic (Al_2O_3) nanoparticles [8-9] and non-metallic RGO[10-11] nanoparticles as separate combinations and ascertaining the better category of nanoparticles and its concentration which can impart the desirable properties and also can give long term usage capability.

2. Materials

Within the present work the base vegetable oil i.e soybean was purchased from local market and nanoparticles such as reduced graphene oxide and aluminum oxide nanoparticles was procured

^{*}Nageswara Rao putta, Department of Mechanical Engineering, VVIT, Guntur, 522508, India, ph.no: +919949001586 & Email: vvitpnr@gmail.com.

from Aarshadhatu green nanotechnologies India private limited, Guntur.

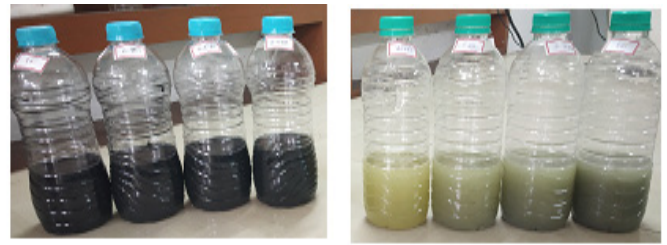
All the procured materials were commercial grade and do not require any further purification.

3. Preparation of nanofluids and evaluation their properties

Initially within the present work it is intended to prepare nanofluid samples containing metallic and non-metallic particles with various concentrations and their properties are to be evaluated day after day. Apart from this it is also intended to observe the suspension phenomena of nanoparticles within the base fluid through sedimentation analysis.

3.1 Nanofluid preparation

Nano fluid samples were prepared using two step method. For this soybean oil as taken as a base fluid for separate dispersion of reduced graphene oxide and Al₂O₃ nanoparticles at different weight percentage (0.25, 0.5, 0.75 and 1%) combinations ranging from 0.25-1% insteps of 0.25%. Afterwards the solution was mixed by bath sonicator upto 60 minutes of duration at room temperature. Prepared samples were shown through figure 1.



(a)

(b)

Figure 1: Prepared nanofluid at different concentrations (a) RGO contained Soybean (b) Al₂O₃ contained Soybean

3.2 Properties evaluation

Since the present oil is intended as lubricating oil for machining applications therefore major important properties i.e thermal conductivity has been found out through KD2 pro thermal conductivity analyzer, viscosity through Brookfield viscometer, density through Anton Parr density meter and specific heat through nanofluid heat capacity apparatus day by day upto 6 days after the preparation of nanofluid. The obtained properties were tabulated through Table 1&2 at alternate two days.

Table 1: Measured Properties of Al₂O₃ suspended nanofluid samples

Properties	Weight percentage concentrations											
	0.25%			0.5%			0.75%			1%		
	Day 2	Day 4	Day 6	Day 2	Day 4	Day 6	Day 2	Day 4	Day 6	Day 2	Day 4	Day 6
Thermal conductivity(w/mk)	0.168	0.168	0.166	0.172	0.172	0.167	0.176	0.176	0.175	0.181	0.180	0.176
Viscosity(cst)	27.8	27.8	27.5	28.1	28.1	28.0	28.5	28.5	28.3	29.0	29.0	28.8
Specific Heat(j/kg.k)	1.979	1.979	1.975	1.981	1.981	1.979	2.108	2.108	2.103	2.119	2.119	2.211
Density(g/cm ³)	0.917	0.916	0.916	0.923	0.923	0.920	0.923	0.923	0.921	0.930	0.929	0.929

Table.2: Measured Properties of RGO suspended nanofluid samples

Properties	Weight percentage concentrations											
	0.25%			0.5%			0.75%			1%		
	Day 2	Day 4	Day 6	Day 2	Day 4	Day 6	Day 2	Day 4	Day 6	Day 2	Day 4	Day 6
Thermal conductivity(w/mk)	0.172	0.172	0.171	0.181	0.181	0.180	0.187	0.186	0.185	0.190	0.190	0.187
Viscosity(cst)	27.9	27.9	27.6	28.3	28.3	28.1	28.7	28.7	28.4	29.2	29.2	29.0
Specific Heat(j/kg.k)	1.982	1.982	1.980	2.018	2.018	2.090	2.185	2.182	2.138	2.228	2.219	2.211
Density(g/cm ³)	0.918	0.918	0.916	0.921	0.921	0.919	0.923	0.923	0.921	0.928	0.928	0.927

From the above obtained results, the variation of properties of the prepared nanofluid samples were analysed in the following sections.

3.3 Variation of properties

From the obtained values of various properties of the different nanofluids at different concentrations the variation properties were plotted with respect to number of days and are shown Figures 2&3.

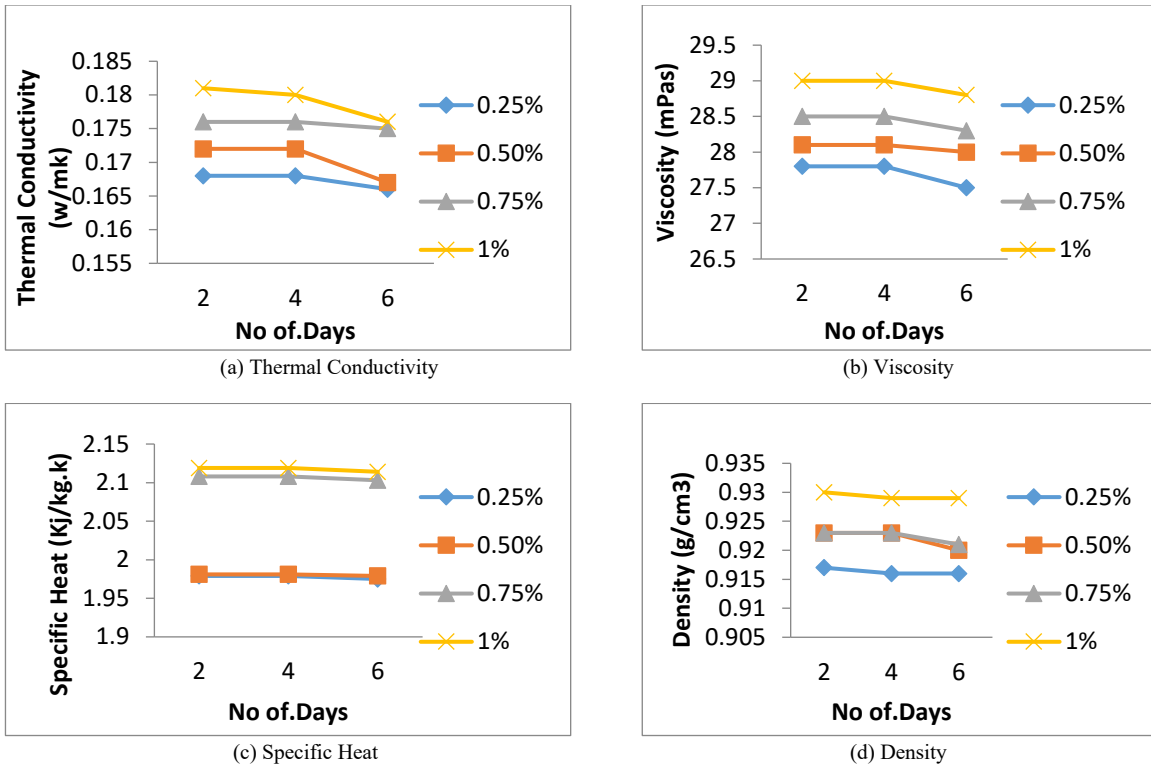


Figure 2: Variation of thermophysical properties of Al_2O_3 suspended soybean oil w.r.t to no. of days

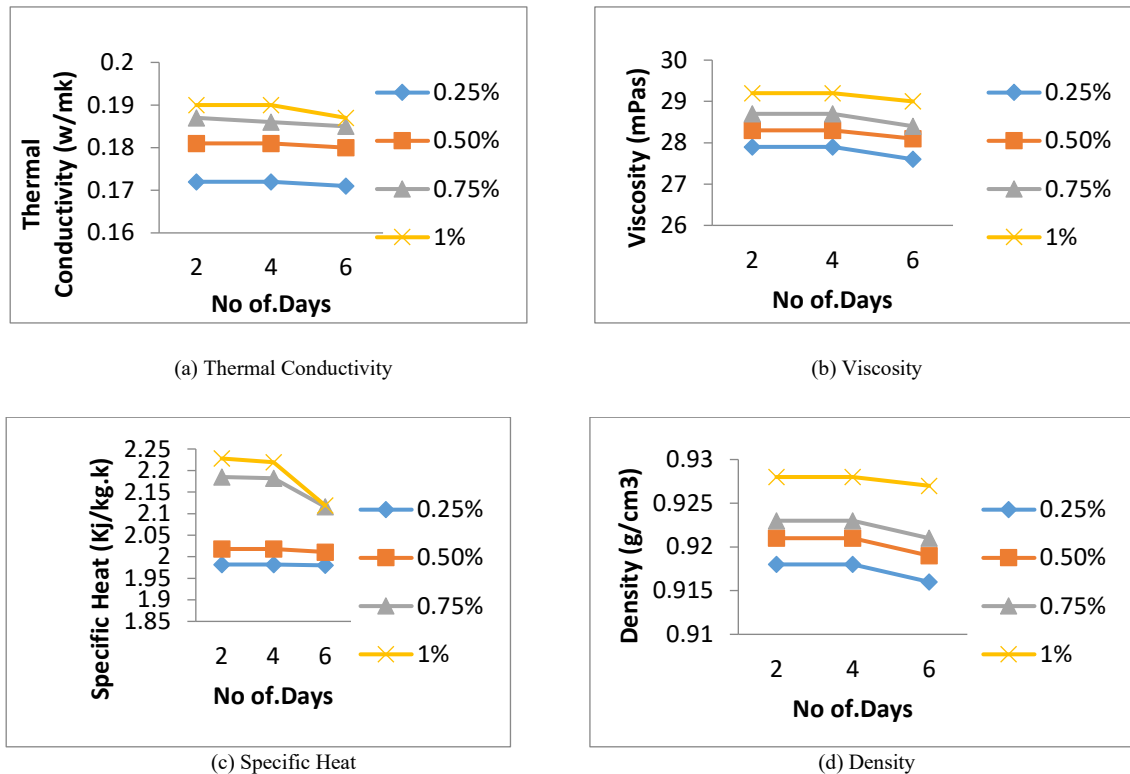


Figure 3: Variation of thermophysical properties of RGO suspended soybean oil w.r.t to no. of days

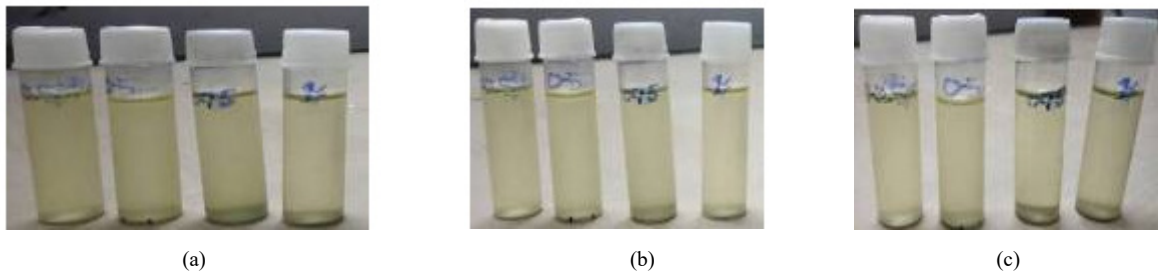


Figure 4: Digital photographs of Al_2O_3 /soybean nanofluid at different concentrations a) 2 days time of preparation b) After 4 days preparation. c) After 6 days of preparation

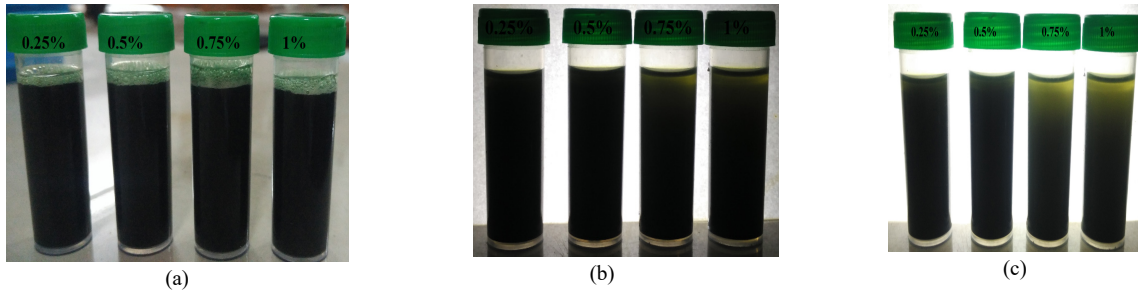


Figure 5: Digital photographs of RGO/soybean nanofluid at different concentrations a) 2 days time of preparation b) After 4 days preparation. c) After 6 days of preparation

Generally, nanofluids has higher properties than the corresponding base oils due to their random drifting of suspended nanoparticles in the base fluid and nanoparticles are compared to more thermophysical properties than liquids which originates from collisions between the nanoparticles and liquid molecules and the same trend was observed in the present work. From the figures 2&3 it was observed that with the increase of particle concentration all the measured properties are found to be increasing after 2nd and up to the 4th day of preparation however after 4 days of preparation all the properties at 0.75 and 1 weight percentage combinations gradually decreased but at 0.5% concentration the properties were found to be constant.

Observing the individual trend of prepared nanofluid samples containing Al_2O_3 and RGO it is observed that RGO suspended nanofluid samples exhibited superior properties compared Al_2O_3 suspended samples. The superiority of RGO suspended nanofluid samples compared to Al_2O_3 suspended samples can be attributed to better stability and interaction of RGO particles with the base fluid compared to Al_2O_3 particles. It is further observed that at 0.5% concentration of RGO the properties tend to be much more stable compared to other concentrations and more importantly the obtained properties were nearer to the optimized values established by Suresh et.al [4].

3.4 Visualization effect

Figures 4&5 shows the photos of different weight percentage concentrations of Al_2O_3 -suspended and RGO suspended soybean oil samples taken after 2, 4 and 6 for visual inspection. From the visual inspection it is observed that the sedimentation was found to very less for all the concentrations after the preparation of oil up to two days after that with the increase in time it was found to be increasing at all concentrations with in

which 0.75 and 1% weight percent of nanoparticles led to high sedimentation compared to 0.25 and 0.5% weight concentration.

4. Stability improvement through ultrasonication

Based on the obtained thermophysical properties and the studies carried out through visualization it is concluded that RGO suspended nanofluid at 0.5% weight concentration has resulted for imparting the desirable properties established by suresh et.al [4]. Since the prepared nanofluid intended for using as a lubricant in machining applications, apart from desirable properties it further required to have long term stability therefore within the present work the prepared 0.5% wt RGO nanofluid samles were subjected to ultrasonication for imparting better stability to the nanoparticles. Ultrasonication was carried out for a durations 60-180 minutes insteps of 30 min. thereby in total five ultrasonicated nanofluid samples have been prepared and the samples were evaluated for their thermal conductivity and viscosity which are most influencing properties.

4.1 Variation of properties after ultrasonication

The thermal conductivity of the nanofluid of prepared nanofluid and their variations at different ultrasonication durations with respect to number of days plotted as shown in Figure 6. With the different ultrasonication durations of RGO-soybean oil thermal conductivity is found to be increasing after 60 minutes to 120 minutes ultrasonication durations upto 20 days of preparation. However, after 120 minutes ultrasonication thermal conductivity at 150 and 180 minutes gradually decreased but at 120 minutes of durations the stability of nanoparticles was found to be stable upto 20 days of preparation. The same trend was observed in viscosity of nanofluid which is from Figure 7.

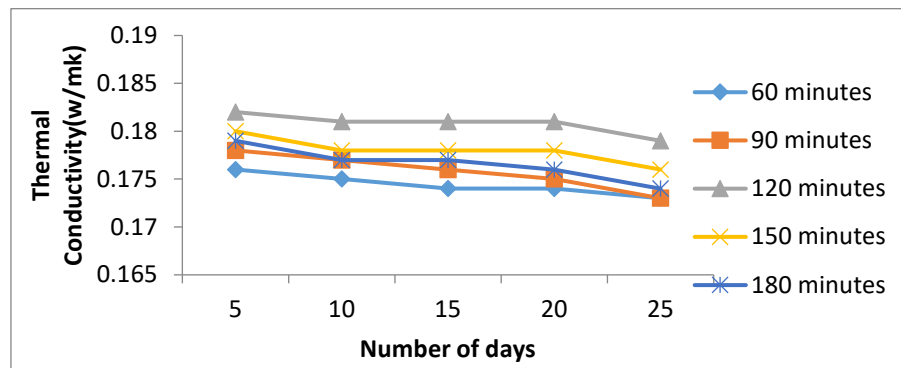


Figure 6: Variation of thermal conductivity with respect to number of days

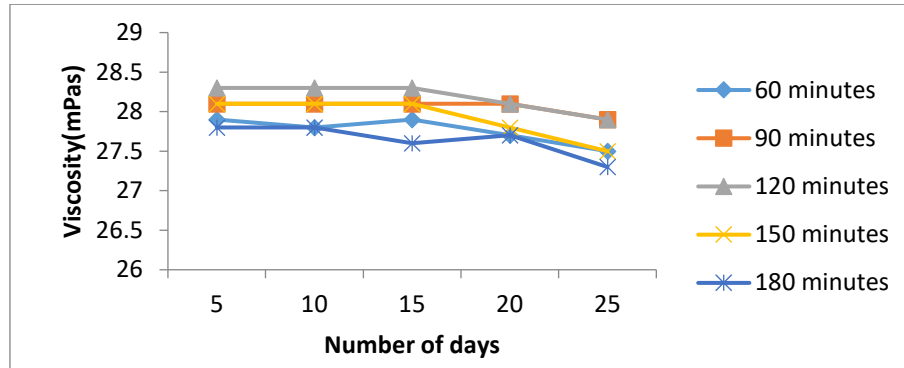


Figure 7: Variation of viscosity with respect to number of days

From the results it is observed that increasing ultrasonication duration of 120 min, dispersity of nanoparticles became more even and homogeneous, and presence of agglomerates diminished significantly. This is attributed due to the fact that ultrasonic energy broke down the nanoparticle agglomerations, and it could yield a more homogeneous nanofluid sample. On the other hand, further increase of ultrasonication period resulted in a slight tendency for re-agglomeration, re-agglomeration of particles with the latter sonication period may be seen.

4.2 Visualization effect of nanofluid with ultrasonication

Figure 8(a)-(b) exhibit the sedimentation pictures of reduced graphene oxide-soybean nanofluid samples prepared at different ultrasonic durations and after the 20 and 25 days of preparation

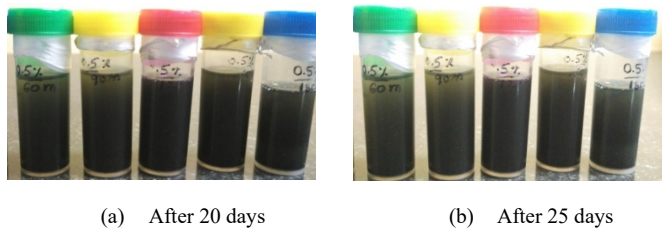


Figure 8: Digital photographs of prepared nanofluid samples at different ultrasonication durations of 60-180 minutes in steps of 30 minutes

From the visual analysis it is observed that the sedimentation rate was found to very less for all the concentrations after the preparation of oil upto 20 days after that with increase in time the rate was found to be increasing at all ultrasonic durations except for 120 minutes where the sedimentation rate was found to be very less.

5. Conclusions

The present work mainly concentration about preparation of soybean based nanofluid containing the thermophysical properties [4] required for improving machining performance and also for long term usage with required properties. Based on the above aims the following conclusion has been drawn.

- Reinforcement of Al_2O_3 and RGO nanoparticles into the base fluid has resulted for enhancement of required properties considerably.
- Compared to metallic Al_2O_3 nanoparticles, non-metallic RGO nanoparticles has resulted for bringing the thermophysical properties upto the desired level compared to metallic nanoparticles.
- It is observed that nanoparticles at 0.5% RGO nanoparticles concentration have resulted for bringing for the properties upto the desired level.
- After that ultrasonic treatment of nanofluid with 0.5% RGO has resulted for long term stability of nanoparticles up to 20 days without having much variation in the thermophysical properties.

References

- [1] Susmitha, Sharan, Jyothi, "Influence of non-edible vegetable based oil as cutting fluid on chip, surface roughness and cutting force during drilling operation of mild steel" in IOP Conf. Series: Materials Science and Engineering, 2016, doi:10.1088/1757-899X/149/1/012037.
- [2] Sunday Albert Lawal, Intiaz Ahmed Choudhury, Ibrahim Ogu, Sadiq, Adedipe, Oyewole, "Vegetable-oil based metalworking fluids research developments for machining processes: survey, applications and challenges" in Manufacturing Rev., Volume 1, 2014.

- [3] Surase, Pawar, Ramkisan, "Performance of vegetable oil based cutting fluid in machining of steel by using MQL -A review" *Lubricants*, Volume 5, Issue 44,2017, doi:10.3390/lubricants5040044.
- [4] Suresh Babu Valeru, Nageswara Rao, K.N.S Suman, "Optimization of vegetable oil properties in machining environment through CFD" in *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Volume-8 Issue-12. 2019.
- [5] Karmakar Gobind, Ghosh Pranab, Sharma Brajendra, "Chemically modifying vegetable oils to prepare green lubricant" in *Lubricants*.5,2017,doi:10.3390/lubricants5040044.
- [6] Shrikant Madiwale, Virendra Bhojwani, "An overview on production, properties, performance and emission analysis of blends of biodiesel" in *Procedia Technology* 25, 2016.
- [7] Putta Nageswara Rao, Suresh Babu Valeru, Koka Naga Sai Suman, "Selection of vegetable oil for MQL as a cutting fluid through MADM methods" Accepted in *industrial engineering journal*, Paper Id:1246.
- [8] Vasheghani Mohammadhassan, Marzbanrad, Ehsan, Zamani Cyrus, Aminy Mohammad, "Effect of Al_2O_3 phases on the enhancement of thermal conductivity and viscosity of nanofluids in engine oil" in *Heat and Mass Transfer* 47,2011.
- [9] Rahman Mostafizur, Rahman Saidur, Abdul Raman, Abdul Aziz, "Thermophysical properties of methanol based Al_2O_3 nanofluids" in *International Journal of Heat and Mass Transfer* 85,2015.
- [10] Syed Nadeem, Abbas Shah, Syed Shahabuddin, Mohd Faizul Mohd Sabri, Mohd Faiz Mohd Salleh, "Experimental investigation on stability, thermal conductivity and rheological properties of rGO/ethylene glycol based nanofluids" in *International Journal of Heat and Mass Transfer*, Volume 150, 2020.
- [11] Kamatchi, Kannan Gopi, "An Aqua based reduced graphene oxide nanofluids for heat transfer applications: synthesis, characterization, stability analysis, and thermophysical properties" in *International Journal of Renewable Energy Research*. 8, 2018.

Socioeconomic and Productive Disparity in Child Stunting in the Central Andes of Peru, Taking as a Model the Community of Tunanmarca, Jauja

Jorge Castro-Bedriñana^{1,*}, Doris Chirinos-Peinado¹, Elva Ríos Ríos²

¹Zootechnic Faculty, Universidad Nacional del Centro del Perú, Huancayo 12000, Perú

²Science Faculty, Universidad Nacional Agraria. Lima 15012, Perú

ARTICLE INFO

Article history:

Received: 12 March, 2020

Accepted: 14 April, 2020

Online: 03 May, 2020

Keywords:

Stunting

Agri-food production

Nutritional status

Breastfeeding

Nutritional survey

Public nutrition

Public Health

ABSTRACT

The differences of stunting through socio-economic and productive indicators in high Andean community of Peru were evaluated ($11^{\circ} 42' 58.16''$ S, $75^{\circ} 37' 31.13''$ W, altitude 3470 m). Cross-sectional study in 52 mothers with children under 5 years old was carried. A validated nutritional survey was applied. Z-scores height for age and nutritional status were determined using anthropometric methods and WHO criteria. The prevalence of stunting was evaluated by maternal educational level, food and health practices, economic level and family food production. Chi-square tests and Spearman correlations were performed in order to establish associations to $P < 0.05$. Prevalence of stunting was 44.2%. The factors associated with stunting ($P < 0.05$) were: Do not use gas for cooking ($r=0.530$), weekly economic income < 50.00 dollars ($r=0.503$), weekly expenditure on family food < 31.00 dollars ($r=0.648$), per capita / day expenditure on food < 1.10 dollars ($r=0.591$), mother without studies ($r=0.454$), no own home ownership ($r=0.413$), consumption of food before 6 months old ($r=0.410$), low frequency of quinoa consumption ($r=0.423$), and fish ($r=0.421$), presence of childhood anemia ($r=0.407$); inadequate venting of smoke in the kitchen ($r=0.491$), not having soap for personal hygiene ($r=0.413$) and not having a bathroom ($r=0.413$). Stunting is associated with various socioeconomic, productive and access factors to food. These results demonstrate socio-economic and productive disparities for stunting in rural high Andean areas of central Peru, taking as a model the community of Tunanmarca in Jauja.

1. Introduction

Food security is assessed at the local, family or individual level, determining indicators of food availability, access and use, and its sustainability [1], in a context of climate change and transition of eating habits [2,3]. In rural areas, availability depends directly on agro-food production and is directly related to access; families with crops and cattle have easy access to food [1]. The biological use of food depends on the capacity of the organism and the health state. Stability affects availability and access [4]. The greatest food insecurity is registered in poor, large families, with limited access to land and in low-income women [5], sustainable agriculture must be promoted to achieve family food security and stunting eradicate [6].

Stunting is associated with non-communicable diseases [7] linked to environmental and social factors that lead to a nutritional transition [7,8]; among them, the lack of education, health or other basic capacities that constitute the well-being of people [9]. The situational diagnosis of food security at the family level can be raised through surveys [10]-[12] and the nutritional evaluation of infants using anthropometric methods [12,13].

The age size index determines stunting and expresses the height of a child with respect to its age, shows if a child is stunted in its growth. Children whose height values for age are below -2 Z-score are classified as stunting; those with less than -3 Z-score are classified as severe stunting [13,14].

Worldwide, 149 million children under 5 years old have stunting. In Peru in 2000, the stunting average in children under 5 years old was 33%, it dropped to 18% in 2012, with an average of 32% in rural areas and 10.5% in urban areas; an average of 12%

*Jorge Castro-Bedriñana, Av. Mariscal Castilla 3909, El Tambo-Huancayo, +51 964408057 & jorgecastrobe@yahoo.com, jcastro@uncp.edu.pe

is reported in 2019; however, the fight against stunting does not end; although national averages show progress, regional data show deep gaps between the coast, highland and jungle regions [15]. There are regions of the highlands whose stunting average exceeds 30%; there are high Andean communities where stunting easily exceeds 60% [11,12]. Another nutritional problem that mainly affects children from 6 and 35 months is anemia, which in 2018 affected 43% [15].

Stunting is closely related to the food insecurity of rural families in the central highlands, especially with factors linked to availability and access [12], [16]; agro-food production levels do not meet its needs.

The production in the plots of the rural families is scarce, a part is stored for a short period, the children being the most vulnerable to a deficient intake of proteins and micro nutrients; their diet is eminently energetic (potato, corn, barley, noodles), the consumption of meat, milk, eggs, fish, fruits and vegetables is rare or null [11]. Similar studies conducted in the province of Jauja, report a stunting prevalence in children under 5 years old, of 21.9% [17].

In this context, the present study aimed to estimate the prevalence of stunting in children under 5 years old and determine the association between some variables of availability, access and use of food of the rural families of the Tunanmarca district, province of Jauja with the stunting prevalence of their children under 5 years old.

2. Materials and methods

2.1. Place and duration of the study

The study was carried out in September 2015, in the Tunanmarca district, province of Jauja, located between Coordinates 11° 42' 58.16" S, 75° 37' 31.13" W, at an average altitude of 3470 m, with an area of 30.07 km² and 1038 habitants, with a density of 34.5 hab/km². In this district is located the community of the same name whose families are engaged in agricultural work in rainfed conditions (Figure 1).

2.2. Aspecto bioético

The research protocol was evaluated and approved by the Research Institute of the Universidad Nacional del Centro del Perú. The objectives of the study were reported to the mothers, who voluntarily participated in the study, signing the informed consent, agreeing to measure the height and weight of their children.

2.3. Study population

60 peasant families are living in the rural community of Tunanmarca, who have at least one child under 5 years old and reside in the district for more than 5 years. The study worked with the entire population, not requiring a sample; therefore, the results are of high reliability. Of the 60 surveys and anthropometric evaluation of children under 5 years old, 52 had complete information, representing 86.6%. If the families had more than one child under 5 years old, the child in the study was the youngest.



Figure 1. Map of the Peasant Community of Tunanmarca - Jauja, Peru. The community is 76.5 km away from the city of Huancayo and 17.9 km away from the city of Jauja by road. In this place, it is the Tunanmarca Archaeological Center, Pre Inca citadel "Siquillapucara", cultural heritage of Peru. It was a fortified city, with circular and rectangular single-story houses made of stone and mud; it is a tourist destination in the center of Peru.

2.4. Techniques Applied in the Collection of Information and Measuring Instruments

A questionnaire was used as an instrument with questions about socio-economic, productive, food and health practices of families with children under 5 years old, considering indicators on educational level, monthly income and expenditure on family food, housing conditions, employment, breastfeeding, feeding frequency, vaccination, presence of diarrheal and respiratory diseases, early stimulation practices, disposal of excreta and garbage, crops and breeding, used in other studies under similar conditions [10,12,16]. The use of questionnaires in nutritional studies has the advantages of speed, moderate cost, and covers a wide variety of aspects and facilitates the analysis of the data [18].

Anthropometric data was taken from children under 5 years old, using 02 height meters and 02 SECA-Unicef scales, following standardized protocols [14,19,20]. Figure 2 shows the sequence of the research process followed.

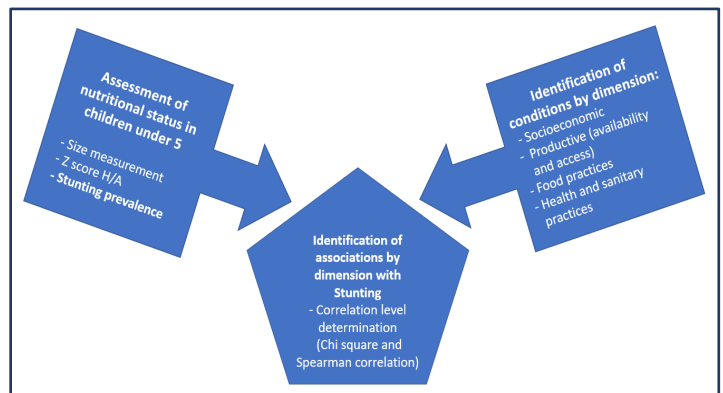


Figure 2. Sequence of the research procedure

2.5. Statistical analysis

The information was tabulated in the SPSS 23 program. The WHO-Anthro 3.2.2 program [14,19,20] was used to determine the Z scores and the nutritional status of the children.

Univariate analysis determines the distribution of data and relative frequencies of each variable studied. The bivariate analysis processed contingency tables with X^2 test and Spearman correlation with 95% confidence.

3. Results

3.1. Socio-economic status of the study families

51.9% of children are male and 48.1% female, aged between 1 and 60 months; 36.5% is the first child. The average age of the mothers is 33 years. The families on average have 5 members, with a range of 2 to 10 members. The weekly economic income of the families on average is S/. 157 soles, with a range of 60 to 250 soles. The average weekly expenditure on food is S/. 106 soles, with a range between 35 and 180 soles; the daily per capita expenditure on food is S/. 3.50 soles (Range: 1.40 to 8.50 soles; Change US dollar = 3.30 soles).

2.1% of parents have no studies, 14.6% have primary, and 72.9% secondary and 10.4% have higher education. In moms, 1.9% have no studies, 42.3% have primary, 51.9% secondary and only 3.8% have higher education. 21.2% of the mothers are single, 23.1% married, 51.9% are living together with the husband and 3.8% are divorced or widowed.

57.7% of families own their homes, 7.7% are renters, and 5.9% are caregivers and 28.8% live in their parents' house. The predominant material of the walls is adobe (84.6%), the rest is concrete and brick. The floor is dirt (88.5%), concrete / ceramic (11.5%) and tile roofs (77%); all homes have electric power. 57.7% of families use firewood for cooking, 5.8% use kerosene and 40.4% use propane gas. 42.3% use only one sleeping room and 48.1% use 2 rooms. 59.6% do not have a place to store firewood and 84.5% do not have a place to store their agricultural tools. 47.7% of kitchens do not have an adequate smoke vent. 53.8% have an exclusive place to store kitchen utensils. In 64% of families, they have animals in the kitchen (guinea pigs, dogs, chickens and cats). 78.8% of families are beneficiaries of some government social support program. In the last 15 days, 40.8% of the parents carried out agricultural work, 32.7% are workers, 16.3% are public / private employees, 8.2% were unemployed and 2% perform outpatient sales. 55.8% of the mothers are engaged in household chores, 32.7% in agriculture and the rest in other independent activities.

3.2. Child malnutrition

7.7% of acute malnutrition (weigh / height) was recorded and in parallel 5.8% overweight, the double burden of malnutrition, due to deficit and excess is observed, which confirms the period of nutritional transition in this community. 9.6% of children have global malnutrition (weight / age).

The stunting prevalence (height / age) determined in the study is 44.2%, including 7.7% of severe stunting (Figure 3), a very critical situation, since the national average is 12% [15]. This result is similar to that reported in Jauja and Janjaillo [11,12], and specifically at the district of Tunamarca, 30.4% of stunting is reported in children under 5 years old [11]. In this study, the prevalence of childhood anemia is 82.7%, an extremely high value associated with conditions of poverty and food insecurity [21].

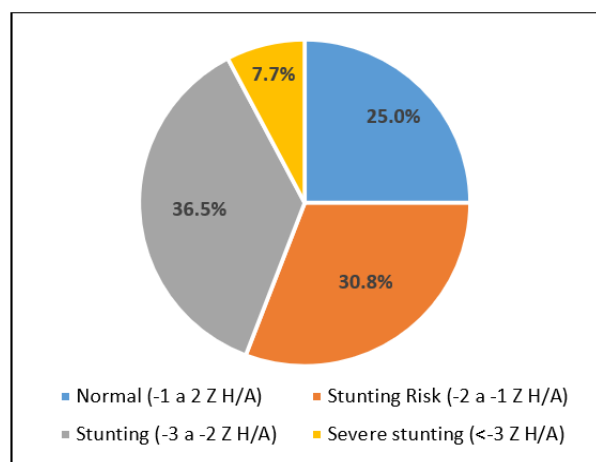


Figure 3. Nutritional status based on Z scores height for age

3.3. Food practices

Of the 52 children under 5-year-old, 88.5% drank breast milk; 65% received before 6 months of age, herbal water, evaporated milk or fresh milk (40%), porridge (30%), juices (5%), family food (5%) and broths (5%). The prevalence of exclusive breastfeeding was 35%.

37% of children consume solid foods more than twice a day and 63% between 3 and 5 times. The most frequently consumed foods are sugar (98%), oil (87.8%), potato (84%), rice (80%), vegetables (57%), fresh milk (47%) and bread (43%). Once a week, they consume olluco (73.5%), beans / lentils (73%), chicken meat (69%), fish (65%), egg (65%), corn (63%), noodles (61, 2%), broad beans / peas (59%), bananas (59%), barley moron (57%), viscera / tripe (55%), quinoa (51%), citrus fruits (51%), guinea pig meat (42.9%). Less frequently, kiwicha, tarwi, maca and red meat are consumed.

3.4. Health practices, early stimulation and sanitation

86.5% of children have ever been vaccinated, but only 75% have all vaccines for their age. The prevalence of diarrhea and acute respiratory diseases was 28.8 and 32.7%, which are attended in the health facility (92.3%), the rest go to the healer or pharmacy. 75% of children are affiliated with the Sistema Integral de Salud del Ministerio de Salud or Seguro Social and 25% do not have any type of health insurance.

63.5, 55.8, 55.8, 46.2 and 38.5% of the mothers to stimulate the development of their children talk to them, caress them, make them play, sing or tell them stories. 55.8, 46.2, 40.4, 32.7, 30.8% of parents make handmade or buy toys, talk to them, play with them, tell them stories or caress them.

48% of families eliminate their feces in the farm / garden and 52% use a bathroom, latrine or pit. 46% throw trash into the farm and 54% burn it or bury it in the farm. 100% of families use previously treated piped water. 75% drink cold boiled water. 73% have a cleanliness place. 55.8% of mothers wash their hands before eating, 57.7% before preparing food, 48.1% after using the bathroom, 40% before breastfeeding and 21.2% after changing the diaper of babies; 71.2% use soap and water and 28.8% use only water.

3.5. Agro-food production

82.7% of families have farmland. 88.4% grow potatoes, 65% board beans, 55.8% barley, 47.6% grow other products. 44.2% of the families reach what is stored for the whole year; at 48.4% for 6 months and 7% for a maximum of 3 months. 23% have a vegetable garden with more than 3 crops (83.3%), including cabbage, lettuce, board beans, onion, carrot, radish, beet, squash, coriander / parsley, caigua, celery, cauliflower, chard, among others. 53.8% have pastures grown for their livestock, the rest only use natural pastures. 92.3% breed some domestic species, including cattle (43%), poultry (61.2%), pigs (47%), guinea pigs (46%) and sheep (41%), all traditionally raised with poor technological level, whose production is generally sold and serves as a family piggy bank.

3.6. Factors associated with stunting

Tables 1, 2 and 3, and figures 4, 5, 6, 7, 8 and 9, show the socioeconomic factors, food and health practices, early stimulation and productive aspects associated with stunting.

Table 1 Socio-economic factors associated with stunting

Factor	X ² : P-value	Spearman's Correlation (r)	P-value of r
Children <30 months of age	0,048	-0,274	0,050
Mother without education or primary	0,001	0,454	0,001
Absence of the father in the home	0,017	0,332	0,016
Not having own home	0,003	0,413	0,002
They use kerosene for cooking	0,045	0,278	0,046
They use gas for cooking	0,000	0,530	0,000
Mothers do not participate in social support	0,032	0,297	0,032
Weekly family income <50 dollars	0,000	0,503	0,000
Weekly food expenditure <31 dollars	0,000	0,648	0,000
Spending per capita / day in food. <.1.10 dollars	0,000	0,591	0,000
Food expenditure <65% of income	0,008	0,370	0,007

P < 0,05 demonstrates significant association and correlation
r = Spearman's correlation

Table 2 Food practices, and health and early stimulation factors associated with stunting

Factor	X ² : P-value	Spearman's Correlation (r)	P-value of r
Do not drink breast milk	0,040	0,284	0,041
Eating food before 6 months of age	0,003	0,410	0,003
Not having had exclusive breastfeeding	0,042	0,290	0,043
Low quinoa consumption frequency	0,003	0,423	0,002
Low frequency of maca consumption	0,010	0,366	0,010
Low frequency of red meat consumption	0,008	0,382	0,007
Low frequency of guinea pig consumption	0,015	0,347	0,014
Low frequency of fish consumption	0,003	0,421	0,003
Presence of childhood anemia	0,003	0,407	0,003
Presence of global malnutrition	0,008	0,366	0,008
Not having all vaccines / age	0,036	0,291	0,037
Presence of diarrhea	0,038	0,288	0,039
Presence of acute respiratory infections	0,038	0,287	0,039
Child without health insurance	0,036	0,291	0,037
Unboiled water consumption	0,036	0,291	0,037
Excreta deposition in open field	0,028	0,306	0,028
Not having a cleanliness place	0,035	0,292	0,035
Mothers do not hands wash, before eating	0,031	0,298	0,032

Mother do not hands wash, after using the bathroom	0,023	0,314	0,023
Washing only with water, without soap	0,038	0,288	0,039
There is no adequate smoke vent in the kitchen	0,000	0,491	0,000
Mother speaks to the child	0,037	0,289	0,038
Mother tells stories to the child	0,027	0,306	0,027
Mother sings to the child	0,043	0,281	0,044
Mother cares for the child	0,031	0,298	0,032
Mother plays with the child	0,031	0,298	0,032
Father speaks to the child	0,043	0,281	0,044

P < 0.05 demonstrates significant association and correlation
r = Spearman's correlation

Table 3 Productive factors associated with stunting

Factor	X ² : P-value	Spearman's Correlation (r)	P-value of r
Not have bio-farm	0,028	0,304	0,028
Do not grow board beans	0,005	0,430	0,004
Do not raise guinea pigs	0,035	0,298	0,036

P < 0.05 demonstrates significant association and correlation
r = Spearman's correlation

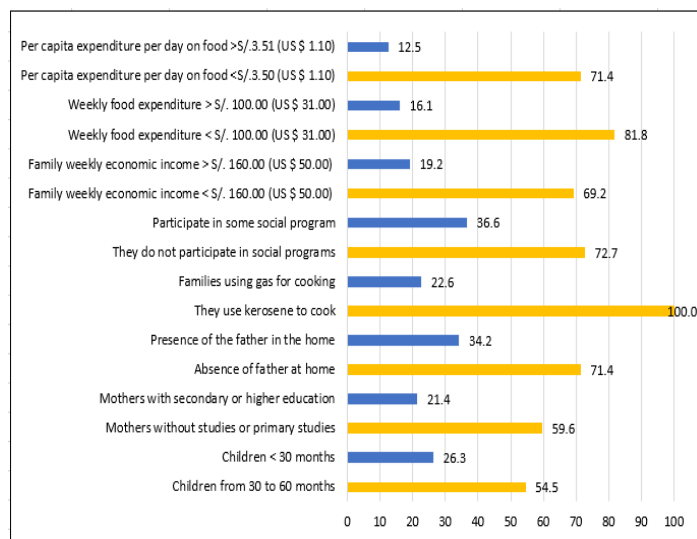


Figure 4. Stunting by socioeconomic variables (%)

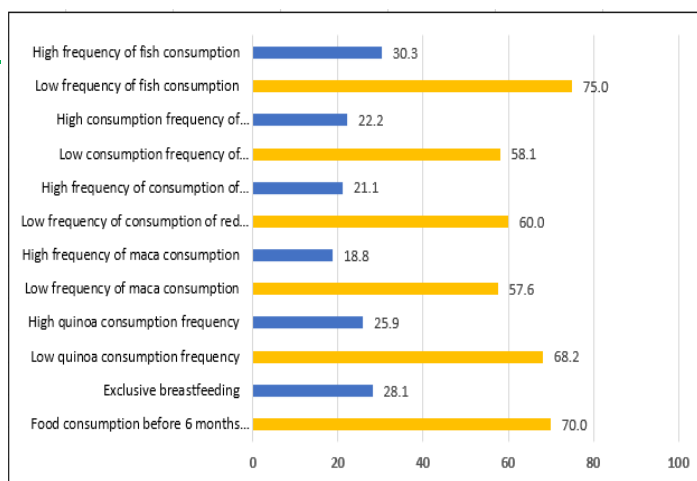


Figure 5. Stunting by food practices (%)

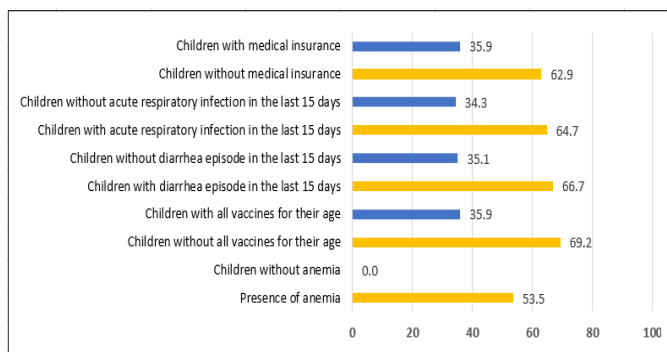


Figure 6. Stunting by health variables (%)

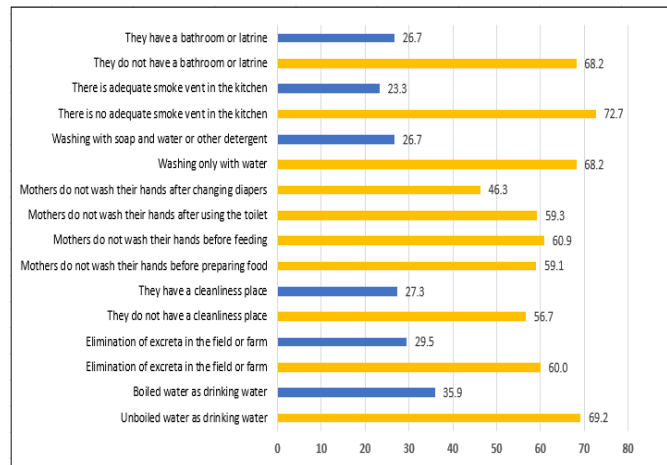


Figure 7. Stunting by health practices (%)

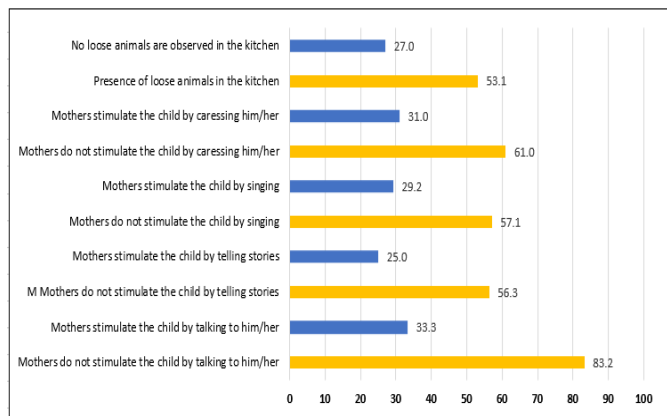


Figure 8. Stunting by development stimulation (%)

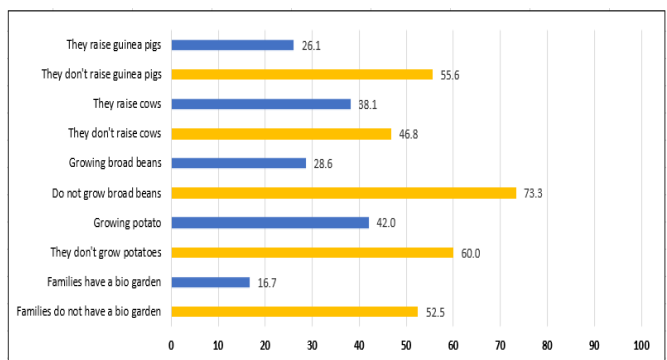


Figure 9. Stunting by agro-food production (%)

4. Discussion

The results of the study are similar to the report of the baseline of the district of Janjaillo-Jauja and other districts of the central highlands of Peru [11], where families produce for self-consumption, the productive yield of nutrient-rich crops necessary for nutrition Healthy is very low [22]. Agricultural communities need to find innovative ways to increase the production and consumption of protective foods to meet their health needs. Similar results were observed in the districts of Pancán, Masma, Masma Chiche, Huamali, Apata, Janjaillo, all from the province of Jauja [12], where the food production of families does not cover their food needs throughout the year and their diet is eminently energetic, being the Main food the potato.

44.2% of children have stunting and were inversely associated with maternal education, immediate consumption of colostrum, increased frequency of protein food consumption, adequate dietary practices, complete vaccination, and early stimulation, boiled water consumption, adequate elimination of excreta, adequate hygiene practices, receiving food support and for raising minor animals. Similarly, in a study carried out in communities in the province of Concepción, Junín, it is observed that agro-food production does not cover the family's food needs, due to the use of inadequate agricultural technologies, with the consumption of eminently energetic diets prevailing; 36.9% of the children had stunting and report that it had an inverse association with the maternal educational level, adequately constituted households, noble material of the walls of the house, drinking water consumption, higher frequency of potato consumption, broad beans, beans, guinea pig, viscera, egg, milk, green vegetables and citrus [1].

In Ricrán-Jauja (community adjacent to the study), the stunting prevalence in children under 5 years old was 49% and was associated with the possession of arable land, technological capacity, educational level and parental employment, food and care of health practices and some aspects related to environmental and sanitation conditions. Children of mothers with higher education do not present a risk of stunting or stunting; with a high stunting prevalence and severe stunting in children whose mothers had only primary or secondary education. In families not properly constituted, the stunting prevalence was higher. The highest prevalence of stunting was recorded in those that have a dirt floor and those that consume well water. The main agro-food variables inversely associated with the high stunting prevalence were the frequency of consumption of chicken meat, eggs, the possession of bio-vegetables with vegetables and the period of food storage, since they store, the stunting percentage for one year was lower compared to those that store for a short time and suffer from marked food insecurity [16].

In our study, the factors that are inversely related to the stunting prevalence are those directly linked to the availability, access, consumption and use of food; as possession of farmland, of bio-gardens and farms that partially cover food needs, under family production systems for self-consumption and sale.

Among the socio-economic factors that show an inverse association with stunting in children under 5 years old in Tunanmarca, there is the maternal educational level, hand washing practices at appropriate times because it reduces the prevalence of

diarrheal diseases [23], having homes properly constituted, be homeowners and not depend on rent.

Among other factors associated with a better nutritional status of children in Tunanmarca is the consumption of breast milk, which protects them against the main causes of infant morbidity and mortality due to their contribution of immunological, nutritional and hormonal factors [24]; reducing the risk of enteric bacterial infections, mainly from *Escherichia coli*, which causes severe diarrhea with increased mortality in the newborn [25] and respiratory diseases such as pharyngitis, a disease that according to one study, it was reduced such as 40% in children who received exclusive breastfeeding [26].

This study recorded a lower prevalence of stunting when the house had adequate smoke vent; smoke exposure is strongly associated with acute respiratory infections in children under five years old [27], reporting that the burden of disease associated with solid fuel use is much more significant in communities with inadequate access to clean fuels, particularly in homes poor and rural areas of developing countries, where it constitutes 2.7% of health risks [28]. In our study, an association was also found between stunting and properly conserving the food produced, since the main factor involved in the origin and prevention of foodborne diseases is food hygiene, ensuring adequate hygiene and conservation [29].

A limitation of the study is the temporal bias that cross-sectional studies have; however, the research shows several strengths, such as the fact of studying several factors simultaneously and providing information that serves as a guide and basis for further studies and for decision-making in comprehensive nutritional interventions [18].

5. Conclusion

The stunting prevalence in children under 5 years old in the high Andean community of Tunanmarca, located in the central Andes of Peru was 44.2% and the factors that are directly associated with the prevalence of stunting were: Do not use gas for cooking, economic income weekly <50.00 dollars, weekly expenditure on family feeding <31.00 dollars, daily expenditure per capita on food <1.10 dollars, low maternal education level, no home ownership, consumption of food before 6 months of exclusive breastfeeding, low frequency of consumption of quinoa and fish, the presence of childhood anemia, inadequate venting of smoke in the kitchen, not having soap for hand washing and personal hygiene, and not having a bathroom.

Stunting is associated with various socioeconomic factors, availability and access to food and some food and health and hygiene practices.

These results demonstrate socio-economic and productive disparities for stunting in rural high Andean areas of central Peru, taking as a model the community of Tunanmarca in Jauja.

Acknowledgments

This study was funded by the Develop Special Fund for the Universidad Nacional del Centro del Perú.

To the mothers and their respective children who voluntarily participated in this study are thanked. To the authorities of the Tunanmarca district.

Conflict of interest

All the authors declare that there is no conflict of interests regarding the publication of this manuscript.

References

- [1] Castro J, Chirinos D, Zenteno F, Situación agroalimentaria y estado nutricional infantil en comunidades de la provincia de Concepción, Junín-Perú. Libro de Resúmenes del XVIII CLN-SALN. 12-14,628. <https://www.slaninternacional.org/congreso2018/aviso/libro-resumenes-congreso.php> (2019).
- [2] Tilman D, Clark M Global diets link environmental sustainability and human health. *Nature* 11/27; 515(7528), 518–522, 2014.
- [3] Gómez MI, Barrett CB, Rancy T et al Post-green revolution food systems and the triple burden of malnutrition. *Food Policy* 42, 129–138, 2013.
- [4] Gross R, Schoeneberger H, Pfeifer H et al The Four Dimensions of Food and Nutrition Security: Definitions and Concepts. *SCN News* 2000 No.20 pp.20-25 ref.13. http://www.fao.org/elearning/course/fa/en/pdf/p-01_rg_concept.pdf (2019).
- [5] Oenema S (2001) La Seguridad Alimentaria en los Hogares. FAO-RLC, Chile. http://www.fao.org/tempref/GI/Reserved/FTP_FaoRlc/old/prior/segalim/nutri/pdf/segalihog.pdf (2019).
- [6] United Nations Transforming our world: the 2030 Agenda for Sustain Dev General Assembly Resolution 70/1, 25. September 2015. <https://sustainabledevelopment.un.org/post2015/transformingourworld> (accessed September 2019).
- [7] Belahsen R Nutrition transition and food sustainability. *Proc Nutr Soc.* 2014, 385–8, 2014.
- [8] Allen LN, Feigl AB Reframing non-communicable diseases as socially transmitted conditions. *Lancet Glob Health.* 2017;5, e644–6, 2017.
- [9] Burchi F, De Muro P From food availability to nutritional capabilities: Advancing food security analysis. *Food Policy* 60, 10-19. DOI: <https://doi.org/10.1016/j.foodpol.2015.03.008>, 2016.
- [10] Castro J, Chirinos D Impact of a Comprehensive Intervention on Food Security in Poor Families of Central Highlands of Peru. *Food and Public Health* 2015, 5(6), 213-219. doi: 10.5923/j.fph.20150506.02.
- [11] Chirinos D Seguridad Alimentaria Nutricional en Poblaciones Vulnerables de la Región Central del Perú. Gráfica JOSIMPRESORES S.A.C. Huancayo, Perú. 270 pp 2015.
- [12] Chirinos D, Castro J Situación agroalimentaria y desnutrición crónica en comunidades de la provincia de Jauja, Junín-Perú. Libro de Resúmenes del XVIII CLN-SALN. 12-14 noviembre 2018, 581-582. <https://www.slaninternacional.org/congreso2018/aviso/libro-resumenes-congreso.php> (2019).
- [13] Gibson RS (2005) Principles of Nutritional Assessment. 2nd ed. Oxford University Press. 908 pp.
- [14] Castro J, Chirinos D Z-Score Anthropometric Indicators Derived from NCHS-1977, CDC-2000 and WHO-2006 in Children Under 5 Years in Central Area of Peru. *Universal Journal of Public Health* 2(2), 73-81, 2014.
- [15] Unicef Estado Mundial de la Infancia 2019 incluye a Perú entre las experiencias exitosas de lucha contra la desnutrición crónica infantil. <https://www.unicef.org/peru/nota-de-prensa/estado-mundial-infancia-nutricion-alimentos-derechos-peru-experiencias-exitosas-desnutricion-cronica-infantil-reporte> (2019).
- [16] Castro J, Chirinos D, Zenteno F Relación entre la producción agroalimentaria y estado nutricional infantil en el distrito de Ricrán-Jauja. Instituto de Investigación de la Universidad Nacional del Centro del Perú, Huancayo 2013.
- [17] DIRESA JUNIN Boletín estadístico de información de salud DIRESA JUNÍN-2017. Dirección Regional de Salud Junín (OITE-2017-RV: 1.0) 2017.
- [18] Gross R, Kielmann A, Korte R et al Guidelines for Nutrition Baseline Surveys in Communities. The Southeast Asian Ministers of Education Organization and GTZ, Jakarta, Bangkok, Thailand, 28p 1997.
- [19] WHO Child growth standards, WHO Anthro (version 3.2.2, 2011) and macros. World Health Organization. <https://www.who.int/childgrowth/software/es/> (2019).

- [20] Chirinos D, Castro J Comparison of NCHS-1977, CDC-2000 and WHO-2006 Nutritional Classification in 32 to 60-Month-old Children in the Central Highlands of Peru. *Universal Journal of Public Health* 1(3), 143-149, 2013.
- [21] Castro J, Chirinos D Prevalencia de anemia infantil y su asociación con factores socioeconómicos y productivos en una comunidad altoandina del Perú. *Rev Esp Nutr Comunitaria*, 25(3) 2019.
- [22] Siegel KR, Ali MK, Srinivasiah A et al Do We Produce Enough Fruits and Vegetables to Meet Global Health Need?. *PLoS One*, 9, Article e104059, 10.1371/journal.pone.0104059, 2014.
- [23] Ejemot-Nwadiaro RI, Ehiri JE, Meremikwu MM et al. Hand washing for preventing diarrhoea. *Cochrane Database of Systematic Reviews* 2008 (1), Art. No: CD004265. doi: 10.1002/14651858.CD004265.pub2.
- [24] González de Cosío-Martínez T, Hernández-Cordero S, Rivera-Dommarco J et al Recomendaciones para una política nacional de promoción de la lactancia materna en México: postura de la Academia Nacional de Medicina. *Salud Pública de México*, 59(1), 106-113, 2017.
- [25] Manthey CF, Autran CA, Eckmann L et al. Human milk oligosaccharides protect against enteropathogenic *Escherichia coli* attachment in vitro and EPEC colonization in suckling mice. *Journal Pediatric Gastroenterology Nutrition*. 58(2), 165-8, 2014.
- [26] Henkle E, Steinhoff M, Omer S et al. The Effect of Exclusive Breast-feeding on Respiratory Illness in Young Infants in a Maternal Immunization Trial in Bangladesh. *Pediatric Infectious Disease Journal*, 2(5), 431-5) 2013.
- [27] Alvis N, De la Hoz F Contaminación del aire domiciliario y enfermedades respiratorias (infección respiratoria aguda baja, EPOC, cáncer de pulmón y asma) evidencias de asociación. *Rev.Fac.Med.* 56, 54-64, 2008.
- [28] Bruce N, Perez-Padilla R, Albalak R Indoor air pollution in developing countries: a major environmental and public health challenge. *Bulletin of the World Health Organization*. 78, 1067-1071, 2008.
- [29] Moreno MG, Alarcón A Higiene alimentaria para la prevención de trastornos digestivos infecciosos y por toxinas. *Rev. Med. Clin. Condes*, 21(5), 749-755, 2010.

Performance Analysis of Joint Precoding and Equalization Design with Shared Redundancy for Imperfect CSI MIMO Systems

Bui Quoc Doanh^{1,2}, Ta Chi Hieu¹, Truong Sy Nam³, Pham Thi Phuong Anh⁴, Pham Thanh Hiep^{*,1}

¹Faculty of Radio-Electronics Engineering, Le Quy Don Technical University, 100000, Vietnam

²Faculty of Professional Telecommunication, University of Telecommunications, 650000, Vietnam

³Center of Informations and Foreign Language, Ha Tinh medical College, 480000, Vietnam

⁴Missile Institute, Institute of military science and technology, 100000, Vietnam

ARTICLE INFO

Article history:

Received: 20 March, 2020

Accepted: 26 April, 2020

Online: 10 May, 2020

Keywords:

Imperfect CSI

Cyclic prefix

Joint Precoding & Equalization

MIMO ISI channel

Protective interval

Unweighted MMSE

ABSTRACT

Analytical researches on a potential performance of multipath multiple-input multiple-output (MIMO) systems inspire the development of new technologies that decompose a MIMO channel into independent sub-channels on the condition of constrained transmit power. Moreover, in current studies of inter-symbol interference (ISI) MIMO systems, there is an assumption that channel state information (CSI) at receivers and/or transmitters is perfect. In this paper, we propose a hybrid design of precoding and equalization schemes based on the unweighted minimum mean square error criterion that not only eliminates the ISI but also improves the system performance. Additionally, the impact of imperfect channel knowledge at receivers on the system performance of MIMO ISI system is investigated. The simulation result shown that the proposed hybrid design of precoding and equalization with shared redundancy outperforms the conventional method in all considered scenarios. Furthermore, the proposed and the conventional schemes are extremely sensitive to the CSI factor; the performance of these systems is quickly deteriorated when the accuracy of channel estimation decreases.

1 Introduction

For satisfying the demand of advanced communications in terms of high data rates, relatively low costs and high-quality services, it is necessary to constantly improve the performance of wireless communication systems [1, 2]. However, protective intervals such as cyclic prefix or zero padding intervals are inserted to block transmission systems because of inter-symbol interference (ISI), which results in a decline in spectral efficiencies, especially in multiple-input multiple-output (MIMO) ISI channels having a long impulse response [3, 4]. In order to overcome the decrease of spectral efficiency and enhance the system performance, which attract lots of attention from the worldwide researchers. For example, redundancies are used instead of protector intervals throughout the signal processing to remove the ISI as in [5]-[9]. Moreover, various approaches based on the precoding scheme or the hybrid of precoding and equalization schemes of MIMO ISI channels were proposed

[10]-[17].

In many approaches, it is assumed that channel status information (CSI) is able to be obtained perfectly. However, in practical scenarios, the perfect CSI is difficult to achieve in MIMO ISI systems because of the imperfection of channel estimation, feedback delay and finite rate channel quantization. Therefore, the research on a negative influences of imperfect CSI in MIMO ISI systems is necessary. A number of researches in this field developing precoders or connecting precoder and equalizer schemes with respect to the CSI at transmitters, receivers or transceivers have been published. There are many surveys expressed to analyze the performance of the MIMO system with imperfect CSI as follows.

Firstly, several surveys with the negative influences of imperfect CSI at transmitters were proposed and analyzed in [18]-[22]. To give a clear example, a MIMO system with joint decoding technique evaluated transmission rates which are affected by imperfect CSI [18]. Multi-user MIMO filterbank multicarrier systems employ-

*Corresponding Author: Pham Thanh Hiep, Faculty of Radio-Electronics Engineering, Le Quy Don Technical University, 236 Hoang Quoc Viet, Hanoi, 100000, Vietnam, Email: phamthanhiep@gmail.com

ing the Zero Forcing technique are affected by imperfect CSI as demonstrated in [19]. The obtainable transmission rate of multipath channel systems with decoding technique was described in [20]. Analysis and investigation of MIMO systems occupying Tomlinson-Harashima precoding are mentioned in [21, 22].

Next, there have been also some papers studied the negative influences of imperfect CSI at the receiver [23]-[27]. For instance, the achievable transmission rate was evaluated in [23] by assumption of knowing the transmitted training symbols. The impact of CSI and feedback quantization error bring to MIMO systems in term of adaptive modulation was investigated in [24]. The effect of CSI to a capacity of MIMO systems was studied in [25]-[27].

Finally, an investigation about the combination of linear precoding and decoding to minimize the total mean-square error of MIMO systems was shown in [28, 29] when there is the imperfect CSI at transceivers. Furthermore, the negative influences of the imperfect CSI on a resource allocation of base stations by using several techniques such as transmit antennas selection, power allocation and beamforming was considered in [30]-[32].

In this work, we focus on combination of both precoding and equalization schemes, which employ redundancies for MIMO ISI systems. The proposed scheme, shared redundancies for both transmitters and receivers are used to perform a combination design of precoder and equalizer. In the next part, an investigation of the negative influence of imperfect CSI on the performance of these systems will be performed. Herein, the length of protective intervals generally defines redundancies as in [8].

The contribution of the research is summarized as follows.

- The proposed joint scheme of precoding and equalization schemes with shared redundancy is mathematical analyzed in both perfect and imperfect CSI.
- The conventional methods, such as training zero and leading zero, are investigated to compare with the proposed scheme.
- The impact of imperfect CSI on the system performance is discussed based on bit error rate and channel capacity.
- The proposed scheme is evaluated based on different parameters, such as the accuracy of channel estimation, the order of the finite impulse response and the transmission block size.

The order of other parts in this paper are as follows: In Section 2, the system model with imperfect CSI is introduced, while both precoding and equalization techniques are combined for the MIMO ISI systems, and then the combination of them are demonstrated. The negative influences of imperfect CSI on the performance of system is analyzed in Section 3. Simulation results and conclusions are shown in Section 4 and Section 5, respectively. Notations in this paper are used as follows: sets of complex numbers are represented by symbol \mathbb{C} , while $(\cdot)^H$ denotes the conjugate transpose, $(\cdot)^T$ is the transpose and boldface font is used for vector and matrix.

2 System Model

In this work, joint scheme of precoding and equalization for MIMO ISI channels of block transmission system models is proposed as

illustrated in Fig.1. M_T and M_R denote the transmitting and receiving antennas, The channel model of this system is assumed to be frequency selective fading. Moreover, the finite impulse response (FIR) of channel has an order of D , in which the channel impulse response (CIR) is expressed by matrices $\tilde{\mathbf{H}}[0], \tilde{\mathbf{H}}[1], \dots, \tilde{\mathbf{H}}[D]$, herein, $\tilde{\mathbf{H}}[d] \in \mathbb{C}^{M_T \times M_R}$, ($d = 0, \dots, D$).

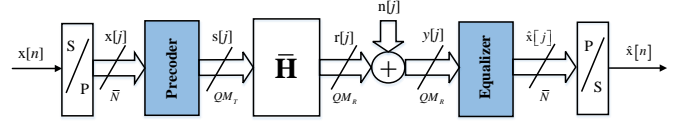


Figure 1: Hybrid scheme of precoder and equalizer in MIMO ISI systems

The system model of the MIMO ISI channel with joint scheme of precoder and equalizer in Figure 1 is expressed as follows. At the transmitter, a single input stream $\mathbf{x}[n]$ is transferred to a serial-to-parallel converter, and then converted into a block of vectors $\mathbf{x}[j]$ with a block size of $\bar{N} \times 1$. Then, the vector $\mathbf{x}[j]$ is sent to the precoder. The precoder performs signal processing procedure so as to perform symbol vector generation $\mathbf{x}[j]$ with the dimension of $QM_T \times 1$ from input vector $\mathbf{s}[j]$. Afterwards, the vector $\mathbf{s}[j]$ is divided into Q vectors, each of which has the size of $M_T \times 1$, and converted by the parallel-to-serial converter, then transmitted over the MIMO ISI channel. At the output side, the received symbol vector $\mathbf{y}[j]$ has two parts because of the affection of noise, including an information vector $\mathbf{r}[j]$ and a noise sample vector $\mathbf{n}[j]$ that is considered as an additive white Gaussian noise sample vector, and $\mathbf{n}[j] \sim CN(0, 1)$.

The signal processing procedure is employed conversely at the receiver. $QM_R \times 1$ symbol vector $\mathbf{y}[j]$ is formed by the Q received vectors in the serial-to-parallel converter. Subsequently, the symbol vector $\mathbf{y}[j]$ is sent to the equalizer in order to regenerate the original symbol vector $\hat{\mathbf{x}}[i]$ with the dimension of $\bar{N} \times 1$. Finally, the parallel-to-serial converter reforms the output symbol stream $\hat{\mathbf{x}}[n]$ from the symbol vector $\hat{\mathbf{x}}[i]$.

Based on the operation and signal processing of system mentioned above, the terms defined as $\mathbf{x}[j]$, $\mathbf{s}[j]$, $\mathbf{y}[j]$, $\hat{\mathbf{x}}[i]$, $\mathbf{r}[j]$ and $\mathbf{n}[j]$ according to the input symbol stream $\mathbf{x}[n]$ and the sampled vector of received signal $\hat{\mathbf{x}}[n]$, can be mathematically shown in the (1) - (6) equations, respectively.

$$\mathbf{x}[j] = [x[j\bar{N}], x[j\bar{N} + 1], \dots, x[j\bar{N} + \bar{N} - 1]]^T \quad (1)$$

$$\mathbf{s}[j] = [s[jQM_T], s[jQM_T + 1], \dots, s[jQM_T + QM_T - 1]]^T \quad (2)$$

$$\mathbf{y}[j] = [y[jQM_R], y[jQM_R + 1], \dots, y[jQM_R + QM_R - 1]]^T \quad (3)$$

$$\hat{\mathbf{x}}[j] = [\hat{x}[j\bar{N}], \hat{x}[j\bar{N} + 1], \dots, \hat{x}[j\bar{N} + \bar{N} - 1]]^T \quad (4)$$

$$\mathbf{r}[j] = [r[jQM_R], r[jQM_R + 1], \dots, r[jQM_R + QM_R - 1]]^T \quad (5)$$

$$\mathbf{n}[j] = [n[jQM_R], n[jQM_R + 1], \dots, n[jQM_R + QM_R - 1]]^T \quad (6)$$

3 Performance Analysis

3.1 Joint Scheme of Precoder and Equalizer

In this section, the proposed joint scheme of precoder and equalizer based on MIMO ISI channels is explained. In addition, the channel

model is assumed to be narrow-band, therefore the Saleh-Valenzuela model [33] can be adopted. In the case of $Q \geq D$, the symbol $\hat{\mathbf{x}}[j]$ is expressed following the ideas in [6]

$$\hat{\mathbf{x}}[j] = \mathbf{G}\tilde{\mathbf{H}}_0\mathbf{F}\mathbf{x}[j] + \mathbf{G}\tilde{\mathbf{H}}_1\mathbf{F}\mathbf{x}[j-1] + \mathbf{G}\mathbf{n}[j] \quad (7)$$

where $\mathbf{n}[j] \in \mathbb{C}^{QM_R \times 1}$; $\mathbf{F} \in \mathbb{C}^{QM_T \times \tilde{N}}$ and $\mathbf{G} \in \mathbb{C}^{\tilde{N} \times QM_R}$ are respectively the precoding and the equalization matrices, and expressed by following equations

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_0[0] & \cdots & \mathbf{F}_{\tilde{N}-2}[0] & \mathbf{F}_{\tilde{N}-1}[0] \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{F}_0[QM_T-2] & \cdots & \mathbf{F}_{\tilde{N}-2}[QM_T-2] & \mathbf{F}_{\tilde{N}-1}[QM_T-2] \\ \mathbf{F}_0[QM_T-1] & \cdots & \mathbf{F}_{\tilde{N}-2}[QM_T-1] & \mathbf{F}_{\tilde{N}-1}[QM_T-1] \end{bmatrix} \quad (8)$$

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_0[0] & \cdots & \mathbf{G}_{(QM_R-2)}[0] & \mathbf{G}_{(QM_R-1)}[0] \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{G}_0[\tilde{N}-2] & \cdots & \mathbf{G}_{(QM_R-1)}[\tilde{N}-2] & \mathbf{G}_{(QM_R-1)}[\tilde{N}-2] \\ \mathbf{G}_0[\tilde{N}-1] & \cdots & \mathbf{G}_{(QM_R-1)}[\tilde{N}-1] & \mathbf{G}_{(QM_R-1)}[\tilde{N}-1] \end{bmatrix} \quad (9)$$

Additionally, $\tilde{\mathbf{H}}_0$ and $\tilde{\mathbf{H}}_1$ are utilized to give a definition of the Toeplitz matrices whose uniform dimension is $QM_R \times QM_T$, and expressed by following equations

$$\tilde{\mathbf{H}}_0 = \begin{bmatrix} \tilde{\mathbf{H}}[0] & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \vdots & \tilde{\mathbf{H}}[0] & \mathbf{0} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \tilde{\mathbf{H}}[D-1] & \vdots & \ddots & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \tilde{\mathbf{H}}[D] & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \ddots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \vdots & \cdots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \tilde{\mathbf{H}}[D] & \tilde{\mathbf{H}}[D-1] & \cdots & \tilde{\mathbf{H}}[0] \end{bmatrix} \quad (10)$$

$$\tilde{\mathbf{H}}_1 = \begin{bmatrix} \mathbf{0} & \cdots & \tilde{\mathbf{H}}[D] & \tilde{\mathbf{H}}[D-1] & \cdots & \tilde{\mathbf{H}}[2] & \tilde{\mathbf{H}}[1] \\ \mathbf{0} & \cdots & \mathbf{0} & \tilde{\mathbf{H}}[D] & \ddots & \cdots & \tilde{\mathbf{H}}[2] \\ \vdots & \ddots & \vdots & \mathbf{0} & \ddots & \ddots & \vdots \\ \mathbf{0} & \ddots & \ddots & \vdots & \ddots & \ddots & \tilde{\mathbf{H}}[D-1] \\ \mathbf{0} & \cdots & \ddots & \ddots & \ddots & \ddots & \tilde{\mathbf{H}}[D] \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (11)$$

The effect of ISI on the MIMO ISI system model shown in Fig. 1 is the second component on the right side of the equation (7), $(\mathbf{G}\tilde{\mathbf{H}}_1\mathbf{F}\mathbf{x}[j-1])$. With an assumption of $QM_T = \tilde{M} + DM_T$, ($\tilde{N} \leq \tilde{M}$), two conditional methods (one is named by the trailing zero (TrZero) and the other one is the leading zero (LeZero) [6]) can be used to cancel the ISI.

Actually, although the performance of both TrZero and LeZero technical schemes is almost similar, the TrZero scheme is solely considered. For the TrZero technical scheme, the equalization matrix keeps the same ($\mathbf{G}_{\text{TrZero}} = \mathbf{G}$), whereas the precoding matrix

changes the last DM_T rows to be zero. The precoding matrix has the form

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_0[0] & \cdots & \mathbf{F}_{(\tilde{N}-2)}[0] & \mathbf{F}_{(\tilde{N}-1)}[0] \\ \mathbf{F}_0[1] & \cdots & \mathbf{F}_{(\tilde{N}-2)}[1] & \mathbf{F}_{(\tilde{N}-1)}[1] \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{F}_0[\tilde{M}-2] & \cdots & \mathbf{F}_{(\tilde{N}-2)}[\tilde{M}-2] & \mathbf{F}_{(\tilde{N}-1)}[\tilde{M}-2] \\ \mathbf{F}_0[\tilde{M}-1] & \cdots & \mathbf{F}_{(\tilde{N}-2)}[\tilde{M}-1] & \mathbf{F}_{(\tilde{N}-1)}[\tilde{M}-1] \\ \mathbf{0} & \vdots & \mathbf{0} & \mathbf{0} \\ \vdots & \cdots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (12)$$

where the definition of $\mathbf{F}_{\text{TrZero}} \in \mathbb{C}^{(Q-D)M_T \times \tilde{N}}$ is as follows

$$\mathbf{F}_{\text{TrZero}} = \begin{bmatrix} \mathbf{F}_0[0] & \cdots & \mathbf{F}_{(\tilde{N}-2)}[0] & \mathbf{F}_{(\tilde{N}-1)}[0] \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{F}_0[\tilde{M}-2] & \cdots & \mathbf{F}_{(\tilde{N}-2)}[\tilde{M}-2] & \mathbf{F}_{(\tilde{N}-1)}[\tilde{M}-2] \\ \mathbf{F}_0[\tilde{M}-1] & \cdots & \mathbf{F}_{(\tilde{N}-2)}[\tilde{M}-1] & \mathbf{F}_{(\tilde{N}-1)}[\tilde{M}-1] \end{bmatrix} \quad (13)$$

Subsequently, the optimal criteria is utilized to design the $\mathbf{F}_{\text{TrZero}}$ precoder and the $\mathbf{G}_{\text{TrZero}}$ equalizer jointly to improve the system performance. Thus, when $\mathbf{G}\tilde{\mathbf{H}}_1\mathbf{F}\mathbf{x}[j-1]$ is disappeared, it means that the ISI is not completely existent, the equation (7) becomes as follows

$$\hat{\mathbf{x}}[j] = \mathbf{G}_{\text{TrZero}}\tilde{\mathbf{H}}\mathbf{F}_{\text{TrZero}}\mathbf{x}[j] + \mathbf{G}_{\text{TrZero}}\mathbf{n}[j]. \quad (14)$$

where the $\tilde{\mathbf{H}}$ matrix comprises $(Q-D)M_T$ first columns of the $\tilde{\mathbf{H}}_0$ matrix and is given by

$$\tilde{\mathbf{H}} = \begin{bmatrix} \tilde{\mathbf{H}}[0] & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \tilde{\mathbf{H}}[0] & \ddots & \cdots & \mathbf{0} \\ \tilde{\mathbf{H}}[D-1] & \ddots & \ddots & \ddots & \vdots \\ \tilde{\mathbf{H}}[D] & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \ddots & \tilde{\mathbf{H}}[0] \\ \mathbf{0} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \tilde{\mathbf{H}}[D-1] \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \tilde{\mathbf{H}}[D] \end{bmatrix} \quad (15)$$

When the transmit power is limited by p'_0 , the optimal MMSE criterion for joint precoder and equalizer schemes is calculated by

$$\mathbf{F}_{\text{TrZero}} = \mathbf{V}\mathbf{\Phi}\mathbf{U}^H \quad (16)$$

$$\mathbf{G}_{\text{TrZero}} = \mathbf{R}_{xx}\mathbf{F}_{\text{TrZero}}^H\tilde{\mathbf{H}}^H(\mathbf{R}_{mm} + \tilde{\mathbf{H}}\mathbf{F}_{\text{TrZero}}\mathbf{R}_{xx}\mathbf{F}_{\text{TrZero}}^H\tilde{\mathbf{H}}^H)^{-1}, \quad (17)$$

where \mathbf{U} and \mathbf{V} are unitary matrices obtained from the eigenvalue decompositions (EVD) algorithm, and $\mathbf{\Phi}$ is a diagonal matrix, whose main diagonal elements $\mathbf{\Lambda}$ matrix can be obtained from a water-filling algorithm,

$$\mathbf{R}_{xx} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H, \quad (18)$$

$$\tilde{\mathbf{H}}^H\mathbf{R}_{mm}^{-1}\tilde{\mathbf{H}} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H, \quad (19)$$

where \mathbf{R}_{xx} and \mathbf{R}_m are the covariance matrices of the input signal and noise, respectively.

Obviously, the signal to interference noise ratios (SINRs) among the decoupled flat sub-channels are not similar because of the unequal eigenvalues obtained from (19). Hence, the sub-channel with low SINR significantly influences the performance of the system, it means that those sub-channels should be discarded to enhance the system BER [17]. In addition, when the protective interval or the redundancy is utilized, the ISI interference can be cancelled in the frequency selective channel, however the addition of the protective interval or the redundancy lets channel energy be lost.

For solving the above issue, the ideas expressed in [6, 9] were combined and both precoder and equalizer technique schemes relied on the *unweighted* MMSE criterion [10] was proposed. This proposal aimed to share redundancies to both the transmitter and the receiver. Especially, instead of setting the last DM_T rows of the \mathbf{F} precoding matrix, only last $\bar{K}M_T = \lfloor \frac{DM_T}{2} \rfloor$ rows are set to zero, while the first $(D - \bar{K})M_R$ columns of the \mathbf{G} equalization matrix are also set to zero at the receiver. Herein, the transmitter and the receiver obtain the shared protective interval, it means that the first $(D - \bar{K})M_R$ rows of the $\tilde{\mathbf{H}}_0$ channel matrix and the last $\bar{K}M_T$ columns of $\tilde{\mathbf{H}}_0$ channel matrix are removed by the \mathbf{G} equalization matrix and the \mathbf{F} precoding matrix, respectively. Consequently, the loss in $\tilde{\mathbf{H}}_0$ channel matrix is able to reduced by the proposed design. In the other words, the partial channel energy loss can be reduced. Basing on the above analysis, in our method, the precoder and equalizer are designed as follows (20) and (21).

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_0[0] & \mathbf{F}_1[0] & \cdots & \mathbf{F}_{(\bar{N}-1)}[0] \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{F}_0[(Q-\bar{K})M_T] & \mathbf{F}_1[(Q-\bar{K})M_T] & \cdots & \mathbf{F}_{\bar{N}-1}[(Q-\bar{K})M_T] \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}, \quad (20)$$

$$\mathbf{G} = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & \mathbf{G}_0[0] & \cdots & \mathbf{G}_{(Q-D+\bar{K})M_R-1}[0] \\ \vdots & \cdots & \vdots & \mathbf{G}_0[1] & \cdots & \mathbf{G}_{(Q-D+\bar{K})M_R-1}[1] \\ \mathbf{0} & \cdots & \mathbf{0} & \vdots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{G}_0[\bar{N}-1] & \cdots & \mathbf{G}_{(Q-D+\bar{K})M_R-1}[\bar{N}-1] \end{bmatrix}. \quad (21)$$

Where $\hat{\mathbf{F}} \in \mathbb{C}^{(Q-\bar{K})M_T \times N}$ and $\hat{\mathbf{G}} \in \mathbb{C}^{N \times (Q-D+\bar{K})M_R}$ are given by following equations (22) and (23).

$$\hat{\mathbf{F}} = \begin{bmatrix} \mathbf{F}_0[0] & \mathbf{F}_1[0] & \cdots & \mathbf{F}_{(\bar{N}-1)}[0] \\ \mathbf{F}_0[1] & \mathbf{F}_1[1] & \cdots & \mathbf{F}_{(\bar{N}-1)}[1] \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{F}_0[(Q-\bar{K})M_T] & \mathbf{F}_1[(Q-\bar{K})M_T] & \cdots & \mathbf{F}_{\bar{N}-1}[(Q-\bar{K})M_T] \end{bmatrix}, \quad (22)$$

$$\hat{\mathbf{G}} = \begin{bmatrix} \mathbf{G}_0[0] & \mathbf{G}_1[0] & \cdots & \mathbf{G}_{(Q-D+\bar{K})M_R-1}[0] \\ \mathbf{G}_0[1] & \mathbf{G}_1[1] & \cdots & \mathbf{G}_{(Q-D+\bar{K})M_R-1}[1] \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{G}_0[\bar{N}-1] & \mathbf{G}_1[\bar{N}-1] & \cdots & \mathbf{G}_{(Q-D+\bar{K})M_R-1}[\bar{N}-1] \end{bmatrix}. \quad (23)$$

The *unweighted* MMSE criterion will be used to design the $\hat{\mathbf{F}}$ precoder and the $\hat{\mathbf{G}}$ equalizer [15]. weak eigenmodes are dropped, so they do not affect the BER of system. The transmit power is redistributed for the remaining eigenvalues. Consequently, power distribution to the higher eigenvalues is larger than that of the MIMO ISI channel.

When the ISI cancellation is complete, the equation (7) can be expressed as follows

$$\hat{\mathbf{x}}[j] = \hat{\mathbf{G}}\hat{\mathbf{H}}\hat{\mathbf{F}}\mathbf{x}[j] + \hat{\mathbf{G}}\mathbf{n}'[j] \quad (24)$$

where the $\hat{\mathbf{H}}$ channel matrix is defined as

$$\hat{\mathbf{H}} = \begin{bmatrix} \tilde{\mathbf{H}}[D-\bar{K}] & \cdots & \tilde{\mathbf{H}}[0] & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \tilde{\mathbf{H}}[D-1] & & & & & & \mathbf{0} \\ \tilde{\mathbf{H}}[D] & \ddots & & & & & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & & & & \tilde{\mathbf{H}}[0] \\ \vdots & \ddots & \ddots & \ddots & & & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \tilde{\mathbf{H}}[D] & \tilde{\mathbf{H}}[D-1] & \cdots & \tilde{\mathbf{H}}[\bar{K}] \end{bmatrix}, \quad (25)$$

herein, the noise sample block with the length of $(Q - D + \bar{K})M_R$ is denoted as $\mathbf{n}'[j]$.

3.2 In the Case of Imperfect CSI

Factually, the perfect CSI cannot be produced at both the transmitter and the receiver. In this work, we assume that the perfect CSI exists at the transmitter, which means that its feedback is immediate and error-free, whereas the receiver has the imperfect CSI. Herein, the accurate channel matrix and the inaccurate channel matrix with a complex Gaussian distribution are denoted as $\tilde{\mathbf{H}}$ and $\hat{\mathbf{H}}_e$, respectively, in which $\hat{\mathbf{H}}_e \in CN(0, 1)$. Thus, the illustration of the channel matrix is as follows [34, 35]

$$\tilde{\mathbf{H}} = \xi\hat{\mathbf{H}} + \sqrt{1 - \xi^2}\hat{\mathbf{H}}_e. \quad (26)$$

The sizes of these matrices are $(Q - D + \bar{K})M_R \times (Q - \bar{K})M_T$, and the accuracy of channel estimation is illustrated by ξ , where $0 \leq \xi \leq 1$. Therefore, the equation (24) is rewritten as

$$\hat{\mathbf{x}}[j] = \tilde{\mathbf{G}}\tilde{\mathbf{H}}\hat{\mathbf{F}}\mathbf{x}[j] + \tilde{\mathbf{G}}\mathbf{n}'[j] \quad (27)$$

The $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{G}}$ matrices of the precoder and the equalizer are designed under the imperfect CSI at the receiver, and optimized in accordance with the *unweighted* MMSE criterion [15]. Therefore, the mathematical description is expressed as

$$\begin{aligned} \min_{\tilde{\mathbf{G}}, \tilde{\mathbf{F}}} : c(\tilde{\mathbf{G}}, \tilde{\mathbf{F}}) &= E\|\mathbf{W}^{1/2}\mathbf{e}\|^2, \\ \text{Subject to: } tr(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H) &\leq p'_0, \end{aligned} \quad (28)$$

where $\mathbf{e} = \mathbf{x}[j] - (\tilde{\mathbf{G}}\tilde{\mathbf{H}}\tilde{\mathbf{F}}\mathbf{x}[j] + \tilde{\mathbf{G}}\mathbf{n}'[j])$, the weight matrix $\mathbf{W} = \mathbf{I}$ and the transmit power is constrained to p'_0 . The expectation (E) relates to the distribution of \mathbf{x} and \mathbf{n} .

$$c(\tilde{\mathbf{G}}, \tilde{\mathbf{F}}) = E\|\mathbf{e}\|^2 = E(tr[\mathbf{e}\mathbf{e}^H]) = tr[\text{Re}(\tilde{\mathbf{G}}, \tilde{\mathbf{F}})], \quad (29)$$

where the error covariance matrix is denoted the $\text{Re}(\tilde{\mathbf{G}}, \tilde{\mathbf{F}})$ and defined as $\text{Re}(\tilde{\mathbf{G}}, \tilde{\mathbf{F}}) := E(\mathbf{e}\mathbf{e}^H)$. Using the expression for \mathbf{e} and the above assumption, we have

$$E(\mathbf{x}\mathbf{x}^H) = \mathbf{I}, \quad (30)$$

$$E(\mathbf{n}'\mathbf{n}'^H) = \mathbf{R}_{\mathbf{n}'\mathbf{n}'}, \quad (31)$$

$$E(\mathbf{x}\mathbf{n}'^H) = 0. \quad (32)$$

Therefore,

$$\text{Re}(\tilde{\mathbf{G}}, \tilde{\mathbf{F}}) = \left[\tilde{\mathbf{G}}(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e)\tilde{\mathbf{F}} - \mathbf{I} \right] \times \left[\tilde{\mathbf{G}}(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e)\tilde{\mathbf{F}} - \mathbf{I} \right]^H + \tilde{\mathbf{G}}\mathbf{R}_{\mathbf{n}'\mathbf{n}'}\tilde{\mathbf{G}}^H. \quad (33)$$

For solving the optimization problem in (28), the method of Lagrange duality and the Karush-Kuhn-Tucker (KKT) conditions are utilized, where η is the Lagrange multiplier.

$$L(\eta, \tilde{\mathbf{G}}, \tilde{\mathbf{F}}) = c(\tilde{\mathbf{G}}, \tilde{\mathbf{F}}) + \eta \left[\text{tr}(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H) - p'_0 \right]. \quad (34)$$

Substituting (29) and (33) into (34), we have

$$L(\eta, \tilde{\mathbf{G}}, \tilde{\mathbf{F}}) = \text{tr} \left\{ \begin{aligned} & \left[\tilde{\mathbf{G}}(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e)\tilde{\mathbf{F}} - \mathbf{I} \right] \times \\ & \left[\tilde{\mathbf{G}}(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e)\tilde{\mathbf{F}} - \mathbf{I} \right]^H + \tilde{\mathbf{G}}\mathbf{R}_{\mathbf{n}'\mathbf{n}'}\tilde{\mathbf{G}}^H \end{aligned} \right\} \\ + \eta \left[\text{tr}(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H) - p'_0 \right] \\ = \text{tr} \left\{ \begin{aligned} & \left[\tilde{\mathbf{G}}(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e)\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e)^H\tilde{\mathbf{G}}^H \right] - \\ & \tilde{\mathbf{G}}(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e)\tilde{\mathbf{F}} - \tilde{\mathbf{F}}^H(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e)^H\tilde{\mathbf{G}}^H \\ & + \mathbf{I} + \tilde{\mathbf{G}}\mathbf{R}_{\mathbf{n}'\mathbf{n}'}\tilde{\mathbf{G}}^H \\ & + \eta \left[\text{tr}(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H) - p'_0 \right] \end{aligned} \right\} \quad (35)$$

By applying the KKT condition, the $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{G}}$ matrices are optimal only if there is an η satisfying the following condition.

$$\nabla_{\tilde{\mathbf{F}}} L(\eta, \tilde{\mathbf{F}}, \tilde{\mathbf{G}}) = 0, \quad (36)$$

$$\nabla_{\tilde{\mathbf{G}}} L(\eta, \tilde{\mathbf{F}}, \tilde{\mathbf{G}}) = 0, \quad (37)$$

$$\eta \geq 0; \text{tr}(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H) - p'_0 \leq 0, \quad (38)$$

$$\eta \left[\text{tr}(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H) - p'_0 \right] = 0. \quad (39)$$

Applying the derivatives of $L(\eta, \tilde{\mathbf{G}}, \tilde{\mathbf{F}})$ with respect to $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{G}}$ as in [36] and substitute (35) into (36) and (37). As a result, the $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{G}}$ matrices can be calculated as follows

$$\left(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e \right) \tilde{\mathbf{F}} = \left(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e \right) \tilde{\mathbf{F}}\tilde{\mathbf{F}}^H \times \left(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e \right)^H \tilde{\mathbf{G}}^H + \mathbf{R}_{\mathbf{n}'\mathbf{n}'}\tilde{\mathbf{G}}^H. \quad (40)$$

$$\tilde{\mathbf{G}} \left(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e \right) = \tilde{\mathbf{F}}^H \left(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e \right)^H \tilde{\mathbf{G}}^H \tilde{\mathbf{G}} \times \left(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e \right) + \eta \tilde{\mathbf{F}}^H. \quad (41)$$

where $\tilde{\mathbf{F}}, \tilde{\mathbf{G}}$ are designed by the *unweighted* MMSE criterion. Consequently, the optimal $\tilde{\mathbf{F}}, \tilde{\mathbf{G}}$ matrices are derived as follows

$$\tilde{\mathbf{F}} = \tilde{\mathbf{V}}\tilde{\Phi}_f, \quad (42)$$

$$\tilde{\mathbf{G}} = \tilde{\Phi}_g \tilde{\mathbf{V}}^H \left(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e \right)^H \mathbf{R}_{\mathbf{n}'\mathbf{n}'}^{-1}, \quad (43)$$

where the $\eta, \tilde{\Phi}_f$ and $\tilde{\Phi}_g$ are given by

$$\eta^{1/2} = \frac{\sum_{i=1}^k (\tilde{\lambda}_{ii}^{-1/2})}{p'_0 + \sum_{i=1}^k (\tilde{\lambda}_{ii}^{-1})}, \quad (44)$$

$$\tilde{\Phi}_f = \left(\eta^{-1/2} \tilde{\Lambda}^{-1/2} - \tilde{\Lambda}^{-1} \right)_+^{1/2}, \quad (45)$$

$$\tilde{\Phi}_g = \left(\eta^{-1/2} \tilde{\Lambda}^{-1/2} - \eta \tilde{\Lambda}^{-1} \right)_+^{1/2} \tilde{\Lambda}^{-1/2}. \quad (46)$$

Applying (44) in (45) and (46), we can calculate as

$$|\tilde{\phi}_{f,jj}|^2 = \left[\frac{p'_0 + \sum_{i=1}^k (\tilde{\lambda}_{ii}^{-1})}{\sum_{i=1}^k (\tilde{\lambda}_{ii}^{-1/2})} \tilde{\lambda}_{jj}^{-1/2} - \tilde{\lambda}_{jj}^{-1} \right], \quad (47)$$

$$|\tilde{\phi}_{g,jj}|^2 = \left\{ \frac{p'_0 + \sum_{i=1}^k (\tilde{\lambda}_{ii}^{-1})}{\sum_{i=1}^k (\tilde{\lambda}_{ii}^{-1/2})} \tilde{\lambda}_{jj}^{-1/2} - \left[\frac{\sum_{i=1}^k (\tilde{\lambda}_{ii}^{-1/2})}{p'_0 + \sum_{i=1}^k (\tilde{\lambda}_{ii}^{-1})} \right]^2 \tilde{\lambda}_{jj}^{-1} \right\} \tilde{\lambda}_{jj}^{-1}, \quad (48)$$

where the main diagonal element of $\tilde{\Lambda}$ is denoted by $\tilde{\lambda}_{jj}$. Moreover, the $\tilde{\Lambda}$ and $\tilde{\mathbf{V}}$ are matrices obtained from the EVD algorithm

$$\left(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e \right)^H \mathbf{R}_{\mathbf{n}'\mathbf{n}'}^{-1} \times \left(\xi\hat{\mathbf{H}} + \sqrt{1-\xi^2}\hat{\mathbf{H}}_e \right) = \tilde{\mathbf{V}}\tilde{\Lambda}\tilde{\mathbf{V}}^H. \quad (49)$$

The precoder and equalizer matrices are planned as equations (42) and (43); therefore, the calculation of (27) can be expressed as follows

$$\hat{\mathbf{x}}[j] = \left(\tilde{\mathbf{G}}\xi\hat{\mathbf{H}} + \tilde{\mathbf{G}}\sqrt{1-\xi^2}\hat{\mathbf{H}}_e \right) \hat{\mathbf{H}} \left(\tilde{\mathbf{F}}\xi\hat{\mathbf{H}} + \tilde{\mathbf{F}}\sqrt{1-\xi^2}\hat{\mathbf{H}}_e \right) \mathbf{x}[j] + \tilde{\mathbf{G}}\mathbf{n}'[j] \\ = \xi^2 \tilde{\mathbf{G}}\hat{\mathbf{H}}\hat{\mathbf{H}}\tilde{\mathbf{F}}\hat{\mathbf{H}}\mathbf{x}[j] + \xi \sqrt{1-\xi^2} \tilde{\mathbf{G}}\hat{\mathbf{H}}_e \hat{\mathbf{H}}\tilde{\mathbf{F}}\hat{\mathbf{H}}\mathbf{x}[j] \\ + \xi \sqrt{1-\xi^2} \tilde{\mathbf{G}}\hat{\mathbf{H}}_e \hat{\mathbf{H}}\tilde{\mathbf{F}}\hat{\mathbf{H}}\mathbf{x}[j] + (1-\xi^2) \tilde{\mathbf{G}}\hat{\mathbf{H}}_e \hat{\mathbf{H}}\tilde{\mathbf{F}}\hat{\mathbf{H}}_e \mathbf{x}[j] + \tilde{\mathbf{G}}\mathbf{n}'[j]. \quad (50)$$

Subsequently, the SINR of sub-channels and the capacity of the system (C_{sys}) are able to be given as

$$\text{SINR}_j = \frac{\xi^2 |\tilde{\phi}_{f,jj}|^2 \tilde{\lambda}_{jj}}{(2\xi \sqrt{1-\xi^2} + (1-\xi^2)) |\tilde{\phi}_{f,jj}|^2 \tilde{\lambda}_{jj} + 1}, \quad (51)$$

$$C_{\text{sys}} = \sum_{j=1}^{j=\bar{N}} \log_2 \left(1 + \frac{\xi^2 |\tilde{\phi}_{f,jj}|^2 \tilde{\lambda}_{jj}}{(2\xi \sqrt{1-\xi^2} + (1-\xi^2)) |\tilde{\phi}_{f,jj}|^2 \tilde{\lambda}_{jj} + 1} \right), \quad (52)$$

where $\bar{N} \leq \text{rank}(\bar{K}M_R, \bar{K}M_T)$, and $\hat{\lambda}_{jj}$ is the main diagonal of $\hat{\Lambda}$ that is able to be obtained by mathematical calculation as follows

$$\hat{\mathbf{H}}^H \mathbf{R}_{\mathbf{n}'\mathbf{n}'}^{-1} \hat{\mathbf{H}} = \hat{\mathbf{V}}\hat{\Lambda}\hat{\mathbf{V}}^H. \quad (53)$$

4 Simulation Results

To evaluate the performance of the proposed scheme under the condition of imperfect CSI at the receiver, the system model is assumed to have four transmit antennas and four receive antennas. The system performance is evaluated according to the channel estimation accuracy ξ , the order of FIR D and transmit block size Q . The Saleh-Valenzuela indoor channel model is used to generate the CIR as in [33] and the 4-QAM modulation is applied. Additionally, the overall transmit power is standardized through transmit antennas ($p_0=1$), and the system performance of the proposed scheme is simulated and calculated according to the Monte Carlo simulation.

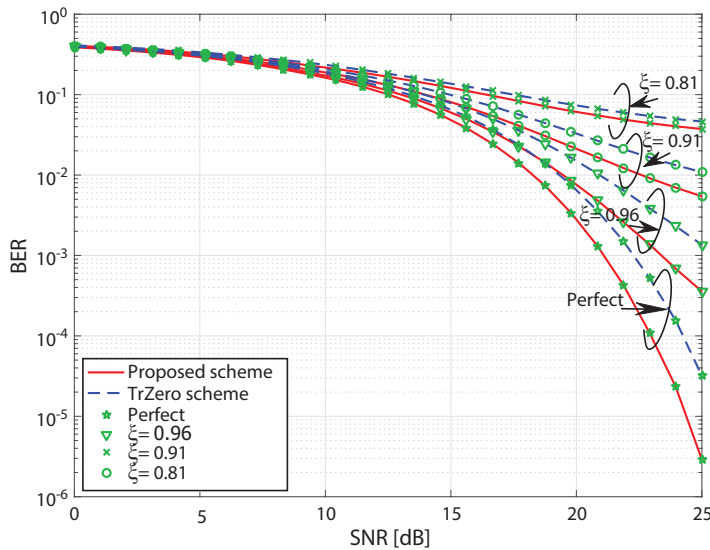


Figure 2: BER performance of the proposed scheme is compared with the TrZero scheme when $D = 12$ and $Q = 26$

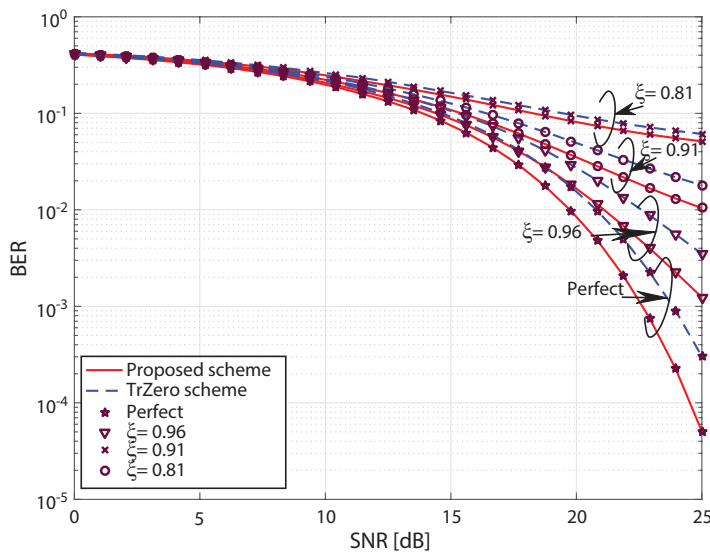


Figure 3: BER performance of the proposed scheme is compared with the TrZero scheme when $D = 12$ and $Q = 30$

On the other hand, the BER performance of the proposed and the TrZero schemes is compared in the different situations of the

order of FIR and the transmission block dimension as illustrated in Figures 2, 3 and 4, respectively. In general, the BER of the TrZero scheme is significantly higher than that of the proposed one at all considered ξ values, such as $\xi = 0.81, 0.91$ and 0.96 , and 1 (the perfect CSI). Furthermore, it is obvious that the accuracy of channel estimation proportionally affects the BER of the system, which means that the BER and the accuracy of channel estimation increase simultaneously, especially for the high SNR range. This is because of the fact that the residual interference due to imperfect CSI increases when the SNR increases. Consequently, the SINR of every sub-channel extremely decreases compared to the SNR of the perfect one.

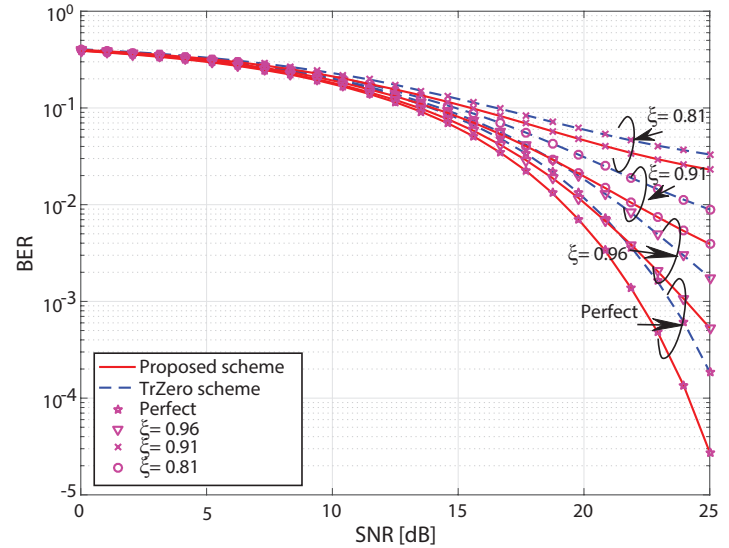


Figure 4: BER performance of the proposed scheme is compared with the TrZero scheme when $D = 10$ and $Q = 26$

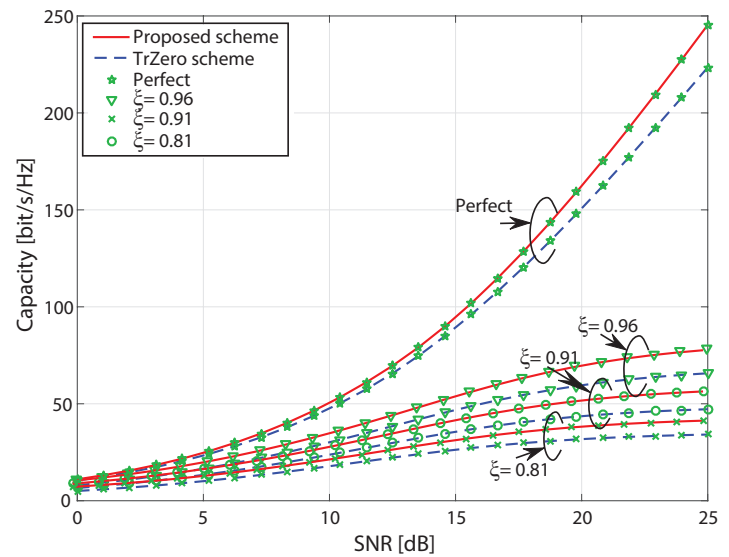


Figure 5: Comparison of system capacity in the proposed and TrZero schemes when $D = 12$ and $Q = 26$

As shown in Figures 2 and 3, the BER of the system is deteriorated when the Q increases, or the D decreases (Figures 2 and

4). It can be explained that the increase of Q makes the number of sub-channels increases while the shared redundancy is fixed. Thus, the channel energy for every sub-channel decreases. In contrast, when the D decreases, the shared redundancy decreases while the number of sub-channels is fixed, it leads the decrease in the channel energy for every sub-channel.

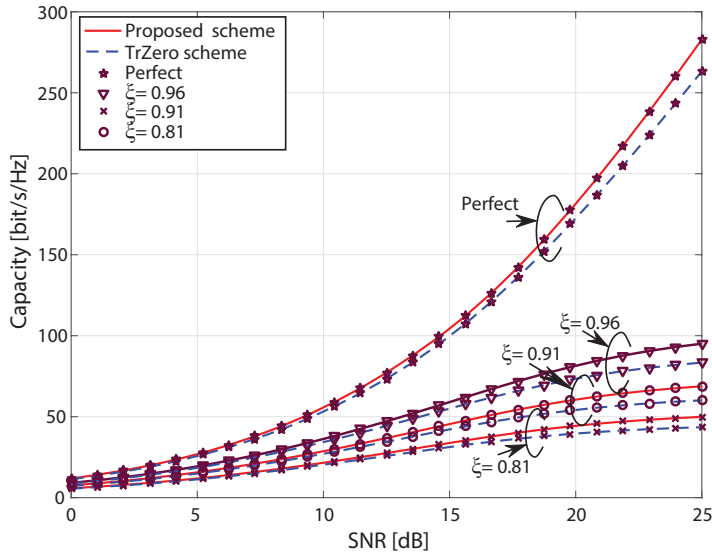


Figure 6: Comparison of system capacity in the proposed and TrZero schemes when $D = 12$ and $Q = 26$

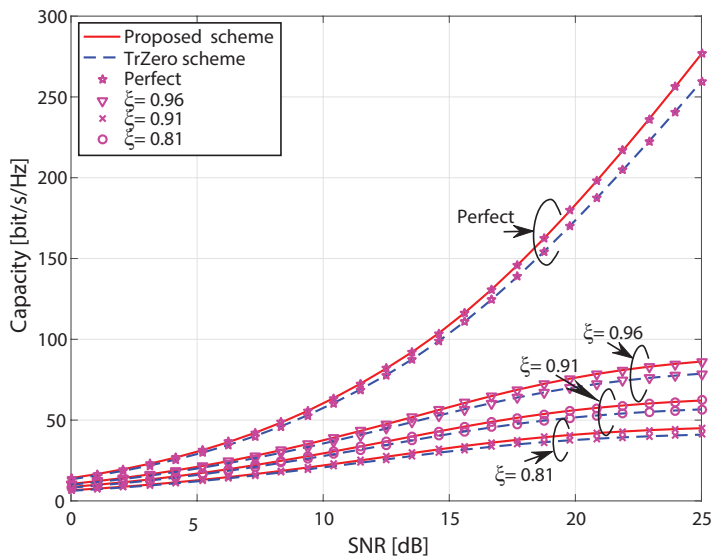


Figure 7: Comparison of system capacity in the proposed and TrZero schemes when $D = 10$ and $Q = 26$

Finally, the capacity of the proposed scheme is slightly higher than that of the TrZero scheme in all analyzed scenarios as in Figures 5, 6 and 7. Moreover, compared to the capacity of system with perfect CSI, the capacity of system with imperfect CSI is decreased dramatically. It is the same for both the TrZero and the proposed technical schemes. For more detail, at the 25 dB of SNR in Figure 5, when the accuracy of channel estimation ξ drops from 0.96 to 0.91, the system capacity witnesses a decrease from 80 to 55 bit/s/Hz.

Furthermore, the capacity of system under the imperfect CSI condition at the receiver increases due to the increase of the transmission block size. On the other hand, the capacity of system decreases once the order of FIR increases in both schemes. The reason is that when the transmission block dimension increases and/or the FIR order decreases, the SINR of every sub-channel decreases, however the number of sub-channels in MIMO ISI channels is increased. As a result, the capacity of the MIMO ISI system under the imperfect CSI condition increases.

5 Conclusion

In this paper, the combination scheme of precoder and equalizer relied on the *unweighted* MMSE criterion for the imperfect CSI of MIMO ISI channels has been developed. The proposed scheme lets the loss of channel energy decrease and eliminates some sub-channels with extremely small eigenvalue in order to improve the system performance. The proposed scheme is evaluated by Monte Carlo simulation, and the simulation result shows that the proposed scheme under limited transmit power can achieve the higher system performance comparing to the conventional scheme in the considered scenarios even when the CSI is imperfect. However, the performance of the MIMO ISI system using redundancies is significantly affected by the CSI factor. Furthermore, the relationship between the performance of the system and several parameters, such as the accuracy of channel estimation, the transmission block size and the FIR order, are also discussed. The result in this research can help us to structure a practical MIMO ISI system which is close to the actual situation. The investigation of the impact of imperfect CSI on the advanced future communication system, such as massive MIMO or MIMO filterbank multicarrier systems, with combining designs of precoder and equalizer is left to future works.

References

- [1] R. Vannithamby and S. Talwar, *New Physical Layer Waveforms for 5G*. Wiley, 2017. [Online]. Available: <https://ieeexplore.ieee.org/document/8043580>
- [2] F. Corno, L. De Russis, and J. Pablo Senz, "On the advanced services that 5g may provide to iot applications," in *2018 IEEE 5G World Forum (5GWF)*, pp. 528–531, July 2018.
- [3] B. Kwon, S. Kim, S. Lee, *Scattered Reference Symbol-Based Channel Estimation and Equalization for FBMC-QAM Systems*. *IEEE Transactions on Communications*, vol. 65, no. 8, pp. 3522 – 3537, 2017.
- [4] J. Wang, Q. Yu ; Z. Li, C. Bi, "Distributed Space Time Block Transmission and QRD Based Diversity Detector in Asynchronous Cooperative Communications Systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 6, pp. 5111 – 5125, 2018.
- [5] C. Dong, J. Lin, K. Niu, Z. He, and Z. Bie, "Block-iterative decision feedback equalizer with noise prediction for single-carrier MIMO transmission," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 8, pp. 3772–3776, 2012.
- [6] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers. i. unification and optimal designs," *IEEE Transactions on Signal Processing*, vol. 47, no. 7, pp. 1988–2006, 1999.
- [7] M.-W. Kwan and C.-W. Kok, "MMSE equalizer for MIMO-ISI channel with shorten guard period," *IEEE transactions on signal processing*, vol. 55, no. 1, pp. 389–395, 2007.
- [8] W. A. Martins and P. S. R. Diniz, "Block-based transceivers with minimum redundancy," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1321–1333, 2010.

- [9] Y.-P. Lin and S.-M. Phoong, "Minimum redundancy for ISI free FIR filterbank transceivers," *IEEE Transactions on Signal Processing*, vol. 50, no. 4, pp. 842–853, 2002.
- [10] B. Q. Doanh, P. T. Hiep, T. C. Hieu *et al.*, "A combining design of precoder and equalizer based on shared redundancy to improve performance of ISI MIMO systems," *Wireless Networks*, pp. 1–10, 2019
- [11] Y. Zhai, J. Tong, and J. Xi, "Precoder design for MIMO visible light communications with decision-feedback receivers," *IEEE Photonics Technology Letters*, 2019.
- [12] Y. Cheng, L. G. Baltar, M. Haardt, and J. A. Nossek, "Precoder and equalizer design for multi-user MIMO FBMC/OQAM with highly frequency selective channels," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*. IEEE, pp. 2429–2433, 2015.
- [13] F. Rottenberg, X. Mestre, and J. Louveaux, "Optimal zero forcing precoder and decoder design for multi-user mimo FBMC under strong channel selectivity," in *ICASSP*, pp. 3541–3545, 2016.
- [14] F. Rottenberg, X. Mestre, F. Horlin, and J. Louveaux, "Single-tap precoders and decoders for multi-user MIMO FBMC-OQAM under strong channel frequency selectivity," *IEEE transactions on signal processing*, vol. 65, no. 3, pp. 587–600, 2017.
- [15] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Transactions on Communications*, vol. 49, no. 12, pp. 2198–2206, 2001.
- [16] H. Yang, C. Chen, W.-D. Zhong, and A. Alphones, "Joint precoder and equalizer design for multi-user multi-cell mimo vlc systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 11 354–11 364, 2018.
- [17] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1051–1064, 2002.
- [18] N. Lee, O. Simeone, and J. Kang, "The effect of imperfect channel knowledge on a MIMO system with interference," *IEEE Transactions on Communications*, vol. 60, no. 8, pp. 2221–2229, August 2012.
- [19] D. Le Ruyet, R. Zakaria, and B. Özbek, "On precoding MIMO-FBMC with imperfect channel state information at the transmitter," in *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*. IEEE, pp. 808–812, 2014.
- [20] I. Dvorakova, A. Malyutin, and Y. Nechaev, "MMSE precoder for multipath channels with imperfectly known state information," in *2015 38th International Conference on Telecommunications and Signal Processing (TSP)*, pp. 195–199, July 2015.
- [21] M. Huang, S. Zhou, and J. Wang, "Analysis of TomlinsonHarashima precoding in multiuser mimo systems with imperfect channel state information," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 5, pp. 2856–2867, Sep. 2008.
- [22] H. K. Bizaki and A. Falahati, "Tomlinson-Harashima precoding with imperfect channel state information," *IET Communications*, vol. 2, no. 1, pp. 151–158, January 2008.
- [23] J. Baltersee, G. Fock, and H. Meyr, "Achievable rate of MIMO channels with data-aided channel estimation and perfect interleaving," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 12, pp. 2358–2368, Dec 2001.
- [24] A. Maaref and S. Aïssa, "Combined adaptive modulation and truncated arq for packet data transmission in mimo systems," in *Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE*, vol. 6. IEEE, pp. 3818–3822, 2004.
- [25] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, April 2003.
- [26] A. Goldsmith, "Capacity and power allocation for fading mimo channels with channel estimation error," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.
- [27] B. Q. Doanh, P. T. Hiep, and T. C. Hieu, "Impact of imperfect CSI on capacity of ISI MIMO systems based on joint precoding and equalization designs," in *2019 19th International Symposium on Communications and Information Technologies (ISCIT)*. IEEE, pp. 334–338, 2019.
- [28] M. Ding and S. D. Blostein, "MIMO minimum total MSE transceiver design with imperfect csi at both ends," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1141–1150, March 2009.
- [29] B. S. Thian, S. Zhou, and A. Goldsmith, "Transceiver design for MIMO systems with imperfect csi at transmitter and receiver," in *2011 IEEE International Conference on Communications (ICC)*, pp. 1–6, June 2011.
- [30] Z. Chang, Z. Wang, X. Guo, Z. Han, and T. Ristaniemi, "Energy-efficient resource allocation for wireless powered massive MIMO system with imperfect CSI," *IEEE Transactions on Green Communications and Networking*, vol. 1, no. 2, pp. 121–130, June 2017.
- [31] P. Aquilina and T. Ratnarajah, "Linear interference alignment in full-duplex MIMO networks with imperfect CSI," *IEEE Transactions on Communications*, vol. 65, no. 12, pp. 5226–5243, Dec 2017.
- [32] J. Cui, Z. Ding, and P. Fan, "Outage probability constrained MIMO-NOMA designs under imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 8239–8255, Dec 2018.
- [33] A. A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE Journal on selected areas in communications*, vol. 5, no. 2, pp. 128–137, 1987.
- [34] S. Serbetli and A. Yener, "MMSE transmitter design for correlated MIMO systems with imperfect channel estimates: power allocation trade-offs," *IEEE Transactions on Wireless Communications*, vol. 5, no. 8, pp. 2295–2304, Aug 2006.
- [35] L. Musavian, M. R. Nakhai, M. Dohler, and A. H. Aghvami, "Effect of channel uncertainty on the mutual information of mimo fading channels," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 5, pp. 2798–2806, Sep. 2007.
- [36] H. Lütkepohl, "Handbook of matrices. wiley," *New York*, 1996.

Non Parallelism and Cayley-Menger Determinant in Submerged Localization

Anisur Rahman*

Department of Computer Science and Engineering, East West University, Dhaka – 1212, Bangladesh

ARTICLE INFO

Article history:

Received: 16 January, 2020

Accepted: 22 April, 2020

Online: 08 May, 2020

Keywords:

Cayley-Menger determinant

Linearization

Mobile beacon

Parallel and non-parallel state

Submerged localization

Underwater wireless sensor network

ABSTRACT

This research paper portrays the technique to determine location of submerged nodes with Cayley-Menger determinant and associated problems with non-parallel states. Cayley-Menger determinant is considered to be the usual means to determine the coordinates of the nodes with a single node where the plane of beacon i.e. beacon's surfing plane and the plane created by the deployed submerged sensors are in parallel state. However, this perfect scenario hardly exists in the submerged world; as a result Cayley-Menger determinant may not be used unless proper measurements are taken. This paper addresses this limitation and analyzed the proposed model to deal with this non-parallel state situation. As precise localization is vital for the validity of the sensed data in Underwater Wireless Sensor Networks (UWSNs); exact distance measurement between nodes is an important and crucial in range based localization methods. Proposed method has addressed uncertainty of nonlinear equations as well as how better immunity could be achieved in multipath fading in propagation. Analyzed simulation and experimental results conclude the accuracy of the proposed model with minimal error, where it has been shown that parallelism of the system is not a factor for using Cayley-Menger determinant. Moreover, a single node has been used in the model to localize submerged nodes where none acquires priori familiarity about its position.

1. Introduction

Underwater wireless sensor networks (UWSNs) are envisioned for exploring the vast underwater world for the profusion of wealth and mystery. Not only the wealth of underwater world, submerged localization necessary for the sustenance of our very own existence and for the marine biome. Lately researchers have shown fervent importance to investigate, explore and analyze the abundance of underwater world; so it has become indispensable to collect accurate environmental data with the help of underwater sensors. Exact localization is not only necessary for underwater applications; it determines the very nature of our own existence at rudimentary and core level as well as helps us to provide a sanctuary for the marine biome. Moreover, autonomous underwater vehicles (AUVs) control and surveillance, monitoring seabed and faults for upcoming natural calamity, searching for lost object, pollutants and nutrients tracking also demand precise localization [1]. Among these applications, some significantly requires accurate localization for meaningful comprehension of gathered data for a practical use [2].

UWSNs may comprise of deployed sensors and surface stations; many a time localization of submerged sensors is done with the help of multiple surfaced nodes. However, localizing using a single beacon is considered to be pragmatic as it requires less provision. Despite numerous underwater applications, the concept and achievement of underwater wireless communication may still seem far-fetched. Mostly, communications is done with acoustic signals as radio signals have very limited propagation in underwater; hence, using acoustic signals for distance measurements in range based localization provides more accuracy that using radio [1,3]. However, multipath fading and synchronization between communicating nodes still remain to be the challenging factors so far.

Duff and Muller solved system of multilateration equations by applying nonlinear square optimization method where positions of the sensors remain unknown [3]. This proposed algorithm is solely based on degree-of-freedom analysis phenomenon, which dictates sufficient numbers of measurements are necessary to generate adequate number of equations for the problem to solve. This proposed method has been validated and showed the improved accuracy by many folds using Kalman filter in [4]. Incorporation

*Anisur Rahman, Dept. of Computer Sci. & Engg., EWU, anis@ewubed.edu

of Kalman filter increases the complexity of the method with respect to degree-of-freedom complexity of system of equations in [5]. It also delineates that the method does not guarantee a unique solution for system of non-linear equations, i.e. trilateration. Moreover, the method requires a specific initial configuration of the nodes which was justified by rigidity theory. 3D positioning system in [6] requires four separate positions to determine the coordinates of the beacon.

This paper analyzed the method proposed in [7] to determine the coordinates of underwater sensors with a single beacon where Cayley-Menger determinant is used for non-parallel situation and validates the method with simulation and experimentation. Recently, localization of submerged sensors for non-parallel states has been proposed in [8]. It also showed that the coordinates could be determined in adhoc basis without any pre-installed infrastructure. Moreover, it has established that the accuracy of the coordinates depends on the distance measurements between surfaced node and underwater deployed sensors, not the state of the planes whether parallel or non-parallel.

The arrangement of the remaining paper is as following where Section 2 focuses on mathematical model of coordinates determination both for parallel and non-parallel state situations; it also talks about linear transformation of the origin to find the coordinates with respect to sensor and beacon separately. Section 3 states simulation to validate the proposed mathematical model, experimental results also shows the accuracy of the model and at last analysis. Section 4 states conclusion with future works.

2. Coordinates Determination

The proposed algorithm in [9] was to determine the coordinates of the submerged sensors with a single beacon; the problem domain was considered to be in parallel state which is only possible in perfect world. However, in [7] Cayley-Menger determinant has been used for non-parallel state situation. Following sections iterates the proposed algorithm for parallel and non-parallel states.

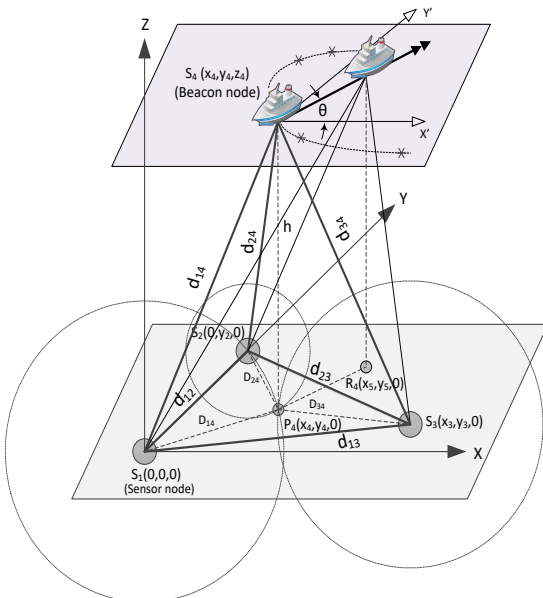


Figure 1: Coordinates Determination

2.1. Coordinates Determination (parallel state)

The problem domain consists of a single beacon $S_j, j=4,5,\dots,9$ which is on the water surface and three underwater sensors $S_i, i=1,2,3$ whose coordinates need to be determined. For simplicity, one of the sensors is considered to the Cartesian origin $(0,0,0)$ and one other placed on another axis leaving the third one on the same quadrant. The distance between the beacon $S_j, j=4,5,\dots,9$ and the deployed sensors $S_i, i=1,2,3$ have been computed according to the process depicted in [10], these measured values are $d_{14}, d_{24}, d_{34}, \dots$ and unknown inter nodes distances are d_{12}, d_{13}, d_{23} . From these values, volume of the tetrahedron V_t which is created by the beacon and the deployed sensors depicted in Figure 1 can be written with Cayley-Menger determinant as in (1).

$$288 V_t^2 = \begin{vmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & d_{12}^2 & d_{13}^2 & d_{14}^2 \\ 1 & d_{12}^2 & 0 & d_{23}^2 & d_{24}^2 \\ 1 & d_{13}^2 & d_{23}^2 & 0 & d_{34}^2 \\ 1 & d_{14}^2 & d_{24}^2 & d_{34}^2 & 0 \end{vmatrix} \tag{1}$$

By expanding (1), we obtain:

$$\begin{aligned} & d_{34}^2 d_{23}^2 - d_{34}^2 d_{12}^2 + d_{34}^2 d_{13}^2 - \frac{d_{14}^4 d_{23}^2}{d_{12}^2} + d_{23}^2 d_{14}^2 + \\ & \frac{d_{13}^2 d_{14}^2 d_{23}^2}{d_{12}^2} - \frac{d_{24}^2 d_{13}^4}{d_{12}^2} + \frac{d_{13}^2 d_{23}^2 d_{24}^2}{d_{12}^2} + d_{13}^2 d_{24}^2 - d_{13}^2 d_{23}^2 - \\ & 144 \frac{V_t^2}{d_{12}^2} + \frac{d_{14}^4 d_{23}^2 d_{24}^2}{d_{12}^2} + \frac{d_{14}^2 d_{23}^2 d_{34}^2}{d_{12}^2} - \frac{d_{23}^2 d_{24}^2 d_{34}^2}{d_{12}^2} - \frac{d_{14}^4 d_{23}^2}{d_{12}^2} + \\ & \frac{d_{13}^2 d_{24}^2 d_{34}^2}{d_{12}^2} - \frac{d_{13}^2 d_{14}^2 d_{34}^2}{d_{12}^2} + \frac{d_{13}^2 d_{14}^2 d_{24}^2}{d_{12}^2} - \frac{d_{13}^2 d_{24}^4}{d_{12}^2} - d_{34}^4 + \\ & d_{24}^2 d_{34}^2 + d_{14}^2 d_{34}^2 - d_{14}^2 d_{24}^2 = 0 \end{aligned}$$

By isolating and grouping known and unknown variables:

$$\begin{aligned} & d_{34}^2 (d_{12}^2 - d_{23}^2 - d_{13}^2) + d_{14}^2 \left(\frac{d_{23}^4}{d_{12}^2} - d_{23}^2 - \frac{d_{13}^2 d_{23}^2}{d_{12}^2} \right) + \\ & d_{24}^2 \left(\frac{d_{13}^4}{d_{12}^2} - \frac{d_{13}^2 d_{23}^2}{d_{12}^2} - d_{13}^2 \right) - \left(d_{14}^2 d_{24}^2 + d_{14}^2 d_{34}^2 - d_{24}^2 d_{34}^2 - d_{14}^4 \right) \frac{d_{23}^2}{d_{12}^2} \\ & - \left(d_{34}^2 d_{24}^2 - d_{14}^2 d_{34}^2 + d_{14}^2 d_{24}^2 - d_{14}^4 \right) \frac{d_{13}^2}{d_{12}^2} + \left(144 \frac{V_t^2}{d_{12}^2} + d_{13}^2 d_{23}^2 \right) \\ & = \left(d_{24}^2 d_{34}^2 - d_{34}^4 + d_{14}^2 d_{34}^2 - d_{14}^2 d_{24}^2 \right) \end{aligned}$$

Here, $\left(\frac{d_{13}^4}{d_{12}^2} - \frac{d_{13}^2 d_{23}^2}{d_{12}^2} - d_{13}^2 \right)$, $\left(\frac{d_{23}^4}{d_{12}^2} - d_{23}^2 - \frac{d_{13}^2 d_{23}^2}{d_{12}^2} \right)$, $\frac{d_{23}^2}{d_{12}^2}$, $\left(d_{12}^2 - d_{23}^2 - d_{13}^2 \right)$, $\frac{d_{13}^2}{d_{12}^2}$, and $\left(144 \frac{V_t^2}{d_{12}^2} + d_{13}^2 d_{23}^2 \right)$ are unknowns.

Here, we rewrite the equation in the following form:

$$\begin{aligned} & d_{14}^2 X_1 + d_{24}^2 X_2 + d_{34}^2 X_3 - (d_{14}^2 - d_{34}^2)(d_{24}^2 - d_{14}^2) X_4 \\ & - (d_{24}^2 - d_{14}^2)(d_{34}^2 - d_{24}^2) X_5 + X_6 = (d_{24}^2 - d_{34}^2)(d_{34}^2 - d_{14}^2) \end{aligned} \tag{2}$$

The above equation has six unknowns as depicted by X_1, X_2, X_3, X_4, X_5 and X_6 ; and the equation resembles with the linear form of

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b_1.$$

Table 1: Calculated Coordinates of the Submerged Sensors

Sensors	Coordinates
S_1	(0,0,0)
S_2	(0, d_{12} , 0)
S_3	$\left(\sqrt{\left(d_{13}^2 - \left(\frac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d_{12}} \right)^2 \right)}, \frac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d_{12}}, 0 \right)$

S_4 , so that in real life the movement would be less between distance measurements. From these six equations reference to the variables have been omitted to get an array of all coefficients, which is represented as augmented matrix. The matrix consists of linear equations and resembles $AX = b$ formula. So, the system of linear equations is articulated as following:

$$A = \begin{bmatrix} d_{14}^2 & d_{24}^2 & d_{34}^2 & -(d_{14}^2 - d_{34}^2)(d_{24}^2 - d_{14}^2) & -(d_{24}^2 - d_{14}^2)(d_{34}^2 - d_{24}^2) & 1 \\ d_{15}^2 & d_{25}^2 & d_{35}^2 & -(d_{15}^2 - d_{35}^2)(d_{25}^2 - d_{15}^2) & -(d_{25}^2 - d_{15}^2)(d_{35}^2 - d_{25}^2) & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ d_{19}^2 & d_{29}^2 & d_{39}^2 & -(d_{19}^2 - d_{39}^2)(d_{29}^2 - d_{19}^2) & -(d_{29}^2 - d_{19}^2)(d_{39}^2 - d_{29}^2) & 1 \end{bmatrix}$$

$$X = \begin{bmatrix} \left(\frac{d_{23}^4}{d_{12}^2} - d_{23}^2 - \frac{d_{13}^2 d_{23}^2}{d_{12}^2} \right) \\ \left(\frac{d_{13}^4}{d_{12}^2} - \frac{d_{13}^2 d_{23}^2}{d_{12}^2} - d_{13}^2 \right) \\ (d_{12}^2 - d_{23}^2 - d_{13}^2) \\ \frac{d_{23}^2}{d_{12}^2} \\ \frac{d_{12}^2}{d_{13}^2} \\ \frac{d_{13}^2}{d_{12}^2} \\ \left(144 \frac{v_i^2}{d_{12}^2} + d_{13}^2 d_{23}^2 \right) \end{bmatrix} \quad b = \begin{bmatrix} (d_{24}^2 - d_{34}^2)(d_{34}^2 - d_{14}^2) \\ (d_{25}^2 - d_{35}^2)(d_{35}^2 - d_{15}^2) \\ \vdots \\ (d_{29}^2 - d_{39}^2)(d_{39}^2 - d_{19}^2) \end{bmatrix}$$

From the above representation, after finding X_1, X_2, X_3, X_4, X_5 and X_6 we calculate d_{12}, d_{13} and d_{23} as follows:

$$d_{12}^2 = \frac{X_3}{(1 - X_4 - X_5)}, \quad d_{13}^2 = \frac{X_3 X_5}{(1 - X_4 - X_5)}, \quad d_{23}^2 = \frac{X_3 X_4}{(1 - X_4 - X_5)}$$

Coordinates of deployed submerged sensors $S_1(0,0,0)$ is considered to be the origin of the Cartesian system, $S_2(0, y_2, 0)$ is on one of the y -axis and $S_3(x_3, y_3, 0)$ is considered on some position not on the axis as depicted in Figure 1. From the above equations, the inter node distances can be calculated as follows:

$$d_{12}^2 = y_2^2, \quad d_{13}^2 = x_3^2 + y_3^2, \quad d_{23}^2 = x_3^2 + (y_3 - y_2)^2$$

As a result, the unknown variables y_2, y_3 and x_3 can be derived with respect to inter node distances:

$$y_2 = d_{12}, \quad y_3 = \frac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d_{12}}, \quad x_3 = \sqrt{\left(d_{13}^2 - \left(\frac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d_{12}} \right)^2 \right)}$$

Here, d_{12}, d_{13} and d_{23} are computed distances between deployed submerged sensors and Table 1 shows the coordinates of the sensors as follows:

2.2. Coordinates with respect to Beacon (parallel state)

So far the previous section has dealt with determination of coordination when beacon's surfing plane and plane of deployed submerged sensors are in parallel. Once the transformation of the origin takes place and taken to the beacon's position, the localization process would be following.

In this method, a pressure sensor is used to measure the depth h of Figure 2, which is elaborated in [11]. Once vertical distance (h) between the planes is known, the projected coordinates of the beacon $S_4(x_4, y_4, z_4)$ can be found on the XY plane and can be denoted as $P_4(x_4, y_4, 0)$. Trilateration can be used to find x_4 and y_4 of the projected coordinates considering D_{14}, D_{24} and D_{34} to be the distances between S_1, S_2, S_3 and P_4 respectively. Once the projected coordinates and the distances are determined, following relations can be devised.

$$D_{14}^2 = x_4^2 + y_4^2 \tag{3}$$

$$D_{24}^2 = x_4^2 + (y_4 - y_2)^2 \tag{4}$$

$$D_{34}^2 = (x_4 - x_3)^2 + (y_4 - y_3)^2 \tag{5}$$

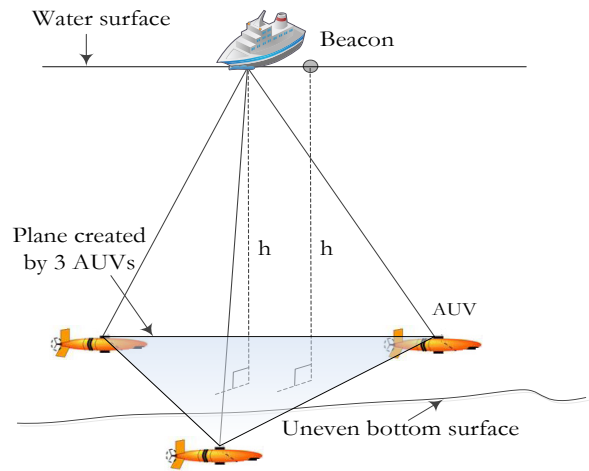


Figure 2: Parallel Plane State Scenario

Equations (3), (4) and (5) help to determine the coordinates of the projected beacon $P_4(x_4, y_4, 0)$, where

$$x_4 = \frac{1}{2d_{12}} \sqrt{\left(4d_{12}^2 D_{14}^2 - (D_{14}^2 - D_{24}^2 + d_{12}^2)^2 \right)},$$

$$y_4 = \frac{1}{2d_{12}} (D_{14}^2 - D_{24}^2 + d_{12}^2)$$

Here, d_{14} , d_{24} and d_{34} are the hypotenuse of the $\Delta S_1 P_4 S_4$, $\Delta S_2 P_4 S_4$ and $\Delta S_3 P_4 S_4$ respectively, so it is possible to obtain D_{14} , D_{24} and D_{34} using simple Pythagorean Theorem. As a result, the coordinate of the beacon node that is on the surface, $S_4(x_4, y_4, z_4)$ would be $S_4(x_4, y_4, h)$; here both x_4 and y_4 have been calculated accordingly and h is known from the installed pressure sensor. So, the coordinates of S_4 would be as follows:

$$\therefore S_4(x_4, y_4, h) = S_4 \left(\begin{array}{l} \frac{1}{2d_{12}} \sqrt{(4d_{12}^2 D_{14}^2 - (D_{14}^2 - D_{24}^2 + d_{12}^2)^2)}, \\ \frac{1}{2d_{12}} (D_{14}^2 - D_{24}^2 + d_{12}^2), h \end{array} \right)$$

Once the origin (0,0,0) of the Cartesian system is transferred on the beacon $S_4(x_4, y_4, h)$, then the coordinates of the deployed sensors with respect to beacon are calculated by linear transformation as depicted in [10].

Table 2: Coordinates of the Submerged Sensors with respect to Beacon for Parallel State

Sensors	Coordinates
S_4	(0,0,0)
S_1	$\left(\frac{\sqrt{4d_{12}^2 D_{14}^2 - (D_{14}^2 - D_{24}^2 + d_{12}^2)^2}}{2d_{12}}, -\frac{1}{2d_{12}} (D_{14}^2 - D_{24}^2 + d_{12}^2), -h \right)$
S_2	$\left(\frac{\sqrt{4d_{12}^2 D_{14}^2 - (D_{14}^2 - D_{24}^2 + d_{12}^2)^2}}{2d_{12}}, \frac{1}{2d_{12}} (d_{12}^2 - D_{14}^2 + D_{24}^2), -h \right)$
S_3	$\left(\left(\frac{\sqrt{\left(d_{13}^2 - \left(\frac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d_{12}} \right)^2 \right) - \frac{\sqrt{4d_{12}^2 D_{14}^2 - (D_{14}^2 - D_{24}^2 + d_{12}^2)^2}}{2d_{12}}}}{2d_{12}}, -h \right)$

2.3. Coordinates with respect to Beacon (non-parallel state)

Having the plane where the beacon moves around i.e. water surface and the plane that is created by the three deployed sensors in parallel state can be found in very limited situations. Previous coordinates determination method could be adjusted to meet all the situations both parallel and non-parallel. As a result, a derivative of parallel state method has been illustrated in [8]; however this paper also shows the relationship between Cayley-Menger and state of the planes. With a little adjustment as depicted in this section where Figure 3 illustrates the scenario when both planes are in non-parallel state situation.

$$\text{Plane A: } \prod_{S_1, S_2, S_3} : (a_1 x + b_1 y + c_1 z + \delta_1 = 0)$$

Here, plane A is created by the submerged deployed sensors S_1 , S_2 , and S_3 as three dots (sensors) can create a plane. Whereas,

$$\text{Plane B: } \prod_{\text{Beacon}} : (a_2 x + b_2 y + c_2 z + \delta_2 = 0)$$

Here, plane B is where the beacon moves around while taking six distance measurements i.e. water surface, are in non-parallel state.

Figure 3 is the most likely the situation found in the world, except in some special cases like, water tank, swimming pools or where the deployed sensors can maintain a predefined height like AUVs or UUVs as in Figure 2. When we consider the water surface as the reference plane then any deployed node having a different height than others would create a non-parallel state situation. The system of equations and the linearization procedure devised in previous section are meant to be used in parallel state; as volume of six different tetrahedrons created by deployed submerged sensors and the surfaced beacon are always equal. Once apex of six tetrahedrons on the same parallel plane having the base same i.e. the height of the tetrahedrons are same having the base fixed, the volume would be same. This fundamental theory creates the scope to use the devised equations in previous section. Vertical distances will vary depending on the dihedral angle α accordingly to the (6).

$$\cos \alpha = \frac{\hat{n}_1 \cdot \hat{n}_2}{|\hat{n}_1| |\hat{n}_2|} = \frac{a_1 a_2 + b_1 b_2 + c_1 c_2}{\sqrt{a_1^2 + b_1^2 + c_1^2} \sqrt{a_2^2 + b_2^2 + c_2^2}} \quad (6)$$

Here, α dihedral angle; $\hat{n}_1 = (a_1, b_1, c_1)$ and $\hat{n}_2 = (a_2, b_2, c_2)$ are normal vectors to plane A and plane B respectively.

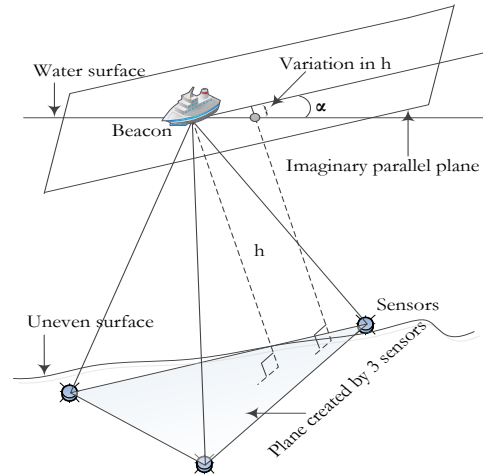


Figure 3: Parallel Planes Effect

It is worth mentioning that all the deployed submerged nodes have several sensors for communication purposes; among them pressure is one of them. Nodes are supposed to read water pressure and communicate to beacon node via acoustic signals where the pressure reading will be converted into depth following method devised in [12]. Once all three sensors' depth is known, following method would be applied to determine the coordinates. Figure 4 illustrates a solvable configuration.

At this point of the procedure all three deployed submerged sensors' depth would be known from the built-in pressure sensors. Let these depth be h_1 h_2 h_3 S_1 , S_2 and S_3

$h_3 > h_2 > h_1$, S_3 is at the lowest point whereas S_1 is at the highest point among all the three sensors with depth h_3 and h_1 respectively. In this situation, it is conspicuous that plane Π_{beacon} where beacon surfs i.e. water surface and the plane $\Pi_{S_1 S_2 S_3}$ created by three submerged sensors i.e. S_1 , S_2 and S_3 would be in non-parallel state.

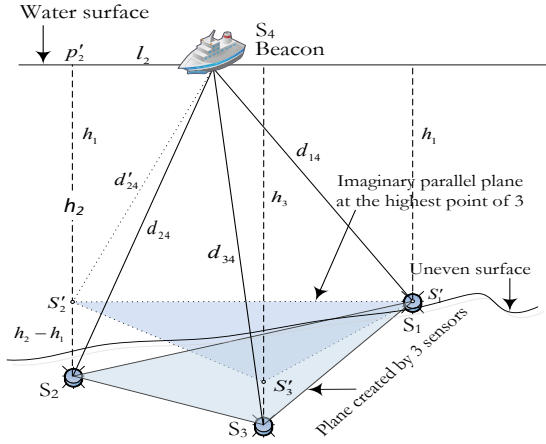


Figure 4: Parallel State in Non-parallel Situation

Let the shortest depth or highest point be $S'_1 = S_1$, considering S'_2 and S'_3 are be two points where exactly right above S_2 and S_3 having equal depth with S'_1 here the coordinates of S'_2 and S'_3 would be same as S_2 and S_3 except z component as these points are right above the nodes. As a result, the plane $\Pi_{S_1 S_2 S_3}$ would be in parallel state with the plane $\Pi_{S'_1 S'_2 S'_3}$.

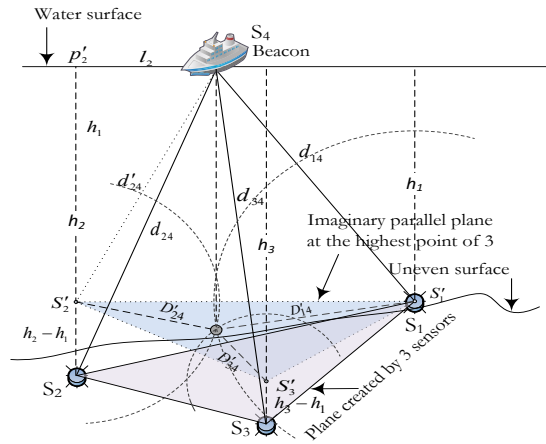


Figure 5: Trilateration and Coordinates Computation

In $\Delta S_4 S_2 p'_2$, where p'_2 is the projection point right above S_2 , and d_{24} is the measured distance what we get with acoustic signal as delineated in [13]. The distance l_2 can be calculated from $l_2 = \sqrt{d_{24}^2 - h_2^2}$. Once l_2 is known, the distance d'_{24} may be calculated from the $\Delta S_4 S'_2 p'_2$ according to (9).

$$d'_{24} = \sqrt{l_2^2 + h_1^2} \quad (7)$$

The aforesaid mentioned technique will be followed to calculate distances between the beacon S_4 and the point S'_3 . Once both distances between beacon S_4 to points S'_2 and S'_3 are determined, coordinates determination procedure will begin for all imaginary points as well as for all the deployed submerged sensors.

Table 2 and 3 show the calculated coordinates for parallel and non-parallel states respectively. So, the devised procedure - Cayley-Menger determinant with a single beacon can be used for both parallel and non-parallel state situations with a little adjustment.

Table 3: Coordinates of the Submerged Sensors with respect to Beacon for Non-parallel State

Coordinates	
S_4	$(0,0,0)$
S_1	$\left(\frac{\sqrt{4d_{12}^2 D_{14}^2 - (D_{14}^2 - D_{24}^2 + d_{12}^2)^2}}{2d'_{12}}, -\frac{1}{2d'_{12}}(D_{14}^2 - D_{24}^2 + d_{12}^2), -h_1 \right)$
S_2	$\left(\frac{\sqrt{4d_{12}^2 D_{14}^2 - (D_{14}^2 - D_{24}^2 + d_{12}^2)^2}}{2d'_{12}}, \frac{1}{2d'_{12}}(d_{12}^2 - D_{14}^2 + D_{24}^2), -h_2 \right)$
S_3	$\left(\left[\sqrt{d_{13}^2 - \left(\frac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d'_{12}} \right)^2} - \frac{\sqrt{4d_{12}^2 D_{14}^2 - (D_{14}^2 - D_{24}^2 + d_{12}^2)^2}}{2d'_{12}} \right], \frac{1}{2d'_{12}}(d_{13}^2 - d_{23}^2 - D_{14}^2 + D_{24}^2), -h_3 \right)$

3. Simulation and Experimental Results

To validate and analyze the proposed model, a simulation has been conducted in Matlab for an aforesaid problem domain; later a stringent experiment has been performed with ultrasonic sensors in terrestrial environment to reinforce the proposed method.

3.1. Simulation Results

The problem domain imitates a scenario for a 200m water column where a single beacon has been used. The method has been devised in [9] elaborately where a group of three sensors are placed arbitrarily and a single beacon is placed to imitate the surfaced sensor. The plane where the sensors have been deployed considered being XY plane and the beacon's surfing area has been considered to be parallel to XY plane. First sensor's coordinate is $(0,0,0)$, which is considered to be the origin of the Cartesian system; whereas the second sensor's is placed on $(0,75,0)$ being on the y -axis. Lastly, the third sensor was placed on $(80,40,0)$ where z component is zero like others. Other environmental variables for the water column are -30°C at the surface and 15°C at the bottom where sensors are deployed having a salinity variation of 0.5ppt between surface and XY plane. Gaussian noise with $(\mu=0, \sigma=1)$ has been added to bottom temperature as well as with flight time of acoustic signals from beacon to submerged nodes.

As the procedure demands, the surfaced sensors has been shifted and measured from six different places; all these six positions were in close proximity to mitigate error incorporation.

However, mobility of the deployed sensors were ignored in the proposed model as this is out of the scope of this paper. Possible errors in determined coordinate of S_2 and S_3 are shown in Figure 6 and 7 for several iterations. As sensor S_1 has been kept at the origin of Cartesian coordinate system; hence generating no error. Having positional distance error of 0.62m with standard deviation of 0.478 for sensor S_2 and 0.75m with standard deviation of 0.603 for sensor S_3 also validate the model.

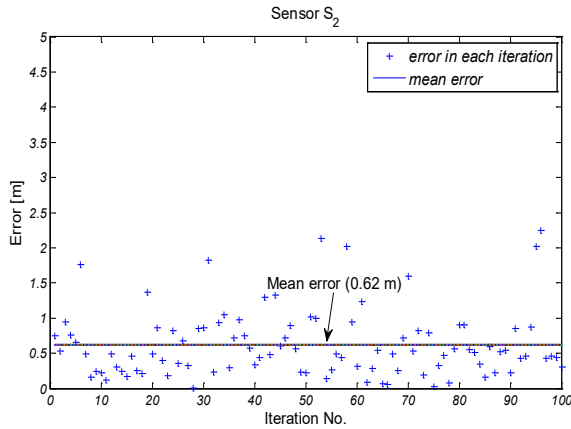


Figure 6: Positional Error from Original Position (0,75,0) for Sensor S_2 (without Gaussian noise)

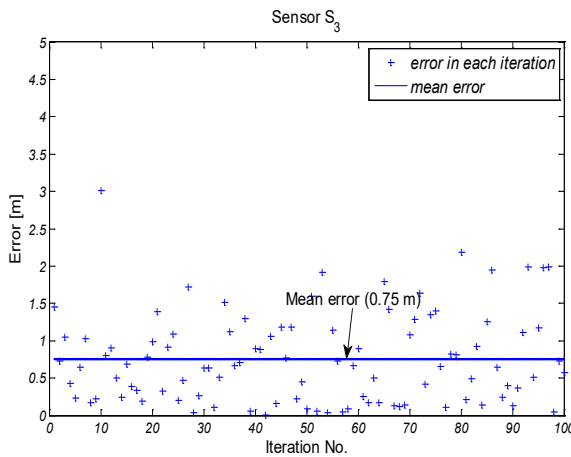


Figure 7: Positional Error from Original Position (80,40,0) for Sensor S_3 (without Gaussian noise)

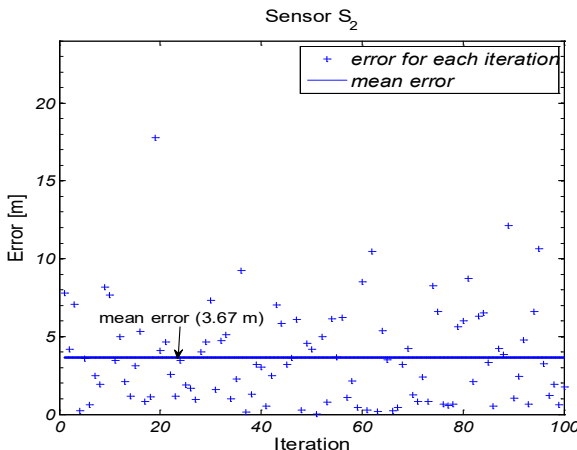


Figure 8: Positional Error from Original Position (0,75,0) for Sensor S_2 (with Gaussian noise)

Errors in determined coordinates with true Euclidean distance i.e. without Gaussian noise are almost negligible; this negligible error in turn validates the proposed mathematical model for coordinates determination with a single beacon. Besides, considering the physical sizes of the sensors deployed underwater, Figure 8 and 9 show errors (with Gaussian noise) are within acceptable range considering for a 200m water column.

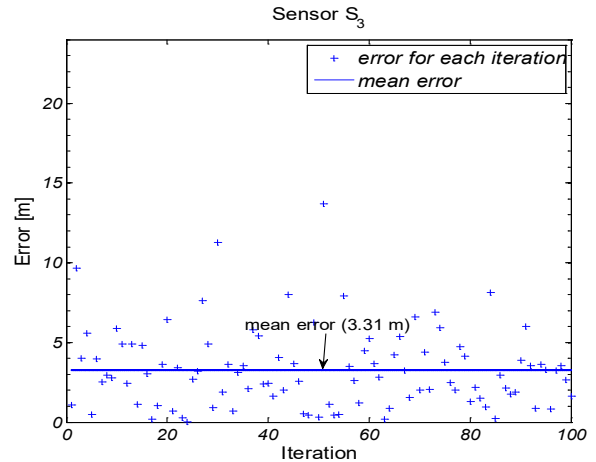


Figure 9: Positional Error from Original Position (80,40,0) for Sensor S_3 (with Gaussian noise)

For sensors with dimension 4.5x2x1.5cm, positional error 0.01-3.81cm is quite outstanding. As simulation in previous sections validates the model with the negligible error considering Euclidean distances, it is now conspicuous that the precision of the distances measured from the beacon to the deployed sensors produce lesser positional errors.

3.2. Experimental Results

A stringent experiment has been conducted in the terrestrial environments i.e. in the lab with compatible ultrasonic sensors to prove the proposed mathematical model for coordinates determination with a single beacon using Cayley-Menger determinant. It also shows that Cayley-Menger determinant could be used for non-parallel state situation with a little adjustment of the proposed model. Flight time i.e., signal's propagation delay of the acoustic signal is used to determine the distances between the surfaced beacon and deployed submerged sensors. A beacon at the ceiling and three other sensors on the table top have been kept to achieve that same scenario as the problem domain described in previous section.

Two different scenarios with four individual tests have been conducted to validate the model. Firstly, in scenario 1 positions of the sensors are S_1 , S_2 and S_3 are (0,0,0), (0,20,0), (30,15,0) respectively. Keeping the origin of the Cartesian system on sensor S_1 , positional error for sensors S_2 and S_3 are within 0.2 and 4cm range. Accuracy in distance measurements with the ultrasonic sensor generates positional error: 0.17cm, whereas in extreme case it is 3.85cm as depicted in Table 4. Secondly, in scenario 1 positions of the sensors are S_1 , S_2 and S_3 are (0,0,0), (0,25,0), (35,10,0) respectively. Keeping the origin of the Cartesian system on sensor S_1 , positional error for sensor S_2 and S_3 are within 0.5 and 6cm range. Accuracy in distance measurements with the ultrasonic sensors generates positional error: 0.17cm, whereas in extreme case it is 3.85cm as illustrated in Table 5.

The ultrasonic sensors used in the experiment have 12° sentry angle; this capacity limits the movement of the surfaced sensor. Besides the built-in Arduino *microseconds To Centimeter* function needed to be changed to ‘double’ to acquire more precise timing as the experiment is taken place where the maximum distances between beacon and deployed sensors are less than 80cm range.

One other challenge we faced while experimenting with Arduino board is to process the generated ultrasonic pulses that are received by the sensors to calculate inter distances. As in [14], the ‘pulseIn’ function usually takes more than 20ms to process the received pulse, whereas it takes only 2.32ms to travel 80cm (approximate max distance for the experimental domain); as a result by the time it finishes pulse processing for the nearest sensor from the beacon and starts processing pulses for other sensors, it is then too late for the pulses to be on the flight. To mitigate this problem we had to generate two other 10µs pulses for rest of the two sensors in 50ms interval. So within around 100ms all three pulses are generated, this fraction of a second will have no effect on the stationary sensor nodes scenario; however, will have negligible effect in cases of mobility.

4. Conclusions

Persistent and accurate positioning is indispensable in various arenas due to necessity as well as research. As specific applications demand more accuracy and convenience, research pushes the boundary to fit and meet the demand in its own course. So, a plethora of localization algorithm has been proposed. This paper illustrates and analyzes a pragmatic approach of localization where a single node can determine deployed submerged nodes using Cayley-Menger determinant; it also portraits associated difficulties of using Cayley-Menger in non-parallel state situations in real time as the original model was designed for parallel states. In nature, parallel state situation is very rare where the plane of the beacon node and the plane of deployed sensors would be parallel. Hence, the necessity of having a pragmatic solution was conspicuous.

A solvable configuration of the domain and its associated model has been simulated to validate as well as to fathom the degree of errors. Once the distances between beacon and the deployed submerged sensors are considered to be true Euclidean, it generates negligible errors. It is also shown that the generated errors are because of erroneous distance determination; so the accuracy of coordinates solely depends on the preciseness of distance measurement method, not the model. This paper also delineates how the model could be used for non-parallel state situation with a little adjustment. Simulation results with Gaussian noise in distance measurements suggest that 0.62-3.67m error for a 200m water column is outstanding as sizes of deployed sensors or AUVs could reach few meters in length. However, mobility of the deployed sensors was not considered in this paper and left for future exploration. On the other hand, mobility of the beacon and span has very limited effect on the coordinates determination of the sensors.

To validate the simulation, a stringent experiment has been conducted in terrestrial environment. So called off-the-shelf ultrasonic sensors have been used imitating the simulated configuration. In the experiment, it has been shown that distance between beacon and deployed sensor can be calculated using acoustic signals, whereas clock synchronization could be performed using electrical signals. Mobility of the sensors was not considered in the experiment as well; however, multiple scenarios

Experimental Scenario 1: Original coordinates of sensors:

$$S_1: (0,0,0); S_2: (0,20,0); S_3: (30,15,0)$$

Table 4: Coordinates of the Sensors S_2 and S_3 according to Scenario 1

Trial 1				Trial 2			
	S_1	S_2	S_3		S_1	S_2	S_3
readings				readings			
R1	54.14	52.31	50.36	R1	62.57	57.27	52.20
R2	57.96	52.65	51.10	R2	60.31	54.78	53.17
R3	51.32	50.56	52.87	R3	53.39	52.60	55.01
R4	61.11	52.49	55.71	R4	63.58	54.62	57.96
R5	59.90	58.57	48.99	R5	62.32	60.94	50.98
R6	60.14	55.04	50.17	R6	56.34	54.42	52.40
Calculated Coordinates:				Calculated Coordinates:			
$S_1: (0,0,0)$				$S_1: (0,0,0)$			
$S_2: (0,19.83,0)$				$S_2: (0,20.55,0)$			
$S_3: (29.58,14.94,0)$				$S_3: (30.88,15.45,0)$			

Trial 3				Trial 4			
	S_1	S_2	S_3		S_1	S_2	S_3
readings				readings			
R1	60.76	55.20	53.57	R1	61.82	53.11	56.36
R2	56.76	54.83	52.79	R2	58.64	53.27	51.70
R3	53.80	53.00	55.43	R3	51.92	51.15	53.49
R4	64.06	55.03	58.40	R4	54.78	52.92	50.95
R5	62.79	61.41	51.36	R5	60.60	59.26	49.57
R6	63.04	57.71	52.60	R6	60.84	55.69	50.76
Calculated Coordinates:				Calculated Coordinates:			
$S_1: (0,0,0)$				$S_1: (0,0,0)$			
$S_2: (0,20.73,0)$				$S_2: (0,20.01,0)$			
$S_3: (30.93,15.42,0)$				$S_3: (29.94,15.01,0)$			

Experimental Scenario 2: Original coordinates of sensors:

$$S_1: (0,0,0); S_2: (0,25,0); S_3: (35,10,0)$$

Table 5: Coordinates of the Sensors S_2 and S_3 according to Scenario 2

Trial 1				Trial 2			
	S_1	S_2	S_3		S_1	S_2	S_3
readings				readings			
R1	65.34	62.51	53.19	R1	59.85	57.26	48.71
R2	54.94	58.15	56.49	R2	50.32	53.25	51.74
R3	55.90	54.13	58.85	R3	53.70	53.04	49.86
R4	58.63	57.91	54.44	R4	51.20	49.57	53.90
R5	62.95	60.46	53.42	R5	57.65	55.37	48.92
R6	60.32	55.26	57.25	R6	55.25	50.62	52.43
Calculated Coordinates:				Calculated Coordinates:			
$S_1: (0,0,0)$				$S_1: (0,0,0)$			
$S_2: (0,26.12,0)$				$S_2: (0,23.72,0)$			
$S_3: (36.55,12.54,0)$				$S_3: (33.26,12.94,0)$			

Trial 3				Trial 4			
	S_1	S_2	S_3		S_1	S_2	S_3
readings				readings			
R1	56.77	52.01	53.88	R1	62.36	59.90	52.92
R2	51.71	54.73	53.17	R2	54.43	57.61	55.97
R3	55.18	54.51	51.24	R3	58.09	57.38	53.94
R4	52.61	50.94	55.39	R4	64.74	61.94	52.69
R5	59.25	56.90	50.28	R5	55.38	53.62	58.31
R6	61.50	58.84	50.06	R6	59.76	54.75	56.72
Calculated Coordinates:				Calculated Coordinates:			
$S_1: (0,0,0)$				$S_1: (0,0,0)$			
$S_2: (0,24.53,0)$				$S_2: (0,25.81,0)$			
$S_3: (34.33,12.63,0)$				$S_3: (36.15,13.64,0)$			

and tests suggest the validity of the model. Acoustic signals are distressed by environmental variables like temperature and humidity as well as propagation speed is low compared to electrical signals. Besides, sensors that have been used in the experiment have limitations and constraints in signal processing; result indicates 0.01-3.81cm positional error for sensors with dimension 4.5x2x1.5cm is within acceptable range.

References

- [1] J. H. Cui, J. Kong, M. Gerla, and S. Zhou, "The challenges of building mobile underwater wireless networks for aquatic applications," *Network, IEEE*, vol. 20, pp. 12-18, 2006.
- [2] L. Hu, B. Liu, K. Zhao, X. Meng, and F. Wang, "Research and Implementation of the Localization Algorithm Based on RSSI Technology," *Journal of Networks*, vol. 9, pp. 3135-3142, 2014.
- [3] P. Duff and H. Muller, "Autocalibration algorithm for ultrasonic location systems," in *Seventh IEEE International Symposium on Wearable Computers, 2003. Proceedings*, pp. 62-68, 2003.
- [4] E. Olson, J. Leonard, and S. Teller, "Robust range-only beacon localization," in *Autonomous Underwater Vehicles, IEEE/OES*, pp. 66-75, 2004.
- [5] J. Guevara, A. R. Jimenez, A. S. Morse, J. Fang, J. C. Prieto, and F. Seco, "Auto-localization in Local Positioning Systems: A closed-form range-only solution," in *IEEE International Symposium on Industrial Electronics (ISIE)*, pp. 2834-2840, 2010.
- [6] J. C. Prieto, A. R. Jiménez, J. Guevara, J. L. Ealo, F. Seco, J. O. Roa, and F. Ramos, "Performance evaluation of 3D-LOCUS advanced acoustic LPS," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, pp. 2385-2395, 2009.
- [7] A. Rahman and V. Muthukkumarasamy, "Embracing localization inaccuracy with a single beacon", *Journal of Advanced Computer Science and Application (JACSA)*, Vol. 10 No. 12, 2019.
- [8] A. Rahman and V. Muthukkumarasamy, , "Localization of Submerged Sensors with a Single Beacon for Non-Parallel Planes State", in *10th International conference on Ubiquitous and Future Networks (ICUFN)*, pp. 525-530, 2018.
- [9] A. Rahman, V. Muthukkumarasamy, and E. Sithirasenan, "Coordinates Determination of Submerged Sensors Using Cayley-Menger Determinant," in *Distributed Computing in Sensor Systems (DCOSS), IEEE*, pp. 466-471, 2013.
- [10] A. Rahman, V. Muthukkumarasamy, and E. Sithirasenan, "Localization of Submerged Sensors Using Radio and Acoustic Signals with Single Beacon," in *Ad-hoc, Mobile, and Wireless Network. LNCS*. vol. 7960, J. Cichoń, M. Gębala, and M. Klonowski, Eds., ed: Springer Berlin Heidelberg, pp. 293-304, 2013.
- [11] I. Vasilescu, K. Kotay, D. Rus, M. Dunbabin, and P. Corke, "Data collection, storage, and retrieval with an underwater sensor network," in *Proceedings of the 3rd international conference on Embedded networked sensor systems*, pp. 154-165, 2005.
- [12] J. Guevara, A. Jiménez, J. Prieto, and F. Seco, "Auto-localization algorithm for local positioning systems," *Ad Hoc Networks*, vol. 10, pp. 1090-1100, 2012.
- [13] A. Rahman, V. Muthukkumarasamy, and X. Wu, "Coordinates and Bearing of Submerged Sensors Using a Single Mobile Beacon (CSMB)", *Journal of Networks*, Vol. 10 No. 8, 2015.
- [14] P. Xie, J. H. Cui, and L. Lao, "VBF: vector-based forwarding protocol for underwater sensor networks," *Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems*, pp. 1216-1221, 2006.

Analysis and Improvement of an Innovative Solution Through Risk Reduction: Application to Home Care for the Elderly

Linda Acosta-Salgado*, Auguste Rakotondranaivo, Eric Bonjour

Research Team on Innovation Process (Laboratoire ERPI, EA N° 3767), Université de Lorraine, Nancy, France

ARTICLE INFO

Article history:

Received: 17 January, 2020

Accepted: 27 March, 2020

Online: 08 May, 2020

Keywords:

Acceptability

Acceptance

Need

Risk

Elderly care

Healthcare innovation

ABSTRACT

The increase in the number of elderly people requires multiple efforts to maintain their well-being and health. A wide variety of products and services have been created to enable seniors to live at home for as long as possible. The market success of these solutions depends on acceptance by the different stakeholders. Older people are reluctant to change their environment, most notably their home. Solutions must provide a benefit and/or reduce risks related to their everyday life in order to be accepted. Design methods are mainly focused on needs analysis, while acceptability assessment models are based on the study of benefits. A need may also correspond to a risk that may be present in the initial situation and that has to be eliminated or reduced. The notion of risk is not sufficiently integrated in these models. This paper proposes a new approach to analyze the actual situation of elderly people at home based on risk analysis. The objective is to contribute to the design of solutions that will be more readily accepted by this population. A model for estimating the risk of falling has been proposed. The probability of two elderly people falling in their home is assessed using this model. The design and improvement of solutions are explored using the results obtained.

1. Introduction

The ageing of the world's population is a key issue for the years to come. Demographic trends reveal an increasing proportion of older people and rising life expectancy. According to projections made by the National Institute of Statistics and Economic Studies (INSEE), in 2070, France should have 76.5 million inhabitants, 10.7 million more than in 2013. Specifically, the population aged 65 and over should increase by 11.2 million and those aged 20–64 decrease by 7.9 million [1]. Thus, by 2070, there will be more seniors than younger people to help them. This situation confronts the health field with a major societal challenge: delaying the age of entry into dependency and improving the well-being of the elderly by controlling risks and expenditure [2].

Public policy encourages the creation of new products and services for keeping people at home as long as possible. For the private industrial and commercial sectors, ageing represents an economic opportunity. Technology companies see this segment of the population as an attractive market for products and services that are easy to use and affordable [3]. In recent years, many

technological and organizational innovations have been developed to improve elderly home care. These include video-vigilance systems, fall detectors, tele-assistance and home automation equipment.

Innovative solutions for the ageing population could increase elderly people's capabilities and help them to reduce their entry into dependency by transforming their living environment [4]. Hence, they have to satisfy the needs and preferences of multiple stakeholders (older people, their families, caregivers, care structures, etc.). However, these stakeholders are not always in favor of integrating changes into their daily lives [5].

The success of an innovative project is evaluated by its acceptability in the design phase and by its acceptance in the use phase [6]. To promote the acceptance of a product or service, it is necessary to accompany its design process, from the early stages, by evaluating its acceptability [7]. To improve innovation acceptance, the project must provide benefits and/or reduce the risks perceived by its future users.

In order to improve the well-being of older people at home and increase the likelihood that innovative solutions will be accepted by their various users, it is important to integrate the notion of risk

*Corresponding Author: Linda Acosta-Salgado, Université de Lorraine, Nancy, France. Email: linda.acosta-salgado@univ-lorraine.fr

into the design and improvement process of solutions. The recognition of the positive and negative points of a situation leads design teams to imagine new products and develop new ways of proceeding [8]. Within this framework, this article proposes an analysis approach for improving innovative solutions based on risk analysis of the real situation of elderly people staying at home.

This paper is an extension of work presented in the International Conference on Engineering, Technology and Innovation (ICE/ITMC) [7]. The approach proposed for analyzing and improving a solution based on its risk study is more detailed. After this introduction, the second section presents a literature review on design methods, acceptability assessment models and risks for older people living at home. The next section proposes a conceptual approach and a model to evaluate the risk of falling. It also describes the methods and materials used to assess the fall risk of two people living in their own home. The fourth section presents the results and explores the value of addressing these results to improve a solution. The final part presents the discussion, conclusions and research perspectives.

2. Literature Review

2.1. Design Methods

The notion and discipline of design has been in constant evolution [9]. The development of new products was initially based on the search for function-based solutions. Over time, user needs were seen as a determining factor in the design process, and several user-centered design currents have developed in recent decades.

Today, design approaches such as user innovation, user experience, emotional design and design thinking, seek to better understand the needs of users and conceive new solutions alongside them [10]. Different design stages have been proposed as well as different ways of identifying them [11]. Generally, these processes include phases for understanding the problem, defining the design project, finding solutions, prototyping and evaluation. At the end of this process, the expected results boost the acceptability of the new solutions.

In various design models, the notion of need is fundamental. Need can be defined as a desire resulting from a lack or a dissatisfaction in the goal achievement process [12]. It is relative to an individual and is perceived differently by each person. Need perception can also be the product of risk perception. This need consists in eliminating or reducing the severity or the probability of occurrence of this risk.

2.2. Acceptability Assessment Models

A lot of research has been carried out to assess the acceptability and acceptance of innovative solutions. Researchers have proposed different theoretical models such as the Theory of Reasoned Action (TRA) [13], the Theory of Planned Behavior (TPB) [14] and the Technology Acceptance Model (TAM) [15]. These models have been modified several times by incorporating new factors and relationships between them [16].

The changes made to these models have been aimed at adapting them to evaluate specific innovations in a given context (actors,

uses) and in a particular field. Many models have been created in this way, such as TAM 2, TAM 3 and UTAUT.

A literature review of models used in the health field [7] revealed that most existing models are not suitable to assess acceptability at the design stage or to incorporate the preferences of multiple stakeholders. In these models, the assessment is mainly based on the analysis of factors such as utility and ease of use. The review identified only a few studies that incorporate "perceived risk" as a factor in acceptance assessment models [17–19]. No documents were identified proposing a risk-based approach to improve an innovative solution or to design new concepts.

2.3. Risks for Elderly People at Home

Risks can be defined as people’s feelings about the potential negative consequences of a situation [20] that may prevent goal achievement. Risk perception is a decisive variable in the choice of behavior [21]. Understanding and managing risk is linked to innovation. Awareness of risks and their factors is a necessary condition for making intelligent decisions in innovation [8]. They are both a threat to the success of a project and an opportunity to make improvements.

Different approaches have been used to diagnose the risks faced by older people at home [22–24]. A literature review conducted by Stuck presents a list of risks of declining functional status for older adults living at home [25]. However, these risks focus on the biological aspects of elderly people such as physical and mental health and social aspects. Other studies have identified additional risks that take into account the impact of the environment on the well-being of elderly people living at home [26–28].

In a previous study [7], risk identification regarding elderly people at home was carried out. This was done through a literature review and the application of expert interviews. Twenty-two risks were identified and classified into eight categories (figure 1).

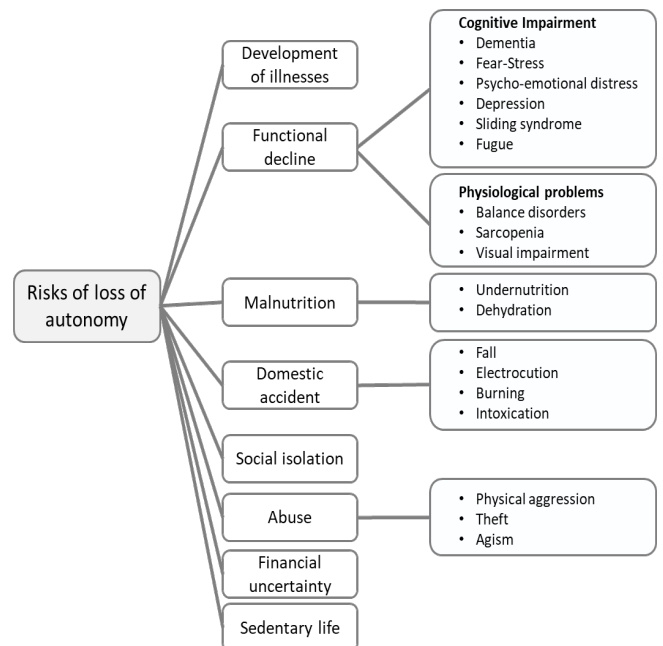


Figure 1: Risks to the elderly at home

The occurrence of such risks is a determining factor in the deterioration of elderly people’s well-being and endangers their ability to stay at home. If one of these risks is detected, it could influence the occurrence of others in a systematic way. In order to estimate the likelihood of a risk occurring, it is necessary to identify the causal factors.

2.4. Risk of Falling

Falls are among the most common and serious problems, the consequences of which mainly affect individuals (functional decline, morbidity, institutionalization, mortality, etc.) and the health system [29]. Most falls by the elderly are the result of interactions between intrinsic (physical and cognitive) and extrinsic (socio-economic and environmental) factors [30]. The identification of risk factors for falls among older adults living at home has been the subject of several studies [31–34]. They have shown that the risk of falling increases as the number of risk factors increases.

3. Material and Methods

3.1. Approach for Analyzing and Improving Innovative Solutions

This article’s assumption is that the notion of risk can contribute to the understanding of a problem situation and impact upon the design of innovative solutions. It may enrich the process of needs analysis as well as the evaluation and improvement of proposed solutions. The approach is depicted in Figure 2.

In the proposed approach, three stages of the design process are considered. The first step is the exploration of the problem. At this stage, different stakeholders (funders, project managers, future users, regulatory institutes) are involved in the needs analysis. Using different engineering and design tools, risks, lacks, dissatisfactions and desires are explored to analyze needs. Thus, the design objective and specifications can be defined to guide the design team in proposing solutions.

The second step is the resolution of the identified problem. Here, different methods are used for the ideation and prototyping of solutions. In the first step, multiple creativity sessions are carried out in order to outline the design possibilities. Subsequently, an alternative solution is chosen and is then further designed. Different intermediate design objects (ICOs) are developed to communicate and validate the characteristics assigned to the solution. ICO is a partial outcome of the project, such as a schematic, a scale model, or a prototype [35]. The objective is to represent a part or the whole of the proposal in order to evaluate it with users and further develop the product.

The third step, solution evaluation, is an important stage in the design process. It allows the solution concept to be tested with future users to identify the benefits and risks they perceive relative to using this concept. In this way, the acceptability of a solution can be estimated. The result of this evaluation brings elements of improvement (to reduce the perceived risks and/or increase the benefits), which can be considered in the problem understanding and in the specification redefinition. This advanced understanding of the problem will allow designers to refine their proposals and better respond to it. We believe that successive evaluations of the solution and integration of gradual improvements would lead to an increase in acceptability. Thus, at the end of the design phase, the final product would have a higher probability of market acceptance.

The next section will apply this approach to the context of home care for the elderly. When we explored the problem of loss of autonomy among the elderly, the choice was made to apply the approach to the risk of falling. This risk has a high probability of occurrence and its consequences are severe.

3.2. Model for Risk Assessment Using Bayesian Networks – Risk of Falling

In order to estimate the probability of elderly people falling at home, we looked for a useful tool to assess the influence between factors and to determine the probability of occurrence. In the scientific literature, statistical methods are often used; however, they are insufficient for the evaluation of imprecise data from various sources. Recently, [36] used the Bayesian network modelling technique to estimate the probability of innovation acceptance based on an analysis of the influence of various factors. Among existing data modelling and analysis techniques, Bayesian networks are the most suitable for knowledge acquisition, representation and manipulation [37]. This technique makes it possible to reason about a problem by evaluating the causality between factors and sub-factors. In addition, it can be used to process uncertain information from different sources: expert knowledge, questionnaire data or the literature.

Bayesian networks have already been used for fall risk analysis [38, 39]. However, the work carried out to date has not sought to be part of an approach to design and improve innovative solutions.

This article proposes to use this technique to assess the probability of elderly people falling at home. It is also assumed that Bayesian networks can be used to make a wider and overall risk assessment of this population.

The construction of the Bayesian network for estimating the risk of falling was carried out following the steps proposed by

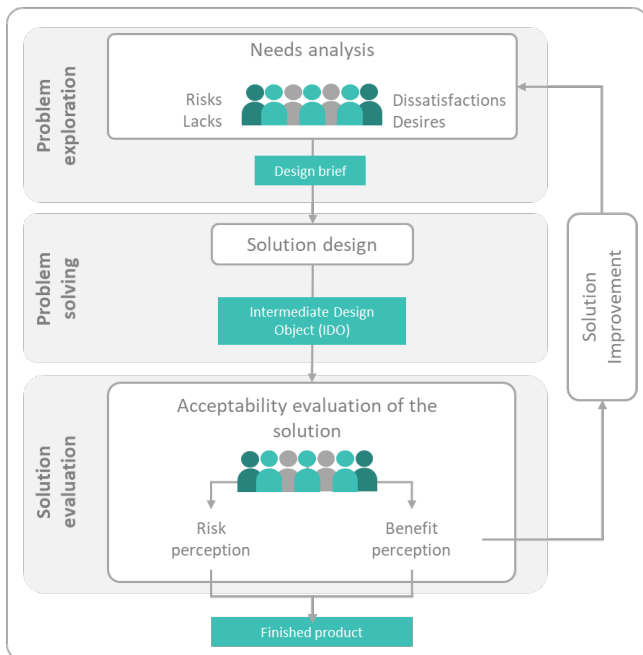


Figure 2: Approach proposed to analyse and improve an innovative solution.

[37]. The first step consists in identifying the risk factors and the statuses that each of them may have. The second is the definition of the Bayesian network structure and the third is the construction of conditional probability tables.

3.3. Identification of Risk Factors

In the first stage, factors were identified mainly through the literature on the risk of falling [30-32,34,40-42]. A list was prepared and presented to an expert gerontologist and two researchers who analyzed and classified these factors (Table 1). Next, two possible statuses were defined for each factor: "yes" and "no". These statuses represent the variables of interest, in other words, the variations that each of the factors may present.

Table 1: Fall Risk Factor Codes

Code	Factor description
AG	Age
GD	Gender
HF	Fall historic
PP	Physiological problems
PP1	Balance disorders
PP2	Visual impairment
PP3	Orthostatic hypotension
CI	Cognitive impairment
CI1	Syncope
CI2	Dementia
CI3	Depression
SI	Social isolation
SI1	Loss of social and family ties
UL	Unhealthy lifestyle
UL1	Alcohol consumption
UL2	Malnutrition
UL3	Lack of physical activity
IE	Inadequate environment
IE1	Inadequate objects
IE2	Inadequate configuration of the living environment
MP	Medication problem
MP1	Poly medication
MP2	Taking psychotropic drugs

3.3.1. Definition of the Bayesian Network Structure

To build the structure of the network, a first proposal of links between factors was made based on the literature consulted. This structure was evaluated and improved with the experts. The final structure is shown in Figure 3. The network is made up of nodes, which correspond to the different factors and sub-factors. These nodes are linked by arcs that represent the causal and conditional dependence between them and the risk to be assessed. The factors have been classified into six groups: physiological problems, cognitive impairment, social isolation, unhealthy lifestyles, inadequate environment and medication problems. Each group is linked to between two and six factors. Four of these factors are related to more than one group: age, gender, history of falls and social isolation.

3.3.2. Construction of Conditional Probability Tables

Conditional probability tables are used to estimate the occurrence probability of one of the node statuses as a function of the parent nodes statuses [37]. A first version of the tables has been proposed from the literature consulted. The information obtained was supplemented by the experts, who used their experience and knowledge to define probability values between factors.

The final version of the network structure and the conditional probability tables by risk factor were modeled using the GeNIe Academic software.

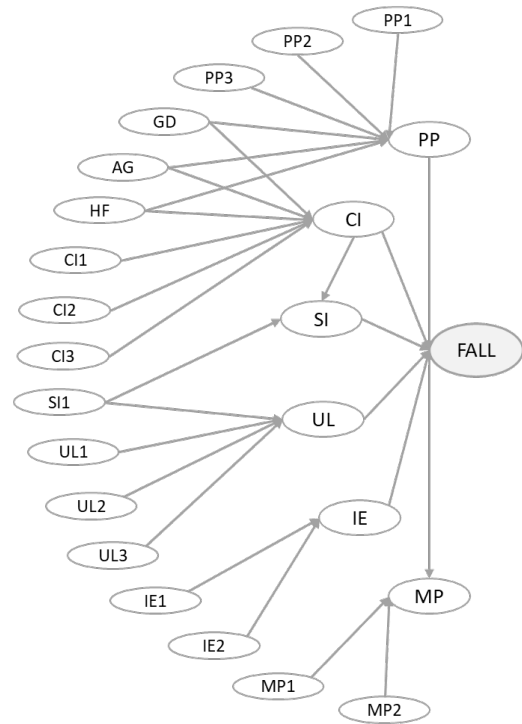


Figure 3: Bayesian network model of home falls risk

3.4. Data Collection Method

In order to make a first assessment of the probability of falling among older people, the proposed model was used. The data collection method chosen was the semi-directive interview. Interviews are one of the most commonly used methods in field studies to collect information about people’s perceptions and preferences in light of their experiences [43]. The semi-directive interview serves to guide the discussion with the interviewee based on a list of themes rather than on direct questions [44].

A semi-directive interview guide was constructed. Following the recommendations of [44], the guide was composed of a list of themes related to the perception of exposure to the different risk factors of the model.

An older couple living independently in their home was selected for interview. Ms. M, 91 years old, and her husband Mr. A, 95 years old, have been married for 70 years. They have been retired for more than 30 years and wish to continue to live in their own home. These people were chosen because of their advanced age and their interest in continuing to live autonomously and safely in their home.

With the consent of these two persons and their children, the interviews were conducted at home. Two members of the research group participated in the development of the interview. One of the researchers led the discussion with each person, while the second researcher wrote down his observations on the development of the interview and on the characteristics of the household. The interviews were recorded and subsequently transcribed.

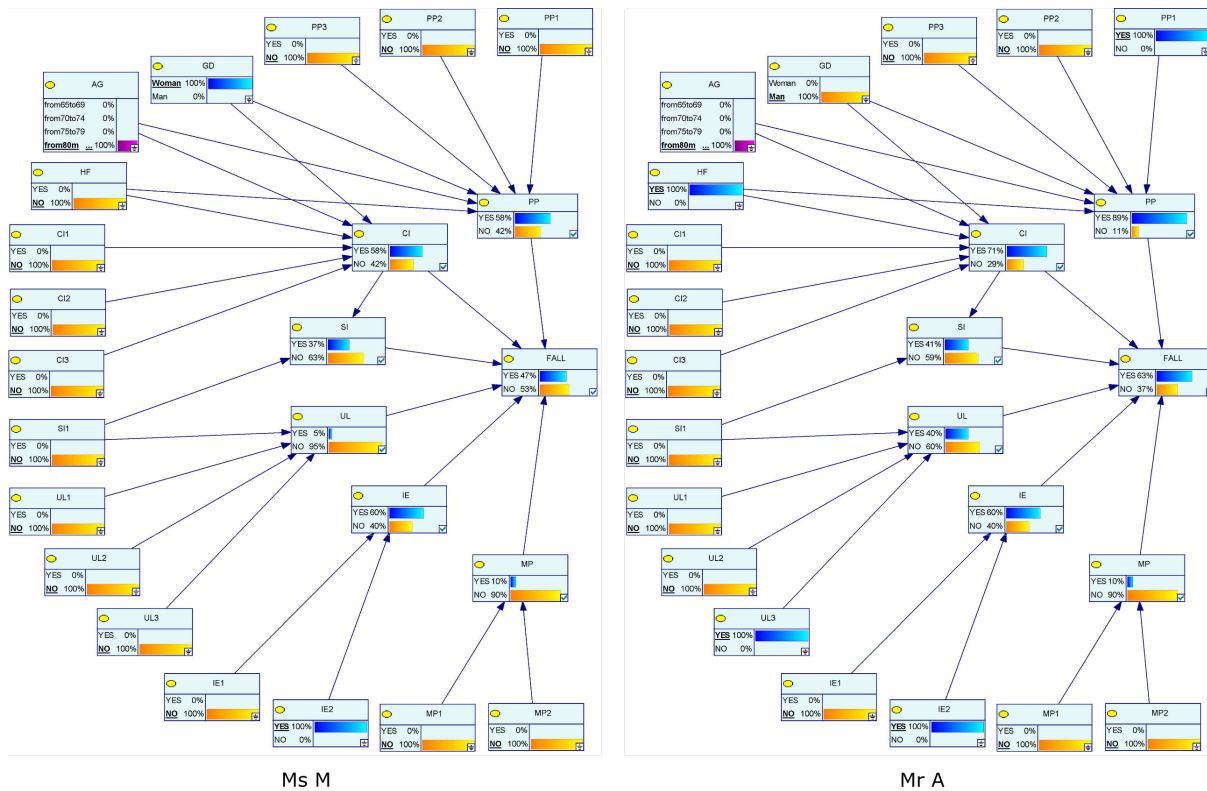


Figure 4: Fall risk estimation for Ms. M and Mr. A.

These elderly person’s perception of the absence or presence of each of the risk factors was interpreted and coded according to the two conditions designated for each factor, namely, yes or no. The results obtained were introduced into the proposed Bayesian network.

The results obtained and the proposed Bayesian network were subsequently used to explore their use in the design and improvement of an innovative solution.

4. Results

4.1. Fall Risk Assessment in the Problem Exploration

In the interview, Ms. M declared that she has no physiological or cognitive problems. She does not consider herself to be in a situation of isolation since she has a large family and maintains strong relationships with them. She says that she does not have any health problems because she eats well, is very active and does not drink alcoholic beverages. In addition, she has home help to prepare meals, clean and care for her. She does not use psychotropic drugs and she does not take multiple medication.

Mr. A, in contrast to Ms. M, was to the victim of a serious fall in the house’s bathroom a few months before the study. After this fall, several modifications were made to the bathroom by the family. However, Mr. A says he has difficulty moving around the house due to the "fear of falling again". This situation has increased his balance problems and decreased his physical activity.

In the observation we remarked that in fact several devices have been installed in the bathroom and in other parts of the house, such as in the bedroom and on the stairs. However, there are areas that may represent a risk: the floors are not non-slip in the kitchen,

corridors and bedroom. Bathroom doors open inwards and locks cannot be opened from the outside. There are no brackets or grab bars in the corridors. They have a remote assistance system, which is installed in the master bedroom, but it cannot be used in the rest of the house.

The information obtained for Ms. M and Mr. A was synthesized in Table 2. Subsequently, this was entered into the Bayesian network and the probability of each person falling was estimated.

Figure 4 shows the networks per person. The existence of a risk factor is represented by the color blue and its absence by the color yellow. For the "age" factor, four bands were used, each represented by a different color.

Table 2. Response by factor

	Ms. M	Mr. A
AG:	91 years old	95 years old
GD:	woman	man
HF	No	Yes
PP1	No	Yes
PP2	No	No
PP3	No	No
CI1	No	No
CI2	No	No
CI3	No	No
SI1	No	No
UL1	No	No
UL2	No	No
UL3	No	Yes
IE1	No	No
IE2	Yes	Yes
MP1	No	No
MP2	No	No

In Ms. M's network, we can identify that the biggest risk factors are inadequate environment (60%), physiological disorders (58%) and cognitive impairment (58%), followed by social isolation (37%). The two other factors, lack of hygiene (5%) and medical problems (10%) are not as representative. Analysis of these factors yields a 47% probability of falling.

In Mr. A's network, the results are different. The risk factors for physiological (89%) and cognitive (71%) problems are very significant. This significance is probably a consequence of the fall he had previously experienced, which caused him balance problems. It may also have been due to the likelihood of isolation risk (41%) and poor hygiene (40%). The occurrence probability of the risk factors "unsuitable environment" (60%) and "medication problems" (10%) were no different from Ms. M's results since they live in the same house. Analysis of these factors results in a greater likelihood of falling, corresponding to 63%.

4.2. Design and Improvement of Innovative Solutions

Using Bayesian networks, the proposed model estimated the probability of two elderly people falling in their homes. An extended model could be useful to detect different risks. On the one hand, the identification of the risk of falling may help project managers to better understand the needs of this population and to define specifications for proposing solutions. On the other hand, based on the risk factors identified, a design team can define areas for improvement of an innovative design solution. These two perspectives for using the results will be explored below.

4.2.1. Designing Innovative Solutions

A variety of solutions can be proposed to reduce the probability of the risk of falling. Some proposals may focus on increasing the person's mobility capabilities according to the characteristics of their home. Other solutions could consist in adapting the living space to the person's abilities.

In the case of Mr. A, who has a high risk of falling, the biggest risk factors are: walking and/or balance disorders, poor configuration of the living space and lack of activity. We imagined a new concept of an exoskeleton that adapts to the person's morphology to help him or her move around at home. In order to find out whether this technology could be accepted by its future users, the perceived risks could be evaluated.

4.2.2. Improvement of an Innovative Design Solution

If an innovative solution is not acceptable, the risks identified could provide insights for improvement. For example, if we take the exoskeleton and assume that future users do not accept this technology, then based on the analysis of the risk of falling, some changes could be made. The impact of these improvements on the reduction of the fall probability can be estimated with the Bayesian network.

Table 3 summarizes the different fall probabilities for Ms. M and Mr. A when introducing the exoskeleton and incorporating modifications to it.

For Mr A, the introduction of the exoskeleton could reduce the probability of falling by 4%. However, if improvements are made to the solution to adapt it to the environmental conditions, such as a contact surface adaptable to floor characteristics, then the risk of

falling could decrease to 54%. In addition, if the exoskeleton incorporates a robotic system to help maintain balance, change posture and move around, then Mr. A's level of physical activity will improve and, as a result, the probability of falling can be reduced to 50%. These improvements could reduce not only the risk of falling but also general risk perception. As a result, the acceptability of exoskeletons may increase.

Table 3. The probability of falling after improvements

	Fall probabilities for Mr. A	Fall probabilities for Ms. M
Initial situation	63%	47%
Exoskeleton	59%	
Exoskeleton with adaptable contact surface	56%	
Exoskeleton with robotic system	54%	43%
Exoskeleton with adaptable contact surface and robotic system	50%	

In the case of Ms. M, the exoskeleton could reduce the risk of falling by approximately 4%. The various changes made to this solution would not contribute significantly to a further reduction of this risk.

5. Discussion

A theoretical approach is proposed for the design, evaluation and improvement of innovative solutions. Unlike other existing methods [12], our proposal focuses on the notion of risk. In representing the approach, we have established three steps for the design process: problem exploration, problem solving, and solution evaluation. In the first phase, we propose that the needs analysis integrates the identification of risks in the initial situation. This would help the design team to identify needs that have not been expressed by stakeholders. In the second phase, these risks, translated into requirements, allow the designers to propose alternative solutions and develop different intermediate design objects. Finally, in the third stage, we propose to evaluate risk perception at the same time as benefit perception in order to estimate the acceptability of the solution. We believe that the results of this evaluation may be useful for a better understanding of the situation of the problem and for improving a solution being designed. Multiple iterations between problem understanding and solution assessment could lead to a product design that will better meet the needs of the target population and be more acceptable to the market.

As part of this approach, the work carried out focused on the risks for elderly people at home.

Risk identification for a person at home implies knowing the causal factors of each risk and relating them to each other. In order to propose a model for estimating the occurrence probability of a risk, we chose one of the most recurrent and significant risks, that of falling. This risk has already been studied [31-34], but no assessment model involving various types of causal factors has been proposed.

We proposed a model for estimating the risk of falling using the Bayesian network technique. Arbelaez already used this method to evaluate the acceptability of innovation [36]. Other

studies have used the same technique to assess the risk of falling [38, 39]. We use this technique to estimate the risk of falling by evaluating different types of causal factors: physiological problems, cognitive impairments, social isolation, environmental maladjustment, poor living conditions and medical problems. The results are intended to contribute to the analysis of needs and the evaluation and improvement of new solutions.

We used this model with two elderly people living at home. Interviews were conducted to obtain information. Using the proposed model, we identified a probability of falling for these two individuals. The results are different for each of them. One of the differentiating factors is the fact that one of these two people has already had a fall, which has affected their mobility and lifestyle.

The risk of falling is only one of the twenty-two identified hazards that can affect the ability to keep elderly people at home. The reasoning used to assess the risk of falling can be replicated to estimate other hazards. To help older adults live well at home, health innovations aimed at reducing risk are needed.

In view of the results obtained, we have imagined the type of solutions that could be proposed to reduce this risk. In addition, in designing an innovative solution, we hypothesized the improvements that could be made to a proposal, based on the risk of falling assessment.

6. Conclusion and Perspectives

This study was conducted with the aim of proposing a new design approach to innovative solutions based on a risk analysis of the real situation of elderly people at home.

In addition to this approach, a risk identification model is proposed, with the purpose of estimating the probability of a person falling. This work is evaluated with a case study concerning the design of an innovative solution for mobility assistance.

The Bayesian network technique was used to build the model. The causal factors of this risk were identified and correlated to define the structure of the model. We believe that this technique could be used for the evaluation of overall risks. This should be validated in future research.

The model developed was applied with two elderly people who were interviewed and observations were made in their homes. The responses obtained were used to feed the network and to determine the probability of each person falling. Based on the high risk of falling presented by one of the people studied, an exploration of the use of these results for the design and improvement of the acceptability of the solution was conducted.

The results presented here are a first exploration of the use of the proposed model. It is important for future studies to validate this approach with a larger number of elderly people.

Further work should consist in assessing risks and benefits globally and with other stakeholders (medical staff, family members, organizations). This evaluation should focus on the development of a fall prevention solution.

When assessing the perception of benefits and risks by multiple actors, various conflicts may arise. It is important to identify and address them. Managing these conflicts could increase the acceptability of an innovative solution.

The method used to build the Bayesian network can be improved. For the construction of the probability tables, several status spaces by factor could be considered in order to obtain more precise results, rather than only yes and no. In addition, field data could be used to complement expert opinions for the definition of the probability tables in a more detailed way.

In the context of a collaborative design approach for innovations in home support for the elderly, this model can be used as a tool for understanding the problem and assisting in decision-making. It will enable project leaders to evaluate and improve the acceptability of a solution, right from the early stages of design.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] N. Blanpain and G. Buisson, "Projections de population à l'horizon 2070, Deux fois plus de personne de 75 ans ou plus qu'en 2013," Paris, France, 2016.
- [2] INSEE, "L'évolution de la dépendance des personnes âgées : un défi en termes de prise en charge et d'emploi", Orléans, 2012.
- [3] S. Peek, "Understanding technology acceptance by older adults who are aging in place: a dynamic perspective," Ph.D Thesis, Universiteit op woensdag, 2017.
- [4] F. Prate, "Evaluation des technologies pour la santé et l'autonomie : Vers un modèle transversal et modulaire. Illustration par la technologie robotique," Ph.D Thesis, Université de Nice Sophia Antipolis, 2015.
- [5] A. M. Sponselee, B. Schouten, D. Bouwhuis, and C. Willems, "Smart home technology for the elderly: perceptions of multidisciplinary stakeholders," in *Constructing Ambient Intelligence: AMI 2007 Workshops*, Darmstadt, Germany, 2007. https://doi.org/10.1007/978-3-540-85379-4_37
- [6] G. Arbelaez Garces, E. Bonjour, and A. Rakotondranaivo, "Towards a method for acceptability analysis: Application to healthcare innovation," in *International Conference on Engineering, Technology and Innovation: Engineering Responsible Innovation in Products and Services (ICE 2014)*, Bergamo, Italy, 2014. <https://doi.org/10.1109/ICE.2014.6871583>
- [7] L. Acosta-Salgado, A. Rakotondranaivo, and E. Bonjour, "Innovation acceptability in elderly care: a risk evaluation approach," in *IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, Sophia-Antipolis, France, 2019. <https://doi.org/10.1109/ice.2019.8792564>
- [8] S. Cleary and T. Malleret, *Risques, perception, évaluation, gestion : Une approche positive des risques globaux auxquels sont confrontés les décideurs*. Maxima, 2006.
- [9] L. Acosta-Salgado, L. Morel, and I. Verilhac, "Towards a better understanding of the concept of design," *Proj. / Proyética / Proj.*, 2(20), 91–114, 2018. <https://doi.org/10.3917/proj.020.0091>
- [10] N. Skiba, "Processus d'innovation centré sur l'utilisateur: identification des besoins et interprétation de données issues de l'intégration de l'utilisateur dans le processus de co-conception," Ph.D Thesis, Université de Lorraine, 2014.
- [11] Carine Lallemand and G. Gronier, *Méthodes de design UX*, EYROLLES, 2016.
- [12] H. BEN Rejeb, V. Boly, and L. Morel, "A new methodology based on kano model for the evaluation of a new product acceptability during the front-end phases," in *IEEE International Computer Software and Applications Conference*, Turku, Finland, 2008. <https://doi.org/10.1109/COMPASAC.2008.94>
- [13] M. Fishbein and I. Ajzen, *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*, Addison-We, 1975.
- [14] I. Ajzen, "The Theory of Planned Behavior," *Organ. Behav. Hum. Decis. Process.*, 50(2), 179–211, 1991. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- [15] F. D. Davis, "A technology acceptance model for empirically testing new end-user information systems: Theory and results," Ph.D Thesis, Massachusetts Institute of Technology, 1986.
- [16] H. Taherdoost, "A review of technology acceptance and adoption models and theories," in *11th International Conference Interdisciplinarity in Engineering, INTER-ENG 2017*, Tirgu Mures, Romania, 2017.

- <https://doi.org/10.1016/j.promfg.2018.03.137>
- [17] Y. Tanaka, M. Kitayama, S. Arai, and Y. Matsushima, "Major psychological factors affecting consumer's acceptance of food additives: Validity of a new psychological model," *Br. Food J.*, 117(11), 2788–2800, 2015. <https://doi.org/10.1108/BFJ-02-2015-0062>
- [18] A. Bearth, M. E. Cousin, and M. Siegrist, "The consumer's perception of artificial food additives: Influences on acceptance, risk and benefit perceptions," *Food Qual. Prefer.*, 38, 14–23, 2014. <https://doi.org/10.1016/j.foodqual.2014.05.008>
- [19] J. Choi, A. Lee, and C. Ok, "The effects of consumers' perceived risk and benefit on attitude and behavioral intention: a study of street food," *J. Travel Tour. Mark.*, 30(3), 222–237, 2013. <https://doi.org/10.1080/10548408.2013.774916>
- [20] M. S. Featherman and P. A. Pavlou, "Predicting e-services adoption: A perceived risk facets perspective," *Int. J. Hum. Comput. Stud.*, 59(4), 451–474, 2003. [https://doi.org/10.1016/S1071-5819\(03\)00111-3](https://doi.org/10.1016/S1071-5819(03)00111-3)
- [21] D. Delignières, "Risque préférentiel, risque perçu et prise de risque," In : *Cognition et performance* [online], Paris, France, 1993. <https://doi.org/10.4000/books.insep.1419>
- [22] S. Vandentorren et al., "August 2003 heat wave in France: Risk factors for death of elderly people living at home," *Eur. J. Public Health*, 16(6), 583–591, 2006. <https://doi.org/10.1093/eurpub/ckl063>
- [23] M. D. L. Á. Vázquez-Sánchez, M. D. C. Gastelu-Cantero, and J. L. Casals-Sánchez, "Valoración de las necesidades de los ancianos que viven solos en una zona básica de salud," *Enferm. Clin.*, 18(2), 59–63, 2008. [https://doi.org/10.1016/S1130-8621\(08\)70699-5](https://doi.org/10.1016/S1130-8621(08)70699-5)
- [24] S. e. Carter, E. m. Campbell, R. w. Sanson-Fisher, S. Redman, and W. j. Gillespie, "Environmental hazards in the homes of older people," *Age Ageing*, 26(3), 195–202, 1997. <https://doi.org/10.1093/ageing/26.3.195>
- [25] A. E. Stuck, J. M. Walthert, T. Nikolaus, C. J. Büla, C. Hohmann, and J. C. Beck, "Risk factors for functional status decline in community-living elderly people: a systematic literature review," *Soc. Sci. Med.*, 48(4), 445–469, 1999. [https://doi.org/10.1016/S0277-9536\(98\)00370-0](https://doi.org/10.1016/S0277-9536(98)00370-0)
- [26] C. Berr, F. Balard, H. Blain, and J. Robine, "Vieillesse et l'émergence d'une nouvelle population," *Med. Sci.*, 28(3), 281–287, 2012. <https://doi.org/10.1051/medsci/2012283016>
- [27] H. J. Swift, D. Abrams, R. A. Lamont, and L. Drury, "The risks of ageism model: how ageism and negative attitudes toward age can be a barrier to active aging," *Soc. Issues Policy Rev.*, 11(1), 195–231, 2017. <https://doi.org/10.1111/sipr.12031>
- [28] S. Clement, C. Rolland, and C. Thoen-Fabre, *Usage, normes, autonomie : Analyse critique de la bibliographie concernant le vieillissement de la population*, Puca, 2007.
- [29] M. E. Tinetti, C. Gordon, E. Sogolow, P. Lapin, and E. H. Bradley, "Fall-risk evaluation and management: challenges in adopting geriatric care practices," *Gerontologist*, 46(6), 717–725, 2006. <https://doi.org/10.1093/geront/46.6.717>
- [30] M.-L. Gaubert-Dahan, A. Cognaud-petit, L. De Decker, C. Annweiler, O. Beauchet, and G. Berrut, "Essai de modélisation des facteurs de risque de chute chez les sujets âgés," *Ger Psychol Neuropsychiatr Vieil*, 9(3), 277–285, 2011. <https://doi.org/10.1684/pnv.2011.0289>
- [31] R. Rajagopalan, I. Litvan, and T. P. Jung, "Fall prediction and prevention systems: Recent trends, challenges, and future research directions," *Sensors*, 17(11), 1–17, 2017. <https://doi.org/10.3390/s17112509>
- [32] C. A. Pfortmueller, G. Lindner, and A. K. Exadaktylos, "Reducing fall risk in the elderly: risk factors and fall prevention, a systematic review," *Minerva Med.*, 105(4), 275–281, 2014.
- [33] B. Leclerc et al., "Risk factors for falling among community-dwelling seniors using home-care services : An extended hazards model with time-dependent covariates and multiple events," *Chronic Dis. Can.*, 28(4), 111–120, 2008.
- [34] L. Z. Rubenstein, "Falls in older people: epidemiology, risk factors and strategies for prevention," *Age Ageing*, 35(2), 37–41, 2006. <https://doi.org/10.1093/ageing/afl084>
- [35] D. Galvez Manriquez, "Évaluation de la capacité à innover : Une approche par auto évaluation et suivi supporté par des analyses multicritères dynamiques," Ph.D Thesis, Université de Lorraine, 2015.
- [36] G. Arbelaez Garces, A. Rakotondranaivo, and E. Bonjour, "An acceptability estimation and analysis methodology based on Bayesian networks," *Int. J. Ind. Ergon.*, 53, 245–256, 2016. <https://doi.org/10.1016/j.ergon.2016.02.005>
- [37] P. Naïm, P.-H. Wuillemin, P. Leray, O. Pourret, and A. Becker, *Réseaux Bayésiens*, 3e édition, Eyrolles, 2007.
- [38] G. Koshmak, M. Linden, and A. Loutfi, "Dynamic Bayesian networks for context-aware fall risk assessment," *Sensors*, 14(5), 9330–9348, 2014. <https://doi.org/10.3390/s140509330>
- [39] H. Wang, M. Li, J. Li, J. Cao, and Z. Wang, "An improved fall detection approach for elderly people based on feature weight and Bayesian classification," in 2016 IEEE/ICMA International Conference on Mechatronics and Automation, 2016. <https://doi.org/10.1109/ICMA.2016.7558609>
- [40] American Geriatrics Society, "Guideline for the Prevention of Falls in Older Persons," *J. Am. Geriatr. Soc.*, 49(5), 664–672, 2002. <https://doi.org/10.1046/j.1532-5415.2001.49115.x>
- [41] P. J. Fos and C. L. McLin, "The risk of falling in the elderly: A subjective approach," *Med. Decis. Mak.*, 10(3), 195–200, 1990. <https://doi.org/10.1177/0272989X9001000306>
- [42] HAS (Haute Autorité de Santé), "Évaluation et prise en charge des personnes âgées faisant des chutes répétées," Paris, 2009.
- [43] S. Q. Qu and J. Dumay, "The qualitative research interview," 8(3), 2011. <https://doi.org/10.1108/11766091111162070>
- [44] L. Garreau and P. Romelaer, *Méthodes de recherche qualitatives innovantes*. Economica, 2019.

A Perturbation Finite Element Approach for Correcting Inaccuracies on Thin Shell Models with the Magnetic Field Formulation

Vuong Dang Quoc^{*1}, Quang Nguyen Duc²

¹School of Electrical Engineering, Hanoi University of Science and Technology, Viet Nam

²Faculty of Electrical Engineering, Electric Power University, Viet Nam, Viet Nam

ARTICLE INFO

Article history:

Received: 29 March, 2020

Accepted: 02 May, 2020

Online: 11 May, 2020

Keywords:

Magnetic field

Eddy current loss

Joule power loss

Thin shell model

Magnetodynamic

Subproblem technique

ABSTRACT

This research makes the improvement of the errors surrounding edges and corners of thin electromagnetic regions by means of the magnetic field finite element formulation. Classical thin shell electromagnetic models are usually replaced by impedance-type interface conditions throughout surfaces that ignore errors in the calculation of the local fields (magnetic vector potentials, magnetic fields, eddy current densities and Joule power loss densities) in the vicinity of borders and corners. In this context, the inaccuracies of the local fields surrounding edges and curvatures related to the thin shell models are improved by a perturbation finite element technique, permitting to solve each subproblem on its own separately mesh and geometry, which makes reducing the computational time.

1 Introduction

Thin shell (TS) electromagnetic models have been presented by many author [1]-[3] to neglect meshing volumic shells (Fig. 1, left) and are introduced by surfaces (Fig. 1, right) with interface conditions (ICs) related to 1-D analytical distributions across the shell thickness.

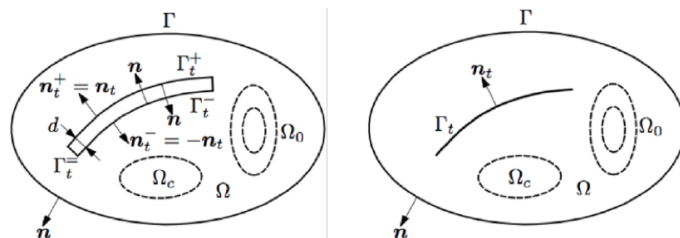


Figure 1: Volumic shell Ω_t (left) and IC Γ_t (right).

In general, the ICs cancel corner and curvature effects, and lead to inaccuracies of magnetic fields, eddy current losses and Joule power losses in the neighbouring regions of geometrical discontinuities around borders and corners, changing with the thickness. So as to overcome with this disadvantage, a subproblem method (SPM) with a magnetic vector potential formulation for improving errors

next to corners and curvatures occurring from the TS has recently presented by many author [4, 5].

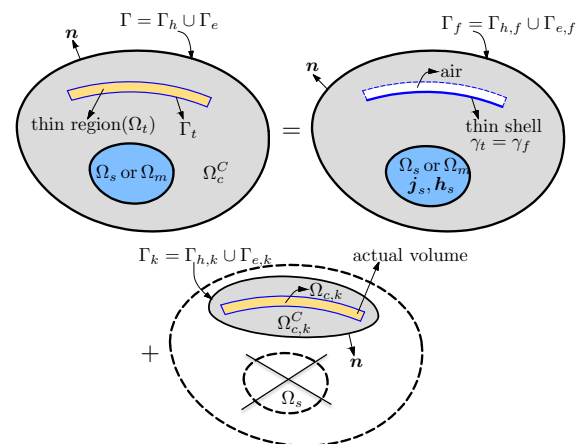


Figure 2: Decomposition of full problem into several SPs.

The idea of this article is to present an extension of the one-process SPM that proposes the magnetic field formulation. The hearth of this technique is based on a SPM, which consists in splitting a complete/full problem (including of stranded inductors and

*Corresponding Vuong Dang Quoc; Email: vuong.danguoc@hust.edu.vn

magnetic or conductive thin structures) into two subproblems (SPs), the complete/full solution being the superposition of all solutions for each SPs.

A problem involving stranded inductors and TS models (Fig. 2, *top right*) is first considered on a simplified mesh. The obtained solution is then adjusted by a improvement problem with the actual volume thin regions (Fig. 2, *bottom*). In this process, each SP is constrained by volume sources (VSs) or surface sources (SSs), where VSs present variations of the permeability and conductivity in volume thin regions and SSs express variations of ICs across surfaces from previous SPs. Each process allows to inherit the previous sources/solutions for new SPs instead of starting a new complete problem for each new geometries as a traditional finite element method (FEM) [6].

2 Sequence of Perturbation Technique

2.1 Magnetodynamic subproblem

In the heart SPM, a canonical magnetodynamic SP p considered at step p is solved in a domain Ω_p , with boundary $\partial\Omega_p = \Gamma_p = \Gamma_{h,p} \cup \Gamma_{e,p}$. The eddy current density is defined in conducting part $\Omega_{c,p}$ of of Ω_p , with $\Omega_p = \Omega_{c,p} \cup \Omega_{c,p}^C$ (where $\Omega_{c,p}^C$ is the non-conducting region. The set of maxwell's equations, constitutive laws and boundary conditions (BCs) of the SPs p give [7, 8]:

$$\text{curl } \mathbf{h}_p = \mathbf{j}_p, \text{ div } \mathbf{b}_p = 0, \text{ curl } \mathbf{e}_p = -\partial_t \mathbf{b}_p, \quad (1a-b-c)$$

$$\mathbf{b}_p = \mu_p \mathbf{h}_p + \mathbf{b}_{s,p}, \mathbf{e}_p = \sigma_p^{-1} \mathbf{j}_p + \mathbf{e}_{s,p}, \quad (2a-b)$$

$$\mathbf{n} \times \mathbf{h}_p|_{\Gamma_{h,p}} = \mathbf{j}_{su,p}, \mathbf{n} \times \mathbf{e}_p|_{\Gamma_{e,p} \cap \Gamma_{b,p}} = \mathbf{k}_{su,p}, \quad (3a-b)$$

where \mathbf{h}_p is the magnetic field (A/m), \mathbf{b}_p is the magnetic flux density (T), \mathbf{e}_p is the electric field (V/m), $\mathbf{j}_{s,p}$ is the electric current density (A/m^2), μ_p is the magnetic permeability (H/m), σ_p is the electric conductivity (S/m) and \mathbf{n} is the unit normal exterior to Ω_p . The source fields ($\mathbf{b}_{s,p}$ and $\mathbf{e}_{s,p}$) in (2a-b) are VSs, and the source fields ($\mathbf{j}_{su,p}$ and $\mathbf{k}_{su,p}$) in (3a-b) are SSs. In general, SSs are defined as zero for classical homogeneous BCs. ICs can express their discontinuities across the negative and positive sides (γ_p^+ and γ_p^-) of any interface γ_p (with the notation $[\cdot]_{\gamma_p} = |\cdot|_{\gamma_p^+} - |\cdot|_{\gamma_p^-}$) in Ω_p . For the idealized thin regions, these ICs are SSs that have to account for special phenomena occurring between two sides of γ_p [4, 5].

2.2 Relations of SPs with SSs and VSs

As proposed in [3], the relation between SPs p (TS models and volume corrections) are respectively defined via SSs and VSs. A volume thin region $\Omega_{t,p}$ in $\Omega_{c,p}$ is initially removed from Ω_i and then defined with the double layer TS surface $\Gamma_{t,p}$. For the magnetic field formulation (\mathbf{h}_p -conformal formulation), the coefficient discontinuity $\mathbf{h}_{d,t,p}$ of the tangential component $\mathbf{h}_{t,p} = (\mathbf{n} \times \mathbf{h}_p) \times \mathbf{n}$ of \mathbf{h}_p is presented across via the TS model, i.e. [3]

$$[\mathbf{n} \times \mathbf{h}_{t,p}]_{\Gamma_{t,p}} = \mathbf{n} \times \mathbf{h}_{d,t,p}, \quad (4)$$

where the discontinuous component ($\mathbf{h}_{d,t,p}$) of the field $\mathbf{h}_{t,p}$ is considered as zero along the TS border, which neglects the present magnetic fields. To present this discontinuity, it gets [3]

$$\mathbf{h}_{t,p}|_{\Gamma_{t,p}^+} = \mathbf{h}_{c,t,p} + \mathbf{h}_{d,t,p}, \mathbf{h}_{t,p}|_{\Gamma_{t,p}^-} = \mathbf{h}_{c,t,p}, \quad (5)$$

where $\mathbf{h}_{c,p}$ is the continuous component of \mathbf{h}_p . The expressions (5) also applies on $\Gamma_{t,p}$ for the tangential components $\mathbf{h}_{t,p}$, $\mathbf{h}_{c,t,p}$ and $\mathbf{h}_{d,t,p}$.

The TS model combined with the ICs and BCs of the impedance BC type [3] are presented through $\mathbf{h}_{c,t,p}$ and $\mathbf{h}_{d,t,p}$, that is

$$[\mathbf{n} \times \mathbf{e}_p]_{\Gamma_{t,p}} = \mu_p \beta_p \partial_t (2\mathbf{h}_{c,t,p} + \mathbf{h}_{d,t,p}), \quad (6)$$

$$\mathbf{n} \times \mathbf{e}_p|_{\Gamma_{t,p}^+} = \frac{1}{2} [\mu_p \beta_p \partial_t (2\mathbf{h}_{c,t,p} + \mathbf{h}_{d,t,p}) + \frac{1}{\sigma \beta} \mathbf{h}_{d,t,p}] - \mathbf{n} \times \mathbf{e}_p|_{\Gamma_{t,p}^-}, \quad (7)$$

$$\beta = \gamma_p^{-1} \tanh\left(\frac{d_p \gamma_p}{2}\right), \gamma_p = \frac{1+j}{\delta_p}, \delta_p = \sqrt{\frac{2}{\omega \sigma_p \mu_p}}, \quad (8)$$

where d_p is the local thickness of the TS, δ_p is the skin depth, $\omega = 2\pi f$ with f is the frequency, j is the imaginary unit and $\partial_t \equiv j\omega$. It should be noted that the term $-\mathbf{n} \times \mathbf{e}_p|_{\Gamma_{t,p}^-}$ in (7) is presented as a SS for the SPs.

As presented, the inaccuracy on the TS solution of a problem p (e.g. $p = u$) is then improved by a volume correction SP k (e.g., $p = k$) which scopes with the TS hypothesis [3]. The changes from a TS region (μ_u and σ_u for SP u) to a volume thin region (μ_k and σ_k for SP k), the field sources ($\mathbf{h}_{s,k}$ and $\mathbf{j}_{s,k}$) in (2a-b) are defined for the total fields [5, 8]

$$\mathbf{b}_{s,k} = (\mu_k - \mu_u) \mathbf{h}_u, \quad \mathbf{e}_{s,k} = (\sigma_k^{-1} - \sigma_u^{-1}) \mathbf{j}_u. \quad (9 a-b)$$

The equations in (9 a-b) are then presented by the updated relations $\mathbf{b}_u + \mathbf{b}_k = \mu_p (\mathbf{h}_u + \mathbf{h}_k)$ and $\mathbf{e}_u + \mathbf{e}_k = \sigma_p^{-1} (\mathbf{j}_u + \mathbf{j}_k)$. Thus, for any number of sources of SPs, one can be written as

$$\mathbf{b}_{s,k} = (\mu_k - \mu_u) \sum_{q \in P, q \neq p} \mathbf{h}_q, \quad (10)$$

$$\mathbf{e}_{s,k} = (\sigma_k^{-1} - \sigma_u^{-1}) \sum_{q \in P, q \neq p} \mathbf{j}_q, \quad (11)$$

where $q \in P$ (total SPs) except the current SP u relates to the sources ($\mathbf{b}_{s,k}$ and $\mathbf{e}_{s,k}$) via their last calculated corrections of \mathbf{h}_q and \mathbf{j}_q .

3 Finite element weak formulation

3.1 \mathbf{h}_p -magnetic field formulation with source and reaction fields

The weak magnetic field formulation (\mathbf{h}_p -conformal formulation) is obtained via the Faraday's law (1c) [5, 8, 9]. The magnetic field \mathbf{h}_p is decomposed into two parts, $\mathbf{h}_p = \mathbf{h}_{s,p} + \mathbf{h}_{r,p}$, where $\mathbf{h}_{s,p}$ is the source magnetic field presented via $\text{curl } \mathbf{h}_{s,p} = \mathbf{j}_{s,p}$, and $\mathbf{h}_{r,p}$ is the reaction magnetic field. In $\Omega_{c,p}^C$, the field $\mathbf{h}_{r,p}$ is expressed through a scalar potential ϕ_p , i.e. $\mathbf{h}_{r,p} = -\text{grad } \phi_p$. The weak forms for SPs p ($p = u, k, \dots$) are [10]

$$\begin{aligned} \partial_t (\mu_p (\mathbf{h}_{r,p} + \mathbf{h}_{s,p}), \mathbf{h}'_p)_{\Omega_p} + (\sigma_p^{-1} \text{curl } \mathbf{h}_p, \text{curl } \mathbf{h}'_p)_{\Omega_{c,p}} + (\mathbf{e}_{s,p}, \text{curl } \mathbf{h}'_p)_{\Omega_p} \\ + \partial_t (\mathbf{b}_{s,p}, \mathbf{h}'_p)_{\Omega_i} + \langle \mathbf{n} \times \mathbf{e}_p, \mathbf{h}'_p \rangle_{\Gamma_{e,p}} + \langle [\mathbf{n} \times \mathbf{e}_p]_{\Gamma_p}, \mathbf{h}'_p \rangle_{\Gamma_p} = 0, \\ \forall \mathbf{h}'_p \in H_p^1(\Omega_p) \end{aligned} \quad (12)$$

where $H_p^1(\Omega_p)$ including the basis functions for \mathbf{h}_p as well as for the test function \mathbf{h}'_i is a curl-conform function space presented in Ω_p . Notation of $(\cdot, \cdot)_{\Omega_i}$ is volume integral in Ω_p and $\langle \cdot, \cdot \rangle_{\Gamma_u}$ is the surface integral on Γ_p of the product of their vector field arguments. The term on surface integral $\Gamma_{e,p}$ gives as a natural BC of type (3 b), usually zero. It should be noted that the source field $\mathbf{h}_{s,p}$ in (12) is performed through a projection technique [11] of a known distribution $\mathbf{j}_{s,u}$, i.e.

$$(\text{curl } \mathbf{h}_{s,p}, \text{curl } \mathbf{h}'_{s,p})_{\Omega_p} = (\mathbf{j}_{s,p}, \text{curl } \mathbf{h}'_{s,p})_{\Omega_{s,p}}, \forall \mathbf{h}'_{s,p} \in H_p^1(\Omega_p). \quad (13)$$

3.2 Weak formulation for the TS model

The TS model [3] occurred in the general equation (12) is defined a free discontinuous $\mathbf{h}_{d,t,u}$ ($p = u$) along the TS and $\mathbf{h}_{d,u}$ in the exterior region related to Γ_u^+ and a BC $\mathbf{n} \times \mathbf{e}_u|_{\Gamma_u^+}$. This discontinuity is used as a test function \mathbf{h}'_u ($\mathbf{h}'_u = \mathbf{h}'_{c,u} + \mathbf{h}'_{d,u}$) in (12), where term contributions in the volume integrals in Ω_u are defined on the positive side of the TS. For that, the term $\langle [\mathbf{n} \times \mathbf{e}_u]_{\Gamma_u}, \mathbf{h}'_u \rangle_{\Gamma_u}$ is analysed as

$$\begin{aligned} \langle [\mathbf{n} \times \mathbf{e}_u]_{\Gamma_u}, \mathbf{h}'_u \rangle_{\Gamma_u} &= \langle [\mathbf{n} \times \mathbf{e}_u]_{\Gamma_u}, \mathbf{h}'_{c,u} \rangle_{\Gamma_u} + \langle \mathbf{n} \times \mathbf{e}_u|_{\Gamma_u^+}, \mathbf{h}'_{d,u} \rangle_{\Gamma_u^+} \\ &\quad + \langle \mathbf{n} \times \mathbf{e}_u|_{\Gamma_u^-}, \mathbf{h}'_{d,u} \rangle_{\Gamma_u^-}, \end{aligned} \quad (14)$$

where with $\mathbf{h}'_{d,u}$ is equal to zero on the negative side of TS Γ_u^- [11]. Moreover, the term $\langle [\mathbf{n} \times \mathbf{e}_u]_{\Gamma_u}, \mathbf{h}'_c \rangle_{\Gamma_u}$ and $\langle \mathbf{n} \times \mathbf{e}_u|_{\Gamma_u^+}, \mathbf{h}'_d \rangle_{\Gamma_u^+}$ are presented by (6) and (7), respectively. Therefore, (3.2) becomes

$$\begin{aligned} \langle [\mathbf{n} \times \mathbf{e}_u]_{\Gamma_u}, \mathbf{h}'_u \rangle_{\Gamma_u} &= \langle \mu_u \beta_u \partial_t (2\mathbf{h}_{c,t,u} + \mathbf{h}_{d,t,u}), \mathbf{h}'_{c,u} \rangle_{\Gamma_u} + \\ &\quad \langle \frac{1}{2} [\mu_u \beta_u \partial_t (2\mathbf{h}_{c,t,u} + \mathbf{h}_{d,t,u}) + \frac{1}{\sigma \beta} \mathbf{h}_{d,t,u}], \mathbf{h}'_{d,u} \rangle_{\Gamma_u^+}. \end{aligned} \quad (15)$$

By substituting the equation (15) into (12), the weak formulation for the TS model is obtained

$$\begin{aligned} \partial_t (\mu_u (\mathbf{h}_{r,u} + \mathbf{h}_{s,u}), \mathbf{h}'_u)_{\Omega_u} + (\sigma_u^{-1} \text{curl } \mathbf{h}_u, \text{curl } \mathbf{h}'_u)_{\Omega_{c,u}} \\ + \langle \frac{1}{2} [\mu_u \beta_u \partial_t (2\mathbf{h}_{c,t,u} + \mathbf{h}_{d,t,u}) + \frac{1}{\sigma \beta} \mathbf{h}_{d,t,u}], \mathbf{h}'_{d,u} \rangle_{\Gamma_u^+} \\ + \langle \mu_u \beta_u \partial_t (2\mathbf{h}_{c,t,u} + \mathbf{h}_{d,t,u}), \mathbf{h}'_{c,t,u} \rangle_{\Gamma_u} + \langle \mathbf{n} \times \mathbf{e}_u, \mathbf{h}'_u \rangle_{\Gamma_{e,u} - \Gamma_{t,u}} = 0, \\ \forall \mathbf{h}'_i \in H_i^1(\Omega_i) \end{aligned} \quad (16)$$

where the surface integral term on $\Gamma_{e,u} - \Gamma_{t,u}$ is usually considered as zero on a natural BC.

3.3 Projection of solutions between sub-models

In the strategy SP technique, the solutions found from a previous problem SP u are considered as sources (SSs and VSs) in a sub-domain $\Omega_{s,k} \subset \Omega_k$ of the SP k (current SP), e.g. from SP u to SP k . At the discrete level, the field \mathbf{h}_u in the mesh of the SP u is transferred to the mesh of the SP k (i.e. $\mathbf{h}_{u,k-proj}$). This will be performed via a projection technique [11] of its curl limited to $\Omega_{s,k}$. The field $\mathbf{h}_{u,k-proj}$ is defined

$$(\text{curl } \mathbf{h}_{u,k-proj}, \text{curl } \mathbf{h}')_{\Omega_{s,k}} = (\text{curl } \mathbf{h}_u, \text{curl } \mathbf{h}')_{\Omega_{s,k}}, \forall \mathbf{h}' \in F_k^1(\Omega_{s,k}) \quad (17)$$

where $H_k^1(\Omega_{s,k})$ is a curl-conform function space containing the projection of \mathbf{h}_u and the test function \mathbf{h}' in the mesh k . Directly projecting \mathbf{h}_u instead of its curl will give numerical errors as estimating its curl.

3.4 Improvement of the TS model

The inaccuracies of the TS solution obtained in SP u is now improved by the actual volume SP k via VSs defined in (10) and (11). They are taken into account the changes from μ_u and σ_u in the TS model to μ_k and σ_k in the volume shell. These VSs also appear in (12) via the volume integrals $(\partial_t (\mathbf{b}_{s,k}, \mathbf{h}'_k))_{\Omega_k}$ and $(\mathbf{e}_{s,k}, \text{curl } \mathbf{h}'_k)_{\Omega_k}$. The weak form for SP k is written as

$$\begin{aligned} \partial_t (\mu_k \mathbf{h}_k, \mathbf{h}'_k)_{\Omega_k} + (\sigma_k^{-1} \text{curl } \mathbf{h}_k, \text{curl } \mathbf{h}'_k)_{\Omega_{c,k}} + (-\mathbf{e}_u, \text{curl } \mathbf{h}'_k)_{\Omega_k} + \\ \partial_t ((\mu_k - \mu_u) \mathbf{h}_u, \mathbf{h}'_k)_{\Omega_{c,k}} + \langle [\mathbf{n} \times \mathbf{e}_k]_{\Gamma_{t,k}}, \mathbf{h}'_k \rangle_{\Gamma_{t,k}} + \langle \mathbf{n} \times \mathbf{e}_k, \mathbf{h}'_k \rangle_{\Gamma_{e,k} - \Gamma_{t,k}} \\ = 0, \forall \mathbf{h}'_k \in H_k^1(\Omega_k) \end{aligned} \quad (18)$$

where the field \mathbf{e}_u is also performed via an electric problem [9]. As a sequence, the field \mathbf{h}_u is transferred to the mesh of SP k via (17), where $\Omega_{s,u}$ is limited to a single layer surrounding of the volumic shell. In the same time to the VSs in (12), SSs related to ICs [3], [5] have to extract the TS discontinuities, i.e. $\mathbf{h}_{d,k} = -\mathbf{h}_{d,u}$ and $[\mathbf{n} \times \mathbf{e}_k]_{\Gamma_{t,k}} = -[\mathbf{n} \times \mathbf{e}_u]_{\Gamma_{t,u}}$. The involved trace $[\mathbf{n} \times \mathbf{e}_u]_{\Gamma_{t,k}}$ is generally defined through other integrals in (12), with $\Gamma_{t,u} = \Gamma_{t,k}$, i.e.

$$\begin{aligned} \langle [\mathbf{n} \times \mathbf{e}_k]_{\Gamma_{t,k}}, \mathbf{h}'_k \rangle_{\Gamma_{t,k}} &= \langle -[\mathbf{n} \times \mathbf{e}_u]_{\Gamma_{t,k}}, \mathbf{h}'_k \rangle_{\Gamma_{t,k}} = (\mu_u \partial_t \mathbf{h}_{s,u}, \mathbf{h}'_k)_{\Omega_u = \Omega_k} \\ &\quad + (\mu_u \partial_t \mathbf{h}_u, \mathbf{h}'_k)_{\Omega_u = \Omega_k}. \end{aligned} \quad (19)$$

The surface integral in (19) is implemented in the step u . At the discrete level, the limitation of the two volume integrals in (19) is defined in a single layer of FEs on both sides of $\Gamma_{t,u}$ touching $\gamma_k^+ = \gamma_u^+ = \Gamma_t^+$, because it involves only the trace $\mathbf{n} \times \mathbf{h}'_k|_{\gamma_k^+}$.

4 Numerical test

The practical test is a TEAM workshop problem (problem 21, model B) [12] including two coils and a magnetic shielding plate (Fig. 3). The test problem is considered in both 2-D and 3-D models.

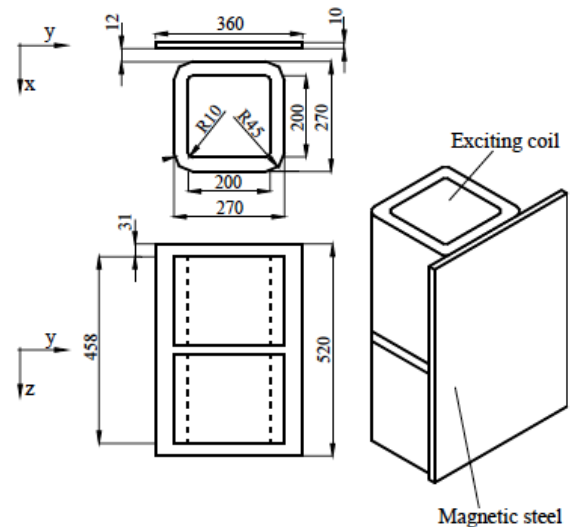


Figure 3: Geometry of the test problem 21 (model B).

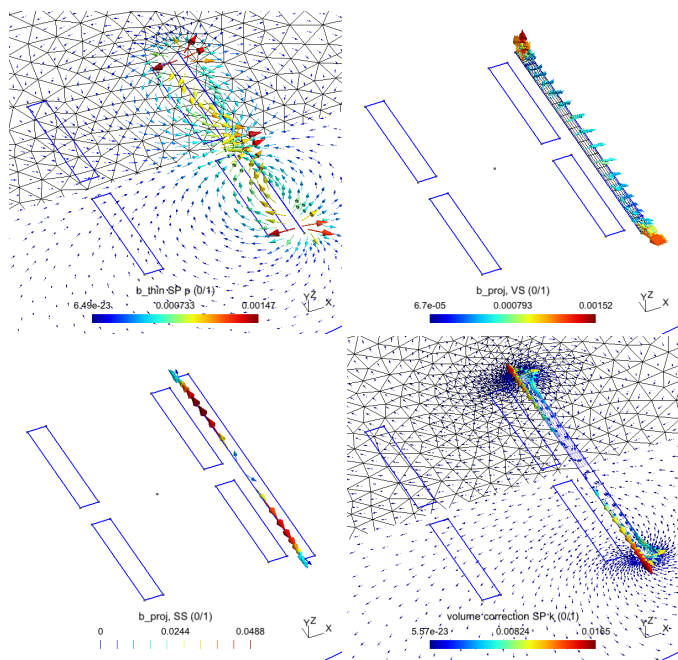


Figure 4: Magnetic flux densities for the TS solution SP u (b_u , top left) and volume correction solution SP k (b_k , bottom right). Projections of SP u solutions ($b_{proj, VS}$, top right) and ($b_{proj, SS}$, bottom left) in the SP k .

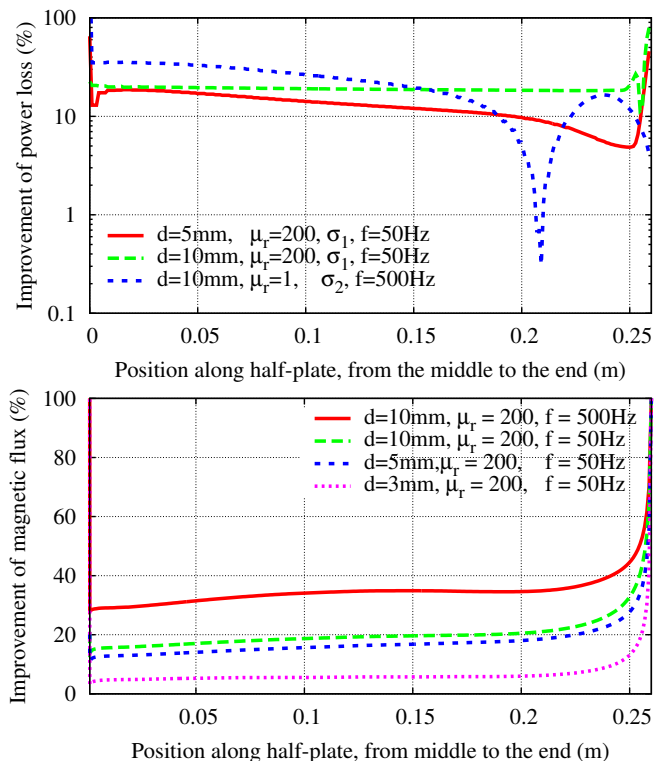


Figure 5: Relative improvements of the joule power loss density (top) and the magnetic flux (bottom) along the thin plate, with $\sigma_{plate} = 6.484$ MS/m.

The first case is tested with a 2-D model: The SP u with the inductors and TS FE magnetodynamic model is solved on a lighter mesh (Fig. 4 (b_u), top left). A volume correction SP k then replaces the TS FEs with the classical volume FEs containing an actual re-

gion of the thin plate and its surrounding, with a suitable refined mesh without including the inductors anymore (Fig. 4 (b_k), bottom right). The projections of TS solution in TS SP u for the VS and SS are respectively pointed out (Fig. 4 ($b_{proj, VS}$), top right) and (Fig. 4 ($b_{proj, SS}$), bottom left). The relative improvements of the joule power loss density and the magnetic flux along the plate are depicted in Fig. 5. They can touch several tens of percents near the end of the TS, such as 80% (Fig. 5, top) and 85% (Fig. 5, bottom), with $d = 10$ mm and $\delta = 1.977$ mm for both cases. For $d = 10$ mm, the inaccuracy is lower than 50% (Fig. 5, top) and 55% (Fig. 5, bottom).

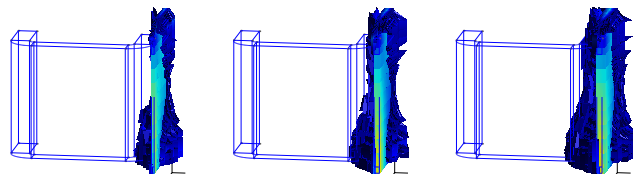


Figure 6: Colored maps of eddy current densities indicating the regions with relative volume improvements greater than 2% for $d = 5$ mm, 7.5 mm and 10 mm (from left to right), $f = 50$ Hz, $\mu_{plate} = 200$ and $\sigma_{plate} = 6.484$ MS/m.

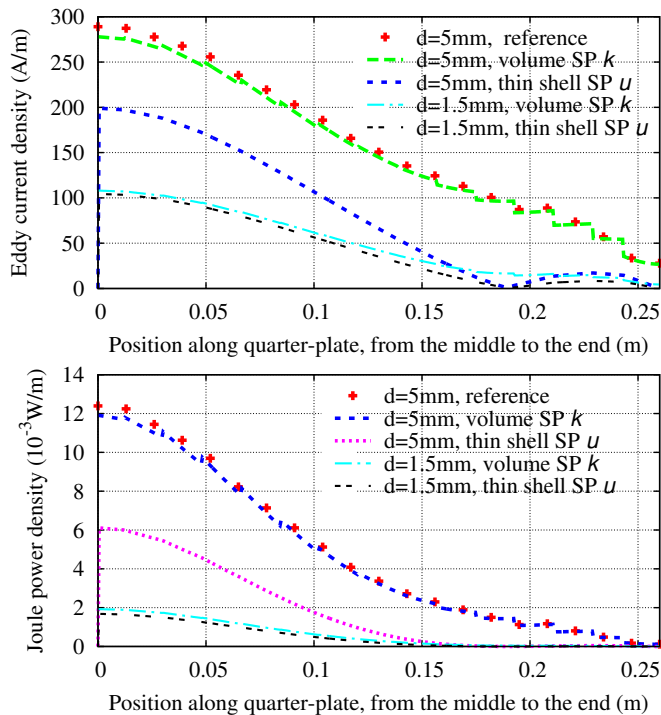


Figure 7: Eddy current density (top) and joule power loss density (bottom) for the TS and the volume correction along the vertical half edge (z -direction) ($f = 50$ Hz, $\mu_{plate} = 200$ and $\sigma_{plate} = 6.484$ MS/m).

The next test is considered with a 3-D model: The errors of the TS model SP u are shown in Fig. 6 (from left to right). They increase with plate thickness, specially near the ends of the plate, are perfectly improved whatever their order of magnitude. The accuracy of the improvement is directly related to the volume correction of the plate and its surrounding.

The inaccuracies on the eddy current density and joule power loss density of TS model (SP u) are improved by the significant volume correction (SP k) (Fig. 7). The inaccuracies on the local fields (eddy current density and joule power loss density) along the vertical half edge (z -direction) near the plate ends reaches 60% (Fig. 7, *top*) and 80% (Fig. 7, *bottom*), respectively, with $\delta = 2.1\text{mm}$ and $d = 5\text{mm}$. The volume corrections are then compared to be similar the reference solutions (obtained from the FEM) for different parameters in both cases.

5 Conclusions

In this study, a perturbation Finite Element Approach has been proposed with a magnetic field formulation for a one-way coupling in order to improve the errors around edges and corners inherent to the TS FE assumptions. The correctly improving on the local fields of magnetic flux densities, eddy current losses and joule power losses is successfully achieved surrounding the borders and corners of the TS structures. The proposed method has been developed for the one-way coupling in two steps. All the processes of the method have been successfully validated on the international practical problem (TEAM workshop problem 21, model B) [12].

The source-codes of the this technique have been being extended based on the source-codes of the SPM that was developing by author and two Prof. Patrick Dular and Prof. Christophe Geuzaine at the Dept of Electrical Engineering and Computer Science, University of Liege, Belgium. They will be then ran and simulated in the background of the Getdp (<http://getdp.info>) and Gmsh (<http://gmsh.info>) softwares developed by Prof. Patrick Dular and Prof. Christophe Geuzaine [13, 14]. These are the open-source codes for anyone to be able to write adapt source-codes for solving studied problems (if possible).

References

- [1] L. Krähenbühl and D. Muller, "Thin layers in electrical engineering. Examples of shell models in analyzing eddy-currents by boundary and finite element methods," IEEE Trans. Magn., vol. 29, no. 2, pp. 1450–1455, 1993.

- [2] Tsuboi, H., Asahara, T., Kobayashi, F. and Misaki, T. (1997), "Eddy current analysis on thin conducting plate by an integral equation method using edge elements," IEEE Trans. Magn., Vol. 33, No. 2, pp. 1346-9.
- [3] C. Geuzaine, P. Dular, and W. Legros, "Dual formulations for the modeling of thin electromagnetic shells using edge elements," IEEE Trans. Magn., vol. 36, no. 4, pp. 799–802, 2000.
- [4] P. Dular, Vuong Q. Dang, R. V. Sabariego, L. Krähenbühl and C. Geuzaine, "Correction of Thin Shell Finite Element Magnetic Models via a Subproblem Method," IEEE Trans. Magn., vol. 47, no. 5, pp. 1158–1161, 2011.
- [5] Vuong Q. Dang, P. Dular, R. V. Sabariego, L. Krähenbühl and C. Geuzaine, "Subproblem Approach for Thin Shell Dual Finite Element Formulations," IEEE Trans. Magn., vol. 48, no. 2, pp. 407–410, 2012.
- [6] S. Koruglu, P. Sergeant, R.V. Sabariego, Vuong. Q. Dang and M. De Wulf "Influence of contact resistance on shielding efficiency of shielding gutters for high-voltage cables," IET Electric Power Applications., Vol. 5, No. 9, pp.715-720, 2011.
- [7] Vuong Quoc Dang and Quang Nguyen Duc, "Coupling of Local and Global Quantities by A Subproblem Finite Element Method Application to Thin Region Models," ASTESJ, vol. 4, no. 2, pp. 40–44, 2019.
- [8] Vuong Quoc Dang and Quang Christophe Geuzaine, "Using edge elements for modeling of 3-D Magnetodynamic Problem via a Subproblem Method," Sci. Tech. Dev. J.; 23 (1), pp. 439–445, 2020.
- [9] P. Dular and R. V. Sabariego, "A perturbation method for computing field distortions due to conductive regions with h-conform magnetodynamic finite element formulations," IEEE Trans. Magn., vol. 43, no. 4, pp. 1293-1296, 2007.
- [10] Vuong Q. Dang, P. Dular, R. V. Sabariego, L. Krähenbühl, C. Geuzaine "Subproblem Approach for Modelling Multiply Connected Thin Regions with an h-Conformal Magnetodynamic Finite Element Formulation," in EPJ AP, vol. 63, no. 1, 2013.
- [11] C. Geuzaine, B. Meys, F. Henrotte, P. Dular and W. Legros, "A Galerkin projection method for mixed finite elements," IEEE Trans. Magn., Vol. 35, No. 3, pp. 1438-1441, 1999.
- [12] Zhiguang CHENG¹, Norio TAKAHASHI², and Behzad FORGHANI, "TEAM Problem 21 Family (V.2009)," Approved by the International Compumag Society Board at Compumag-2009, Florianopolis, Brazil.
- [13] C. Geuzaine and J.-F. Remacle "Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities," International Journal for Numerical Methods in Engineering 79(11), pp. 1309-1331, 2009.
- [14] P. Dular and C. Geuzaine "GetDP reference manual: the documentation for GetDP, a general environment for the treatment of discrete problems," University of Liege, 2013.

Digestibility, Digestible and Metabolizable Energy of Earthworm Meal (*Eisenia Foetida*) Included in Two Levels in Guinea Pigs (*Cavia Porcellus*)

Jorge Castro-Bedriñana^{*1}, Doris Chirinos-Peinado¹, Hanz Sosa-Blas²

¹Faculty of Zootechnics, Universidad Nacional del Centro del Perú. Huancayo, 12000, Perú

²FONCODES - Technical advisor - Chanchamayo, 12840, Perú

ARTICLE INFO

Article history:

Received: 05 April, 2020

Accepted: 05 May, 2020

Online: 15 May, 2020

Keywords:

Food nutritional value

Total digestible nutrients

Voluntary food intake

Gross energy

Digestibility assays

Digestibility coefficients

Animal protein production

Proximal chemical analysis

Metabolic cages

ABSTRACT

The increasing cost of fishmeal and soybean meal forces us to look for unconventional sources of protein to feed guinea pigs, being able to use earthworm meal *Eisenia foetida* (EF) with 60-80% high quality raw protein; for this, the contribution of digestible nutrients and metabolizable energy must be known to formulate rations. The objective is to evaluate the nutritional quality of EF used in 10 and 20% in guinea pig diets. The research procedure considered the preparation of EF, proximal chemical analysis, digestibility tests "in vivo" by the direct method for the reference diet of barley meal (BM) and indirect tests for EF, using 5-month-old male guinea pigs and homogeneous weights (700-750 g), placed in individual metabolic cages and randomly distributed in 3 groups of 3 animals per group, (G1): Reference diet (BM), (G2): EF 10% + 90% BM and (G3): EF 20% + 80% BM. To determine the mean difference of digestibility, total digestible nutrients (TDN), digestible energy (DE) and metabolizable energy (ME) between G2 and G3. The average content of dry matter, crude protein, fat, nitrogen-free extract and organic matter of the EF was 77.16, 66.90, 10.0, 21.1 and 91.0%, and the average digestibility coefficients of these components were 68.01, 92.96, 72.40, 41.34 and 71.68%; the ME content was 3125.31 Kcal / kg. As the EF level increased from 10 to 20%, the digestibility coefficients of dry matter, protein, fat, nitrogen-free extract and organic matter increased by 7.75; 2.18; 5.45; 16.20 and 4.83%, the ME value increased by 7.25% ($P < 0.05$). Increasing the inclusion of EF from 10 to 20% in guinea pig diets improves digestibility nutrients and ME content. Reduction in the cost of animal protein production, added value for the cultivation of *Eisenia foetida* and contribution to environmental health.

1. Introduction

The guinea pig, originally from the Andes Mountains of Peru, Colombia, Ecuador and Bolivia, due to its nutritional quality and healthy properties, is important in the nutritional food security of the Andean population [1,2].

The high protein content (21.4%), B vitamins (15 mg / 100 g), low cholesterol (65 mg / 100 g), low saturated fat (3%), low sodium, linoleic and linolenic acid presence (absent or low concentration in other meats) qualify guinea pig meat as a healthy food for any population group [3,4].

According to estimates of the Cámara Peruana del Cuy [5], about 18 million guinea pigs are raised in Peru, 50% more compared to the IV National Agricultural Census of 2012, and national and international demand requires guinea pigs of a standard size and quality. Exports went from two tons in 2002 to twenty in 2015; United States is the main destination.

This productive level demands rapid growth, better rates of food conversion and higher yield of guinea pig meat, using diets of high nutritional value; but, if conventional protein foods (fish meal or soybean meal) are included, production costs are raised, and unconventional protein sources such as California redworm meal *Eisenia foetida* should be evaluated [6].

The earthworm meal *Eisenia foetida* (EF) contains 18.6% dry matter and depending on the processing, a dry basis it contains 60-

*Corresponding Author. Jorge Castro-Bedriñana, Av. Mariscal Castilla N° 3909 El Tambo, Huancayo. Perú. +51 964408057 & : jorgecastro@yaho.com

80% low-cost crude protein [7]-[9], 1.3% crude fiber (CF); 7-10% fat, 8-20% nitrogen free extract, 3-6% ash, and 4000 kcal / kg gross energy [10]. Essential amino acids outperforms to the fishmeal and soybean meal, and contains lysine (2.7-6.8%), methionine (0.76-1.2%), phenylalanin (1.8-3.5%), leucine (3.1-5.0%), tryptophane (0.12-1.73%), threonine (1.8-5.2%), histidine (1.4-2.6%), arginine (2.8-4.4%), valine (1.3-4.7%), isoleucine (1.2-6.2%) among others [11-15]. EF also contains niacin, riboflavin, thiamine, pantothenic acid, pyridoxin, cyanocobalamin, folic acid, being considered as a biotechnological resource of high nutritional and ecological interest of some developing countries [9,15], and by its pleasant smell and good palatability [16] can be included in animal and human diets [8,9].

The EF can use manure of pigs, guinea pigs, rabbits, poultry, sheep and cattle as substrates, combined with organic and agricultural waste, which are commonly thrown away and are a source of contamination [7,11,17,18]. The EF adapts to a wide range of soil and climate conditions, and reproduces better when the substrate temperature is between 14 and 27 °C, being the optimum 21 °C, but can survive between 0-42 °C [9].

In the vermicompost, different organic materials are transformed into useful products for humans, such as organic fertilizers (with multiple benefits for agriculture) [19], earthworm meal, earthworm tea, cookies, cakes, among others [20], obtaining 100% natural and organic products. The EF is a powerful natural anabolic and because of the rapid absorption of amino acids, the immune system does not recognize them as foreign elements, as with other proteins of animal origin that are rejected by the immune system [14].

The research objective is to evaluate the nutritional quality of *Eisenia foetida* worm meal (Figure 1) for guinea pigs, included in 10 and 20% to diets based on ground barley, through proximal chemical analysis, digestibility tests and digestible and metabolizable energy estimation.

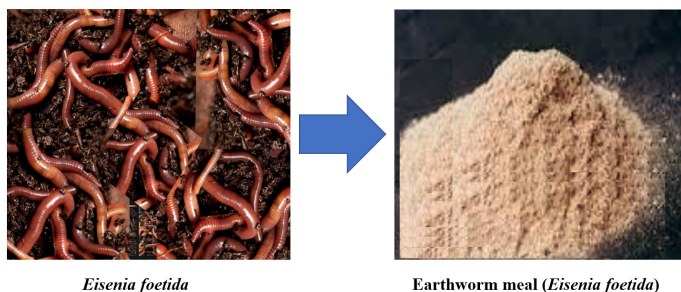


Figure 1. *Eisenia foetida* and earthworm meal

2. Materials and methods

2.1. Site Study

The research was conducted in the Digestibility Room of the “Yauris” Agricultural Farm of the Universidad Nacional del Centro del Perú (UNCP), Huancayo-Junín, altitude 3253, at 12 ° 03 '14' 'SL and 75 ° 12' 55 ' 'WL. The proximal analyzes were carried out in the Animal Nutrition Laboratory, Faculty of Zootechnics-UNCP.

The digestibility room was well ventilated and illuminated. All procedures related to the handling and treatment of animals

followed the ethical standards of animal welfare in research, and the care and use of laboratory animals. After the study, the guinea pigs returned to their breeding system.

2.2. Experimental samples

For digestibility tests, nine 5-month-old male guinea pigs of the Wanka breed were used, of similar weights (700 g), in good health and distributed in three groups of three guinea pigs each, randomly arranged in individual metabolic cages.

The number of animals per group responds to international recommendations for the care and use of research animals, which suggest minimizing the number of animals used [21, 22].

The process followed in the investigation is summarized in seven phases (Figure. 2).

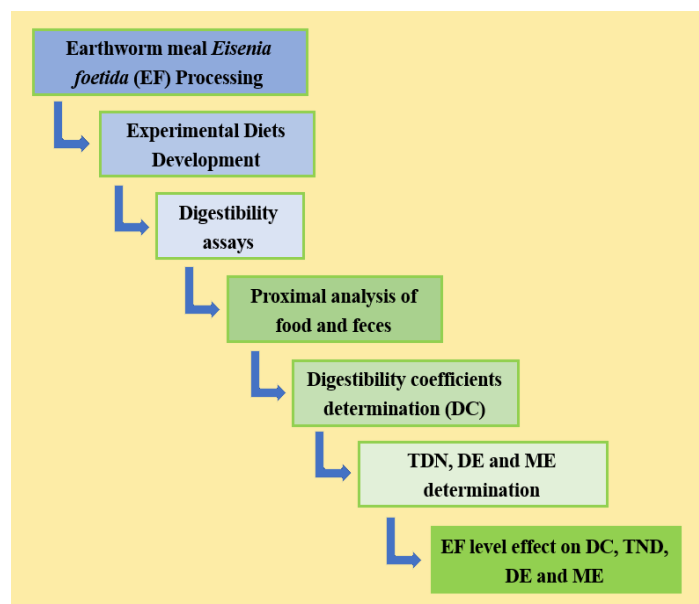


Figure 2. Sequence of the investigation procedure

2.3. Earthworm meal *Eisenia foetida* (EF) manufacturing process

Eisenia foetida is a product of the vermicomposting; process involving mutual action of earthworm and microorganisms to transform biodegradable organic matter to humus-like vermicast and vermiliquid [23]; thus, the symbiotic activity of the microorganisms contributes to improving the quality of the products and increases their nutritional value [24].

Earthworms were fed a diet of organic waste compost twice a week specifically forage residues and guinea pig manure, that are high in organic matter and readily available, as suggested by [11,17,18]; manure is the most preferred by earthworms, and providing organic matter, it stimulates biodegradation and raises pH within the culture substrate.

The feed was spread on top of compost and water was sprinkle on it and then feed was thoroughly mixed with compost. In order to guarantee optimum growth conditions, optimum temperature 12-24 °C, moisture 80-90% and <5 pH <9 of the compost was kept under control [25].

20 kg of earthworms *Eisenia foetida* were used to make 2.6 kg of meal (13% yield). The worms were placed in a bowl for 3 h to evacuate the intestinal contents [26], washed with clean water and placed in a hot 4% salt solution for 10 minutes [27], rinsed and dried in an oven at 90°C. ° C for 6 h [28]; then, they were cooled, ground and sieved on a 60 mesh, obtaining a light brown flour. The EF was packaged in clean, dry plastic bags with an airtight seal to prevent excess moisture and attack by fungi or other pathogens that rapidly contaminate and degrade [9].

2.4. Study diets and pre-experimental stage of the digestibility assays

At the beginning of the pre-experimental phase all animals received 300g / day of *Lolium multiflorum*, gradually replaced by the experimental diets and drinking water + vitamin C “ad libitum”, in nipple drinking fountains (Figure 3), until the fifteenth day they exclusively consumed the study diets.

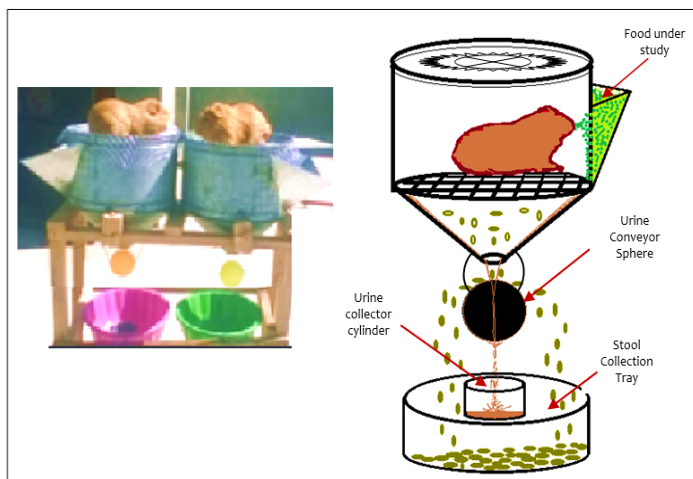


Figure 3. Metabolic cage for guinea pigs (Granja Agropecuaria de Yauris - UNCP). Fuente: [29]

Digestibility tests were performed for three diets:

- (D1): Reference diet (100% barley meal)
- (D2): 10% FE + 90% reference diet
- (D3): 20% EF + 80% reference diet

For D1, digestibility was determined by the direct "in vivo" method, assuming that the indigestibility of barley flour (BM) is the same when combined with EF. And from their indigestion coefficients, subsequent calculations are performed to determine the digestibility of EF by the indirect method [29,30].

2.5. Experimental stage and chemical analysis

The experimental phase considered the exact measurement of consumption and individual production of guinea pig feces, assuming that feces collected on a given day correspond to the indigestible food consumed the previous day. This stage had a period of seven days.

The study diets and their corresponding feces were taken to the stove to determine the dry matter content, then were ground, homogenized and taken to the laboratory.

To determine the chemical composition of BM, EF, and Guinea pig feces, the proximal analysis was used, following the AOAC protocols [31]. The proximal analysis considered crude protein (N x 6.25), ether extract, crude fiber, ash and nitrogen free extract.

The TDN content, which describes the energy available in food [32] was calculated from the digestibility tests [33,34].

The values of DE and ME were estimated with validated equations. DE content was estimated from the TDN concentration, it was considered that one kg of TDN is equivalent to 4,400 Kcal of DE [35-37]. The ME content was corresponded to the average of two estimates, the first where, $ME = DE * 0.82$ [36-40], and the second, where, $1 \text{ kg TDN} = 3,560 \text{ kcal / kg ME}$ [41].

2.6. Statistical analysis

The level of statistical significance was set at $P < 0.05$, and all analyses were two-sided. The digestibility coefficients, TDN, DE, and ME of EF used in 10 and 20%, were compared by means of a “t” test for the difference of independent means, using SPSS 23. Intake data were expressed as quantity of diet consumed (in grams) per day per animal, and as percentage of live weight [42].

3. Results

3.1. Chemical composition of barley meal (BM) and earthworm meal (EF)

The proximal chemical composition of BM and EF used in the digestibility tests (Table 1) is within the range found in other studies [43]. The analyzes of each food sample were in triplicate.

Table 1. Proximal chemical composition of barley meal and earthworm meal

Food	M	DM	CP	EE	CF	NFE	A	OM
Barley meal (BM)	11.38	88.63	11.81	3.50	6.10	76.6	2.0	98.0
Earthworm meal (EF)	22.84	77.16	66.90	10.00	0.00	14.1	9.0	91.0

Averages of all the components of the BM and EF were highly significant ($P < 0.01$).

M: Moisture, DM: Dry matter, CP: Crude protein, EE: Ether extract, CF: Crude fiber, NFE: Nitrogen free extract, A: Ash, OM: Organic matter.
Unit: Percentage

BM has been used as a reference diet in digestibility tests by the indirect method, and generally constitutes 90% of the diet with 10% of the ingredient under study [29].

The crude protein and fat content of the EF found in this study is similar to that reported by [43] 66.2%.

The crude protein and fat content of EF was similar to the reports by [44]. Other studies reported lipid contents on dry

matter of EF were 7.34% [12], 11.3% [27], 6.6% [45] and 18% [46].

3.2. Digestibility of barley meal (BM)

The digestibility coefficients of the BM components (Table 2), were used to determine the digestibility of EF by the indirect method [29].

Table 2. Digestibility coefficients and total digestible nutrients of barley meal

Digestibility coefficient (%)	N	Min	Max	Mean	SD	Variance (%)
Dry matter	3	81.9	85.53	83.75	1.82	3.30
Crude protein	3	58.56	67.28	62.56	4.40	19.40
Ether extract	3	63.8	79.33	71.81	7.78	60.47
Crude fiber	3	75.37	78.65	77.26	1.70	2.87
Nitrogen free extract	3	86.15	89.17	87.69	1.51	2.28
Organic matter	3	83.2	86.41	84.92	1.62	2.62
Total Digestible Nutrients	3	83.06	87.29	84.92	2.16	4.66

The digestibility of crude protein from BM determined in this study is similar to that reported in other studies in which barley flour was used as a reference diet in digestibility trials in guinea pigs [29]; also, was similar to that observed in growing pigs [47].

3.3. Digestibility, TDN, DE and ME of earthworm meal with 10 and 20% inclusion

The digestibility coefficients of the EF were higher when entering the diet in 20% than with 10% (Tables 3 and 4).

Table 3. Digestibility coefficients and total digestible nutrient of the earthworm meal included in 10%

Digestibility coefficient (%)	N	Min	Max	Mean	SD	Variance (%)
Dry matter	3	59.78	70.45	64.14	5.60	31.33
Crude protein	3	89.90	93.57	91.87	1.85	3.42
Ether extract	3	63.72	77.07	69.67	6.79	46.13
Nitrogen free extract	3	23.51	45.14	33.24	10.98	120.51
Organic matter	3	61.97	74.69	69.26	6.56	43.04
Total digestible nutrients	3	82.91	85.50	84.15	1.30	1.69

Table 4. Digestibility coefficients and total digestible nutrients of earthworm meal included in 20%

Digestibility coefficient (%)	N	Min	Max	Mean	SD	Variance (%)
Dry matter	3	71.23	72.29	71.88	0.57	0.33
Crude protein	3	91.87	97.77	94.05	3.24	10.48
Ether extract	3	68.43	83.27	75.12	7.53	56.64
Nitrogen free extract	3	39.92	54.80	49.43	8.26	68.25
Organic matter	3	72.32	75.46	74.09	1.61	2.59
Total digestible nutrients	3	88.59	91.76	90.25	1.59	2.53

In the practical feeding of guinea pigs, depending of EF percentage in the diets, the digestibility coefficients determined in this study could be used (Figure 4).

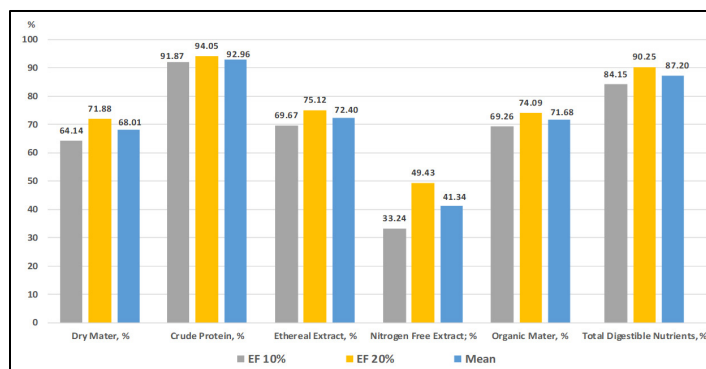


Figure 4. Digestibility coefficients and total digestible nutrient content of earthworm meal included in two levels

The results of the "t" tests for the difference of means of samples independent of the digestibility coefficients, TDN, DE and ME of EF included in 10 and 20%, show that the inclusion of 20% of EF significantly improves the availability of EF energy (Tables 5 and 6).

Table 5. "T" tests for the digestibility coefficients and TDN content of the earthworm meal included in 10 and 20%

Components	T test	p-value	Sig
Dry mater	-2.397	0.075	ns
Crude protein	-1.015	0.367	ns
Ether extract	-0.929	0.405	ns
Nitrogen free extract	-2.036	0.111	ns
Organic material	-1.247	0.281	ns
Total digestible nutrients	-4.961	0.008	*

* Statistical difference (P < 0.05).
ns: No statistical difference (P > 0.05)

When EF inclusion increased from 10% to 20%, EF's DE and ME contribution increased by 7.25% (Figure 5).

Table 6. "T" test for the digestible and metabolizable energy of earthworm meal included in 10% and 20%

Energy value	N	Mean	SD
Digestible energy, HL 10%	3	3702.60b	57.13
Digestible energy, HL 20%	3	3971.15a	69.99
Metabolizable energy, 10% HL	3	3015.93b	46.54
Metabolizable energy, HL 20%	3	3234.68a	57.01

a,b, Average values of digestible and metabolizable energy with different letters vary statistically (P < 0.05)

3.4. Earthworm meal Eisenia foetida consumption by guinea pigs

When EF 10% was used, the daily fresh consumption per animal per day was 1.78 g, equivalent to 1.37 g of dry matter and

0.21% of live weight; while when using 20%, fresh consumption per animal per day was 4.36 g, equivalent to 3.37 g of dry matter and 0.47% of live weight (Table 7).

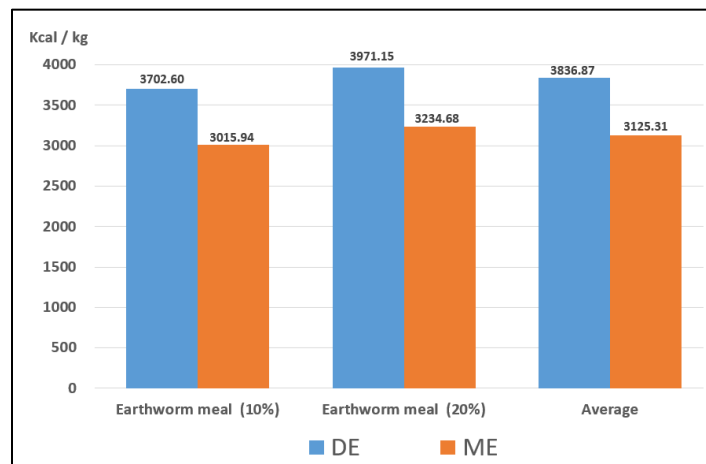


Figure 5. Digestible and metabolizable energy content of worm meal, included in 10 and 20%

Table 7. Average consumption of earthworm meal per animal per day

	Evaluated diets	Mean	SD
Average consumption / animal / day, fresh base, g	EF (10%)	1,78b	0,020
	EF (20%)	4,36a	0,313
Average consumption / animal / day, dry basis, g	EF (10%)	1,37b	0,015
	EF (20%)	3,37a	0,241
Average consumption / animal / day, live weight %	EF (10%)	0,21b	0,021
	EF (20%)	0,47a	0,015

Average consumption in fresh, dry basis and as live weight percentage by inclusion levels, with different letters vary statistically ($P < 0.05$).
EF: Earthworm meal (*Eisenia foetida*)

Average consumption in fresh, dry basis and as live weight percentage by inclusion levels, with different letters vary statistically ($P < 0.05$).

4. Discussion

Although the statistical analysis does not show significant differences ($P > 0.05$) between the digestibility coefficients of dry matter, crude protein, ether extract, nitrogen free extract, and organic matter, when 10 and 20% of EF was used; with 20% of EF the digestibility coefficients were 7.75, 2.18, 5.45, 16.20 and 4.83% higher than with 10%.

Inclusion 20% of EF allowed a higher content of TDN than when included in 10% ($P < 0.05$). The TDN content was improved by 5.6% (82.91 to 88.59); Similarly, digestible and metabolizable energy were higher when 20% of EF was used ($P < 0.05$).

The improvement in digestibility and energy content of EF is due to the greater contribution of proteins and essential amino acids [11,18] by using 20% EF, which increases the consumption, digestion and absorption of nutrients; similar result was observed when using 15% EF [44], which reports digestibility coefficients

of dry matter, protein, fat, nitrogen free extract and organic matter of 65.23, 89.97, 68.36, 37.45 and 70.08%.

Digestibility coefficients are not constant for a given food and are influenced by many factors, including chemical composition, protein and amino acids levels. When dietary protein increases, the digestibility of the whole diet increases [29].

Digestibility of a food mixture is not necessarily the average of the values of its components determined separately; each food can influence the digestibility of others [29, 30].

When the inclusion of EF raised from 10% to 20%, the DE and ME contribution of EF increased by 7.25% (3702.60 to 3971.15 and 3015.94 to 3234 kcal/kg); In this regard [9] reports more than 95% digestibility of Californian red worm flour, which allows high nutrient absorption and high energy contribution.

When the inclusion of EF raised from 10% to 20%, the consumption of EF increased by 2.45 times, which could have been due to the greater contribution of proteins and other nutrients; when the chemical composition of the diet is improved, it also improves digestibility and consumption [29, 39].

Another study on the digestibility of EF-71.2% CP, in weaned guinea pigs, included in 15%, reports digestibility coefficients for MS, PT, EE, ELN, MO of 65.23, 89.97, 68.36, 37.45 and 70.08%, and when was included in 20%, the coefficients were 69.87, 92.09, 72.55, 43.89 and 73.01%, the higher level of inclusion improves the absorption of nutrients [44].

A recent study indicates that the inclusion of the EF in the proportions (85:15) in animal diets increased ($P < 0.05$) the CP, EE, and A, mainly when mixed with rice powder, corn meal and soy cake meal [48].

Our results indicate the EF is among the non-conventional protein sources, with promising results, thanks to its high protein levels, proper amino acid profile, high reproduction rate, low mortalities, fast growth and ease of production [6].

Studies have shown that EF has recommendable levels of protein, essential amino acids and lipids, which are similar to those found in fishmeal and, are in line with the nutritional requirements of many species [11,18], and for its high biological value, is a possible solution to nutritional problems that humanity has. The richness and quality of amino acids and vitamin contribution of EF, not only satisfies the requirements of monogastric animals, but also of children between 2-5 years recommended by FAO / WHO; being a biotechnological resource of high nutritional and ecological interest [9].

5. Conclusion

Earthworm meal (*Eisenia foetida*) has a high nutritional quality for guinea pigs, significantly increasing its digestibility, energy value and consumption when included in the reference diet by 20% compared to 10%.

The digestible and metabolizable energy content of earthworm meal included by 20% increased by 6.76% compared to 10%.

Due to the high content of digestible nutrients and metabolizable energy, earthworm flour could be included in 20% of guinea pig diets.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

To the Guinea Pig Program of the Yauris Agricultural Farm, Universidad Nacional del Centro del Perú, for the facilities provided in the breeding and processing of earthworm meal and digestibility tests. This research did not receive any specific funding.

References

- [1] Avilés DF, Martínez AM, Landi V, Delgado JV. El cuy (*Cavia porcellus*): un recurso andino de interés agroalimentario. *Animal Genetic Resources*, 55, 87–91, 2014. <https://doi.org/10.1017/s2078633614000368>.
- [2] Chirinos D. Seguridad alimentaria nutricional en poblaciones vulnerables de la región central del país. Primera edición. Gráfica JOSIMPRESORES SAC. Huancayo. Perú. 2016.
- [3] Zumárraga S. Innovaciones gastronómicas del cuy en la provincia de Imbabura. Universidad Técnica del Norte (UTN), Ecuador. 2011. <http://repositorio.utn.edu.ec/bitstream/123456789/1139/2/06%20GAS%20014%20TESIS.pdf>
- [4] Crespo N. La Carne de Cuy: Nuevas propuestas para su uso. Tesis Gastronomía y Servicio de Alimentos y Bebidas. Universidad de Cuenca. Ecuador. 2012. <https://dspace.ucuenca.edu.ec/bitstream/123456789/1563/1/tgas26.pdf>.
- [5] AAN. Producción de cuy en Perú creció 50% en los últimos cinco años. Agencia agraria. Grupo Camposur. 2016. Available: <https://agraria.pe/noticias/produccion-de-cuy-en-peru-crecio-50-12352>.
- [6] Musyoka SN, Liti DM, Ogello E, Waidbacher H. Utilization of the earthworm, *Eisenia foetida* (Savigny, 1826) as an alternative protein source in fish feeds processing: A review. *Aquaculture Research*. 2019. <https://doi.org/10.1111/arc.14091>.
- [7] Bravo CM, Angulo LM, González YA, Martínez MM, Carmona JC, Garay OV. Evaluación reproductiva de la lombriz roja californiana (*Eisenia foetida*) alimentada con diferentes sustratos en el trópico bajo colombiano. *Livestock Research for Rural Development*, 30(2), 2018. <https://www.lrrd.cipav.org.co/lrrd30/2/over30036.html>
- [8] Alcívar-Cedeño U, Dueñas-Rivadeneira A, Sacon-Vera E, Bravo-Sánchez L, Villanueva-Ramos G. Influencia de los tipos de secado para la obtención de harina de Lombriz Roja californiana (*Eisenia foetida*) a escala piloto. *Tecnología Química*, 36(2), 187-196, 2016. <https://www.redalyc.org/comocitar.oa?id=445546335007>
- [9] Pires MB. Harina de lombriz: una alternativa saludable para nuestra alimentación. Universidad FASTA, Facultad de Ciencias Médicas. Licenciatura en Nutrición. 2013. <http://redi.ufasta.edu.ar:8080/xmlui/handle/123456789/299>.
- [10] Durán L, Henríquez C. Crecimiento y reproducción de la lombriz roja (*Eisenia foetida*) en cinco sustratos orgánicos. *Agronomía Costarricense*, 33(2), 275-281, 2009. <https://www.redalyc.org/html/436/43613279011/>
- [11] Vodounnou DS, Kpogué DN, Apollinaire MG, Didier FE. Culture of earthworm (*Eisenia foetida*), production, nutritive value and utilization of its meal in diet for *Parachanna obscura* fingerlings reared in captivity. *International Journal of Fisheries and Aquatic Studies*, 4(5), 01–05, 2016a.
- [12] Zhenjun S, Xianchun L, Lihui S, & Chunyang S. Earthworm as a potential protein resource. *Ecology of Food and Nutrition*, 36(2–4), 221–236, 1997. <https://doi.org/10.1080/03670244.1997.9991517>.
- [13] Dynes RA. Earthworms: technology information to enable the development of earthworm production. In: Dynes, R.A., editor/s. Canberra: Rural Industries Research and Development Corporation. 2003. <http://hdl.handle.net/102.100.100/190414?index=1>
- [14] Bahadori Z, Esmailzadeh L, Karimi-Torshizi MA, Seidavi A, Olivares J, Rojas S. The effect of earthworm (*Eisenia foetida*) meal with vermi-humus on growth performance, hematology, immunity, intestinal microbiota, carcass characteristics, and meat quality of broiler chickens. *Livestock Science*, 202, 74–81, 2017. <https://doi.org/10.1016/j.livsci.2017.05.010>
- [15] Rondón RV, Ovalles-Durán JF, León-Leal A, Medina A. Valor nutritivo de la harina de lombriz (*Eisenia foetida*) como fuente de aminoácidos y su estimación cuantitativa mediante cromatografía en fase reversa (HPLC) y derivatización precolumna con o-ftalaldehído (OPA). *Ars Pharmaceutica*, 44(1):43-58, 2003.
- [16] Ferruzzi C. Manual de Lombricultura. Madrid: Ediciones Mundi Prens. Pp. 138. 2003.
- [17] Sharma K, Garg V. Management of food and vegetable processing waste spiked with buffalo waste using earthworms (*Eisenia foetida*). *Environmental Science Pollution Resources*, 24, 7829–7836, 2017. <https://doi.org/10.1007/s11356-017-8438-2>.
- [18] Vodounnou DS, Kpogué D, Tossavi NS, Mensah GA, Fiogbe ED. Effect of animal waste and vegetable compost on production and growth of earthworm (*Eisenia foetida*) during vermiculture. *International Journal Recycling Organic Waste Agriculture*, 5, 87–92, 2016b. <https://doi.org/10.1007/s40093-016-0119-5>
- [19] Paco G, Loza-Murguía M, Mamani F, Sainz H. Efecto de la Lombriz Roja Californiana (*Eisenia foetida*) durante el composteo y vermicomposteo en predios de la Estación Experimental de la Unidad Académica Campesina Carmen Pampa. *J Selva Andina Res Soc.* 2(2),24-39, 2011. <http://www.scielo.org.bo/pdf/jsars/v2n2/a04.pdf>.
- [20] García M, Fajardo V. Cría de la lombriz de tierra: Una alternativa ecológica y rentable. Fundación Hogares Juveniles Campesinos. Ed. San Pablo, Bogotá, Colombia. pp. 193. 2005.
- [21] Schofield J, Noonan D, Chen Y, Penson P. 2014. Laboratory Animals Regulations and Recommendations for Global Collaborative Research: Australia and New Zealand. In: *Laboratory Animals*, 333-376. <http://dx.doi.org/10.1016/B978-0-12-397856-1.00012-X>
- [22] Aller MA, Rodríguez J, Rodríguez G. (2020). Normas éticas para el cuidado y utilización de los animales de experimentación. *Cirugía Española*, 67(1), 10-13. <https://www.elsevier.es/es-revista-cirugia-espanola-36-articulo-normas-eticas-el-cuidado-utilizacion-8848>.
- [23] Hussain N, Abbasi SA. Efficacy of the vermicomposts of different organic wastes as “clean” fertilizers: State-of-the-art. *Review Sustainability*, 10(4), 1205, 2018. <https://doi.org/10.3390/su10041205>.
- [24] Niamah, A. K., Sahi, A. A., & Al-Sharifi, A. S. (2017). Effect of feeding soy milk fermented by probiotic bacteria on some blood criteria and weight of experimental animals. *Probiotics and antimicrobial proteins*, 9(3), 284-291. <https://doi.org/10.1007/s12602-017-9265-y>
- [25] Bou-Maroun E, Loupiae C, Loison A, Rollin B, Cayot P, Cayot N, Marquez E, Medina A. Impact of preparation process on the protein structure and on volatile compounds in *Eisenia foetida* protein powders. *Food Nutr Sci*. 2013; 4: 1175-1183, 2013. <https://doi.org/10.4236/fns.2013.411151>
- [26] Akpodiete OJ, Okagbire GN. Feed accessories from animal production. In: *Issue on Animal Sciences. A compendium of ideas, fact and methods in the science and technology of Animal Agriculture* Ram Kemmedy City, Nigeria. *Animal Sciences*, 71-82, 1999.
- [27] Medina AL, Cova JA, Vielma RA, Pujic P, Carlos MP, Torres JV. Immunological and chemical analysis of proteins from *Eisenia foetida* earthworm. *Food and Agricultural Immunology*, 15(3-4):255–263, 2003. <https://doi.org/10.1080/09540100400010084>.
- [28] Suárez-Hernández L, Barrera-Zapata R, Forero-Sandoval A. Evaluación de alternativas de secado en el proceso de elaboración de harina de lombriz. *Corpoica Cienc & Tecnol Agropecuaria*, 17(1), 55-71, 2016. https://doi.org/10.21930/rcta.vol17_num1_art:461.
- [29] Castro J, Chirinos D. Nutrición Animal. tercera edición. Huancayo, Perú: Gráfica José Impresores. Huancayo. Perú .265 p. 2017.
- [30] Kassa A. The Different Methods of Measuring Feed Digestibility: A Review. *EC Nutrition* 14(1), 68-74, 2019. <https://www.econicon.com/ecnu/pdf/ECNU-14-00542.pdf>
- [31] AOAC. Official Methods of Analysis of AOAC International: AOAC International, Maryland, USA, 2011.
- [32] Tharel L. Total Digestible Nutrients and Protein per Acre Produced by Five Indiangrass Cultivars. Technical Report. Booneville PMC Study. Booneville, AR. CP 512 Pasture and Hay. Plant Materials Technical Report, CP 512, 2008. <https://www.nrcs.usda.gov/wps/portal/nrcs/publications/plantmaterials/pmcp/southeast/arpmpub/>
- [33] Stein HH, Sève B, Fuller MF, Moughan PJ, de Lange CF. Invited review: Amino acid bioavailability and digestibility in pig feed ingredients: terminology and application. *J Anim Sci*. 85(1), 172-180, 2007.
- [34] Kassa A. The Different Methods of Measuring Feed Digestibility: A Review. *EC Nutrition* 14(1), 68-74, 2019. <https://www.econicon.com/ecnu/pdf/ECNU-14-00542.pdf>.

- [35] Crampton E, Harris L. Nutrición animal aplicada. Editorial Acribia. segunda edición Zaragoza – España. 756pp. 1974.
- [36] NRC. The Nutrient Requirements of Horses, 4th revised edition. National Academy Press, Washington DC. 1978.
- [37] Weiss W, Tebbe A. Estimating digestible energy values of feeds and diets and integrating those values into net energy systems, Translational Animal Science, 3(3), 953–961, 2019. <https://doi.org/10.1093/tas/txy119>.
- [38] NRC. Nutrient Requirements of Beef Cattle: Seventh Revised Edition: Update 2000. National Academy Press, Washington DC. 2000.
- [39] Galycan ML, Cole NA, Tedeschi LO, Branine ME. BOARD-INVITED REVIEW: Efficiency of converting digestible energy to metabolizable energy and reevaluation of the California Net Energy System maintenance requirements and equations for predicting dietary net energy values for beef cattle. Journal of Animal Science, vol 94, No. 4, 1329–1341, 2016. <https://doi.org/10.2527/jas.2015-0223>.
- [40] Barakat NA, Laudadio V, Cazzato E, Tufarelli V. Potential Contribution of Retama raetam (Forssk.) Webb & Berthel as a Forage Shrub in Sinai, Egypt. Arid Land Research and Management, 27(3), 257–271, 2013. <https://doi.org/10.1080/15324982.2012.756561>
- [41] Sainz RD, Fernández C, Baldwin RL. Valoración de alimentos para rumiantes en cebo: el Sistema Americano NRC. X Curso de Especialización FEDNA. Madrid. 1994. http://www.ucv.ve/fileadmin/user_upload/facultad_agronomia/NRC_Bovinos.pdf.
- [42] Castro J, Chirinos D, Calderón J. Calidad nutricional del rastrojo de maca (*Lepidium peruvianum* Chacón) en cuyes. Rev. investig. vet. Perú, 29(2), 410-418. ISSN 1609-9117. <http://dx.doi.org/10.15381/rivep.v29i2.13405>.
- [43] Gunya B, Masika P, Hugo A, Muchenje V. Nutrient Composition and Fatty Acid Profiles of Oven-dried and Freeze-dried Earthworm *Eisenia foetida*. Journal of Food and Nutrition Research, 4(6), 343-348, 2016. <https://doi.org/10.12691/jfnr-4-6-1>
- [44] Segovia E. Obtención y caracterización de la Harina de Lombriz (*Eisenia Foetida*) para Consumo Animal. Tesis. Universidad Nacional Federico Villarreal. Lima. Perú. 1997.
- [45] De Chaves RC, Paula RQ, Gücker B, Marriel IE, Teixeira AO, Boëchat IG. An alternative fish feed based on earthworm and fruit meals for tilapia and carp postlarvae. R. bras. Bioci., Porto Alegre, vol. 13(1), 15–24, 2015. <https://www.researchgate.net/publication/307547140>.
- [46] Mohanta KN, Subramanian S, Korikanthimath VS. Potential of earthworm (*Eisenia foetida*) as dietary protein source for rohu (*Labeo rohita*) advanced fry. Cogent Food & Agriculture, 2(1), 2016. <https://doi.org/10.1080/23311932.2016.1138594>.
- [47] Wang H, Shi M, Xu X, Ma X, Liu L, Piao X. Comparative energy content and amino acid digestibility of barley obtained from diverse sources fed to growing pigs. Asian-Australas J Anim Sci. 30(7), 999-1005, 2017. <https://doi.org/10.5713/ajas.16.0775>
- [48] Pérez-Corria K, Botello-León A, Mauro-Félix A, Rivera-Pineda F, Viana M, Cuello-Pérez M, Botello-Rodríguez A, Martínez-Aguilar Y. Chemical Composition of Earthworm (*Eisenia foetida*) Co-Dried with Vegetable Meals as an Animal Feed. Ciencia y Agricultura, 16(2),79-92, 2019. <https://doi.org/10.19053/01228420.v16.n2.2019.9130>

Piezoelectric Teeth Aligners to Accelerate Orthodontics Treatment

Muath Bani-Hani^{*1}, M. Amin Karami²

¹Department of Aeronautical Engineering, Jordan University of Science and Technology, Irbid, Jordan

²Department of Mechanical and Aerospace Engineering, The State University of New York at Buffalo, 14260, U.S.A

ARTICLE INFO

Article history:

Received: 13 April, 2019

Accepted: 28 August, 2019

Online: 15 May, 2020

Keywords:

Biomedical materials

Piezoelectric actuators

Orthotics & patient treatment

ABSTRACT

In this paper, we are proposing a device that generates vibration to possibly reduce the duration required for the orthodontic treatment by enhancing the tooth movement. This is achieved by harmonically exciting a bio-compatible material, namely polyvinylidene fluoride (PVDF) to generate vibration and consequently a cyclic force at a low frequency of 30 Hz. (PVDF) is a very common piezoelectric polymer due to the high elasticity it can provide, biocompatibility, and its low cost compared to other piezoelectric materials. Applying a cyclic loading (vibration) in general will reverse bone loss, stimulate bone mass, induce cranial growth, and consequently accelerate tooth movement. This has a major effect on the patient's treatment experience by reducing the associated pain and discomfort throughout the treatment process and hence improves the patient's compliance with the treatment with negligible side effects compared to therapeutic treatments. The proposed device is fitted to a positioner or tooth aligner. The proposed device can accommodate a voltage harmonic function generator, a casing to house the battery and the micro-processor. Modern methods require a bulky and extra-oral device that is quite cumbersome to the patient. In this paper, an analytical model based on the distributed parameter approach is employed. Finite Element Analysis (FEM) is also employed to study and analyze the Piezoelectric actuation behavior.

1. Introduction

The work presented in this paper is an extension of what was presented in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) [1].

Orthodontics is a dental domain that deals with enhancing the movement of teeth to treat malocclusion. This mainly deals with the correction of malposition of the teeth and the jaws. Malocclusion suggests that the upper and lower jaws are not in complete alignment for biting or chewing which leads to uneven bite, cross bite, or overbite. Therefore, this might negatively affect a person's physical appeal, communication and eating capabilities.

Malocclusion is traditionally adjusted through a steady/constant mechanically applied force to encourage bone modelling, which enhance the teeth movement to a preferred position. For this method, an orthodontic device applies a constant and stagnant force on the teeth through a bent wire attached to brackets fixed on the teeth or an aligner. This continuous force is

slowly dissipated as the teeth move into position. The wires or the retainer are tweaked accordingly to ensure the teeth move to the optimal position.

Swartz [2] shows the optimum force to move a tooth is just enough to overcome blood vessels pressure (20-26 grams/cm²). Any extra force used to move a tooth can significantly affect blood circulation, leading to root resorption. In this method, as known as fixed orthodontic treatment, 2 to 3 years are required for the completion of the treatment [3, 4]. It also poses a high risk of caries [5, 6], root resorption [7, 8], and reduce patient compliance over time [9]. Multiple studies have found a certain connection between extended treatment and the root resorption experienced by a patient [7, 10, 11]. However, even when no radiographic signs of root resorption can be visible, it is acceptable that most teeth undergoing orthodontic tooth movement will experience some degree of root resorption followed by repair [12]. The first person to reveal a radiographic proof of root resorption amid an orthodontic treatment was Ketchum [13]. He found that out of five hundred patients, 21% were found to have root resorption. For this reason, it is imperative to study the acceleration of tooth movement

*Muath A.Bani-Hani, Jordan University of Science and Technology, Jordan, Mabanihani@just.edu.jo

www.astesj.com

<https://dx.doi.org/10.25046/aj050324>

to reduce treatment duration time. Several methods to accelerate teeth movement have been studied, including laser therapy [14, 15], pulsating electromagnetic fields [16], electrical currents [17], corticotomy [18, 19], distraction osteogenesis, [20-22] and mechanical vibration [23].

Rubin et al 2001 revealed that the use of low amplitude mechanically vibrating forces/pulses at a high frequency could help in stimulating bone growth. This type of therapy, also referred to as vibration therapy, has been used to improve or maintain bone and muscle mass in rehabilitation of mobility impaired patients [24]. It has been also utilized to combat decreased bone density [25, 26] and improving post-surgical healing [27, 28]. During his animal studies, Rubin found out that bone density in the proximal femur showed a 43% increase after sessions of 20 minutes/day of high frequency (30Hz) and low amplitude force (0.3g) [29-31].

The small strains generated by high frequency mechanical vibrations are implemented to speed up periodontal and bone tissue remodeling through speeding up tooth movement. With the minimal side effects, this method of orthodontic treatment has the advantage over the current medicinal treatments. It is also a safer alternative for enhancing bone remodeling in the medical field [32, 33].

Researchers have suggested that a pulsating could possibly move teeth faster with less discomfort associated with the traditional orthodontics. Nishimura and Chiba M in 2008 [23], used intermittent stimulation of the periodontal tissue by resonance mechanical vibration to accelerate experimental tooth movement in rats. His study was performed on a sample of 42 wistar rats with an experimental duration of 21 days. It was shown a major speed up in the rate of the tooth malocclusion treatment in comparison to non-vibration control sample. Nevertheless, Mao was the pioneer in his field to actually prove that the use of pulsating/cyclic forces compared to static loading, could significantly improve dental straightening in rabbits [34].

Currently, a wide range of teeth vibration devices are available and used in orthodontic treatment. Craven H. Hurz [35] uses a hefty exterior power source that supplies electricity to one to four small motors. His device [36, 37] is modified to use pulsating fluids that are produced due to the chewing motion of a jaw. A radio device is also included that oscillates in response to a radio wave and generates a pulsating/vibrating force on the teeth. The devices are fitted to a headband and attached to the teeth by its intraoral components. This configuration made the device complex, expensive to build, and made it uncomfortable for the patient to wear.

In 1991, Branford [38] proposed a hand-held device used to potentially treat the periodontal disease. It used a mouth-piece made of malleable dental brass adjustable to the patient's bite and an external vibrator to induce oscillations in the teeth. However, this device required an external power source to power the vibrations.

Michael J. Powers [39] also proposed a hand-held tooth vibrator in 1999. Power's device employed a motor linked to a vibrating intraoral mouthpiece. The drawback to his proposed device was user discomfort due to increased blood flow from the vibrations. The bulkiness of the power actuator also affected user

comfort and patient compliance, as defined by John C. Voudouris in 2001 [40]. The device is mounted on brackets which reduce the effectiveness of conveying the vibration forces to the teeth. Matsushita Electric Works, Ltd. [41] created a multipart intraoral pulsating device, intended to be attached or fitted over the lower and upper teeth to apply vibrating/pulsating forces. However, the device complexity and cost to manufacturing were a major drawback. Note that what reduces the effectiveness of these aforementioned devices is that the vibrational forces are not perfectly optimized in terms of frequency and magnitude for bone remodeling.

Following these findings, in 2006 Mao presented a way to enhance bone remodeling with pulsating forces by achieving a quicker rate for the tooth movement [42, 43]. The device, AcceleDent by OrthoAccel, describes an intra and extra oral dental vibrator combined with microchip to collect, process and convey patient usage data. Due to its improved bite plate, this device is more effective in transmitting pulsating forces to the corresponding teeth. The device is also optimized in terms of force magnitude and frequency which enhances the comfort level and patient compliance. AcceleDent is prescribed to be used for two daily 10 minute sessions during orthodontic treatment and could be used with a secured appliance or aligner treatment. University of Texas-Houston in Texas, US, examined the safety and effectiveness of vibrations for orthodontic tooth movement. The study showed a nearly 70% enhancement in the tooth movement in comparison to previous work [44, 45] and reduced root resorption. No injurious cases were recorded from the corresponding treatment. This validated the safety of vibration-based treatment in addition to the AcceleDent device. In comparison to previous approaches, AcceleDent is also more compact, since it does not require the patient to wear a head-dress mount. Based on the study, an output force of 25 grams and 30Hz frequency were equally scattered on the upper and lower jaws. However, the corresponding device offers vibrations for all the teeth, causing it to be bulky. For this reason, further improvements may be necessary.

The tooth retainer or positioner are appliances often used in the closing stages of orthodontic treatment to aid in the achieving of a complete treatment. It was created by Kesling [46] in the 1940s and fabricated using rubber-based materials. Silicone was the favored material for the retainer for its transparent look, heat resistivity, and hypoallergenic properties. However, it is less effective in treating malocclusions due to inaccuracies from fabrication.

Kurz in 1982 [47] introduced a device composed of tooth positioner and a vibrator with a hydraulic pump or an ultrasonic motor. However, the device vibrational forces were not optimized in terms of frequency and magnitude. In addition, Kurz's device involved the use of straps wrapped around a patient's head during sleeping, leading to discomfort and reduced patient compliance. OrthoAccel [48] claims that their vibratory tool that can be combined with an existing orthodontic device, namely Invisalign, but their device is mainly not designed as an aligner. Currently, there are many accessible options for a vibrating motor that is small enough to be housed within the mouth. Such motors include piezoelectric motors which can be manufactured at the nano scale. Utilization involves an unbalanced motors which may

be preferred and more affordable. Planar motors can also be utilized where the vibration is in parallel to the substrate [49-52].

This paper presents an innovative intraoral pulsating/vibrating device that is confined within the mouth. The proposed device is expected to potentially decrease the time and enhance the effectiveness of orthodontic treatments which leads to less employed resources to each individual patient and also increases patient compliance.

The proposed device is composed of piezoelectric actuators and a voltage generator to excite the piezoelectric material at a certain amplitude and frequency. The vibrations are produced by a bio-compatible smart material, namely (PVDF) piezoelectric actuator, that can be oscillated at a specific frequency and amplitude. Due to its many physical properties including its great flexibility, biocompatibility, and low cost, PVDF is preferred in many biomedical and mechanical/electricity conversion applications including micro electric-mechanical devices, electro-mechanical actuators, and energy harvesters [53-56]. For our design, the piezoelectric actuator is attached to a specific part of a retainer and provides vibrations to the targeted teeth. The major advantage of the device is that it can be relocated to various locations on the teeth and can be adjusted to a patient's needs. The intra-oral compromises of a function generator, a processor, and battery are all contained into one unit allowing for a small design in comparison to the other available designs. Therefore, drooling can be minimized due to lips are partially kept open by the designated extra-oral components.

2. Device configuration

2.1. Electromechanical model and configuration of the proposed PVDF apparatus

The proposed device is designed to be combined with the existing approach of tooth alignment, a retainer or aligner, for instance. Generally, tooth aligners are created using complex software to create a digital plan for each individual patient by predicting the movement of teeth throughout the treatment. This means that several aligners are used over the course of the treatment. These aligners are created to apply a constant pressure to the corresponding teeth, forcing the teeth to the desired position over time. This process can be accelerated by applying a cyclic force, or vibration, to the teeth for at approximately 20 minutes every day divided over two 10 minutes sessions. Dental remodeling can be accelerated as much as 70% with this method [44, 45]. The proposed device is fitted to an aligner, such as Invisalign, and a vibration is applied to the teeth through said device. Medical grade silicone rubber is preferred as a tooth positioner due to its transparency, strength, comfort, and lack of taste. The proposed device has the ability to be equipped with an micro-processor chip to collect and transmit the treatment data, which delivers vital statistics for monitoring patient compliance.

Figure 1 shows a 3D view of an tooth aligner matching the shape of the targeted teeth. It is generally U-shaped. The tooth aligner is designed to contact facial, lingual and occlusal surfaces of the targeted teeth by applying a corrective pressure to corresponding targeted teeth. Only the lower jaw is shown here.

Our teeth vibration device compromises only three parts due to the restrictions of attaching the device to an aligner. The parts are: 1) a harmonic function generator and micro-chip processor, 2) an intra-oral low current and high voltage battery and 3) couple of PVDF vibrators fitted over the teeth aligner, see Figure 2.a, and rigid steel wires connect the battery with the piezoelectric actuators and hold both the function generator processor and battery, Figure 2.a. The function generator processor and battery are housed in a waterproof unit. This unit can be accessed to allow for collection of usage information, which gives clinicians insight on patient compliance.

The micro-chip processor is also adjustable to allow the patient to adjust frequency and forces. The enclosure unit can also contain a charging port, on/off switch, lights, or access to the battery. In our analysis introduced in this paper, the vibrating actuator is set to produce a frequency of 30 Hz and a force of 8.5 grams produced at 100 volts.

Our device is designed to allow for movement of the actuators on the upper or lower jaw along different locations of a tooth aligner. The function generator is utilized to excite the PVDF actuators to generate a cyclic force to the targeted teeth. Polyvinylidene fluoride (PVDF) is ideal for this device due to its ease of handling and acceptable mechanical properties. Stainless steel was chosen for its malleability, stiffness, resilience, biocompatibility, and low cost.



Figure 1: 3D view image of a tooth aligner fitted on a lower jaw

The PVDF actuator is primarily demonstrated as a flexible U-shaped rod fitted to the top of an aligner. It is excited and polarized along the actuator thickness restricting vibrations in the z -direction. During use, the patient bites down on the actuators, causing the actuators to only stretch axially to the surface of the aligner. The cyclic forces mentioned previously are applied longitudinally to produce directional vibrations to the structure. Figure 2a shows a plain model of the proposed device fitted to the aligner and the axial force direction.

Piezoelectric constitutive formulas are used for modeling the piezoelectric devices and are based on assumptions where the transducer's total mechanical strain is the sum of the mechanical strain and that the controlled actuation strain results from the applied electrical voltage.

Instead of using letters to label axis, numbers will be used in its place. 1 is the x -axis, 2 will be the y -axis, and 3 is the z -axis. The third axis is parallel to the direction of polarization of the piezo ceramic device, and axis 1 and 2 are normal to axis 3.

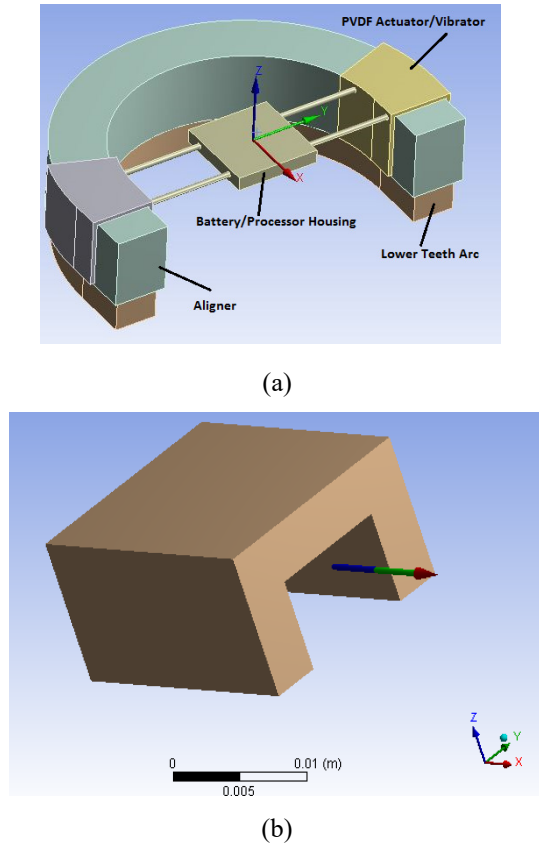


Figure 2: (a) Proposed PVDF device attached to the tooth aligner. (b) U-shape PVDF actuator/vibrator

By employing the Piezoelectric constitutive equations, the coupled electromechanical model of the piezoelectric material can be presented as:

$$\begin{aligned} S_i &= s_{ij}^E T_j + d_{mi} E_m \\ D_m &= d_{mi} T_i + \epsilon_{ik}^T E_k \end{aligned} \quad (1)$$

The corresponding model is mainly presenting the stress components (T), mechanical strain (S), electric field (E) and the electric displacement components (D). ϵ^T is the Permittivity (F/m). (d) is the piezoelectric matrix of the strain constants (m/V). The indexes $i, j = 1, 2, \dots, 6$ and $m, k = 1, 2, 3$ are labeled for the multi-directions coordinate system, as illustrated by:

Index	Coordinate
1	x
2	y
3	z
4	Shear about x
5	Shear about y
6	Shear about z

Equations (1) can be given in the matrix form as:

$$\begin{bmatrix} S \\ D \end{bmatrix} = \begin{bmatrix} s^E & d \\ d^t & \epsilon^T \end{bmatrix} \begin{bmatrix} T \\ E \end{bmatrix} \quad (2)$$

Superscripts E and T define the corresponding variables that are calculated at constant electric field and stress, respectively. The

superscript t is a mathematical indication for a vector transpose. The fully expanded formula of Equation (2) is outlined as:

$$\begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \end{bmatrix} = \begin{bmatrix} s_{11}^E & s_{12}^E & s_{13}^E & s_{14}^E & s_{15}^E & s_{16}^E \\ s_{21}^E & s_{22}^E & s_{23}^E & s_{24}^E & s_{25}^E & s_{26}^E \\ s_{31}^E & s_{32}^E & s_{33}^E & s_{34}^E & s_{35}^E & s_{36}^E \\ s_{41}^E & s_{42}^E & s_{43}^E & s_{44}^E & s_{45}^E & s_{46}^E \\ s_{51}^E & s_{52}^E & s_{53}^E & s_{54}^E & s_{55}^E & s_{56}^E \\ s_{61}^E & s_{62}^E & s_{63}^E & s_{64}^E & s_{65}^E & s_{66}^E \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \end{bmatrix} + \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{bmatrix}^t \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix}$$

The shear stresses are expressed as:

$$\begin{aligned} T_4 &= \tau_{23} \\ T_5 &= \tau_{31} \\ T_6 &= \tau_{12} \end{aligned}$$

and the shear strains are expressed as:

$$\begin{aligned} S_4 &= \gamma_{23} \\ S_5 &= \gamma_{31} \\ S_6 &= \gamma_{12} \end{aligned}$$

The electric displacement is expressed as:

$$\begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix} + \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ \tau_{23} \\ \tau_{31} \\ \tau_{12} \end{bmatrix} + \begin{bmatrix} \epsilon_{11}^T & \epsilon_{12}^T & \epsilon_{13}^T \\ \epsilon_{21}^T & \epsilon_{22}^T & \epsilon_{23}^T \\ \epsilon_{31}^T & \epsilon_{32}^T & \epsilon_{33}^T \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix} \quad (4)$$

Equation (1) can be re-written to relate the stress and strain as:

$$\{T\} = [c^E]\{S\} - [e]\{E\} \quad (5)$$

$[c^E]$ represents the stiffness matrix that is calculated at constant electric field. $[e]$ represents the piezoelectric constants matrix that relates stress to the electric field.

Equation (1) can be transformed to Equation (5) by implementing the set of equations of:

$$\begin{aligned} \{S\} &= [s^E]\{T\} + [d]\{E\} \\ [s^E]\{T\} &= \{S\} - [d]\{E\} \\ \{T\} &= [s^E]^{-1} \{S\} - [s^E]^{-1} [d]\{E\} \end{aligned} \quad (6)$$

And hence,

$$\begin{aligned} [c^E] &= [s^E]^{-1} \\ [e] &= [s^E]^{-1} [d] \end{aligned} \quad (7)$$

Y (Young's modulus of elasticity) and ν (Poisson's ratio) are indicated in Table 1. At this point and at constant electric field, the final shape of the PVDF compliance matrix $[S^E]$ can be obtained as:

$$[S^E] = \begin{bmatrix} 1/Y & -\nu/Y & -\nu/Y & 0 & 0 & 0 \\ -\nu/Y & 1/Y & -\nu/Y & 0 & 0 & 0 \\ -\nu/Y & -\nu/Y & 1/Y & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/G & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/G & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/G \end{bmatrix} (m^2/N)$$

G represents the modulus of rigidity ($G = \frac{Y}{2(1+\nu)}$).

The stress constants matrix of the PVDF actuator, $[d]^t$ can be represented as:

$$[d]^t = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 23 & 3 & -33 & 0 & 0 & 0 \end{bmatrix} \times 10^{-12} (m/V)$$

PVDF dielectric Matrix $[\epsilon^T]$:

$$[\epsilon^T] = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix} \times 8.854 \times 10^{-12} (F/m)$$

Employing Equation (7) and the piezoelectric constant matrix, the PVDF stiffness matrix is generated as indicated in Equations (8) and (9).

Table 1 indicates the PVDF material physical properties.

$$[c^E] = \begin{bmatrix} 2.7 & 1.154 & 1.154 & 0 & 0 & 0 \\ 1.154 & 2.7 & 1.154 & 0 & 0 & 0 \\ 1.154 & 1.154 & 2.7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.77 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.77 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.77 \end{bmatrix} \times 10^9 (N/m^2) \quad (8)$$

$$[e] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0.0273 \\ 0 & 0 & 0 & 0 & 0 & -0.0035 \\ 0 & 0 & 0 & 0 & 0 & -0.0588 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (9)$$

$\left(\frac{N}{mV}\right)$

From Equation(9), we obtain:

$$\begin{aligned} e_{31} &= 0.0273 \left(\frac{N}{mV}\right) \\ e_{32} &= -0.0035 \left(\frac{N}{mV}\right) \\ e_{33} &= -0.0588 \left(\frac{N}{mV}\right) \end{aligned} \quad (10)$$

Table 1: PVDF material properties

Density (kg/m ³)	1,780	ν	0.3	Y (GPa)	2.0
------------------------------	-------	-------	-----	---------	-----

2.2. PVDF actuator electromechanical model at e_{31} mode

Figure 3 visualizes a PVDF actuator where the electric field (+E) is in the direction of the PVDF thickness. The positive sign of the electric field indicates a same direction as for the polarization, and this is analytically analyzed by implementing the distributed parameter approach. Finite Element Method (FEM) software such as ANSYS is utilized to verify the introduced analytical model.

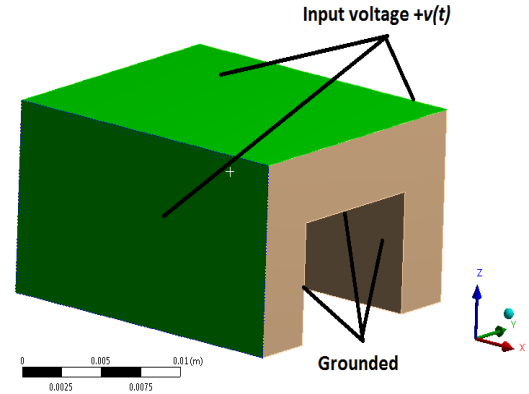


Figure 3: U-shape PVDF actuator/vibrator at e_{31} actuation mode showing an electric field E in the direction of the material thickness

It must be noted that the electromechanical model is derived based on the assumption that the PVDF actuator is solely poled along the z-axis (the PVDF thickness direction) where the stress and strain are undertaken axially along the direction of actuator length (x-direction). At this point, the energy method can be employed to generate the dynamic model of the PVDF actuator. Consequently, the coupled electromechanical model is generated by utilizing the so called Hamilton's principle that takes the variation of a function with respect to time. According to Hamilton's principle the Lagrangian (L) function is defined as:

$$L = K + W_e^* \quad (11)$$

K is the kinetic energy. W_e^* is the stored electrostatic energy. The PVDF actuator will be represented as a rod. The physical properties of the proposed elastic PVDF bar are labeled as: length l, width b, mechanical stiffness at constant electric field c_{11}^E , mass density ρ and the cross-sectional area A.

At this point, the absolute displacement of the PVDF actuator as a function of x and time is indicated by $u(x, t)$. Using the principle of separation of variables, the solution is written as:

$$u(x, t) = \sum_{n=1}^{\infty} U_n(x)\eta_n(t) \quad (12)$$

The function $U(x)$ is the orthogonal mode shape and $\eta(t)$ is the temporal function which are related to the Lagrangian function as the following:

$$K = \frac{1}{2} \int_0^l \rho A \left(\frac{du}{dt}\right)^2 dx \quad (13)$$

$$W_e^* = -\frac{1}{2} \int_0^l c_{11}^E A \left(\frac{du}{dx}\right)^2 dx + \int_0^l A \left(\frac{du}{dx}\right) e_{31} E_3 dx$$

e_{31} is the piezoelectric constant at a constant stress and the electric field as a function of voltage is ($E_3 = \frac{-v(t)}{h}$) [57]. Based on Hamilton's principle, the variation of the function with respect to respect to time equals to zero:

$$\delta \int_{t_1}^{t_2} \int_0^L L dx dt = 0 \tag{14}$$

Which becomes:

$$\delta \int_{t_1}^{t_2} Q dt = 0, \tag{15}$$

Where $Q = \frac{1}{2} \int_0^l \rho A \left(\frac{du}{dt}\right)^2 dx - \frac{1}{2} \int_0^l c_{11}^E A \left(\frac{du}{dx}\right)^2 dx + \int_0^l A \left(\frac{du}{dx}\right) e_{31} E_3 dx$

Solving the above equations yields to:

$$\int_{t_1}^{t_2} \int_0^l \rho A \frac{du}{dt} \delta \left(\frac{du}{dt}\right) dx dt + \int_{t_1}^{t_2} \int_0^l c_{11}^E A \frac{du}{dx} \delta \left(\frac{du}{dx}\right) dx dt - \int_{t_1}^{t_2} \int_0^l \frac{Ae_{31}}{h} \delta v(t) \left(\frac{du}{dx}\right) dx dt - \int_{t_1}^{t_2} \int_0^l \frac{Ae_{31}}{h} v(t) \delta \left(\frac{du}{dx}\right) dx dt = 0 \tag{16}$$

The Euler-Lagrange formulas can now be assembled, obtaining the dynamic governing equation:

$$\rho A \frac{\partial^2 u(x,t)}{\partial t^2} - c_{11}^E A \frac{\partial^2 u(x,t)}{\partial x_i^2} - \frac{Ae_{31}}{h} v(t) \frac{d}{dx} [H(x) - H(x-l)] = 0 \tag{17}$$

$H(x)$ represents the Heaviside function.

At this stage, we evaluate the cyclic forces produced by the proposed PVDF actuators excited over a certain range of constant voltages by implementing of statically indeterminate structures principle.

The equilibrium equations relate structural deformations in such way to obtain the reaction forces. Consequently, due to the linearity of the system, the superposition method is implemented to obtain the corresponding forces by assuming that one of the reaction forces is redundant. This force is considered as an unknown load, which together with the other loads must produce

deformations that match the original constrains (see Figure 4 for details). δ indicates the elongation of the rod.

The mode shapes at this point can be generated for a fixed-fixed rod to estimate the axial forces produced by the PVDF actuator.

The mode shapes presented as a function of the natural frequency ω_n are written as:

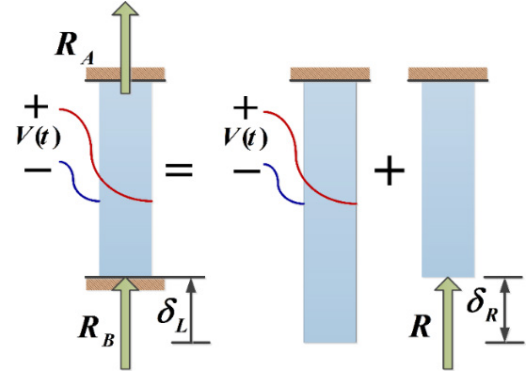


Figure 4: Free body diagram of a statically indeterminate rod

$$U(x) = C_n \sin\left(\frac{(2n+1)\pi x}{2l}\right), \quad \omega_n = \frac{(2n+1)\pi c}{2l}, \quad n = 1, 2, \dots \tag{18}$$

$$c = \frac{\sqrt{c_{11}^E}}{\sqrt{\rho}}$$

c_n represents the modal constant, and by employing the orthogonality condition, the modal shapes $U(x)$ can be mass normalized (i.e. solving for the modal constant) as the following:

$$\rho A \int_0^l U_n(x)^2 dx = 1 \tag{19}$$

$$C_n = \frac{\sqrt{2}}{\sqrt{L\rho A}}$$

By plugging equation (12) into the governing equation (17), multiplying by the mode shape, then integrating over the length of the rod, and then imposing the orthogonality conditions, the electromechanically coupled ordinary differential equation for the modal response of the rod is obtained as such:

$$\ddot{\eta}_n + 2\zeta\omega_n\dot{\eta} + \omega_n^2 \eta + \frac{Ae_{31}}{h} v(t)[U(l) - U(0)] = 0 \tag{20}$$

At this point, and by equation (12), the axial force at the tip of the rod (in Newtons) is obtained as:

$$f(l,t) = \frac{\sum_{n=1}^{\infty} U_n(l)\eta_n(t)c_{11}^E A}{l} \tag{21}$$

The dimensions of the PVDF actuator used in the study are shown in Table 2:

Table 2: Piezoelectric PVDF actuator physical dimensions

PVDF Dimensions	Quantity
Length (L)	20 mm
Cross section's outer edges length (b)	31 mm
Thickness (h)	4 mm

Figure 5 illustrates the estimated axial forces generated by the PVDF actuator. that shows the data of the actuator when it is stimulated by a wide range of Direct Current (DC) voltage.

The results obtained analytically are numerically validated at actuation mode of e_{31} with the help of a finite element FEM solver, namely ANSYS. In addition, Figure 5 indicates that the optimum external force produced by the PVDF actuator which is approximately equal to 8.5 grams at an excitation voltage of 100 volts. This resulted force can be transmitted to a group of four teeth.

As can be clearly observed from Figure 5, the axial force is proportional with the voltage due to the linearity of the system.

Applying a harmonic voltage (which is represented by $V\sin(\omega t)$), with the requirements of achieving a high amplitude cyclic forces, the PVDF actuators should normally have a fundamental natural frequency matching the harmonic voltage excitation frequency.

The damping has a major effect on the response of the system. Therefore, we can include the viscoelastic linear effect and hence the modulus of elasticity (the stiffness constant) under a constant electric field can be defined as:

$$c_{11}^E = (E' + iE'') \tag{22}$$

E' is the storage modulus and E'' is the loss modulus in viscoelastic model which are obtained by A.M. Vinogradova and F. Hollowayb experimentally as presented in their paper. This means that the natural frequency can be re-written as follows:

$$\omega_n^2 = \frac{\pi^2(E' + iE'')}{4l^2\rho}, n = 1 \tag{23}$$

If both the Laplace and Fourier transformations of equation (20) are taken and then used with the expression of the natural frequency in equation (23) we obtain:

$$\left(-\omega^2 + \frac{\pi^2(E')}{4l^2\rho} + \frac{\pi^2(iE'')}{4l^2\rho}\right)\eta(\omega) + \theta v(\omega) = 0 \tag{24}$$

Comparing the terms in equation (20) and equation (24), the damping ratio is obtained by the expression:

$$2\zeta\omega_n = \frac{\pi^2(E'')}{4l^2\rho} \tag{25}$$

A frequency response function (FRF) of the force over a wide range of excitation frequencies is shown below, and Figure 6 contains the plot of this function:

$$\frac{F(l, \omega)}{V(\omega)} = \sum_{n=1}^{\infty} \left| \frac{(U_n(l) - U_n(0))Ae_{31}}{h(\omega_n^2 - \omega^2 + 2j\zeta\omega\omega_n)} \right| \tag{26}$$

Figure 6 indicates the maximum axial force which is normally obtained at the resonance. However, because of the compact size of the PVDF actuator, the PVDF actuator natural frequencies are presented in kilohertz. Consequently, matching the exciting frequency with one of the PVDF actuator fundamental natural frequencies is not attainable. That being said, based on a study performed by Rubin revealed that a low frequency vibration (30 Hz to be exact) generated forces with a preferred magnitude to rapidly accelerate the tooth bone remodeling in comparison to a high magnitude and frequency force. Therefore, the cyclic forces are generated at 30 Hz harmonic voltage.

As can be seen in Figure 7, the cyclic forces in grams per unit voltage evaluated analytically are compared to their equivalents obtained by the FEM analysis. Comparison is performed only at the e_{31} actuation mode. As shown in Figure 7, both methods are in perfect agreement.

When examining Figure 6 and Figure 7, it can be seen that the absolute value of the harmonic forces are equivalent to the ones obtained in Figure 5. The reasoning behind this is that the harmonic voltage frequency is way smaller than the corresponding 1st fundamental natural frequency of the PVDF piezoelectric actuator. As a result, the system can be assumed to be a quasi-static (i.e. the system changes so slowly and hence can be treated to be always at equilibrium).

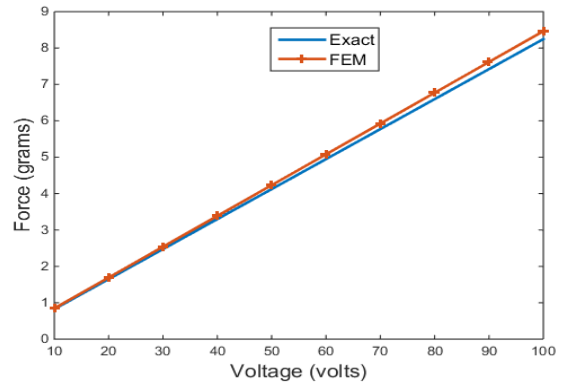


Figure 5: Exact and FEM axial force Vs voltage at e_{31} actuation mode

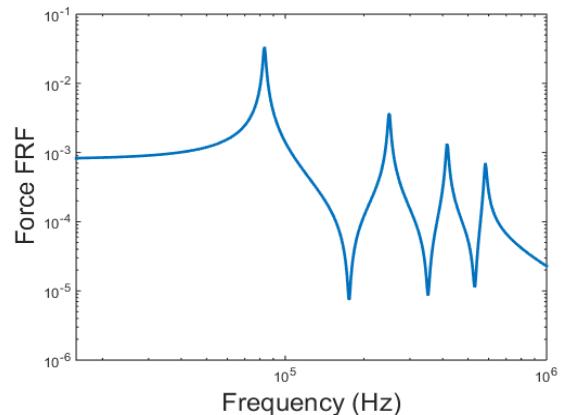


Figure 6: Force frequency response function e_{31} actuation mode

In order to improve the total generated axial force, the actuator design parameters denoted by the length, thickness and width of the actuator are analyzed.

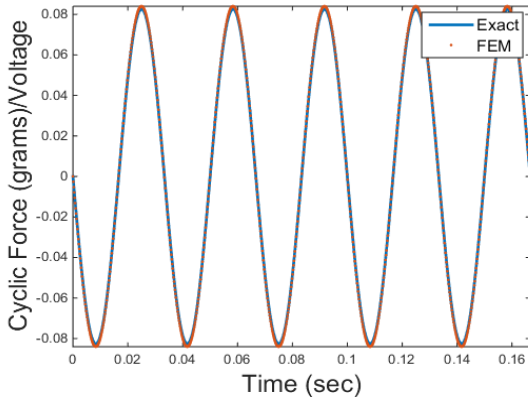


Figure 7: Analytically and numerically obtained cyclic forces in grams per unit voltage at 30 Hz at e_{31} actuation mode

By using equation (21) with sinusoidal harmonic input voltage denoted as $V \sin(\omega t)$, where V being the amplitude and ω being the excitation frequency ($= 30$ Hz), the following expression for the force (in *Newtons*) is obtained:

$$f(l, t) = \frac{2e_{31}c_{11}^E b}{l^2 \rho} \sum_{n=1}^{\infty} \left(\sin\left(\frac{(2n+1)\pi}{2}\right) \right)^2 \frac{V \sin(\omega t + \phi_n)}{\sqrt{(\omega_n^2 - \omega^2)^2 + 4\zeta^2 \omega_n^2 \omega^2}} \quad (27)$$

In a polarized system in the z -axis in the e_{31} actuation mode, it can be seen from Equation (27) that the actuator thickness has no effect on resulting force. In addition, in reference to the natural frequency in Equation (18), a conclusion can be made from Equation (27) that the PVDF actuator length has a minimal effect on the resulting axial force, as observed in Figure 8. For this particular case, the actuator thickness of 4 mm is chosen to be consistent with the aligner thickness and hence keeps the device lightweight and compact.

From Equation (27) and Figure 9, an observation can be made that the actuator width is linearly proportional to the resulted axial force.

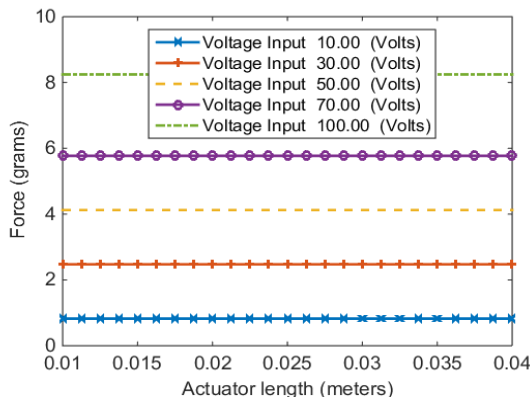


Figure 8: The variation of axial force with PVDF actuator length at different voltage inputs at e_{31} actuation mode

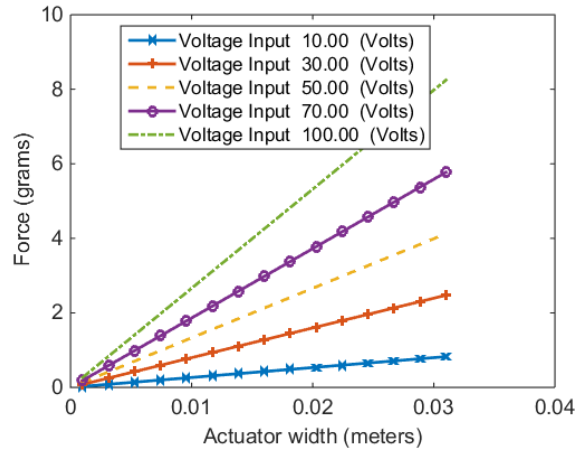


Figure 9: The variation of axial force with PVDF actuator width at different voltage inputs at e_{31} actuation mode

2.3. Electromechanical model of PVDF actuator at full actuation mode, Equation (3)

In this analysis, e_{31} , e_{32} and e_{33} actuation modes are all utilized. Basically, the effect of the e_{32} implies that a strain and a stress are produced in the lateral direction namely, the y -axis. This behavior is usually observed in plates. This behavior can be analytically modelled when the actuator is excited in e_{31} and e_{32} actuation mode. In this case, the normal stresses are parallel to the x and y axis (T_1, T_2), with the shear stress acting in the xy -plane (T_6).

The plate's bending deformation has no pairing with the shear deformation, and by assuming that the normal stress to the middle plane of the un-deformed plate remains straight and perpendicular to the mid plane after deformation, the stresses that are within the plane of any point through the thickness of the PVDF plate in a state plane stress can be written as:

$$\begin{bmatrix} T_1 \\ T_2 \\ T_6 \end{bmatrix} = \frac{Y}{1-\nu^2} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \frac{1-\nu}{2} \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \\ S_6 \end{bmatrix} \quad (28)$$

$$-\frac{Y}{1-\nu^2} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \frac{1-\nu}{2} \end{bmatrix} \begin{bmatrix} 0 & 0 & d_{31} \\ 0 & 0 & d_{32} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix}$$

The expressions for the strains in Equation (28) can be found in the literature. However, we use the Finite Element Technique to implement the full actuation matrix in Equation (3) using ANSYS. This is because implementing the full piezoelectric stress constants matrix $[e]$ and analytically deriving the analytical electromechanical model of the U-shape PVDF actuator complicates the model significantly. ANSYS provides a convenient way to obtain the results using the FEM.

In ANSYS, we are able to study the effect of the different piezoelectric constants at constant stress as discussed previously.

Figure 10 compares the maximum axial force generated by the PVDF actuator at two actuation modes; full matrix actuation and e_{31} modes. The maximum axial force that is obtained at full matrix

actuation mode is perfectly matches the force obtained in the analytical and FEM models in e_{31} actuation mode discussed previously. Therefore, the analytical model in e_{31} actuation mode is sufficient to depict the system.

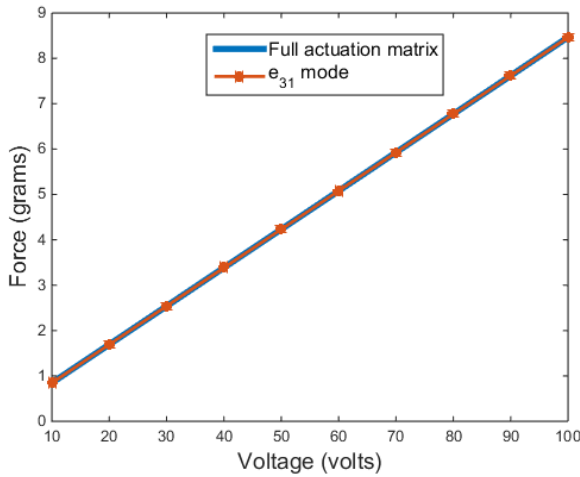


Figure 10: Axial force comparison by FEM at full matrix actuation and e_{31} actuation mode

It must be noted that, the axial force is applied to four teeth; two in the bottom and two on the top of the actuator as we will see that later in this section. Since it is a linear system, the force and the applied voltage have a linear relationship, as shown in Figure 10.

The material friction coefficient is defined as the measure of the sliding resistance of one material over another. In this case, the PVDF actuator’s contact part is a standard aligner which is usually made of industrial plastic, as shown in Figure 2.

To calculate the required force that is required to prevent sliding the actuators and aligner across each other, we need to calculate the required perpendicular/normal biting force between the mating sliding faces as the following:

$$F_n = \frac{f(l, t)}{\mu} \quad (29)$$

where, F_n is the normal force, μ is the co-efficient of friction and equals to 0.18 and $f(l, t)$ is the axial force by the PVDF actuators. Therefore, the required biting force should be at least equal or larger than 48 grams’ force.

2.4. Actual model of the PVDF actuator at full actuation mode, Equation (3)

In this section we discuss the actual U-shaped model of the PVDF actuator as illustrated in Figure 11.

Figure 11 is clearly annotated to indicate the targeted teeth, the direction of the axial force and the side wings of the U-shaped PVDF actuator.

A single actuator exerts forces on the upper and lower teeth that are in contact with the it. The forces generated by the actuator are equally spread among the teeth above and below the actuator.

The axial force is calculated by implementing the FEM principle with the help of commercial FEM software, namely

ANSYS. The results are then compared to the analytical model introduced previously in this paper for the U-shaped PVDF actuator.

Figure 12 shows the total force generated by the actuator to the front/back and upper/lower teeth in comparison to that obtained analytically. The deviation of the forces are due to the side wings of the actuator.

The single layer PVDF actuator with now side wings, is shown in Figure 13.

In Figure 14, the total force generated by the single layer actuator on the front/back upper and lower teeth is compared that obtained analytically for single layer actuator with no wings.

Figure 14 clearly reveals that the results are in good agreement. As a result, the conclusion can be made that the side wings of the actual model don’t have any effect on the total resulted axial force.

We can simulate the stress build-up at the fixed ends and the total deformation at the free ends of both; the actual U-shaped and single layer PVDF actuators as we can see in Figure 15.

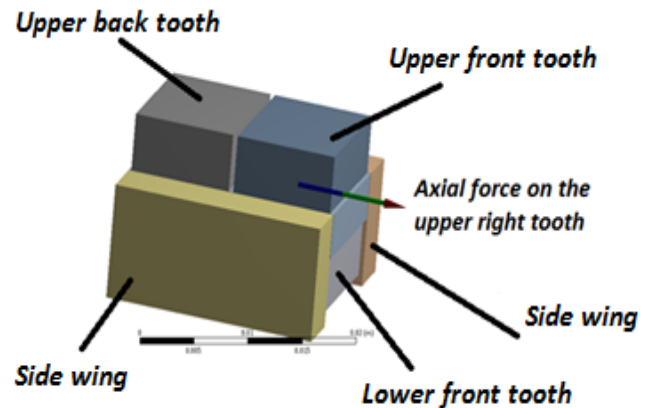


Figure 11: The actual U-shaped PVDF actuator

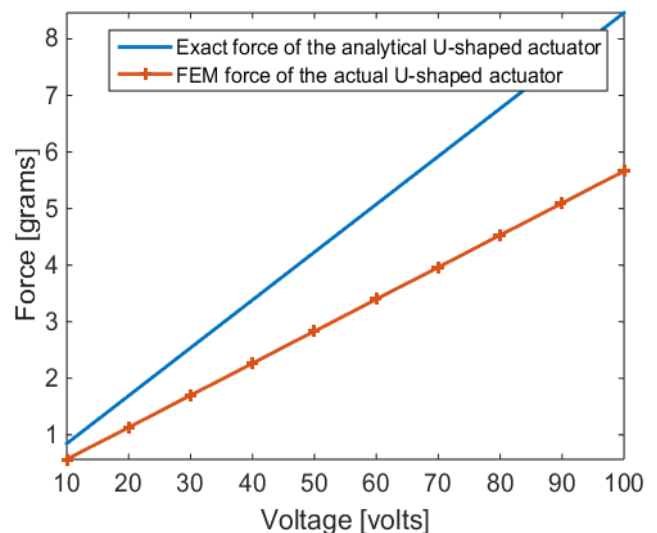


Figure 12: Total axial force comparison between actual and analytical U-shaped PVDF actuator

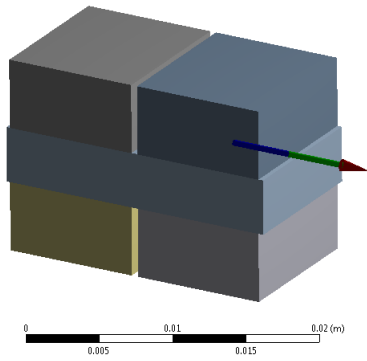


Figure 13: The actual single layer PVDF actuator

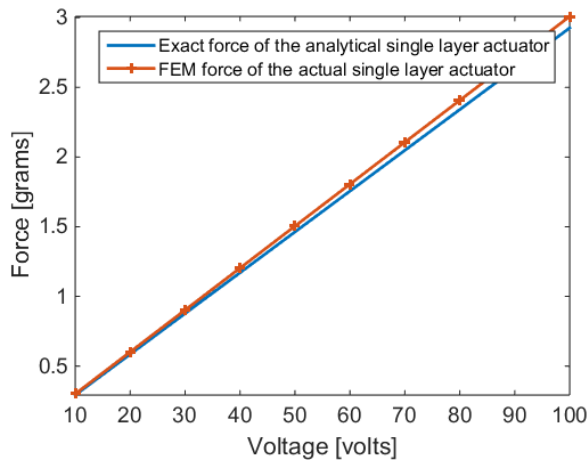


Figure 14: Total axial force comparison between actual and analytical single layer PVDF actuator

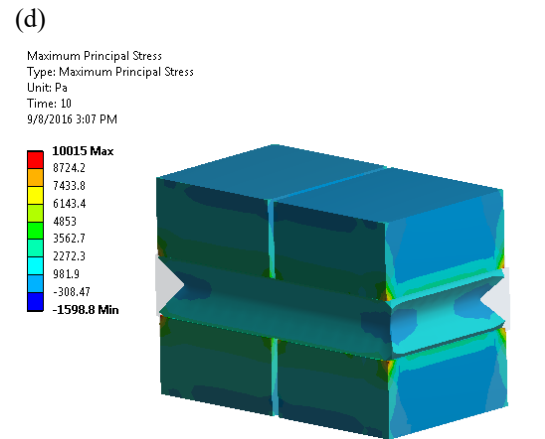
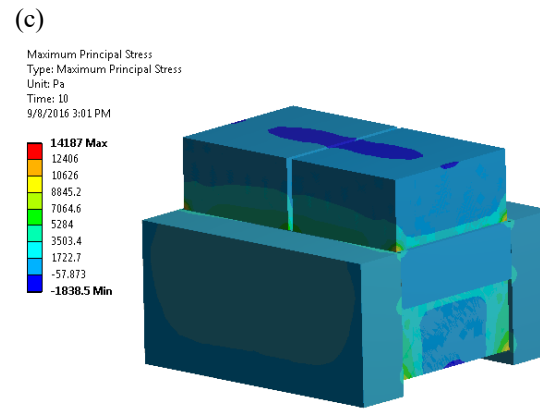
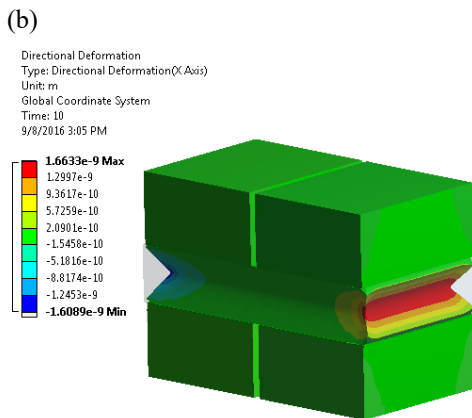
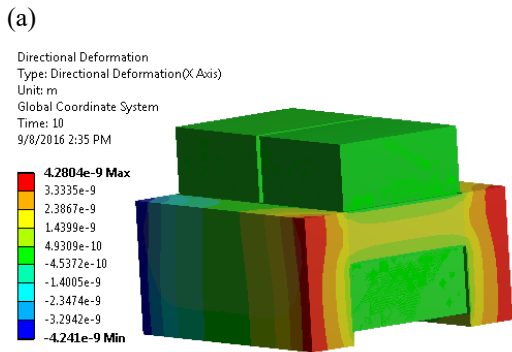


Figure 15: Total deformation in meters at the free ends of (a) U-shaped PVDF and (b) Single layer PVDF actuators. Normal stresses building up at fixed-fixed ends in Pascals (c) U-shaped PVDF and (d) Single layer PVDF actuators.



2.5. AcceleDent® versus the proposed PVDF vibrator

In 2006, Mao introduced a dental appliance which utilized pulsating forces to improve tooth movement rate, namely AcceleDent by OrthoAccel. The device is optimized in terms of force magnitude and frequency which enhances the comfort level and patient compliance to the treatment. The appliance is used twice a day for a duration of 10 minutes during the orthodontic treatment process and can be utilized along with a fixed appliance or an aligner treatment device. This device also provided an average of 23 to 25 grams of output force at 30 Hz frequency. Figure 16.a shows a one-time cycle of force output at 23 grams and at a low frequency of 30 Hz, while Figure 16.b displays an illustration of the AcceleDent device.

In Figure 16.b, the arrow shows the produced force by the extra-oral component of the vibrator. OrthoAccel made the claim that the resulting axial force of their device is equally dispersed among the upper and lower teeth, but that is not the case. In order to defend our claim, ANSYS was utilized to find the applied forces to the jaws produced by the AcceleDent device. The forces are then compared to the ones produced by the proposed PVDF piezoelectric actuator to each corresponding tooth based on the outcome and findings obtained in previous sections of this paper. Table 3 shows the material properties of the jaw bone, AcceleDent and the tooth.

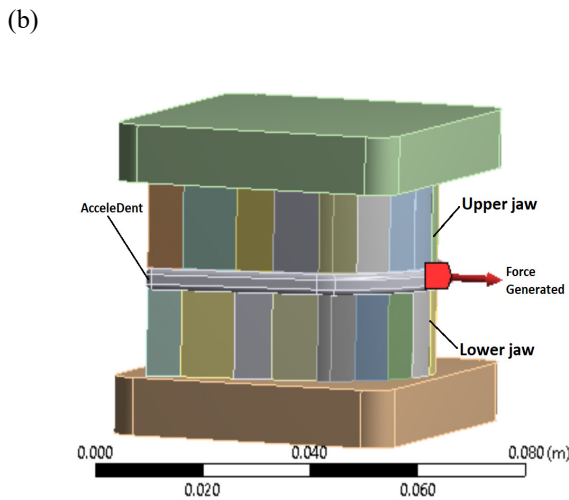
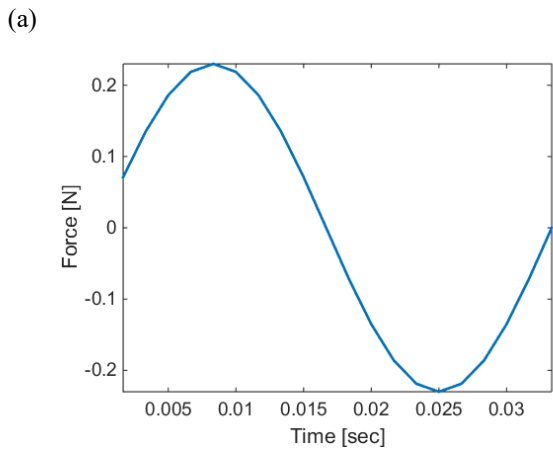


Figure 16: (a) One cycle of force by AcceleDent at 23 grams and 30 Hz (b) Simplified illustration of the AcceleDent with the lower and upper jaws assembly.

Table 3: Material properties of piezoelectric PVDF material

Component	Material	Density(kg/m ³)	Y (GPa)
AcceleDent	Versaflex™ CL2250 ~clear/transparent thermoplastic elastomer (TPE) grade1	888	0.0139
Jaw Bone	Cortical (surrounding), Cancellous (core)	1180	18.3
Teeth	Dentine	1200	84.1

¹Source: The University of Texas Health Science Center at San Antonio

It can be seen from Figure 17.a that the force created by the AcceleDent device is not evenly dispersed among all of the teeth. Most part of the generated force by the AcceleDent device was concentrated at the top front teeth of the jaws. While, as indicated Figure 17.b, the proposed PVDF device at a specific voltage magnitude and frequency, provides precise cyclic forces to a precise location of the tooth aligner, which then channels the forces to the targeted teeth. As a result, this allows flexibility of the tooth treatment for patients for the treatment process as the device can be adjusted.

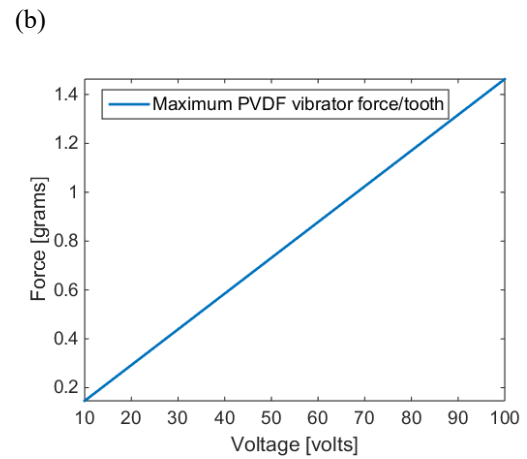
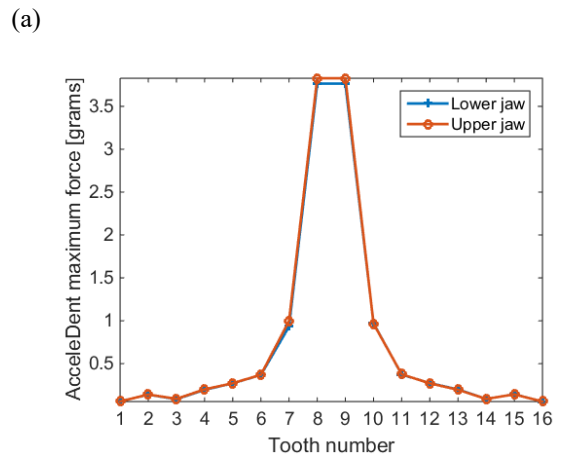


Figure 17: (a) Force distribution among all the teeth by AcceleDent (b) Force generated by the proposed single layer PVDF actuator at each targeted tooth versus voltage input.

3. Conclusion

In this paper, a novel dental vibratory apparatus was proposed in order to enhance the patient compliance to a vibration based orthodontic treatment as well as to minimize the time required for the treatment while simultaneously maintaining an acceptable level of comfort. The vibrating component of the device consisted of a bio-compatible smart material, namely (PVDF) piezoelectric actuators excited via voltage function generator at a frequency of 30 Hz over wide range amplitudes. The proposed device could be fitted to a tooth aligner and consequently provided recurring/cyclic forces to a particular part of the aligner, which then transmits these produced forces to the corresponding teeth.

The proposed model of the PVDF actuator was obtained both analytically and numerically with the help of ANSYS. The ultimate force obtained was 7.3 grams at the front and back teeth. This proposed device in particular is compact in size in comparison to its competition. The vibrating component of the device can be relocated and repositioned along the tooth aligner for customized fits, and it also compromises: the vibrating component and the intra-oral voltage function generator, battery, and processor are in a single device. This device is also predicted to minimize the drooling associated with the current orthodontic treatments. It ultimately enhances the overall experience for the

orthodontic treatment process and increases the chances of patient compliance.

Acknowledgment

Sincere thanks and appreciation to students Steven Lin, Raymond Huang and their effort of proofreading the paper. Special thanks to Olivia Licata for her effort in correcting related grammatical errors and providing the 3D image of a tooth aligner,

References

- [1] M. B.-H. a. M. A. Karami, "Piezoelectric Tooth Aligner for Accelerated Orthodontic Tooth Movement," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI, 2018.
- [2] A. M. Schwarz, "Tissue changes incidental to orthodontic tooth movement," *International Journal of Orthodontia, Oral Surgery and Radiography*, vol. 18, pp. 331-352, 1932.
- [3] D. F. Fink and R. J. Smith, "The duration of orthodontic treatment," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 102, pp. 45-51, 1992.
- [4] M. A. Fisher, R. M. Wenger, and M. G. Hans, "Pretreatment characteristics associated with orthodontic treatment duration," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 137, pp. 178-186, 2010.
- [5] A. M. Geiger, L. Gorelick, A. J. Gwinnett, and B. J. Benson, "Reducing white spot lesions in orthodontic populations with fluoride rinsing," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 101, pp. 403-407, 1992.
- [6] S. E. Bishara and A. W. Ostby, "White spot lesions: formation, prevention, and treatment," in *Seminars in Orthodontics*, 2008, pp. 174-182.
- [7] G. Segal, P. Schiffman, and O. Tuncay, "Meta analysis of the treatment-related factors of external apical root resorption," *Orthodontics & craniofacial research*, vol. 7, pp. 71-78, 2004.
- [8] N. Pandis, M. Nasika, A. Polychronopoulou, and T. Eliades, "External apical root resorption in patients treated with conventional and self-ligating brackets," *American journal of orthodontics and dentofacial orthopedics*, vol. 134, pp. 646-651, 2008.
- [9] A. Royko, Z. Denes, and G. Razouk, "[The relationship between the length of orthodontic treatment and patient compliance]," *Fogorvosi szemle*, vol. 92, pp. 79-86, 1999.
- [10] G. T. Sameshima and P. M. Sinclair, "Predicting and preventing root resorption: Part I. Diagnostic factors," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 119, pp. 505-510, 2001.
- [11] L. Linge and B. O. Linge, "Patient characteristics and treatment variables associated with apical root resorption during orthodontic treatment," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 99, pp. 35-43, 1991.
- [12] M. Harry and M. Sims, "Root resorption in bicuspid intrusion: a scanning electron microscope study," *The Angle orthodontist*, vol. 52, pp. 235-258, 1982.
- [13] A. H. Ketcham, "A preliminary report of an investigation of apical root resorption of permanent teeth," *International Journal of Orthodontia, Oral Surgery and Radiography*, vol. 13, pp. 97-127, 1927.
- [14] D. R. Cruz, E. K. Kohara, M. S. Ribeiro, and N. U. Wetter, "Effects of low-intensity laser therapy on the orthodontic movement velocity of human teeth: A preliminary study," *Lasers in surgery and medicine*, vol. 35, pp. 117-120, 2004.
- [15] M. Yamaguchi, M. Hayashi, S. Fujita, T. Yoshida, T. Utsunomiya, H. Yamamoto, et al., "Low-energy laser irradiation facilitates the velocity of tooth movement and the expressions of matrix metalloproteinase-9, cathepsin K, and alpha (v) beta (3) integrin in rats," *The European Journal of Orthodontics*, vol. 32, pp. 131-139, 2010.
- [16] R. Showkatbakhsh, A. Jamilian, and M. Showkatbakhsh, "The effect of pulsed electromagnetic fields on the acceleration of tooth movement," *World J Orthod*, vol. 11, pp. e52-e56, 2010.
- [17] D.-H. Kim, Y.-G. Park, and S.-G. Kang, "The effects of electrical current from a micro-electrical device on tooth movement," *Korean Journal of Orthodontics*, vol. 38, pp. 337-346, 2008.
- [18] A. H. Hassan, A. A. Al-Fraidid, and S. H. Al-Saeed, "Corticotomy-assisted orthodontic treatment: review," *The open dentistry journal*, vol. 4, 2010.
- [19] S. M. B. E.-D. Aboul, A. R. El-Beialy, K. M. F. El-Sayed, E. M. N. Selim, N. H. EL-Mangoury, and Y. A. Mostafa, "Miniscrew implant-supported maxillary canine retraction with and without corticotomy-facilitated orthodontics," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 139, pp. 252-259, 2011.
- [20] E. J. Liou and C. S. Huang, "Rapid canine retraction through distraction of the periodontal ligament," *American journal of orthodontics and dentofacial orthopedics*, vol. 114, pp. 372-382, 1998.
- [21] S. Sayin, A. O. Bengi, A. U. Gürton, and K. Ortaoğlu, "Rapid canine distalization using distraction of the periodontal ligament: a preliminary clinical validation of the original technique," *The Angle Orthodontist*, vol. 74, pp. 304-315, 2004.
- [22] H. İşeri, R. Kişnişci, N. Bzizi, and H. Tüz, "Rapid canine retraction and orthodontic treatment with dentoalveolar distraction osteogenesis," *American journal of orthodontics and dentofacial orthopedics*, vol. 127, pp. 533-541, 2005.
- [23] M. Nishimura, M. Chiba, T. Ohashi, M. Sato, Y. Shimizu, K. Igarashi, et al., "Periodontal tissue activation by vibration: intermittent stimulation by resonance vibration accelerates experimental tooth movement in rats," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 133, pp. 572-583, 2008.
- [24] I. Bautmans, E. Van Hees, J.-C. Lemper, and T. Mets, "The feasibility of whole body vibration in institutionalised elderly persons and its influence on muscle performance, balance and mobility: a randomised controlled trial [ISRCTN62535013]," *BMC geriatrics*, vol. 5, p. 1, 2005.
- [25] V. Gilsanz, T. A. Wren, M. Sanchez, F. Dorey, S. Judex, and C. Rubin, "Low-level, high-frequency mechanical signals enhance musculoskeletal development of young women with low BMD," *Journal of Bone and Mineral Research*, vol. 21, pp. 1464-1474, 2006.
- [26] N. Gusi, A. Raimundo, and A. Leal, "Low-frequency vibratory exercise reduces the risk of bone fracture more than walking: a randomized controlled trial," *BMC musculoskeletal disorders*, vol. 7, p. 1, 2006.
- [27] H. Omar, G. Shen, A. S. Jones, H. Zoellner, P. Petocz, and M. A. Darendeliler, "Effect of low magnitude and high frequency mechanical stimuli on defects healing in cranial bones," *Journal of Oral and Maxillofacial Surgery*, vol. 66, pp. 1104-1111, 2008.
- [28] A. E. Goodship, T. J. Lawes, and C. T. Rubin, "Low-magnitude high-frequency mechanical signals accelerate and augment endochondral bone repair: Preliminary evidence of efficacy," *Journal of orthopaedic research*, vol. 27, pp. 922-930, 2009.
- [29] C. Rubin, A. S. Turner, S. Bain, C. Mallinckrodt, and K. McLeod, "Anabolism: Low mechanical signals strengthen long bones," *Nature*, vol. 412, pp. 603-604, 2001.
- [30] J. Rubin, C. Rubin, and C. R. Jacobs, "Molecular pathways mediating mechanical signaling in bone," *Gene*, vol. 367, pp. 1-16, 2006.
- [31] C. Rubin, S. Judex, and Y.-X. Qin, "Low-level mechanical signals and their potential as a non-pharmacological intervention for osteoporosis," *Age and Ageing*, vol. 35, pp. ii32-ii36, 2006.
- [32] Z. Kalajzic, E. B. Peluso, A. Utreja, N. Dymant, J. Nihara, M. Xu, et al., "Effect of cyclical forces on the periodontal ligament and alveolar bone remodeling during orthodontic tooth movement," *The Angle Orthodontist*, vol. 84, pp. 297-303, 2013.
- [33] S. Yadav, T. Dobie, A. Assefnia, H. Gupta, Z. Kalajzic, and R. Nanda, "Effect of low-frequency mechanical vibration on orthodontic tooth movement," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 148, pp. 440-449, 2015.
- [34] M. C. Meikle, J. J. Reynolds, A. Sellers, and J. T. Dingle, "Rabbit cranial sutures in vitro: a new experimental model for studying the response of fibrous joints to mechanical stress," *Calcified Tissue International*, vol. 28, pp. 137-144, 1979.
- [35] C. H. Kurz, "Pulsating orthodontic appliance," ed: Google Patents, 1981.
- [36] C. H. Kurz, "Pulsating orthodontic appliance," ed: Google Patents, 1982.
- [37] C. H. Kurz, "Radio wave vibrational orthodontic appliance," ed: Google Patents, 1983.
- [38] W. G. Branford, "Vibratory dental mouthpiece," ed: Google Patents, 1991.
- [39] M. J. Powers, "Hand held device for reducing the discomfort associated with the adjusting of orthodontic appliances," ed: Google Patents, 1999.
- [40] J. C. Voudouris, "Powered orthodontic bracket," ed: Google Patents, 2003.
- [41] M. K. Lowe, "Vibrating dental devices," ed: Google Patents, 2015.
- [42] C. T. Rubin, W. B. Tarver, and M. K. Lowe, "Vibrating compressible dental plate for correcting malocclusion," ed: Google Patents, 2009.
- [43] J. J. Mao, "Use of cyclic forces to expedite remodeling of craniofacial bones," ed: Google Patents, 2006.
- [44] N. Mandall, C. Lowe, H. Worthington, J. Sandler, S. Derwent, M. Abdi-Oskoui, et al., "Which orthodontic archwire sequence? A randomized

- clinical trial," *The European Journal of Orthodontics*, vol. 28, pp. 561-566, 2006.
- [45] K. O'Brien, D. Lewis, W. Shaw, and E. Combe, "A clinical trial of aligning archwires," *The European Journal of Orthodontics*, vol. 12, pp. 380-384, 1990.
- [46] "Tooth positioner," ed: Google Patents, 1968.
- [47] C. H. Kurz, "Vibrational orthodontic appliance," ed: Google Patents, 1982.
- [48] M. K. Lowe, C. L. Wasden, and W. B. Tarver, "Systems, methods, and adjunctive procedures for correcting malocclusion," ed: Google Patents, 2007.
- [49] H. M. Layson Jr, "Body worn active and passive tracking device," ed: Google Patents, 2000.
- [50] M. G. Strugach and A. Szilagyi, "Piezoelectric vibrating device," ed: Google Patents, 1998.
- [51] S. H. An, "Flat Vibration Motor," ed: Google Patents, 2006.
- [52] M. K. Lowe, "Vibrating orthodontic remodelling device," ed: Google Patents, 2013.
- [53] X. Chen, S. Xu, N. Yao, and Y. Shi, "1.6 V nanogenerator for mechanical energy harvesting using PZT nanofibers," *Nano letters*, vol. 10, pp. 2133-2137, 2010.
- [54] A. Holmes-Siedle, P. Wilson, and A. Verrall, "PVdF: An electronically-active polymer for industry," *Materials & Design*, vol. 4, pp. 910-918, 1983.
- [55] T. Sato, H. Ishikawa, and O. Ikeda, "Multilayered deformable mirror using PVDF films," *Applied optics*, vol. 21, pp. 3664-3668, 1982.
- [56] J. Ha, H. O. Lim, and N. J. Jo, "Actuation behavior of CP actuator based on polypyrrole and PVDF," in *Advanced Materials Research*, 2007, pp. 363-366.
- [57] A. Meitzler, H. Tiersten, A. Warner, D. Berlincourt, G. Couqin, and F. Welsh III, "IEEE standard on piezoelectricity," ed: Society, 1988.

Generating a Blockchain Smart Contract Application Framework

Arif Furkan Mendi^{1,*}, Tolga Erol¹, Emre Şafak²

¹*Training and Simulation Technologies, HAVELSAN, 06510, Turkey*

²*R&D Technology, Product Management Department, HAVELSAN, 06510, Turkey*

ARTICLE INFO

Article history:

Received: 19 February, 2020

Accepted: 19 March, 2020

Online: 19 May, 2020

Keywords:

Blockchain

Smart Contract

Smart Contract Framework

ABSTRACT

Blockchain is a new generation technology that allows the central control mechanism or trusted authority to be removed, spreading the encrypted data across all participants in the network in a distributed database structure instead of central trust. The Smart Contract structure, which defines the rules and flow that allow the things we value to operate automatically as determined without the need for an external trigger mechanism, is the core element of this technology. Blockchain has gained popularity with its most famous application, Bitcoin. After Bitcoin became popular, it turned out that Blockchain might have new uses due to the advantage of technology such as security, brokerage, and transparency, and these areas are being investigated. Many big companies have started to invest in this technology in the face of the opportunities brought by Blockchain. HAVELSAN is a large-scale software company that studies and adapts new generation technologies. Blockchain technology has become of the new generation technologies that HAVELSAN is interested in due to its impressive advantages. HAVELSAN has a wide range of activities, so the company can develop various Blockchain-based applications depending on these areas. Combining this diversity with the importance of the Smart Contract concept, which can be considered as the basis for most Blockchain applications, it is decided to create a strong Smart Contract framework before starting to build different applications with Blockchain technology. The creation of HAVELSAN Blockchain Smart Contract Framework; which infrastructures are used during the development phase, the problems encountered during development and the structure of the most suitable applications to be created with the framework to be developed will be explained in this article.

1. Introduction

Blockchain is one of the trending technologies of today, but it was found in 1992. Although it was found at that time, its name began to be heard more frequently in 2009 with the use of technology as the basis of Bitcoin. Initially, the advantages of Blockchain were overshadowed by the popularity of Bitcoin, and what these technologies promised was not clearly understood. Bitcoin managed to place the concept of crypto money in the literature, and later many similar cryptocurrencies began to take its place in the market. The advantages of Blockchain, the technology that is the basis of cryptocurrencies, are much more than that. Blockchain is not only for financial transactions; it is a technology that can be programmed to record anything of value and importance, and where transactions are stored in a digital

distributed ledger that is intact. This value can be anything that can be expressed in code. Values such as supply chain management, food tracking, land registration can be easily managed this way. It is envisaged that concepts such as driver license, marriage certificate, birth certificate and smart property exchange can be managed easily and securely with Blockchain technology. Having such a wide range of uses makes Blockchain technology attractive. With its increasing popularity, Blockchain has started to be described as "New Internet". Just as the internet provides a structure for communication, it is argued that a similar structure for information sharing will be provided by Blockchain. The ground for this ambitious argument stems from the revolutionary attributes of Blockchain technology. Blockchain distributed ledger structure comes first among these revolutionary qualities. The distributed ledger is the structure in which all transactions performed on the system are recorded. A copy of this ledger is

*Arif Furkan Mendi, Ankara, Turkey, afmendi@havelsan.com.tr

shared with all parties in the system. All transactions on the network are recorded and stored in the nodes where all participants are located. With the distributed ledger structure, the security of the data is ensured within the system parties. Thus, the need for a third-party central authority such as bank and notary is eliminated. Even though it is a relatively new technology, due to these impressive advantages, the number of studies on it has increased significantly. Many large-scale companies are investing in Blockchain technology and conducting projects using this technology.

Blockchain technology allows clients and providers to operate securely with each other directly, without the need for approval by any third party. All transactions are stored in a distributed database using cryptography so that this exchange between client and provider can be done securely. To be able to make changes to the data in the system, that change has to be recorded in the majority of the parties in the system. To succeed in any chain of cyberattacks, it is necessary to verify over at least 50% of the computers, which makes the probability almost impossible.

Five principles are at the heart of Blockchain technology [1]:

- **Distributed Database Structure:** Each member of the network has access to data with the entire history. There are no single control points or central authority.
- **Peer to Peer Communication:** The communication takes place between the nodes directly, instead of being controlled from a central control point. Each party in the system stores the information and transmits it to other participants.
- **Pseudonymity:** Pseudonymity definition comes with Blockchain. Each transaction can be seen by the participants who have access to the system within the framework of the authority granted. In the Blockchain network, each user has a unique address that identifies it in the system. With this idea, users can choose to remain anonymous or prove their identity to others.
- **Irreversibility of Records:** In the system, each record is stored by associating it with the previous data. In case any record registered in the system is changed or a new record is added to the system, this link will be broken, so manipulation is detected. Thus, it is not possible to manipulate data on the system, which makes the system cyber-attack resistant. Various computational algorithms are used to create interconnected chains on the system. By using this calculation algorithm, the system is created in chronological order and avoids manipulation. To succeed in any chain of cyberattacks, it is necessary to verify over at least 50% of the computers, which makes the probability almost impossible.
- **Computational Logic:** With the automated structure of Blockchain, all operations performed are carried out based on calculation algorithm logic and programming. Therefore, with this established automated structure, users perform the transactions between them within the framework of pre-determined rules.

After evaluating all these advantages, many companies start to develop new applications on Blockchain. When we combine the reliability of Blockchain technology against cyber threats and the demands of clients and providers to make a secure exchange, Blockchain application areas are emerging. Applications are developed and still being developed in many different areas using Blockchain technology. Health, Digital Right Management, the Internet of Things (IoT) are the most popular of these areas. Up to date a lot of application has already been developed and it is predicted that the number of applications will increase in the coming period.

Gartner, which is one of the most important companies performing the current condition and market analysis of new generation technologies, has not been insensitive to Blockchain technology and has evaluated this technology within the scope of emerging technologies research. They published a report called "Emerging Technologies". In this report, "Hypecycle" curve, which shows the roadmap of emerging technologies, it could be seen that Blockchain looks to exceeded the "Peak of Inflated Expectations" and moving to "Trough of Disillusionment" phase (Fig. 1). It is also emphasized that it will take five to ten years for the technology to reach its efficiency. When we examine the hype curve, the fact that Blockchain technology is in a downward trend shows that it cannot exceed the expected threshold in less than 5 years. With the works to be carried out in the next 5-10 years, the capabilities of the technology will be fully revealed, and the applications developed using this technology will eliminate the shortcomings in the current situation and ensure that the technology will reach maturity.

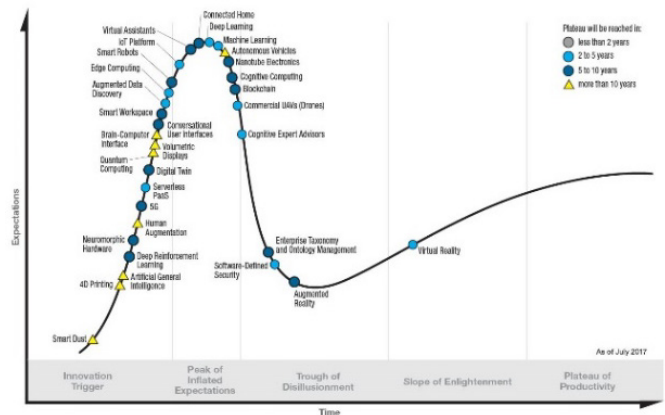


Figure 1: GARTNER Emerging Technologies Hypecycle [2]

Harvard has not been indifferent to the popularity and advantages of Blockchain technology and has conducted research. In the Harvard Business Review, they have defined Blockchain technology as a facilitator that enables contracts to be embedded in digital code. This digital contract environment enables data to be stored transparently with a shared database structure that protects data from tampering. This digitalization makes every deal, process, task and payment to have a digital record and signature that enables identification, authentication, storage and sharing. Hence, the mandatory need for a trusted third party will no longer be valid. Blockchain technology enables people, organizations and machines to interact individually with each other quickly, easily and securely without requiring central trust. It can be said that the

enormous potential of Blockchain comes from here [1]. Turkey is one of the emerging countries are investing in new generation technology. Especially, ensuring digitalization across the country is one of the main short and medium-term goals set by the government. In line with this goal, technology companies across the country are encouraged to invest in emerging technologies and contribute to digitalization to accelerate the country's development. HAVELSAN adopts this goal of the government regarding digitalization and operates as a large-scale software defense company that develops and applies emerging technologies in its field of activity. After evaluating that Blockchain technology is a new generation technology with high potential that cannot be ignored and evaluating that we can develop projects and products suitable for our fields of activity, we determined to create HAVELSAN Smart Contract framework. With this framework to be developed, we aimed to develop Blockchain-based applications suitable for the company's fields of activity. The creation of HAVELSAN Blockchain Smart Contract Framework; which infrastructures are used during the development phase, the problems encountered during development and the structure of the most suitable applications to be developed with the framework to be created will be explained in this article.

2. Work Done

2.1. Infrastructure Selection

Concepts such as Smart Contract, Smart Property are the terms where Blockchain technology is used together, which indicates the value ownership in the digital environment. Smart contracts define the rules and flow that enable the cases that we attach value to operate automatically as determined without the need for an external trigger mechanism. The Smart Contract structure is a key concept to achieve all the advantages offered by Blockchain technology. Smart Contracts not only define the rules and penalties around a contract like a traditional contract does, but also automatically enforce those obligations. The flow is carried out completely automatically and no manual intervention is possible. This means that there is no need for any authority to manage the flow, that is, to run the flow with the consensus between the parties. With this extraordinary structure, Blockchain technology stands out and becomes the reason for preference. When we look at the idea of smart contracts, we see that it was proposed by Nick Szabo in 1997 [3]. Nick Szabo argues that through smart contracts, a structure in which digital contracts can be triggered can be established with the structure to be set up digitally. Although the idea of the Smart Contract was put forward by Nick Szabo, we see that the idea of smart property was proposed by Mike Hearn in 2012 [4]. Mike Hearn proposed the idea of smart property, thereby arguing that the property status and the transfer process can be monitored. Blockchain applications are developed using smart contracts infrastructure and it is not possible to develop Blockchain applications in a different structure. With the smart contract structure to be designed, it is possible to develop products in a wide range of desired areas. Therefore, it is wrong to think of the smart contract structure as a single application. The smart contract structure is a framework in which applications can be developed on. Based on this point; after evaluating that HAVELSAN has a wide area of activity, we decided to develop a smart contract framework that enables the development of applications in different areas instead of developing a single Blockchain-based

application. In order to create such a frame structure, it is important to construct the structure carefully from the beginning. It is the first option to construct the services and infrastructure to be offered from scratch. However, there are many infrastructure providers currently offering smart contract infrastructure and allowing us to create our framework. These organizations have many experiences in providing such services. We can obtain similar experience and knowledge by working, but it does not seem logical when considering the need for rapid action in the age of information sharing. Instead of rediscovering the same things again, taking such services from the service providers who have experienced and present mature solutions and establishing our framework on this will contribute to creating a more effective and strong structure. The starting point of the study; when we consider the purpose of creating a strong framework in which HAVELSAN can develop applications with different qualities suitable for its fields of activity, it seems reasonable to receive service from such a service infrastructure provider. For this reason, as the first step of our road map, basic services will be obtained using open source infrastructure providers, and we will create our Smart Contract Infrastructure. When we examine who are the open-source infrastructure providers, Ethereum, Hyperledger, Corda, which are the most powerful organizations in this field, appear.

Ethereum is one of the most popular among smart contract providers. Ethereum was announced by Vitalik Buterin in 2015 and is a smart contract infrastructure platform where many cryptocurrencies currently operating in the market are developed. Apart from cryptocurrencies, it offers its users a structure to create a smart contract protocol and provides the opportunity to develop different applications using this infrastructure [5]. Ethereum allows its users to develop their smart contract-based applications using Solidity language.

R3 Corda is another smart contract infrastructure service provider. Similar to Ethereum, Corda offers smart contract infrastructure for the development of Blockchain applications. Corda is a Blockchain platform developed by R3 and its first version was released in 2016. R3 has entered the Blockchain world with a different product by working with more than two hundred members and partners in the development process, revealing the permissioned Blockchain architecture for financial services, insurance and the business world. With its smart contract infrastructure, R3 Corda allows its users to develop their smart contract-based applications using Java and Kotlin languages. With the infrastructure provided by R3 Corda, users can establish their permissioned type Blockchain networks [6].

Thirdly, Hyperledger is another alternative considered. Hyperledger was founded by the Linux Foundation in 2016, adopting the principle of developing an open-source infrastructure and setting standards that could not be achieved for Blockchain-based application development. Within this organization, it is an organization with more than 50 members in total, with many companies from the financial sector, as well as important companies in the world of technology such as IBM and Fujitsu. With the basic principle of establishing a standard for distributed databases, they aim to develop Blockchain technology by creating an open standard platform across industries and institutions for distributed databases. It is advocated that with this platform, the way of developing applications in different fields can be opened,

thus, a new setup will emerge that will include different actors on a global basis. Unlike its competitors, which currently offer Blockchain-based smart contract infrastructure, the Hyperledger Project has targeted areas of application outside of cryptocurrencies since its first appearance. It is seen that Hyperledger produces a wide range of solutions for different needs for incubation at all development stages for enterprise-class Blockchain [7] (Figure 2). Fabric is one of the most popular products offered by Hyperledger. It has been the reason for the preference of users especially with its modular structure and flexible architecture. Fabric offers its users a smart contract infrastructure suitable for various usage scenarios. It also offers the flexibility to apply the consensus model suitable for the applications to be developed. Hyperledger Fabric is the first Blockchain network created in a specific standard using programming languages, without any cryptocurrency dependency. Besides, having a structure that supports digital identity enables the development of applications in a wide range of products. Currently, the Proof of Work (PoW) consensus algorithm used in many Blockchain applications, such as Bitcoin and other cryptocurrencies, does not have to be used, so it offers important solutions to overcome problems such as resource consumption and performance [8].



Figure 2: Hyperledger Product Tree [9]

When the infrastructure that can be used in building our structure in Blockchain technology is examined, these three organizations stand out seriously. The truth is different, although it is thought that all three alternatives are provided with a similar quality of service, that is, it allows us to define smart contract rule protocol. There are differences such as the type of service provided, usage area. When we evaluate that we will establish our smart contract infrastructure; it is of great importance for us to determine the need criteria correctly and to create the structure healthily.

As the first criterion, the usage area of the structure to be created must be determined correctly. Will the applications be developed in the financial field or will the financial applications be excluded from our field? The answer to this question is important for determining the smart contract service provider. Corda can be considered as the optimum solution if it is desired to create a Blockchain framework for financial applications. However, for our case, choosing Corda does not make sense because we do not intend to develop any financial applications.

After determining the application area, how the nodes are included in the system to be installed and their authorization is another criterion. As is known, basically 2 types of Blockchain networks can be established; public and permissioned. Public Blockchain network is the network where the nodes in the system

can freely join the system as a participant or authenticator without going through any control mechanism. Bitcoin and other cryptocurrencies generally aim to set up such a network and ensure that the participants are freely included in the system. On the other hand, there is a need for the parties to join the system to participate in the network as a result of having a certain check under certain criteria. Private Blockchain networks are being established for those in this situation. In order to join the system in the private Blockchain network and be able to read the data on the network, certain authorization must be made [10]. At this point, the solution offered by service providers differ. When we compare Ethereum and Hyperledger solutions, it is seen that there is a difference in terms of public and private networking opportunities. Ethereum provides its users with an infrastructure to establish a public Blockchain, while Hyperledger allows users to set up a private Blockchain according to their needs. In addition, we can say that Hyperledger offers its users a wider choice in terms of setting up the network.

When our needs are evaluated; since we aim to develop applications with specific needs for specific customers, private Blockchain and offering more flexible options seem more appropriate to meet our needs. Therefore, it appears that Hyperledger fully meets these criteria. As a result, the Hyperledger was chosen as the infrastructure to be used when developing our Blockchain application framework, and the study was carried out using this infrastructure. The screenshot of the login page of the developed Blockchain Smart Contract Application Framework is given in Figure 3.



Figure 3: Login Page of Developed Blockchain Smart Contract Application Framework

2.2. Problems

After choosing the Hyperledger infrastructure for the framework application to be developed, it was necessary to choose the most suitable tools. In the next step after deciding to use Hyperledger, various Hyperledger solution alternatives have been evaluated to find the appropriate solution for which tool to install the network, the distribution of the network and the installation of the smart contract structure. Different solutions such as Hyperledger Fabric, Sawtooth, Burrow and Indy are available in the marketplace [11]. Hyperledger Fabric is the most preferred and most popular among these options. Hyperledger Fabric offers an infrastructure that does not have a cryptocurrency dependency and enables the development of enterprise-level Blockchain applications. Due to the factors such as enabling the use of different consensus methods, enabling the operations to be shown

in a grouped manner safely thanks to its channel structure, its modular structure and ease of use, Fabric overrides other options. After considering all these advantages, it was decided to use it within the smart contract framework project. A sub-module is also needed for the distribution and display of the network formed with fabric. At this point, the use of Cello and Composer, among the existing solutions, are the highlights of the options. Hyperledger Cello offers an advanced interface for networking and deployment to meet this need. Using Cello, it is possible to securely distribute and display the Blockchain network. It allows us to create a private and highly secure Blockchain network that allows us to instantly see changes in the network regardless of the size of the network. For all these reasons, it was decided to use Hyperledger Fabric & Cello. After this decision, some difficulties had to be overcome to get a useful framework.

Lack of resources is the first problem encountered, since Hyperledger is a new organization established in 2016. The number of projects developed and used by using Hyperledger infrastructure is limited, and there are not enough projects completed to transfer the experiences gained to new projects. The development of the Hyperledger project is still ongoing. The organization has released many versions. As a result of the experiences and feedbacks, the release of advanced versions continues periodically. User manuals for Hyperledger services are available to everyone over the internet. However, since the history of the services provided is not very old, the number of studies carried out using Hyperledger in the literature is very limited. When a problem is encountered during the development phase, research is done on the internet, and the problem encountered is solved quickly by using resources that encounter similar problems and produce solutions. Since there was a limited number of Hyperledger applications in the literature, it causes difficulties in quickly overcoming the problems encountered during the application development phase. During the development process, it was not possible to solve some problems with the existing documents, it was possible to find solutions to the problems by examining the source code and performing various trials. This process took more time than expected and caused high effort.

The second important problem encountered concerns the display of blocks to users. We can say that we had difficulty to show detailed information about the newly added block. Data on the network created by Hyperledger Fabric is kept in a closed box structure with high security. In this structure, it is possible to access the data on the blocks only with the commands provided by Fabric. Due to this restriction, limitations are encountered in accessing block data through the developed application layer. This led to insufficient flexibility during development. If we cannot access detailed information about the blocks, it will cause us to have problems at the point of the display to the end-users. If the user cannot have enough information about the blocks, he/she will have problems controlling the system and it will be difficult to adopt the system. If the data held on the blocks remain hidden, this will create a mystery. As a result of this mystery, users may tend to avoid the system due to uncertainty. For users to be familiar with the system, they must have sufficient access to the blocks and the information they contain. Otherwise, we would have missed the transparency advantage offered by Blockchain, which causes the system's success to be questioned. In order to avoid this uncertainty, we had to research how detailed information about the

blocks could be shown. We wrote different functions from the display layer of the data for detailed access to the data on the layer on which it was kept. With the prepared methods, queries were prepared, in which we can access the details of the block-related address and the data kept, such as Hash Number, Previous Hash and Transaction ID.

The third problem encountered is the problems experienced at the point of integration. Thanks to the high security smart contract infrastructure offered by Hyperledger Fabric, there is no shortage in developing a closed box smart contract infrastructure. Fabric and Cello tools used for network distribution and network representation enable us to obtain a closed box system. It helps to protect the system at a high level and to make transactions safer and without intermediaries. However, when it is desired to get an application-level infrastructure with all these advantages, it is observed that this planned structure is limited in terms of providing integration points. We need a different solution to overcome this problem. At this point, it has been observed that Hyperledger Composer offers a more flexible solution for the establishment of integration points and communication in high integration applications. Hyperledger Fabric offers a flexible solution for the lower-level design and architectural installation. Hyperledger composer offers an additional layer over the structure offered by Hyperledger Fabric. With this created layer, integration and application development are carried out by Composer.

Hyperledger Composer offers its users an interface that facilitates and speeds up the development phase. The components offered through this interface provide a simpler understanding of the development. Hyperledger Composer, written in JavaScript, paves the way for integration with external systems thanks to the possibility of creating the Rest API. Blockchain applications developed are easily integrated with external systems using these Rest APIs. Since it has emerged as a new solution, the Hyperledger Composer solution was introduced while we were in development, so it was not possible to plan it at the beginning of development. As a result of all these experiences, it can be said that it is more appropriate to use Hyperledger Composer in high integration applications. The problem we face is solved by transferring operations in the work to Hyperledger Composer using smart contracts based on Hyperledger Fabric, especially at the point of integration. As a result of all these experiences, it has been seen that it is appropriate to use Hyperledger Composer tool for applications where intervention is required and applications that do not require independence from external systems. Composer selection will be reasonable especially in systems that require integration due to the possibility of communicating with Rest APIs. Although we did not use Composer during the system development period, we managed to solve the problem by moving the existing structure to Composer after this problem we encountered during the development period.

2.3. Structure of the Applications to be Developed with the Framework to be Created

Blockchain is a new generation technology that has gained popularity in recent years due to its advantages and the sensation created by cryptocurrencies such as Bitcoin in the financial field. With this popularity, we see that many companies have started to show interest in this technology, they have moved their existing

systems to Blockchain technology or developed their new projects based on Blockchain. Unfortunately, it is even seen that some companies force themselves to use Blockchain technology because of this popularity. This creates a very unhealthy situation. Gartner predicted that 90% of the corporate projects started in 2015 using Blockchain technology will fail in 18 to 24 months [12]. The fact that such uncontrolled elections are held with the effect of popularity returns to companies as a waste of time and money. Based on this, it is of great importance to determine the use of the smart contract framework work we have done in the right areas and to prevent such waste of resources. When deciding on application scenarios, HAVELSAN's field of activity and project needs should be taken into consideration, and the benefit/loss analysis should be done correctly.

Before deciding on Blockchain applications, determining the capabilities of the infrastructure on which the application will be developed is an important factor affecting the success of the application. If a closed box Blockchain application that does not need to be integrated with external systems will be developed, the use of Hyperledger Fabric-Cello will be the right choice for ensuring high security. However, if an application to be integrated with external systems is to be developed, the use of Hyperledger Fabric-Cello will not be suitable. As Composer enables and facilitates integration through Rest APIs, the issue of integration with the use of Fabric-Composer will not be a problem anymore. Choosing the appropriate infrastructure is important for developing successful applications. The experiences we have obtained from the problems we encountered during the development phase of the framework support this claim. Therefore, this criterion should be taken into account when determining the framework. In addition, different solution proposals of Hyperledger and Ethereum are tried as proof of concept work to enhance the knowledge about this topic. The results are obtained at the end of all these works.

Correct addressing the two issues described above is a prerequisite for determining the appropriate infrastructure. Therefore, the most appropriate applications that match the field of activity and strategy of HAVELSAN were decided first, and then the infrastructure type and services to be used in the line with these choices were determined. While developing such a Blockchain application framework, it was observed that a high-security, flexible, closed-box system is needed, in which we can adjust the workflow according to customer needs. The products that will meet these needs best are Hyperledger Fabric and Cello. Therefore, it was considered that this would be the most feasible use of this framework.

3. Conclusion

Blockchain is one of the most popular emerging technologies that HAVELSAN is interested in due to its impressive advantages. HAVELSAN has a wide area of activity, so we decided to develop a smart contract framework that enables the development of applications in different areas instead of developing a single Blockchain-based application. Before moving to the development phase of such a Blockchain-based smart contract framework, the infrastructure used to create our framework had to be determined first. There is an open code or paid services from which the infrastructure service to be used can be obtained. Open source is

one of the most preferred cost-effective methods in today's technology age. Open-source service providers are made up of wider audiences. With this broad participation, the service is provided faster and thus, faster solutions to the problems are provided. In addition, applications can be developed with open participation without the need to pay any fees. After evaluating these advantages, we decided to choose an open-source provider. Organizations that are providing such services have been examined. It was decided to use Hyperledger after evaluating the factors such as allowing application development in areas other than cryptocurrencies, having private network structure, and configuration flexibility. As Blockchain is a relatively new technology, we encountered many different problems. Since the number of completed Blockchain projects is limited, the number of articles and other resources in the literature is also very low, making it difficult to quickly find solutions to problems encountered during the development phase. It could be predicted that this problem will be eliminated as the number of completed projects in literature is increased. Another problem encountered is; concerns the display of blocks to users. We can say that we had a problem showing detailed information about the newly added block. Data on the network created by Hyperledger Fabric is kept in a closed box structure with high security. In this structure, it is possible to access the data on the blocks only with the commands provided by Fabric. Due to this restriction, limitations are encountered in accessing block data through the developed application layer. To avoid this uncertainty, we had to research how detailed information about the blocks could be shown. We wrote different functions from the display layer of the data for detailed access to the data on the layer on which it was kept. Maybe the most important problem that has encountered during the development of the application is about integration abilities. Although it is possible to create high security, flexible, modular closed box Blockchain network suitable for customer needs by using Hyperledger Fabric and Cello tools, there are shortcomings in terms of integration with external systems. If there is a need for such an external system integration, the use of Hyperledger Fabric-Composer would more suitable. It makes it easier to integrate with Rest API with other systems. Besides all the problems encountered, it is to decide which type of framework to build as a result of our studies on another subject evaluated. If there is no need to integrate the applications to be developed with external systems, the use of Hyperledger Fabric-Cello may be preferred, but if there is a need for integration with external systems, the use of Hyperledger Fabric-Composer needs to be determined. A serious effort was made to overcome all these problems, different proof of concept work was completed. However, important experiences were obtained at the end. All these were obtained while generating HAVELSAN Smart Contract Framework. In the coming period, it is planned to develop different applications using this developed smart contract infrastructure. Thus, the accumulated knowledge can be further increased.

4. References

- [1] C. Mooney, "the Truth About ," *Sci. Am.*, no. August, pp. 80-85, 2011.
- [2] Kasey Panetta, "Top Trends in the Gartner Hype Cycle for Emerging Technologies, 2017 - Smarter With Gartner," 2017. [Online]. Available: <https://www.gartner.com/smarterwithgartner/top-trends-in-the-gartner-hype-cycle-for-emerging-technologies-2017/>. [Accessed: 02-Mar-2020].
- [3] N. Szabo, "The Idea of Smart Contracts," 2018. [Online]. Available: <http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Lite>

- rature/LOTwinterschool2006/szabo.best.vwh.net/idea.html. [Accessed: 31-Mar-2018].
- [4] Mike Hearn, "Smart Property - Bitcoin Wiki," 2018. [Online]. Available: https://en.bitcoin.it/wiki/Smart_Property. [Accessed: 31-Mar-2018].
 - [5] G. W. Founder and E. Gavin, "Ethereum: A Secure Decentralised Generalised Transaction Ledger," pp. 1–32, 2017.
 - [6] R. G. Brown, J. Carlyle, I. Grigg, and M. Hearn, "Corda: An Introduction," pp. 1–15, 2016.
 - [7] C. Cachin, "Architecture of the hyperledger blockchain fabric," Pdfs.Semanticscholar.Org, 2016.
 - [8] C. Cachin et al., "Hyperledger fabric," pp. 1–15, 2018.
 - [9] T. Blummer et al., "An Introduction to Hyperledger," p. 33, 2018.
 - [10] K. Wust and A. Gervais, "Do you need a blockchain?," Proc. - 2018 Crypto Val. Conf. Blockchain Technol. CVCBT 2018, no. i, pp. 45–54, 2018.
 - [11] Hyperledger, "Hyperledger Projects - Hyperledger," 2018. [Online]. Available: <https://www.hyperledger.org/projects>. [Accessed: 18-Jun-2018].
 - [12] GARTNER, "Top 10 Mistakes in Enterprise Blockchain Projects - Smarter With Gartner." [Online]. Available: <https://www.gartner.com/smarterwithgartner/top-10-mistakes-in-enterprise-blockchain-projects/>. [Accessed: 08-Mar-2019].

Factors Influencing Social Knowledge Management in Social Society: A Systematic Literature Review

Erick Fernando^{1,*}, Meyliana¹, Achmad Nizar Hidayanto², Harjanto Prabowo³

¹School of Information Systems, Information Systems Department, Bina Nusantara University, 11480, Indonesia

²Faculty of Computer Science, Universitas Indonesia, 16424, Indonesia

³BINUS Graduate Program - Doctor of Computer Science, Computer Science Department, Bina Nusantara University, 11480, Indonesia

ARTICLE INFO

Article history:

Received: 15 April, 2020

Accepted: 21 April, 2020

Online: 19 May, 2020

Keywords:

Social society

Knowledge Management

Social Knowledge Management

Systematic literature review

ABSTRACT

Knowledge is important now for the development of social society; it is necessary for knowledge management. Knowledge management (KM) aims to support the creation, transfer, and application of knowledge in social societies. This fact illustrates that in the management of social knowledge, the role of social communities is very important and is influenced by factors in the process. With this, this study will look for theories and factors that influence KM social interactions that occur in social societies. The method used is a systematic literature review. The results of this study found theories and factors that influence KM in social societies.

1. Introduction

In the development of organizations and social societies, the need for knowledge becomes the most important [1][2][3]. The need is seen in transactions of knowledge through social ties and organizations that will provide developments within the organization or social community. The importance of acquiring knowledge and combinations as a source of value creation and competitive advantage [1]. Thus knowledge needs to be better managed [4][5]. Knowledge management (KM) aims to support the creation, transfer, and application of knowledge in organizations and social communities [3][6]. Social knowledge management is a management based on KM that involves social relationships [7]. This KM In management, there are two main aspects: one that refers to the management of general and other knowledge, which is particularly included in its social character.

A social character can be seen in social network relationships and knowledge management, to explain how it is acquired, transferred, exchanged, and generating knowledge [5] about the basic social processes and learning of the organization; Knowledge science and technology management, which aims to promote research and development and use of ICT; A holistic knowledge management model, including models not seen in previous models and integrating new subdisciplines [7][8].

Knowledge management that occurs in the social community influenced interaction relationships that occur therein. The interaction that occurs will have the factors that affect an individual to interact. With this in the study wanted to find these factors. This study was conducted by systematic review literature. Therefore, this study tries to define "Theory dan factor influencing social knowledge management of social society?". This research will focus on looking for factors that influence social character within each social networking relationship for knowledge management. Where knowledge management begins making, transferring, making, and producing the knowledge gained in social society.

2. Theoretical foundations

2.1. Knowledge Management

KM is a collection of tools, techniques, and strategies to maintain, analyze, organize, enhance, and share insights and experiences that justify the belief that knowledge is an asset to enhance the capacity of the organization to be able to work more effectively[1][9].

The existence knows knowledge management of Tacit Knowledge and Explicit Knowledge[10], [11]. Tacit Knowledge means the science or experience one gets through a daily activity in doing a field of work. Tacit knowledge will simply disappear if the person concerned does not share his / her knowledge with

*Erick Fernando, Email: erick.fernando001@binus.ac.id

www.astesj.com

<https://dx.doi.org/10.25046/aj050326>

others (transfer knowledge) or is not well documented in hard/soft copy[11]. Explicit Knowledge is more to how the science is well documented, so it can be stored ideally and does not just disappear[11]. Knowledge management model, Nonaka, and Takeuchi are renowned for their SECI (Socialization, Externalization, Combination, and Internalization) methods [11].

2.2. Social Community

The community is a small or large social unit that has similarities that are described as norms, religions, values, or identities. This community is located in a specific geographic region or in cyberspace through a platform. Sometimes the community can be defined as a social bond. It is just as important as their identities, practices, and roles in social institutions such as family, home, work, government, community, or humanity [12].

2.3. Social Knowledge Management

Social Knowledge Management is production management and dissemination of knowledge, research, and socially promoted epistemological models of the class[13]. Knowledge development in society involves the participation of social actors themselves and Trans Discipline (which is not equivalent to interdisciplinarity). The knowledge formed in this environment is the responsibility of the social individual to disseminate that has been facilitated in the exercise of critical thinking (social accessibility of knowledge) and has adequate social outreach for community development [13].

3. Methodology

This study uses a systematic literature review (SLR) approach. SLR is used to identify, evaluate, and interpret all relevant research from research questions, phenomena, and topic areas [14]. SLR is carried out with several processes, including search process, inclusion criteria and exclusion, data extraction, and analysis of findings to answer research questions.

3.1. Search process

The first systematic literature review (SLR) process was carried out by searching for relevant articles and find with research in a reliable, reputable, and up-to-date journal database. The database includes:

1. ACM Digital Library (dl.acm.org)
2. IEEE Xplore Digital Library (http://ieeexplore.ieee.org)
3. JStor (www.jstor.com)
4. Science Direct (www.sciencedirect.com)
5. Emerald Insight (www.emeraldinsight.com)
6. Springer Link (link.springer.com)
7. Taylor and Francis (tandfonline.com)

Search process in this research using the Boolean operator. This search can filter search data better, so search gets priority data. the Boolean operators used are AND and OR. the composition used by keyword is as follows:

1. ('Social' AND 'knowledge 'AND 'Management') OR ('Theory')

2. ('Knowledge' AND 'Management' OR ' Social')
3. ('Knowledge' AND 'Management' OR ' Social Theory')
4. ('Knowledge' AND 'Management' OR ' Social Society')

3.2. Inclusion and exclusion criteria

The search inclusion process has three criteria, namely (1) Founded Study is the process of finding documents based on keywords, (2) Candidate Study is the process of selecting documents based on titles and abstracts that are relevant to the research objectives, (3) Selected Study is the process of filtering documents by reading carefully consider all that is used to answer research questions.

The exclusion criteria process has provisions including determining the time of publication of a paper used so that in this study using a time period before 2001, a complete article identity structure (title, author, journal name, etc.), ensuring the article used is not duplicate and The built SLR can answer the questions in the research consistently.

3.3. Data extraction

The literature search process began in November 2017 by finding 278 article documents from the database, and the search criteria that have been determined are founded studies. The second process found 152 article documents from the screening process of relevant documents based on titles and abstracts as candidate studies. the final process found 39 article documents which through the process of careful reading of the contents of the document as candidate studies and used to answer research questions

Table 1. Number studies in selected sources

Source	founded Studies	Candidate studies	Selected studies
Emerald	60	45	3
IEEE	40	16	6
Jstor	30	17	3
ACM	25	10	3
Science Direct	35	15	5
Taylor and Francis	48	31	9
springer	40	18	10
	278	152	39

4. Results and discussions

4.1. Demographic and trend characteristics

4.1.1. Publishing outlets

From the search process, there is a great deal of research in various conferences and journals. This study found among others: journal Knowledge Management Research & Practice (#5), Hawaii International Conference on System Sciences (#2), International Journal of Information Management (#2), Total Quality Management & Business Excellence (#2), and the other

amounted to 1. The total number of 39 from conference and journal, can be seen in the following table 2

Table 2. Source of publications

Journal/ conference	Journal/conference name	#	%
Journal	Knowledge Management Research & Practice	5	12.82
conference	2011 44th Hawaii International Conference on System Sciences	2	5.13
Journal	International Journal of Information Management	2	5.13
Journal	Total Quality Management & Business Excellence	2	5.13
conference	2010 International Conference on E-Product E-Service and E-Entertainment, ICEEE2010	1	2.56
conference	2014 IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA)	1	2.56
conference	Adaptation and Value Creating Collaborative Networks: 12th IFIP WG 5.5 Working Conference on Virtual Enterprises, PRO-VE 2011	1	2.56
Journal	Advances in Intelligent and Soft Computing	1	2.56
Journal	Computational and Mathematical Organization Theory	1	2.56
Journal	Development in Practice	1	2.56
Journal	Healthcare Knowledge Management SE - 8	1	2.56
Journal	Housing, Care and Support	1	2.56
conference	IFIP Advances in Information and Communication Technology	1	2.56
Journal	Information & Management	1	2.56
Journal	Information Processing and Management	1	2.56
conference	International Conference on Context-Aware Systems and Applications	1	2.56
conference	International Conference on Knowledge Management in Organizations	1	2.56
conference	International Symposium IUKM 2013	1	2.56
Journal	Journal of Business Ethics	1	2.56
Journal	Journal of Information Technology	1	2.56
Journal	Journal of Integrated Care	1	2.56
Journal	Journal of International Business Studies	1	2.56
Journal	Journal of Knowledge Management	1	2.56
conference	Procedia Computer Science	1	2.56

Journal/ conference	Journal/conference name	#	%
conference	Proceeding OSDOC '13 Proceedings of the Workshop on Open Source and Design of Communication	1	2.56
conference	Proceedings of the 2010 International Conference on Information Technology and Scientific Management	1	2.56
conference	Proceedings of the 2012 iConference on - iConference '12	1	2.56
conference	Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)	1	2.56
conference	Proceedings of The 3rd Multidisciplinary International Social Networks Conference on Social Informatics	1	2.56
Journal	Public Performance and Management Review	1	2.56
Journal	Technology Analysis and Strategic Management	1	2.56
Journal	The Academy of Management Journal	1	2.56
Total source publication		39	

4.1.2. Most productive institutions

The most productive institution is a National Central University, National Central University, Ewha Woman's University with two paper each. An others institution each have one paper. Detail data can be seen in table 3 which in total there are 53 institutions.

Table 3 Source of publications

Institutions	# papers	%
National Central University,	2	3.64
Hebei University of technology Tianjin	2	3.64
Ewha Womans University	2	3.64
Vrije Universiteit Amsterdam	1	1.82
University of Zagreb	1	1.82
University of Westminster,	1	1.82
University of Twente,	1	1.82
University of Turkey,	1	1.82
University of Toronto	1	1.82
University of Science and Technology	1	1.82
University Of Salford	1	1.82
University of Melbourne	1	1.82
University of Manchester	1	1.82
University of Las Palmas de Gran Canaria	1	1.82
University of Groningen	1	1.82
University of Caxias do Sul	1	1.82
University of Auckland	1	1.82

Institutions	# papers	%
Universiti Teknologi Malaysia	1	1.82
Universiti Teknologi Malaysia	1	1.82
universitas liverpool	1	1.82
Tilburg University,	1	1.82
The Hong Kong Polytechnic University	1	1.82
Technical University of Bari,	1	1.82
Royal Tropical Institute,	1	1.82
Renmin University of China,	1	1.82
North Carolina State University ,	1	1.82
National University of Singapore,	1	1.82
National Taiwan Ocean University	1	1.82
National Sun Yat-sen University,	1	1.82
National Quemoy University,	1	1.82
National Kaohsiung First University of Science of Technology	1	1.82
Nanyang Technological University,	1	1.82
Jiangsu University of Science and Technology	1	1.82
Jiangsu University of Science and Technology	1	1.82
Instituto Universitario de Lisboa	1	1.82
Hong Kong Baptist University	1	1.82
FCET Staffordshire University,	1	1.82
Dalian University of Technology Dalian	1	1.82
City University of Hong Kong	1	1.82
Center for Innovation Research,	1	1.82
Bei Hang University,	1	1.82
Autonomous University of Baja California	1	1.82
Arizona State University	1	1.82
Åbo Akademi University	1	1.82
University of Siegen	1	1.82
The University of Sydney Business School	1	1.82
National Chiao Tung University	1	1.82
National Central University,	1	1.82
Harrisburg University of Science and Technology	1	1.82
Gazi University	1	1.82
Australian National University (ANU)	1	1.82
Al Ghurair University,	1	1.82
Peking University,	1	1.82
Total institution : 53 institutions	56	

4.2. Authors' academic backgrounds

The academic background author that can be seen in table 3 consists of 18 backgrounds of 91 authors that match the research.

Table 4. Discipline of authors

Department	#	%
Department of Information Systems	14	12.74
Information Management	13	11.83

Business Administration	11	10.01
Industry	10	9.1
Management	9	8.19
Economics and Management	9	8.19
Architecture and Urban Studies	4	3.64
Department of Informatics	3	2.73
Chemical Sciences and Engineering	3	2.73
Center for Sustainable Innovation	3	2.73
Technology and Management	2	1.82
Shipping and Transportation Management	2	1.82
Business and Finance	2	1.82
Business and Economics	2	1.82
Science and Technology	1	0.91
Development Policy Management	1	0.91
Department of Management and Marketing	1	0.91
Computing Drive	1	0.91
Total	91	

4.3. Background of authors

The author's background consists of 81 academic and 32 from the industry. Those who do related research on knowledge management.

Table 5. Background of authors

Background author	#	%
Academic	81	90.9
industry	10	9.1
Total	91	

4.4. University affiliation according to country

University affiliation of the 68 countries, China has 16 authors with six institutions, Taiwan has 13 authors with six institutions, the USA has nine authors with five institutions, Netherland has eight authors with five institutions, the UK has seven authors with four institutions, and detail data can see in table six. it is a country that contributes in the development research in knowledge management.

Table 6. University affiliation according to country

Country	# authors	% authors	# institutions	% institutions
China	16	17.58	6	12.77
Taiwan	13	14.29	6	12.77
USA	9	9.89	5	10.64
Netherlands	8	8.79	5	10.64
UK	7	7.69	4	8.51
Hongkong	5	5.49	3	6.38
Italy	4	4.40	1	2.13
Korea Selatan	4	4.40	1	2.13
Australia	3	3.30	2	4.26
Portugal	3	3.30	1	2.13
Brazil	3	3.30	1	2.13

Country	# authors	% authors	# institutions	% institutions
Finland	2	2.20	2	4.26
Singapore	2	2.20	2	4.26
Spain	2	2.20	1	2.13
Turkey	2	2.20	1	2.13
Malaysia	2	2.20	1	2.13
New Zealand	2	2.20	1	2.13
United Arab Emirates	1	1.10	1	2.13
Australia	1	1.10	1	2.13
México	1	1.10	1	2.13
Croatia	1	1.10	1	2.13
Total country: 21 countries	91		47	

4.5. Most prolific authors

From the analyst's point of view, there are 90 authors with 39 papers. The author who actively found the author on KM in social society is Marleen Huysman (# 2). Another writer averages one article. It can be seen clearly in Table 7 below.

Table 7. most prolific authors

Author	#	%
Marleen Huysman	2	2.25
Albert A. Cannella Jr	1	1.12
Gang Qu	1	1.12
Rolando Vargas Vallejos	1	1.12
Amanda Edwards	1	1.12
Andrew Long	1	1.12
Annette Boaz	1	1.12
Anne-wil Hazing	1	1.12
AydÄntan Belgin	1	1.12
Bosen Li	1	1.12
CarlaC.J.M. Milla	1	1.12
Carlos J. Costa	1	1.12
Caterina De Lucia	1	1.12
Chen Yen Yao	1	1.12
Chen Yijia	1	1.12
Cheng Yang Lai	1	1.12
Chia Fen Chung	1	1.12
Chin-Chung Tsai	1	1.12
Chongju Choi	1	1.12
Chun-Wei Choo	1	1.12
David Sundaram	1	1.12
Dino Borri	1	1.12
Fan Yi-Wen	1	1.12
G.Widen	1	1.12
Goksel Aykut	1	1.12
Guido Sechi	1	1.12
He Wei	1	1.12
Hsing Kuo Wang	1	1.12

Author	#	%
Hsiu-Fen Lin	1	1.12
Janaina Macke	1	1.12
Javier Osorio	1	1.12
Jay Liebowitz	1	1.12
Jin Hui	1	1.12
Jingjing Han	1	1.12
Jin-Xing Hao	1	1.12
Jo van Engelen	1	1.12
Jordan Lewis-Pryde	1	1.12
Jui Pattnayak	1	1.12
Julia Nieves	1	1.12
Jung Feng Tseng	1	1.12
Kadgia Faccin	1	1.12
Kang Kai	1	1.12
Kelly Lyons	1	1.12
Kwok-Kee Wei	1	1.12
Lesley Gray	1	1.12
Li Wang	1	1.12
Lorna Uden	1	1.12
Lv Jingyin	1	1.12
M. Ann McFadyen	1	1.12
Manuela Aparicio	1	1.12
Margaret Sheng	1	1.12
Mark W.McElroy	1	1.12
Markus Schatten	1	1.12
Matti Mantymaki	1	1.12
Mohamed Khalifa	1	1.12
Niels Noorderhaven	1	1.12
Nuno Sousa	1	1.12
Ping Chuan Chen	1	1.12
Qian Qian	1	1.12
Rafael Pimienta-romo	1	1.12
Rendi Hartono	1	1.12
Rene J.Jorna	1	1.12
Reyes Juarez-ramirez	1	1.12
Richard David Evans	1	1.12
Richard Heeks	1	1.12
Riemer kai	1	1.12
Rose Alinda Alias	1	1.12
Sabyasachi	1	1.12
Sarah Cummings	1	1.12
Shih-Wei Chou	1	1.12
Shiu Wan Hung	1	1.12
Steven Chuang	1	1.12
Tzu Fong Liao	1	1.12
Valeria Sadovykh	1	1.12
Viesturs Celmins	1	1.12
Viesturs Celmins	1	1.12
Violeta Ocegueda-miramontes	1	1.12
Volker Wulf	1	1.12
Wu Cheng-Chieh	1	1.12
Xiao Ying Dong	1	1.12

Author	#	%
Yan Yu	1	1.12
Yen-Chiang Fang	1	1.12
Yu Fang Yen	1	1.12
Yu-Chieh Chang	1	1.12
Yujin Choi	1	1.12
Yung Ming Li	1	1.12
Zainal Wardah Abidin	1	1.12
Zhang Jie	1	1.12
Total	89	

4.6. Mapping to Theories, Factors, and Paper

The comprehensive review process classifies the factors that are mostly done in knowledge management in social societies. This factor classification mapping is based on the theory, factors, and authors of the articles used in this study. The following results are shown in Table 8

Table 8. Theories, Factors, and Paper

Theory	Factor	ID Paper
Social Capital	Cognitive social capital	[15],[16],[17]
	Relational social capital	[18],[19],[20],[21],[22],
	Structural social capital	[23],[24],[25],[26],[27],[28],[29],[8],[30],[31],[32],[33],[34],[35],[36],[37],[38]
Social network-based Markov Chain (SNMC) models	Semantic Similarity	[39]
	Profession, Reliability,	
	Social intimacy	
	Popularity,	
Enterprise Social Networking	Work Discussion	[40]
	Input Generation	
	Problem-solving, Social Praise	
	Idea Generation	
	Status Updates Informal	
	Task Management	
	Talk Event, Notifications	
Theory of Planned Behavior (TPB)	The intention, Perceived behavior toward	[19]
Transactive Memory System (TMS) Theory	Credibility	[18]
	Specialization	
	Coordination	
SECI model	Socialization, Externalization,	[2]

Theory	Factor	ID Paper
	Combination, and Internalization	
Social Networks	People-related, Process-related, Technical-related, Adaptability/ Agility, Creativity, Institutional memory building, Organizational Internal Effectiveness, Intangibility, Heterogeneity, Perishability	[41],[42],[43], [44]
Semantic Social Networks	Basic typed Semantic Social networks and Trust annotated semantic social networks	[45]
Social Interaction	Employee motivation (Intrinsic motivation and Extrinsic motivation), Social interaction (Interpersonal trust, Openness in Communication, and Social reciprocity), and Knowledge Management (KM) strategy (Codification Knowledge Strategy and Personalization Knowledge Strategy)	[46]

4.7. Keyword Analysis

Keyword Analysis used in searching papers related to the study of social knowledge Management. The keyword used yields 278 papers from 7 reputable journal database sources. Among the frequently used keywords are ‘Knowledge Management Social Theory’. Article data found on other keywords that occur overlap.

Table 9. Most frequently used keywords

Keyword	Paper						
	Emerald	IEEE	Jstor	ACM	Science Direct	TandF	springer
Knowledge Management Social	17	10	4	8	10	15	8
Knowledge Management Social Theory	12	8	10	9	8	12	15

Knowledge Management Society	15	1 5	8	5	12	14	8
Social knowledge Management Theory	16	7	8	3	5	7	9

4.8. Mapping Theories for Factors to uses in Social Knowledge Management

- Theory of Planned Behavior (TPB)
Ajzen explains that TRA is an individual's attitude toward behavior positively influencing intention to participate in that behavior[19]. The main factor of planned behavior theory is the individual's intention to perform certain behaviors. An intent is assumed to capture the motivating factors that influence the behavior of the individual. This is an indication of how hard it is for people to try, how much effort they plan to do that behavior. As a general rule, the stronger the intention to engage in the behavior, the more likely it is to perform. However, it should be clear that behavioral intentions can only express behavior only if the behavior in question is under complete control [47].
- Transactive memory system (TMS) theory
The concept of TMS is the specific division of labor processes that relate to taking, storing, and retrieving knowledge from different environments[48]. The TMS process occurs in individual/group transactions that are aware of the need to develop unique knowledge and member expertise so that a group can rely on member knowledge. This TMS consists of three main components (1) specialization shows different members' knowledge structures, (2) credibility shows members' beliefs about the accuracy and reliability of other members' knowledge, (3) coordination shows the effective and orderly storage of knowledge [18].
- Social capital theory
Social capital is a knowledge resource that comes from social networks that people can use to make a behavior. This is illustrated through the ability to access and exchange the knowledge resources of individuals who are in the social structure. [22],[23],[24],[25], [26][49]. This makes Social capital has been recognized as an important factor for social interaction[15],[16], [17],[18], [19], [20], [21]. Interpersonal networks provide channels for the exchange of tangible and intangible resources needed. Nahapiet & Ghoshal suggest that social capital is a multifaceted concept and can be divided into three dimensions: capital structure, relationship capital, and cognitive capital.

- SECI model
The process of socialization refers to the transfer of knowledge. The externalization process refers to documenting their knowledge so that it is possible to share it with others [1], [10], [11]. The Combination Process refers to combining knowledge with other knowledge to rearrange new knowledge [2]. Internalization Process. During the process of making new knowledge that everyone can share with each individual through

further consultation, training, and assimilating this knowledge [2], [11].

- Social Networks
Social Network Theory is a theory or study that studies how people, organizations, or groups interact with others in the network that exists in them[50], [51].In understanding this theory, it is easier when you examine individual pieces that start with the most significant element, i.e., the network, and work up to the smallest element, i.e., the actor[19]. Social networks are social structures that have a set of social actors (such as individuals or organizations), a collection of ties, and other social interactions between actors [52]. The view of the social network provides a set of methods used to analyze the structure of the entire social entity as well as theories that explain the patterns observed in this structure [45].

5. Conclusions

Results obtained from the discussions that have been described. Described the knowledge management occurs within the social environment can be seen from the development of research conducted with several theories. Social capital theory[15][16][17] shows that 3 very strong factors can influence social relationships that occur in individuals within the social community is structural capital, cognitive capital, and relationship capital, in addition to this Theory of Planned Behavior (TPB) illustrates that the main factor for determining the intention of individuals to perform certain behaviors is the attitude of individuals to behave positively [19] [47]. it also affects knowledge management.

After that, individual transactions described in the Transactive memory system (TMS) theory [48] have major components: specialization, credibility, coordination, so that teams can develop a shared awareness of each member's unique knowledge and expertise[18].

Maintain relationships and transactions that occur requires a good social network. Social network theory explains by increasing interpersonal trust, informal communication, and reciprocal relationships, People-related, Process-related, Technical-related, Adaptability/agility, Creativity, Institutional memory building, Organizational internal effectiveness, Intangibility, Heterogeneity, Perishability, in turn, increasing KM maturity. The theory of social interaction affects the improvement of social networking relations. Factors are employee motivation (intrinsic motivation and extrinsic motivation), social interaction (interpersonal trust, openness in communication, and social reciprocity), and knowledge management (KM) strategy (codification knowledge strategy and personalization knowledge strategy). From the results obtained, this study illustrates these factors that affect KM in social society.

6. Implication

Based on the findings of structural capital, cognitive capital, and relationship capital on social capital theory, it becomes an essential factor to be seen in KM in social society. Thus KM can

well by looking at the individual social interactions that occur. By calculating or reviewing the interaction. The implications are given in science is that social knowledge management requires developed social media today that can facilitate interactions that occur within a social community.

7. Limitation and Future research

This study has a limited database used; this is due to limited access. The number of articles to be added is mainly extracted from credible and published databases in the last five years. Future research, researchers will perform statistical analysis of these factors on the social community to determine the influence of these factors. so, it can be seen with concrete factors that influence in KM

References

- [1] I. Nonaka, "The knowledge-creating company," *Harv. Bus. Rev.*, no. December 1991, 1991.
- [2] N. Sousa, C. J. Costa, and M. Aparicio, "IO-SECI: A Conceptual Model for Knowledge Management," in *Proceeding OSDOC '13 Proceedings of the Workshop on Open Source and Design of Communication*, 2013, pp. 9–17.
- [3] N. Houari and B. H. Far, "Application of intelligent agent technology for knowledge management integration," *Proc. Third IEEE Int. Conf. Cogn. Informatics*, 2004., pp. 240–249, 2004.
- [4] M. Asrar-ul-Haq and S. Anwar, "A systematic review of knowledge management and knowledge sharing: Trends, issues, and challenges," *Cogent Bus. Manag.*, vol. 3, no. 1, pp. 1–17, 2016.
- [5] M. Alavi and D. E. Leidner, "Knowledge Management and Knowledge Systems : Conceptual Foundations and Research Issue," *MIS Q.*, vol. 25, no. 1, pp. 107–136, 2001.
- [6] M. Huysman and V. Wulf, "IT to support knowledge sharing in communities, towards a social capital analysis," *J. Inf. Technol.*, vol. 21, no. 1, pp. 40–51, 2006.
- [7] S. G. López, J. L. S. Benítez, and J. M. A. Sánchez, "Social Knowledge Management from the Social Responsibility of the University for the Promotion of Sustainable Development," *Procedia - Soc. Behav. Sci.*, vol. 191, pp. 2112–2116, 2015.
- [8] Y. Choi, "The Impact of Social Capital on Employees' Knowledge-Sharing Behavior: An Empirical Analysis of U.S. Federal Agencies," *Public Perform. Manag. Rev.*, vol. 39, no. 2, pp. 381–405, 2016.
- [9] H. Prabowo, "Information Science and Applications 2017," *Inf. Sci. Appl.*, vol. 424, 2017.
- [10] I. Nonaka and H. Takeuchi, "A theory of organizational knowledge creation Ikujiro Nonaka and Hirotaka Takeuchi Katsu hiro Umemoto," *IJTM, Spec. Publ. Unlearning Learn.*, vol. 11, no. 7/8, pp. 833–845, 1996.
- [11] I. Nonaka, R. Toyama, and N. Konno, "SECI, Ba and Leadership: A Unified Model of Dynamic Knowledge Creation," *Long Range Plann.*, vol. 33, no. 1, pp. 5–34, 2000.
- [12] R. James, "Postcolonial Development and Sustainability," pp. 1–14, 2012.
- [13] S. Rivas-Gomez et al., "MPI windows on storage for HPC applications," in *Proceedings of the 24th European MPI Users' Group Meeting on - EuroMPI '17*, 2017, pp. 1–11.
- [14] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering – A systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, 2009.
- [15] W. He, Q. Qiao, and K.-K. Wei, "Social relationship and its role in knowledge management systems usage," *Inf. Manag.*, vol. 46, no. 3, pp. 175–180, Apr. 2009.
- [16] Y. Yu, J.-X. Hao, X.-Y. Dong, and M. Khalifa, "A multilevel model for effects of social capital and knowledge sharing in knowledge-intensive work teams," *Int. J. Inf. Manage.*, vol. 33, no. 5, pp. 780–790, Oct. 2013.
- [17] J. Pattanayak and S. Pattnaik, "Integration of Web Services with E-Learning for Knowledge Society," *Procedia Comput. Sci.*, vol. 92, pp. 155–160, 2016.
- [18] B. Li and G. Qu, "Relationship between team social capital and knowledge transfer: The mediated effect of TMS," *2010 Int. Conf. E-Product E-Service E-Entertainment, ICEEE2010*, pp. 1–4, 2010.
- [19] S.-W. Chou and Y.-C. Chang, "An Empirical Investigation of Knowledge Creation in Electronic Networks of Practice: Social Capital and Theory of Planned Behavior (TPB)," *Proc. 41st Annu. Hawaii Int. Conf. Syst. Sci. (HICSS 2008)*, pp. 340–340, 2008.
- [20] G. Sechi, D. Borri, C. De Lucia, and V. Celmins, "Social capital as knowledge facilitator: Evidence from Latvia," *Knowl. Manag. Res. Pract.*, vol. 9, no. 3, pp. 245–255, 2011.
- [21] H. P. Zhang, X. G. Wei, and K. Kang, "A Study of the Effects of Social Capital on Inter-Firm Knowledge Transfer and Innovation Performance," *Proc. 2010 Int. Conf. Inf. Technol. Sci. Manag. Vols 1-2*, no. 2006, pp. 819–823, 2010.
- [22] Yi-Wen Fan and Cheng-Chieh Wu, "The Role of Social Capital in Knowledge Sharing: A Meta-Analytic Review," in *2011 44th Hawaii International Conference on System Sciences*, 2011, pp. 1–10.
- [23] S. P.-M. Law and M.-K. Chang, "Social Capital and Knowledge Sharing in Online Communities: A Mediation Model," *2012 45th Hawaii Int. Conf. Syst. Sci.*, pp. 3530–3539, 2012.
- [24] Wang Li and Han Jingjing, "The relationship study of social capital, knowledge management and performance in logistics enterprises based on SEM," in *2014 IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA)*, 2014, no. 1998, pp. 1249–1251.
- [25] C. Choi and C. C. J. M. Milla, "Networks ,Social Norms and Knowledge Sub-Network," *J. Bus. Ethics*, vol. 90, no. 2009, pp. 565–574, 2014.
- [26] M. A. McFayden and A. A. Cannella, "Social capital and knowledge creation: Diminshing returns of the number and strenght of exchange," *Acad. Manag. J.*, vol. 47, no. 5, pp. 735–746, 2004.
- [27] S. Cummings, R. Heeks, and M. Huysman, "Knowledge and learning in online networks in development: A social-capital perspective," *Dev. Pract.*, vol. 16, no. 6, pp. 570–586, 2006.
- [28] S. W. Hung, P. C. Chen, and C. F. Chung, "Gaining or losing? The social capital perspective on supply chain members' knowledge sharing of green practices," *Technol. Anal. Strateg. Manag.*, vol. 26, no. 2, pp. 189–206, 2014.
- [29] J. Nieves and J. Osorio, "The role of social networks in knowledge creation," *Knowl. Manag. Res. Pract.*, vol. 11, no. 1, pp. 62–77, 2013.
- [30] Y. F. Yen, J. F. Tseng, and H. K. Wang, "The effect of internal social capital on knowledge sharing," *Knowl. Manag. Res. Pract.*, vol. 13, no. 2, pp. 214–224, 2015.
- [31] A. Göksel and B. Aydıntan, "How can tacit knowledge be shared more in organizations? A multidimensional approach to the role of social capital and locus of control," *Knowl. Manag. Res. Pract.*, vol. 15, no. 1, pp. 34–44, 2017.
- [32] A. Edwards, "What is 'Knowledge' in Social Care?," *Housing, Care Support*, vol. 4, no. 4, pp. 2–3, Nov. 2001.
- [33] L. Grayson, A. Boaz, and A. Long, "Organising Social Care Knowledge: In Search of a 'Fit for Purpose' Classification," *J. Integr. Care*, vol. 12, no. 1, pp. 42–48, Feb. 2004.
- [34] M. W. McElroy, R. J. Jorna, and J. van Engelen, "Rethinking social capital theory: a knowledge management perspective," *J. Knowl. Manag.*, vol. 10, no. 5, pp. 124–136, Sep. 2006.
- [35] W. Z. Abidin, L. Uden, and R. A. Alias, "Knowledge Management in Organizations," vol. 224, pp. 754–769, 2015.
- [36] J. Hui and C. Yijia, "The research on how social capital facilitates knowledge sharing between individuals," *Adv. Intell. Soft Comput.*, vol. 110, pp. 261–270, 2011.
- [37] R. V. Vallejos, J. MacKe, and K. Faccin, "Establishing Knowledge Management as an important factor to develop Social Capital for collaborative networks," *IFIP Adv. Inf. Commun. Technol.*, vol. 362 AICT, pp. 58–65, 2011.
- [38] G. Widen, "Social capital and knowledge sharing: Lessons learned," *Adapt. Value Creat. Collab. Networks 12th IFIP WG 5.5 Work. Conf. Virtual Enterp. PRO-VE 2011*, vol. 362, pp. 48–57, 2011.
- [39] Y. M. Li, T. F. Liao, and C. Y. Lai, "A social recommender mechanism for improving knowledge sharing in online forums," *Inf. Process. Manag.*, vol. 48, no. 5, pp. 978–994, 2012.
- [40] M. Mäntymäki and K. Riemer, "Enterprise social networking: A knowledge management perspective," *Int. J. Inf. Manage.*, vol. 36, no. 6, pp. 1042–1052, 2016.

- [41] R. Juárez-ramírez, R. Pimienta-romo, and V. Ocegueda-miramontes, "Using Social Networks for Integrating a Tacit Knowledge Repository to Support," *Int. Symp. Iuk.* 2013, pp. 167–179, 2013.
- [42] J. Liebowitz, "The Hidden Power of Social Networks and Knowledge Sharing in Healthcare," *Healthc. Knowl. Manag. SE* - 8, pp. 104–111, 2007.
- [43] V. Sadovykh and D. Sundaram, "Context-Aware Systems and Applications," vol. 193, pp. 22–31, 2017.
- [44] J. Lewis-Pryde and R. D. Evans, "A Social Networking Strategy for Improving Knowledge Management and Communication in the Travel Industry," *Proc. 3rd Multidiscip. Int. Soc. Networks Conf. Soc. 2016, Data Sci. 2016 - MISNC, SI, DS 2016*, pp. 1–5, 2016.
- [45] M. Schatten, "Knowledge management in semantic social networks," *Comput. Math. Organ. Theory*, vol. 19, no. 4, pp. 538–568, 2013.
- [46] H.-F. Lin, "The effects of employee motivation, social interaction, and knowledge management strategy on KM implementation level," *Knowl. Manag. Res. Pract.*, vol. 9, no. 3, pp. 263–275, 2011.
- [47] I. Ajzen, "The theory of planned behavior," *Organizational Behav. Hum. Decis. Process.*, vol. 50, pp. 179–211, 1991.
- [48] D. M. Wegner, "Summary for Policymakers," in *Climate Change 2013 - The Physical Science Basis*, vol. 13, no. 3, Intergovernmental Panel on Climate Change, Ed. Cambridge: Cambridge University Press, 1995, pp. 1–30.
- [49] J. Nahapiet and S. Ghoshal, "Social Capital, Intellectual Capital, and the Organizational Advantage," *Acad. Manag. Rev.*, vol. 23, no. 2, p. 242, Apr. 1998.
- [50] W. Liu, A. Sidhu, A. M. Beacom, and T. W. Valente, "Social Network Theory," *Int. Encycl. Media Eff.*, no. September, pp. 1–12, 2017.
- [51] K. S. K. Chung and L. Crawford, "The Role of Social Networks Theory and Methodology for Project Stakeholder Management," *Procedia - Soc. Behav. Sci.*, vol. 226, no. October 2015, pp. 372–380, 2016.
- [52] R. A. Costa, R. Y. Oliveira, E. M. Silva, and S. R. Meira, "A.M.I.G.O.S: Knowledge Management and Social Networks," *Spec. Interes. Gr. Des. Commun.* 2008, pp. 235–242, 2008.

Automated Abaca Fiber Grade Classification Using Convolution Neural Network (CNN)

Neptali Montañez, Jomari Joseph Barrera*

Department of Computer Science and Technology, Visayas State University, Baybay City, Leyte, 6521, Philippines

ARTICLE INFO

Article history:

Received: 02 April, 2020

Accepted: 06 May, 2020

Online: 19 May, 2020

Keywords:

Artificial intelligence

Deep learning

Convolutional neural network

Abaca fiber grade

VGGNet-16

CNN classifier

ABSTRACT

This paper presents a solution that automates Abaca fiber grading which would help the time-consuming baling of Abaca fiber produce. The study introduces an objective instrument paired with a system to automate the grade classification of Abaca fiber using Convolutional Neural Network (CNN). In this study, 140 sample images of abaca fibers were used, which were divided into two sets: 70 images; 10 per grade, each for training and testing. The input images were then scaled to 112x112 pixels. Next, using a customized version of VGGNet-16 CNN architecture, the training set images were used for training. Finally, the performance of the classifier was evaluated by computing the overall accuracy of the system and its Cohen kappa value. Based on the result, the classifier achieved 83% accuracy in correctly classifying the Abaca fiber grade of a sample image and obtained a Cohen kappa value of 0.52 — Weak, Level of Agreement. The implementation of this study would greatly help Abaca producers and traders ensure that their Abaca fiber would be graded fairly and efficiently to maximize their profit.

1. Introduction

Abaca (*Musa textilis*) or Manila hemp in international trade is indigenous to the Philippines. The Philippines is the world's largest producer of abaca fiber accounting for about 85% of the global production. Abaca plants are cultivated in 130 thousand hectares across the island by over 90 thousand farmers. It is one of the major export products of the country together with the banana, coconut oil, and pineapple. Because of this, the Philippines' agricultural sector is very important for the economy as it contributes about 16% of Gross Domestic Product (GDP). The abaca industry continues to make a stronghold in domestic and international markets providing a yearly average baling (in bales of 125 kgs) of 424, 212 annually from 2012 – 2017 [1].

Philippine Fiber Industry Development Authority (PhilFIDA), is a government agency tasked to grade different fiber grades of the Abaca fibers, due to the use of different stripping knives serrations and spindle/machine-stripper. There are 13 fiber classifications in grading in which eight in normal grades (see Table 1), four in residual and one in wide strips that are currently used in the market based on its texture, color, length and strength [2]. Many manufacturing industries require large qualities of different classifications of fiber such as pulp, paper products, and for automotive applications.

Owing to the lack of instruments for an objective measurement of fiber quality, grading and classification have to be done by visual inspection, which is time-consuming and costly hence, the purpose of this study.

2. Related Work

Abaca fiber is considered as one of the strongest among natural fibers which are three times stronger than sisal [3]. It can be planted through disease-free tissue cultured plantlets, corm cut into 4 pieces with one eye each (seed piece) and sucker (used for replanting missing hills) [4]. It is mostly grown in the upland parts and grow on light textured soils under the shed of coconut trees. Abaca is harvested by cutting the stem using a sharp machete within 18–24 months of planting. After which, fibers are stripped by hand or spindle stripping. The fibers are then sun-dried and sold on an 'all in' basis [5]. Abaca fiber bundles are then delivered by local traders to the Grading Baling Establishment (GBE) that will undergo grading and classification.

In a study by [6], the grade of the Abaca fiber image in the bundle was predicted by two types of learning algorithms in neural networks based only on the color of the eight different normal grades. The learning algorithm in neural networks was the Self-Organizing Map (SOM) and Back Propagation Neural Network (BPNN). When implemented, the accuracy of SOM is 46% while

*Jomari Joseph Barrera, Philippines, jomarijoseph.barrera@vsu.edu.ph

www.astesj.com

<https://dx.doi.org/10.25046/aj050327>

Table 1: Normal grades of hand stripped abaca fiber [6]

Grade		Description			
Name	Alphanumeric code	Fiber strand size (in mm)	Color	Stripping	Texture
Mid current	EF	0.20 – 0.50	Light ivory to a hue of very light brown to very light ochre	Excellent	Soft
Streaky Two	S2	0.20 – 0.50	Ivory white, slightly tinged with very light brown to red or purple streak	Excellent	Soft
Streaky Three	S3	0.20 – 0.50	Predominant color – light to dark red or purple or a shade of dull to dark brown	Excellent	Soft
Current	I	0.51 – 0.99	Very light brown to light brown	Good	Medium soft
Soft seconds	G	0.51 – 0.99	Dingy white, light green and dull brown	Good	
Soft Brown	H	0.51 – 0.99	Dark brown	Good	
Seconds	JK	0.51 – 0.99	Dull brown to dingy light brown or dingy light yellow, frequently streaked with light green	Fair	
Medium brown	M1	0.51 – 0.99	Dark brown to almost black	Fair	

BPNN is 88%. BPNN is then the efficient alternative for the identification of Abaca fiber grades based on color classification provided by PhilFIDA. This study did not determine the grade of the Abaca fiber based on its texture which is the primary basis in grading its quality.

In [7] the study shows that by combining hand-crafted (color and texture) and convolutional neural network (CNN) features, the classification of images acquired in different lighting conditions, greatly improved. The same study also shows that VGGNet-16 CNN architecture outperforms other CNN architecture.

Furthermore, CNN has achieved human-like performance in several recognition tasks such as handwritten character recognition, face recognition, scene labeling, object detection, and image classification among others [8–16].

3. Material and Method

Figure 1 illustrates the system architecture applied in this study. The Abaca fiber grade classifier was composed of two phases. Each phase started with image acquisition and followed by a segregation of samples into folders for training and testing dataset. Then rescaling the image dimension from 2976x2976 pixels into 112x112 pixels then followed by converting each image to RGB channel.

3.1. Image Acquisition and Pre-processing

The abaca fiber samples (see Figure 2) were obtained by requesting Ching Bee Trading Corporation in Brgy. Hilapnitan, Baybay City, Leyte for Abaca fiber samples of different grades and the actual grading was done by a certified PhilFIDA inspector. Samples of abaca fiber for seven grades (S2, S3, I, G, H, JK, M1) only were segregated for image acquisition. The EF grade fiber was not available in the current season because of the ongoing El Niño phenomenon and other climate change factors.

In [7] it is further explained that the different viewpoint and illumination of the acquired digital image will greatly affect the classification accuracy of the system. So, in this case, Abaca fiber sample image must be acquired in a controlled environment using an objective instrument — a customized photo box and a 16.0-megapixel Samsung S6 camera with a distance of 19 cm above the sample and taken in the middle part only as shown in Figure 3.

3.2. Training Phase

The CNN architecture implemented in this study is shown in Figure 4. It consists of five convolutional layers followed by a rectified linear unit as activation function, and three pooling layers. After the last pooling layer is flatten into a single column vector, the concatenated 1024 data values are inputs to the neural network and it needs to be trained using back propagation algorithm. Then, the training optimization used was stochastic gradient descent optimization technique in finding the set of weights and biases between the neurons that determine the global minima of the loss function. All weights and bias values are set to random. Drop-outs and batch normalization were also applied in the CNN training phase.

3.3. Testing Phase and Performance Evaluation

The generated model and label files after the training phase will be used in the testing phase. Ten sample abaca fibers per grade were used as testing data to determine the performance of the system. The results will be plotted in a confusion matrix and from this, we can acquire the classifier’s accuracy using (1) and Cohen kappa value using (2). Thereafter, we can evaluate the classifier’s level of agreement based on the Cohen kappa value (see Table 2).

$$\text{Accuracy} = \frac{\text{sum of correctly predicted grade}}{\text{total number of predictions}} \quad (1)$$

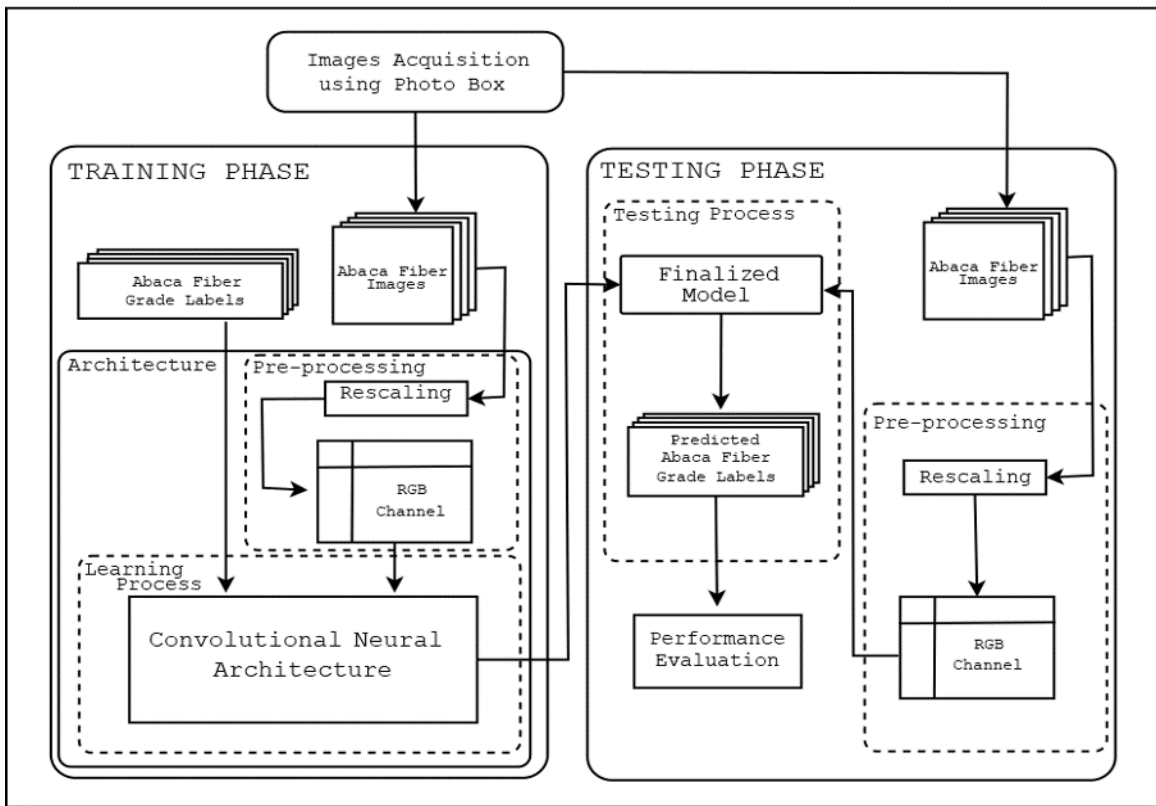


Figure 1: System architecture of the Abaca fiber grade classifier.

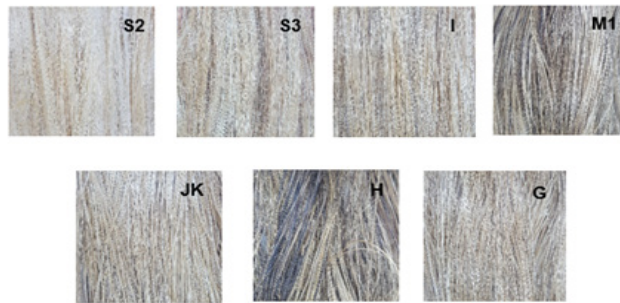


Figure 2: Sample images of abaca fiber with their corresponding grade.

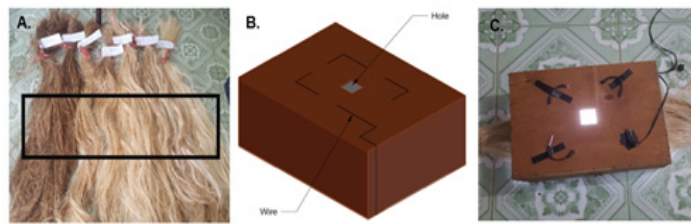


Figure 3: (A) Location of the Abaca fiber, (B) CAD design of photo box, and (C) acquisition set-up of abaca fiber image.

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (2)$$

where,

- p_o = the relative observed agreement among raters.
- p_e = the hypothetical probability of chance agreement

4. Experimental Results

After conducting the training phase using the training set abaca fiber sample images (see Table 3), the parameters of the

instantiated model based from Figure 4 were adjusted. The adjusted model will be used in the testing phase as the Abaca grade classifier.

In the testing phase, the user needs to load an Abaca fiber captured image from the objective instrument, and trigger the “Predict Grade” button where a matching percentage per grade class will be generated (see Figure 5). After testing the model with the testing dataset of 70 images (10 images per grade), the sample

results were shown in Table 4. Confusion matrix (see Table 5) were then derived from Table 4.

Based on the results, the classifier achieved an 83% overall accuracy in correctly classifying the Abaca fiber grade of a sample image and attained a Cohen kappa value of 0.52.

The overall accuracy rate of the system indicates that the objective instrument paired with the application of a customized VGGNet-16 convolutional neural network is sufficient enough to support an accurate classification of the Abaca fiber grade based on an Abaca fiber sample image.

No grade G Abaca fibers were correctly identified. Instead they were mistakenly classified as either grade M1 or I. Only grade G’s recall achieved 0% (see Table 6) while other grades achieved greater than or equal to 90%. This means that CNN did not draw a distinct difference between features extracted from grade G Abaca fibers, and M1 or I Abaca fibers. This is a highly likely scenario since grades M1 and I Abaca fibers have a 100% accurate prediction.

The Cohen kappa value of the classifier conveys that the percentage of data that are only reliable ranges from 15% – 35% only. The low Cohen’s Kappa value stems from having no correct classification of G grade due to ambiguity or low distinct features between grade G Abaca fibers and grade M1 and I Abaca fibers, in 2-dimensional texture processing.

Table 2: Cohen's Kappa Value - Level of Agreement Equivalency

Value of Kappa	Level of Agreement	Percentage of Data that are Reliable
0-.20	None	0-4%
0.21-0.39	Minimal	4-15%
0.40-0.59	Weak	15-35%
0.60-0.79	Moderate	35-63%
0.80-0.90	Strong	64-81%
Above .90	Almost Perfect	82-100%

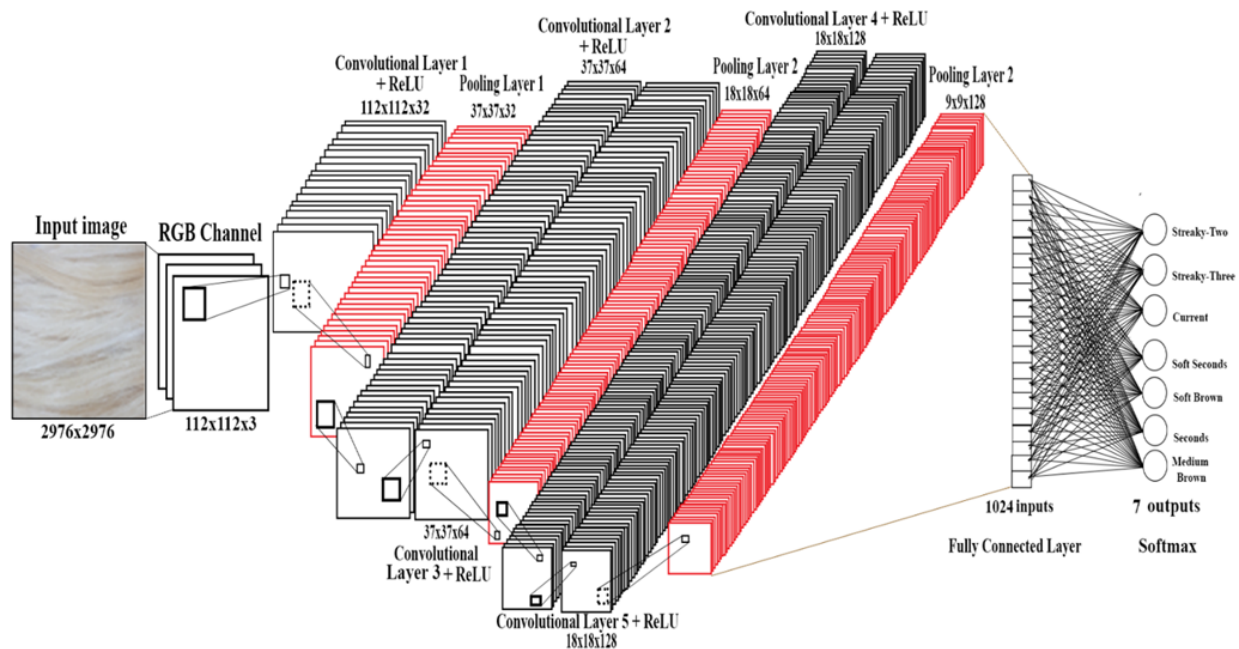


Figure 4: Customized VGGNet-16 CNN Architecture.

Table 3: Sample Abaca Fiber Training Set Images per Grade

Grade	Sample Images				
S2					
S3					
I					

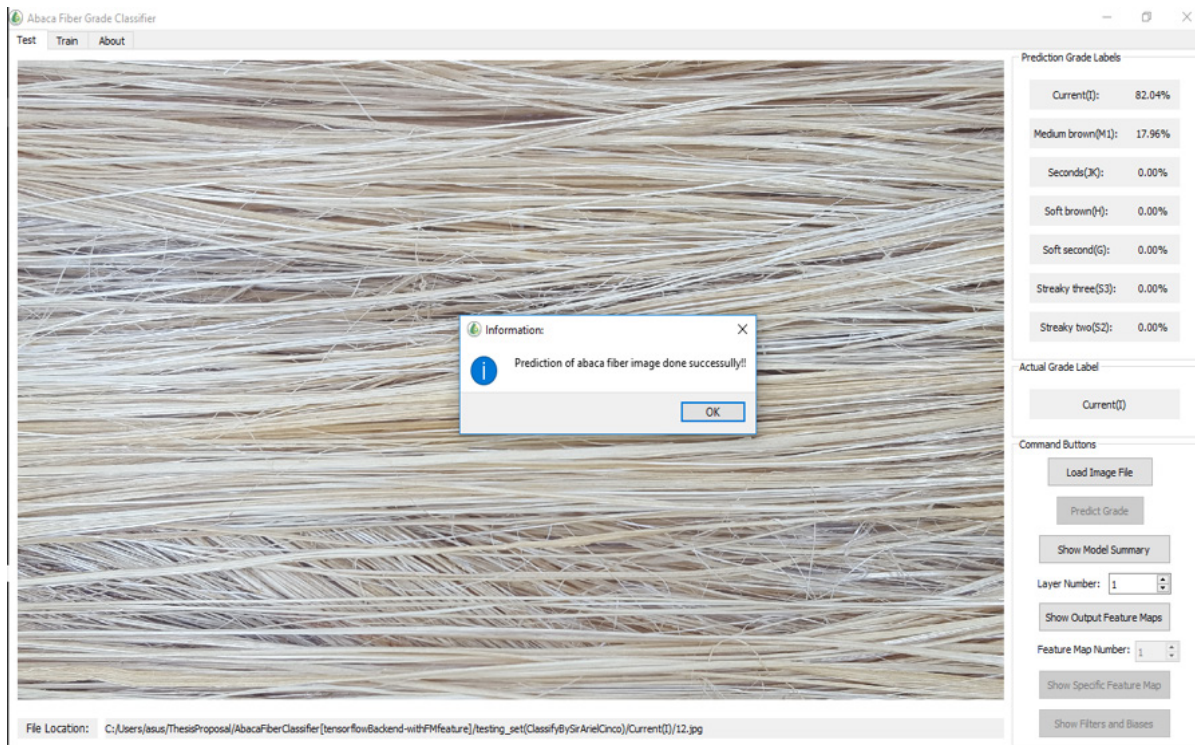


Figure 5: Prediction results of the Abaca fiber image.

Table 4: Prediction Results of Testing Dataset Abaca fiber images

Sample No.	Actual Grade	Predicted Grade	Remarks
1	Current(I)	Current(I)	Correct
2	Current(I)	Current(I)	Correct
3	Current(I)	Current(I)	Correct
4	Current(I)	Current(I)	Correct
5	Current(I)	Current(I)	Correct
6	Current(I)	Current(I)	Correct
7	Current(I)	Current(I)	Correct
8	Current(I)	Current(I)	Correct
9	Current(I)	Current(I)	Correct
10	Current(I)	Current(I)	Correct
11	Medium Brown(M1)	Medium Brown(M1)	Correct
12	Medium Brown(M1)	Medium Brown(M1)	Correct

13	Medium Brown(M1)	Medium Brown(M1)	Correct
14	Medium Brown(M1)	Medium Brown(M1)	Correct
15	Medium Brown(M1)	Medium Brown(M1)	Correct
16	Medium Brown(M1)	Medium Brown(M1)	Correct
:	:	:	:
70	Streaky Two(S2)	Streaky Two(S2)	Correct

Table 5: Confusion matrix based on Table 4.

		Actual Abaca Fiber Grades						
		S2	S3	I	G	H	JK	M1
System Predicted Grades	S2	10	0	0	0	0	0	0
	S3	0	9	0	0	0	0	0
	I	0	0	10	2	0	0	0
	G	0	0	0	0	0	0	0
	H	0	0	0	0	10	1	0
	JK	0	0	0	0	0	9	0
	M1	0	1	0	8	0	0	10

Table 6: Recall for each Abaca fiber grade from Table 5.

Grade	Recall
S2	100%
S3	90%
I	100%
G	0%
H	100%
JK	90%
M1	100%

5. Conclusion and Future Work

This study addresses the lack of objective instrument in acquiring unbiased datasets of Abaca fiber images. The implemented system used CNN as the main classifier and its performance was evaluated. Though the system correctly classified most of their Abaca fiber grades, it fails to classify any of the grade G Abaca fibers. Thus, adding other features aside from ones extracted from the image or trying other AI systems to be used as a classifier would be recommended.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

We would like to thank our consultant, Mr. Ariel C. Cinco, a certified PhilFIDA inspector assigned to Ching Bee Trading www.astesj.com

Corporation in Brgy. Hilapnitan, Baybay City, Leyte at the time of conducting the study, for lending his expertise on Abaca fiber trading.

References

- [1] PhilFIDA, "Fiber Statistics," PhilFIDA, 2018. [Online]. Available: <http://www.philfida.da.gov.ph/index.php/2016-11-10-03-32-59/2016-11-11-07-56-39>. [Accessed 28 October 2018].
- [2] PhilFIDA, "Abaca fiber: Grading and Classification – Hand-stripped and Spindle/ Machine stripped.," Diliman, Quezon City, Philippine National Standard. Bureau of Agriculture and Fisheries Standards, 2016.
- [3] F. Göltzenboth and W. Mühlbauer, "Abacá – Cultivation, Extraction and Processing," in Industrial Applications of Natural Fibres: Structure, Properties and Technical Applications, John Wiley & Sons, Ltd, 2010, pp. 163-179. <https://doi.org/10.1002/9780470660324.ch7>
- [4] P. P. Milan and F. Göltzenboth, Abaca and Rainforestation Farming: A Guide to Sustainable Farm Management, Baybay, Leyte: Leyte State University, 2005.
- [5] R. Armeccin, F. Sinon and L. Moreno, "Abaca Fiber: A Renewable Bio-resource for Industrial Uses and Other Applications," in Biomass and

- Bioenergy: Applications, Springer International Publishing, 2014, pp. 107-118. https://doi.org/10.1007/978-3-319-07578-5_6
- [6] B. Sinon, Development of an Abaca Fiber Grade Recognition System using Neural Networks, Visayas State University (VSU), Baybay City, Leyte: Undergraduate Thesis (Department of Computer Science and Technology), 2013.
- [7] C. Cusano, P. Napoletano and R. Schettini, "Combining multiple features for color texture classification," *Journal of Electronic Imaging* 25(6), p. 10, 2016. <https://doi.org/10.1117/1.JEI.25.6.061410>
- [8] L. Tobias, A. Ducourmau, F. Rousseau, G. Mercier and R. Fablet, "Convolutional Neural Networks for Object Recognition on Mobile Devices: A Case Study," in 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 2016. <https://doi.org/10.1109/ICPR.2016.7900181>
- [9] F. Chollet, Deep Learning with Python, San Diego, USA: Manning Publications, 2017.
- [10] M. Ji, L. Liu and M. Buchroithner, "Identifying Collapsed Buildings Using Post-Earthquake Satellite Imagery and Convolutional Neural Networks," *MDPI*, p. 20 pages, 2018. <https://doi.org/10.3390/rs10111689>
- [11] A. Rosebrock, Deep Learning for Computer Vision with Python, California, USA: Pyimagesearch, 2017.
- [12] A. Zisserman and K. Simonyan, "Very Deep Convolutional Neural Networks For Large-Scale Image Recognition," *ICLR*, p. 14 pages, 2015. <https://arxiv.org/abs/1409.1556>
- [13] F. Bianconi, "Theoretical and experimental comparison of different approaches for color texture classification," *Journal of Electronic Imaging*, p. 20, 2011. <https://doi.org/10.1117/1.3651210>
- [14] Y. Lecun, L. Bottou, Y. Bengio and H. P., "Gradient-Based Learning Applied to Document Recognition," *IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998. <https://doi.org/10.1109/5.726791>
- [15] J. Li, G. Liao, Z. Ou and J. Jin, "Rapeseed Seeds Classification by Machine Vision," in Workshop on Intelligent Information Technology Application (IITA 2007), Zhang Jiajie, China, 2007. <https://doi.org/10.1109/IITA.2007.56>
- [16] T. Mäenpää and M. Pietikänen, "Classification with color and texture: jointly or separately?," *Pattern Recognition*, vol. 37, no. 8, pp. 1629-1640, 2004. <https://doi.org/10.1016/j.patcog.2003.11.011>

Machine Learning Model to Identify the Optimum Database Query Execution Platform on GPU Assisted Database

Dennis Luqman*, Sani Muhamad Isa

BINUS Graduate Program-Master in Computer Science, Computer Science Department, Bina Nusantara University Jakarta, 11480, Indonesia

ARTICLE INFO

Article history:

Received: 20 February, 2020

Accepted: 13 May, 2020

Online: 21 May, 2020

Keywords:

GPU Database

Query Processing

GPU co-processor

Machine Learning

ABSTRACT

With the current amount of data nowadays, the need for processing power has vastly grown. By relying on CPU processing power, current processing power is depending on the frequency and parallelism of the current CPU device. This means this method will lead to increased power consumption. Current research has shown that by utilize the power of GPU processing power to help CPU to do data processing can compete with parallel CPU processing design but in a more energy-efficient way. The usage of GPU to help CPU on doing general-purpose processing has stimulated the appearance of GPU databases. GPU databases have gained its popularity due to its capabilities to process huge amount of data in seconds. In this paper we have explored the open issues on GPU database and introduce a machine learning model to enhance the GPU memory usage on the system by eliminating unnecessary data processing on GPU as on certain queries, CPU processing still outperforms the GPU processing speed. To achieve this, we develop and implement the proposed approach machine learning algorithm using python 3 languages and OmniSci 4.7 for the database system. The applications are running on Ubuntu Linux environment as the GPU environment and Docker as the CPU environment and the results we find that KNN algorithm performs well for this setup with 0.93 F1-Score value.

1 Introduction

The amount of data has proliferated every day, which strengthened the need for a high-speed database management system. With the current amount of data that big, very powerful processing powers need also increased. On the other hand, real-time data processing needs have pushed the conventional database into its limits. Digital Universe & EMC estimates that data collected in 2020 will have nearly 44 trillion gigabytes [1]. This situation stimulates the born of another database system such as *Hadoop system, Big Query, Apache Spark, ClickHouse, Amazon Athena, etc.* All of these databases are born with enormous processing power which makes big data processing much faster. However, their processing capabilities depend on the number of nodes and parallelism of the current CPU device, which mean this leads to increased power consumption [2, 3].

In 2013, there is a new database created which gets the attention of some researchers, this kind of database is using graphics processing units (GPUs) to help central processing unit (CPU) on processing the data which makes it very powerful but in an affordable way. The name of this new database is OmniSci. GPU is very famous for its parallel processing [4]. However, due to different

memory architecture between CPU and GPU, data under the main memory cannot directly be accessed by GPU. Hence the data need to be transferred into GPU memory to do data processing on GPU [5]. The data transfer between GPU device and main memory is going through PCI Express bus slot, Nowadays, the latest PCI Express bus on the market is version 5.0 which have a maximum transfer speed of 63.02 GB/s, this behaviour caused the huge processing amount of data on GPU will have I/O bottleneck. Some researchers found the side effect of using GPU as a co-processor which can make a query run slower than CPU only processing. The main challenge of doing data processing on GPU is the data transmission bottleneck while doing non-numerical type data processing [6]. Hence, the query execution time using GPU co-processor not guarantee the processing will be faster compared to the CPU only. Until today, We cannot find research that fully identifies the components of query which still not optimized on the GPU nor use a machine learning model to switch query execution platform between CPU and GPU in a hybrid way.

In this research, we introduce a hybrid approach to select the optimum execution platform processing platform between CPU and GPU with a machine learning model helps. The main focus of this

*Dennis Luqman, Jl. Kebon Jeruk Raya No. 27 Kebon Jeruk Jakarta Barat 11530, +6281807185590, dennis.luqman@binus.ac.id

research is to find out which query is faster on CPU and which on GPU co-process, to do that, we will create an automatic query parser to parse a single query to determine machine learning parameters, and based on obtained parameters, the machine learning model will determine which platform is the best to execute the parsed query. By do query processing platform management, we can also manage the usage of GPU memory to ensure a GPU type query can be executed on GPU.

2 Background & Related Works

2.1 GPU Architecture

GPU is a device that commonly used on a computer or notebook. GPU's primary purpose is to do intensive graphical functions such as watching videos, gaming, or video rendering. As the time being, GPU was started to be used as general-purpose processing [7]. GPU become very popular on general-purpose processing due to its Single Instruction Multiple Data (SIMD) characteristics which will help to boost processing performance on data-intensive computations [5, 8]. However, apart from its processing power, there is a challenge that needs to be faced while utilizing the GPU processor as a CPU co-processor due to its memory architecture.

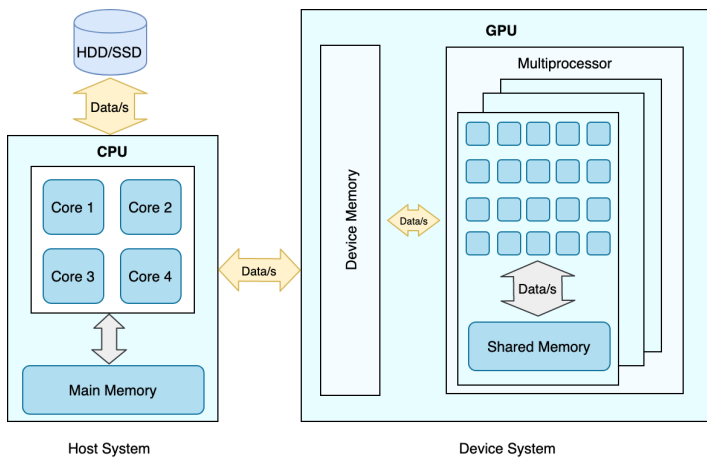


Figure 1: CPU - GPU Architecture [9]

As shown on Figure 1, GPU usually called as a device system, while CPU is called as a host system, the device system is connected to the host system using PCI express bus. Each host and device has its own memory and processors, typically the host and device memory do not share the same address space, which means the device system cannot directly access the host's memory and vice versa. Therefore, to do data processing on GPU, the data need to be transferred into the device memory first. In general, data are stored on a hard drive or solid-state drive, this resulting the data need to pass through host memory then device memory and after the device has finished on processing the data, it will send the data back to host memory in order to show the data to the users [9]. Due to these I/O procedures, the GPU will not help much to improve the processing speed if there is an I/O bottleneck on the system [10].

2.2 OmniSci

When first released, OmniSci is named as Mapd, OmniSci is an open-source SQL-based, relational and columnar type of database which developed to leverage the data processing needs nowadays by utilizing the use of GPU processing power. Its ability to outperform the current big data platform makes it popular to be an alternative solution for big data processing.

A vital component of the OmniSci SQL engine performance advantage is the hybrid or parallelized execution of queries. A parallelized code allows a processor to compute multiple data items simultaneously. This is necessary to achieve optimal performance on GPU, which contains thousands of execution units.

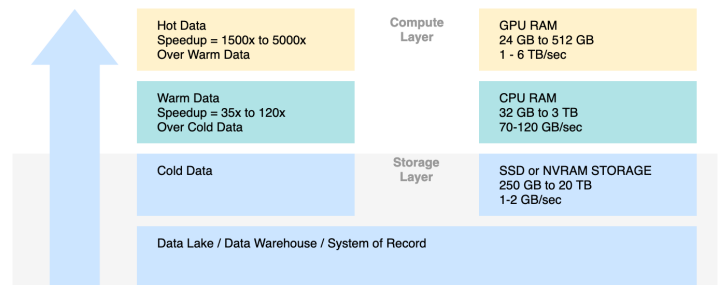


Figure 2: OmniSci Advanced Memory Management [11]

Figure 2 shows that OmniSci has advanced memory management with three-tier caching in their system. The caching contains two layers of computation, storage layer, and compute layer. On the storage layer, it is started with the data itself. The data can be sourced from various sources, i.e. data lake, data warehouse, or system of records, then continue to the third tier of caching called cold data, in this tier the data will be cached on SSD or NVRAM Storage. Moving to the compute layer, there lay second-tier caching called warm data, in this tier the caching will happens in host memory, or usually, we call it RAM, and the first tier called warm data where the caching is happens on the device memory. All three tiers meant to eliminate the transfer overhead between CPU and GPU.

During the query execution, OmniSci system adapts query vectorization and hybrid execution system. This feature allows the system to vectorize the code and compute multiple data items simultaneously across multiple GPUs and CPUs [11].

2.3 Scikit-learn

For machine learning, this paper will use help from Scikit-learn, Scikit-learn is a python framework which provides many popular machine learning algorithm implementations. It is easy to use interface, and well-integrated with python language make this framework can easily be used by data analysis who not specialized in the software and web industries [12]. In this paper, we will test our model using Random Forest, Nave Bayes, Logistic Regression, KNN, and Adaptive Boosting classifier algorithm.

2.3.1 Random Forest

Random Forest is an ensemble machine learning algorithm that can be used to do data classification and regression. Random For-

est is prevalent due to its excellent performance on many data sets. In many cases, Random Forest achieved the best in class performance with higher accuracy than another machine learning algorithm [13, 14], hence we wanted to try this machine learning algorithm with our model to see how it performs.

2.3.2 Nave Bayes

Nave Bayes classification is a straightforward probabilistic model. The model is based on Bayes rule along with a robust assumption of independence. The main characteristic of the Nave Bayes Classifier is a powerful assumption (naive) of independence from each condition or event. The advantage of using Nave Bayes method is that it only requires a small amount of training data to determine the estimated parameters needed in the classification process, and words are conditionally independent of each other. As the drawback, this assumption will slightly affect the accuracy of text classification. However, as an advantage, it will make the high-speed classification algorithm applicable to the problem [15, 16]. Nave Bayes characteristics are matched with what we are looking for in this paper, which is a high-speed classification machine learning algorithm with the highest accuracy.

2.3.3 Logistic Regression

Logistic Regression is a technique that can be used for traditional statistics as well as machine learning. Logistic Regression will work by predicts if something is true or false, 0 and 1, or Yes and No. Logistic Regression is widely used on some classification tasks due to its simplicity and lightweight, and it does not need many computational resources to operate. We choose this algorithm to test with is because it fits our model, whereas there is only two decision that needs to be made, CPU or GPU [16, 17].

2.3.4 K-Nearest Neighbor

K-Nearest Neighbour is a supervised learning algorithm where the result of a new instance classified based on the majority of the nearest K-neighbor category. KNN become popular among classifier algorithms is because of its simplicity, practical, robust, and conceptual clarity. It also can achieve higher accuracy on unknown or non-normal distributed data set [18]. We are choosing KNN as one of 5 machine learning algorithms we test in this paper is because it can perform well in unknown or non-normal distributed data sets, which will fit on our model where usually most user's ad-hoc query is non-predictable.

2.3.5 Adaptive Boosting

Adaptive boosting or in short AdaBoost is a machine learning algorithm introduced in 1995 by Freud and Schapire. The advantage of this algorithm is it fast, simple, and easy to program due to there is only one parameter that needs to be tuned which is the number of rounds [14, 19].

2.4 TPC-H Dataset & Query set

The TPC Benchmark H (TPC-H) is a benchmark dataset for a decision support system. The queries and information provided by TPC-H were selected to make the dataset can have excellent relevance with the industry-wide but in ease of implement manner. TPC-H dataset size can freely be customized based on the user's needs. TPC-H also contains query set for system testing; the query set consist of 22 query type with various complexity. As shown on Figure 3, TPC-H has eight tables, and every relation for each table is one to many relationship[20].

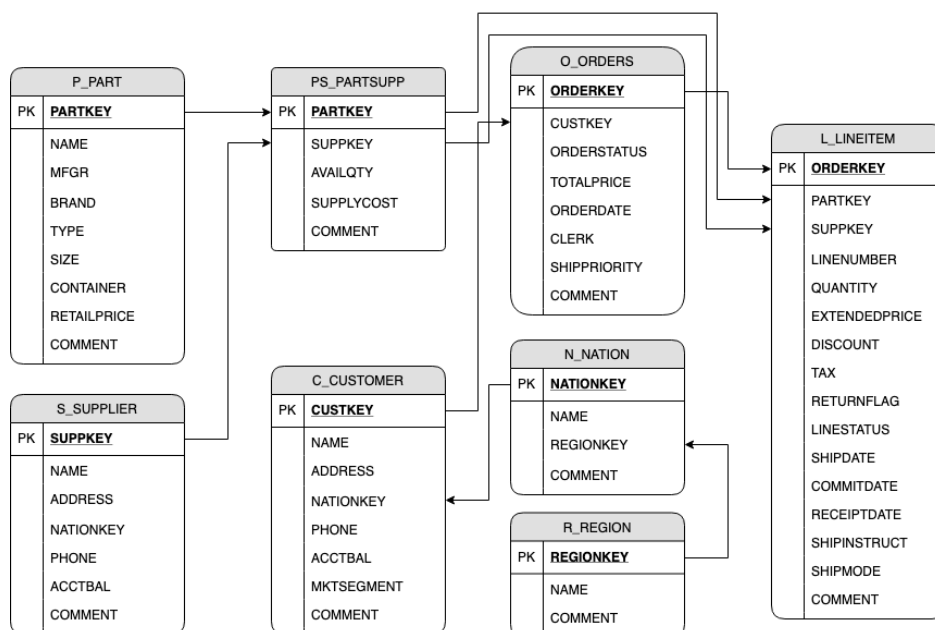


Figure 3: TPC-H Schema [11]

2.5 Related Works

The research on GPU processing sector started in 2004, in this year [21] do research to create graphics card as a co-processor to do data processing. In their research, they focus on speeding up basic database queries such as select and aggregation and got 108x speedup over CPU. This research got more attention from other researchers to continue to investigate the possibility to create a fully operational GPU database. The research on this sector is continued by [4]; they create and implement the use of GPU processing power to SQLite. In their research, they demonstrate the main principle of the GPU database that data processing efficiency on a GPU card depends on the I/O cost. They explain that GPU cannot help to increase the data processing performance if the data need to be processed from physical hard drive due to I/O limitation from hard drive to GPU. This research got 35x speedup over CPU. This research then also continued by [22] by enhancing the Bakkum model to be able to do a Relational database join by translate SQL code into opcode, and as a result, they get 20 to 30 processing times speedup over the CPU processing time.

In 2011, [10] did research on the data transfer between main memory to GPU memory; they explain even though GPU can do a very vast data processing, we also need to pay attention to the transfer time between main memory to GPU. If the transfer time overshadowing the processing time, there will be no speedup obtained from GPU. Hence, they suggest on PCI-Express port usage.

Another research done by [23], they recommend the use of column-major storage is very recommended to maintain the data transfer between host memory and device memory. In their research, they also found that the efficiency of data processing on GPU will depend on the amount of data that will be processed. In the case of small data, the CPU will be the best platform to process the data.

[1] In 2016 conduct an experiment to test GPU as a query accelerator, they are testing geospatial data computation on GPU with help from Mi-Galactica as a GPU accelerator. As the results, they found if the framework execution time is outperforming the Spark one, and they state that GPU based data processing can be an alternative to Big Data.

In 2017, [6] trying to find the problem of GPU speedup which hugely depends on the amount of data processed by creating an algorithm to make a hybrid approach to select the processing platform between CPU and GPU. They call the algorithm as Hybrid Query Processing algorithm.

As shown in Figure 4, the algorithm works by splitting 1 query into subquery block, from this block system will identify which operator being used, what the data type, how big the data and the aggregation type, then the system will do speed comparison to see which platform complete the query faster. However, in this algorithm, the data used to compare the speed is only a sample of 10 top rows which might not represent all rows in the table.

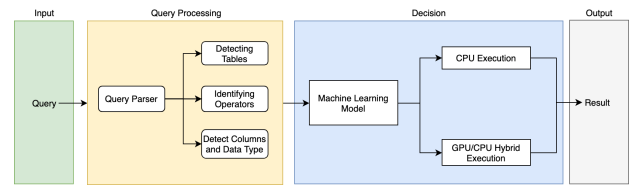


Figure 5: Proposed Framework

3 Proposed framework

In this section, we discuss the detail of the proposed framework. Figure 5 show the overall of the proposed framework. There are four main sections: input, Query Processing, Decision, and Output. The process started by query inputted into the system, then the system will process any inputted query by split the query and extract the information from it. The extracted information then will be pass to the machine learning model, which has a crucial role to determine which platform will be used to run the query. Finally, the system will show the query result in the output section.

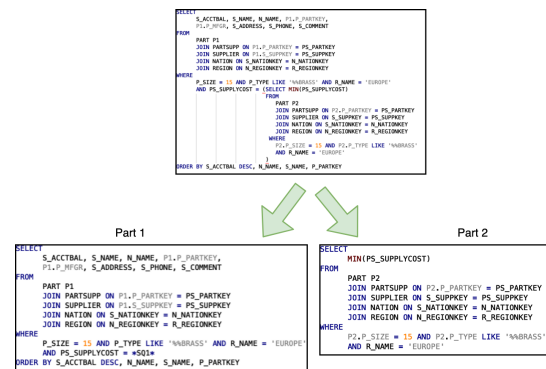


Figure 6: Sub Query Split Illustration

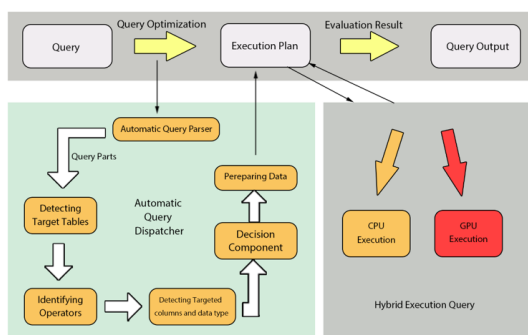


Figure 4: Related Work Framework [6]

3.1 Query parser

After queries are submitted, the system will process the query by parsing it into several parts and extract the information. Query parser part is the most crucial part of the framework because if the system cannot identify a query correctly, the result accuracy will also not be convinced. To do that, we need to identify if there are any subqueries or not. Figure 6 show a scenario where subqueries exist on a query. In this scenario, the system will split the query into two parts, the main query, and the subqueries.

After the query become several parts, the system then will identify query components from each part of inputted queries starting from the number of columns, aggregation type, number of joins, join type, number of filters, number of wild cards, having filter, and

group by and order by columns. Finally, the system will merge the analyzing result of each part of the query then pass it into machine learning.



Figure 7: Identical Feature Generated Sample

In some scenarios illustrated in Figure 7, we found that the query parser result might generate some same result while the selected columns are different. The difference between 2 queries on Figure 7 is only on the column selection. The first query is selecting L_ORDERKEY while the bottom query is selecting S_ADDRESS, but the execution result is showing the bottom query runs faster on GPU co-processor.

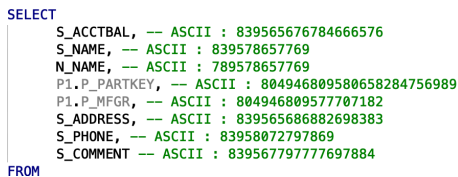


Figure 8: Column to ASCII Conversion Illustration

To overcome this kind of situation, we make the system identify the uniqueness of the selected column by converting the selected column into ASCII number, as shown in Figure 8.

The number then will be summed up and converted into a 5-digit float, i.e., 30873177894140434 will become 30873.177894140434. By using this method, we can get a unique feature combination of the generated data. The real calculation result is shown in Figure 7 on the column_variance feature, while the number_of_column feature is the same, but the ASCII calculation is showing different results.

3.2 Machine Learning Model Methodology

The machine learning model is the second most crucial part of the framework. This part will become the brain of the framework, which will decide the best query execution platform for each inputted query. As illustrated on Figure 9 The machine learning model will work after the system has successfully extract information from the inputted query. The extracted queries become dataset. 75% of the dataset will be sent to the machine learning model to train the model, while another 25% of the dataset will be sent to the trained

machine learning model to decide which platform will be selected to run the query.

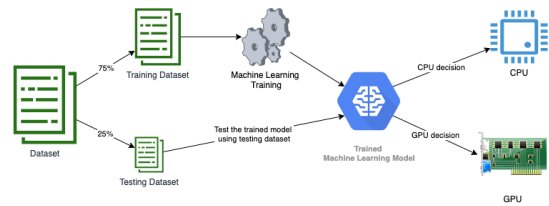


Figure 9: Machine Learning Methodology

For machine learning algorithms itself, we are testing five machine learning algorithms, Random Forest, Nave Bayes, logistic Regression, KNN, and adaptive boost classifier module from scikit-learn. The reason we choose these five machine learning algorithms is due to their simplicity, effectiveness, robustness, and performance reputation on doing the classification. Out of five algorithms, we will choose a machine learning which has the least testing time and high accuracy. Hence the machine learning decision timing will not overshadow the query processing time itself. The machine learning model will make a decision based on the inputted parameters listed in Table 1. These parameters were chosen based on the previous paper [6].

3.3 Generate Training Data

To make machine learning model can effectively select the best platform to run a query, the machine learning need to be trained first. To train the machine learning, we need to create a dataset which will be inputted into the machine learning model and become its knowledge.

To create a dataset, we use the TPC-H queries test set as a base and use a binary model to generate alternatives query. The TPC-H queries test set contains 22 queries, and after using a binary model to generate alternative for each query type, it contains 3365 queries as shown in Table 2.

This binary model was created to make the system can cover all query possibilities which usually user use, the binary model as described in Figure 10.

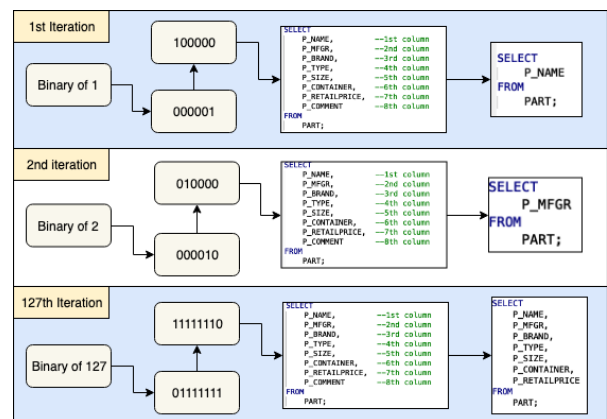


Figure 10: Query Generation Example

Table 2: Total Number of Generated Dataset

Row Labels	CPU	GPU	Total
Type 1	3	5	8
Type 2	254	0	254
Type 3	68	443	511
Type 4	0	14	14
Type 5	164	90	254
Type 6	1	15	16
Type 7	503	8	511
Type 8	511	0	511
Type 9	24	6	30
Type 10	6	219	225
Type 11	30	0	30
Type 12	0	30	30
Type 13	0	12	12
Type 14	1	127	128
Type 15	0	14	14
Type 16	14	0	14
Type 17	2	126	128
Type 18	2	508	510
Type 19	21	107	128
Type 20	0	15	15
Type 21	0	15	15
Type 22	7	0	7
Grand Total	1611	1754	3365

The binary model will work to generate the column selection, grouping and sorting. While for the table name, relation, and aggregation will be determined based on the TPC-H query set. After the query set has been generated, the queries will be run on GPU and CPU sequentially to get the GPU_TIME and CPU_TIME information. The result then will be used to determine the label on each

query. The query labelling will use the following rules:

- $CPU_TIME > (GPU_TIME - GPU_TIME_REBATE) \rightarrow GPU$ Execution
- $CPU_TIME \leq (GPU_TIME - GPU_TIME_REBATE) \rightarrow CPU$ Execution

On the GPU there is a variable called GPU_TIME_REBATE, this variable is defined to give advantage for CPU, in this paper we give GPU_TIME_REBATE value of 0.5 seconds. This value may be variable depends on each user's tolerance. We give CPU advantage due to during our experiment is because we found that many queries have almost similar execution time for both CPU and GPU. Meanwhile, GPU memory was considered more expensive than host memory, therefore if a query runs on GPU but there is no much time gained over CPU, it will be waste of GPU memory, the OmniSci was using caching in their system, and if there is not enough GPU memory for a query to be executed on GPU, those queries will be passed to host system to be processed using CPU. In this case, if the GPU memory was full because of a query that did not have much gain over CPU while there is a query which will have more time gain over CPU need to be executed on GPU but failure due to memory is full, it will be very unfortunate.

3.4 Machine learning model Training

To train the machine learning model, the dataset will be split into a training set and test set with 75 - 25 ratio, 75 % used to train the model, and 25 % used to validate. As shown in Table 3 for machine learning able to cover all query scenarios, the data is divided on the query type level. To validate the results, we also use cross-validation with amounts of 5 folds.

Table 1: Extracted Feature

Feature	Data Type	Description
NUMBER_OF_JOINS	Integer	Showing total number of joins
NUMBER_OF_LEFT JOINS	Integer	Showing number of left joins
NUMBER_OF_INNER JOINS	Integer	Showing number of inner joins
NUMBER_OF_FILTER	Integer	Showing number of filters
NUMBER_OF_COUNT	Integer	Showing number of count
NUMBER_OF_SUM	Integer	Showing number of sum
NUMBER_OF_MIN	Integer	Showing number of min
NUMBER_OF_AVG	Integer	Showing number of avg
NUMBER_OF_COLUMN	Integer	Showing number of columns
NUMBER_OF_GROUPBY	Integer	Showing number of groupby
NUMBER_OF_ORDERBY	Integer	Showing number of orderby
NUMBER_OF_HAVING	Integer	Showing number of having
NUMBER_OF_WILDCARD	Integer	Showing number of wildcards
NUMBER_OF_SQUERY	Integer	Showing Number of Sub Queries
COLUMN_VARIANCE	Float	Result of ASCII value from selected column
GPU_TIME	Float	Showing GPU execution time
CPU_TIME	Float	Showing CPU execution time
DECISION	varchar (3)	Showing label for machine learning

Table 3: Training and Test data Split Result

Row Labels	Total	Training	Test
Type 1	8	6	2
Type 2	254	190	64
Type 3	511	383	128
Type 4	14	10	4
Type 5	254	190	64
Type 6	16	12	4
Type 7	511	383	128
Type 8	511	383	128
Type 9	30	22	8
Type 10	225	168	57
Type 11	30	22	8
Type 12	30	22	8
Type 13	12	9	3
Type 14	128	96	32
Type 15	14	10	4
Type 16	14	10	4
Type 17	128	96	32
Type 18	510	382	128
Type 19	128	96	32
Type 20	15	11	4
Type 21	15	11	4
Type 22	7	5	2
Grand Total	3365	2517	848

4 Experimental Evaluation

4.1 Experimental Setup

The proposed approach has been conducted on a PC with Core i7 3770 processor, 16gb of Ram, Nvidia GTX 970 with 4gb GDDR5 memory. The running OS is Ubuntu 18.04 with CUDA 10, on the software side, we are using scikit-learn version 0.21.3, OmniSci 4.7, and the code is written under python 3 languages. For the database setup, we are using JDBC connection to connect to OmniSci database. Each OmniSci database has been injected with TPC-H dataset; the dataset size was generated with amount of 3gb. We are choosing 3gb to ensure each query can be processed on GPU without memory limit restriction. While configuring database connection, we found that OmniSci GPU/CPU setup cannot be changed dynamically using JDBC, hence to overcome this, we install another set of OmniSci database on docker and configure it to always using CPU as the main processing power, on the other hand OmniSci installed on the Linux is configured to use GPU as the main processing power, the port also needs to be different between them otherwise there will be port conflict and the docker version of OmniSci services won't start. As there is an environmental difference between Linux and docker, we conduct a preliminary study to compare if there is any performance difference between OmniSci running on native Linux and OmniSci running on docker under Linux environment. Furthermore, the result is that there is no performance difference between them.

4.2 Experimental Results

To benchmark the machine learning quality, we are using F1-Score to measure the quality of each machine learning algorithm, in order to validate the result, we are using a cross-validation method with five folds. As the main objective is to find the fastest machine learning algorithm with high accuracy, we also show the training time and testing time. As shown in Table 4 we found that the Random Forest has the best accuracy, while the Nave Bayes algorithm has the best training time Table 5, and logistic Regression algorithm has the best testing time. However, it also has the least accuracy. Based on this result, we choose Random Forest as the most compatible machine learning algorithm, although its training time is higher than Nave Bayes and KNN, the testing time and accuracy is considerably good.

Table 5: Machine Learning Training and Testing Time

Algorithm	Train Time	Test Time
Random Forest	0.033s	0.004s
Nave Bayes	0.003s	0.094s
Logistic Regression	0.073s	0.002s
KNN	0.004s	0.003s
Adaboost	0.722s	0.024s

To illustrate how we calculate f1-score on Table 4, we use the first fold of Random Forest number as an example. To calculate f1-score, we need precision and recall value. We get recall value using the ratio of true positive / (true positive + false positive) which resulting :

$$Precision = \frac{401}{401 + 15} = 0.96394$$

Next, we calculate recall value using ratio of true positive / (true positive + false negative) which resulting :

$$Recall = \frac{401}{401 + 21} = 0.95024$$

After we get precision and recall value, we can calculate f1-score using following formula, $F1 = 2 * (precision * recall) / (precision + recall)$ which resulting :

$$F1 - Score = \frac{2 * 0.96394 * 0.95024}{0.96394 + 0.95024} = 0.95704$$

4.3 Experimental Analysis

The usage of GPU to help CPU on doing general-purpose processing has stimulated the appearance of GPU databases. GPU databases have gained its popularity due to its capabilities to process a massive amount of data in seconds. However, in this research, we found that not all query not always run faster on GPU, as shown on Figure 11, we found that on query type 5, 7, 8, 11, 16, and 22, GPU execution time is not giving much time gain over CPU, while query type 6, 10, 15, 19 have much higher time gain on GPU.

Table 4: Machine Learning Results

Random Forest						
	1st Fold	2nd Fold	3rd Fold	4th Fold	5th Fold	Avg
Precision	0.96	0.95	0.95	0.94	0.95	0.95
Recall	0.95	0.95	0.95	0.94	0.95	0.95
F1-score	0.96	0.95	0.95	0.94	0.95	0.95
Nave Bayes						
	1st Fold	2nd Fold	3rd Fold	4th Fold	5th Fold	Avg
Precision	0.92	0.91	0.92	0.91	0.92	0.92
Recall	0.91	0.89	0.92	0.89	0.91	0.90
F1-score	0.91	0.89	0.91	0.89	0.9	0.90
Logistic Regression						
	1st Fold	2nd Fold	3rd Fold	4th Fold	5th Fold	Avg
Precision	0.22	0.24	0.22	0.24	0.23	0.23
Recall	0.47	0.49	0.47	0.49	0.48	0.48
F1-score	0.3	0.32	0.3	0.32	0.31	0.31
KNN						
	1st Fold	2nd Fold	3rd Fold	4th Fold	5th Fold	Avg
Precision	0.94	0.93	0.92	0.93	0.93	0.93
Recall	0.94	0.93	0.92	0.93	0.93	0.93
F1-score	0.94	0.93	0.92	0.93	0.93	0.93
AdaBoost						
	1st Fold	2nd Fold	3rd Fold	4th Fold	5th Fold	Avg
Precision	0.95	0.93	0.94	0.94	0.94	0.94
Recall	0.95	0.93	0.94	0.94	0.94	0.94
F1-score	0.95	0.93	0.94	0.94	0.94	0.94

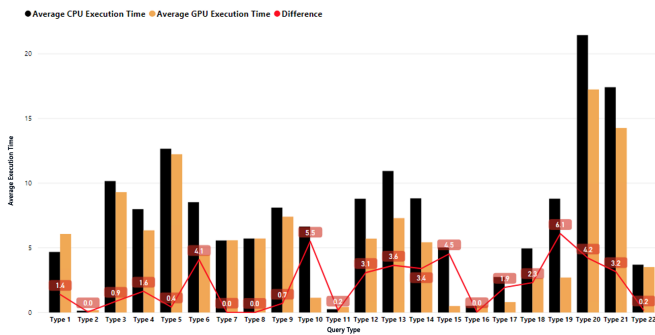


Figure 11: CPU vs GPU Average Run Time

shown that GPU can help on speed up the query processing speed if the query has a lesser amount of group by the operator used.

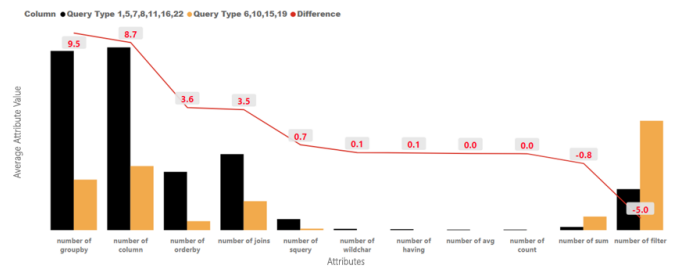


Figure 12: Average Attribute Value by Query Groups

To get closer information, we group the query type 1, 5, 7, 8, 11, 16, 22 and query type 6, 10, 15, 19 and breakdown into the attribute information as shown on Figure 11 There is three information provided by this figure, The black bars, orange bars, and red line. The black bars are representing an average of the attribute value for the query type 1,5,7,8,11,16,22, while the orange bars are representing an average of the attribute value for query type 6,10,15,19. The red line is showing the variance between 2 categorys value.

4.3.1 number_of_groupby

Based on Figure 12, to get the optimal use of GPU processing speed we need to pay attention of this attribute, this attribute represents the number of group by used on the queries, the numbers on Figure 12

4.3.2 number_of_column

The number of columns is related to the number_of_groupby whereas if a query wants to generate an aggregated result within some column information, then group by is mandatory. The difference of this attribute with number_of_groupby is this attribute also calculate the aggregation used. Based on Figure 13, the numbers proved that GPU co-process is not optimized to deal with a query that has a tremendous amount of column selected.

On Figure 13, the system figured that the more column we try to display on a query, the more ineffective GPU coprocessor works, this confirm if there is a transfer overhead between main memory into GPU memory, as the more columns are trying to be processed the more data also need to be transferred into GPU memory.

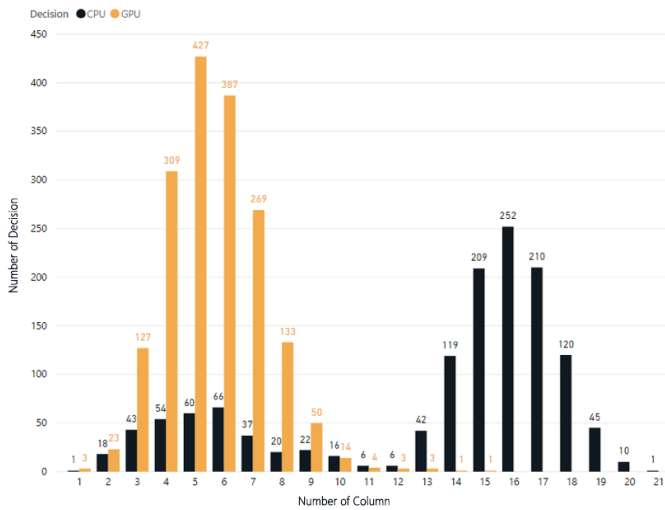


Figure 13: Query Processing Platform Decision Based on Column Count

4.3.3 number_of_orderby

This attribute is affecting the GPU device processing performance, Figure 12 is showing that a higher amount of order by feature used on the query the more ineffective the GPU becomes. However, these numbers might be affected by the number of columns as in general, order by usually applied to a column.

4.3.4 number_of_joins

This attribute indicates how many joins used on a query. On Figure 14 we can see, that GPU only optimized if the query join is not more than three tables. The processing speed on GPU starting to increase on five or more join are involved, which makes the system labelled the query to run on CPU only execution.

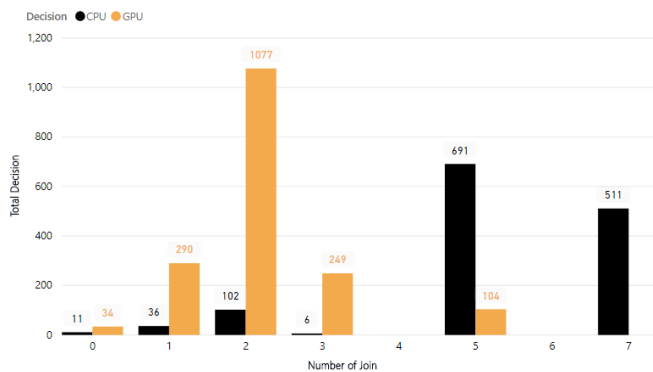


Figure 14: Query Processing Decision Based on Number of Join Used

4.3.5 number_of_subquery

From the chart, we can see if the current GPU database still not optimized for queries with subquery in it. The query type 1, 5, 7, 8, 11, 16, 22 have a higher amount of subqueries in it, and as the results, the system is determining these types of this query to be run on CPU only mode. Therefore, with this result, we can conclude

that if a query has no more than one subquery, it can be run on GPU but, if it has more than 1, the GPU might be not a correct platform to run those queries.

4.3.6 number_of_wildcard

From the chart, we can conclude that usage of wildcard operators on a query is still not optimized to be run on GPU Co-process database. This result also strengthens by Figure 15. In this figure, we can see that there is the least decision made for a query to run on GPU co-processor if there is a wildcard operator used on those queries.

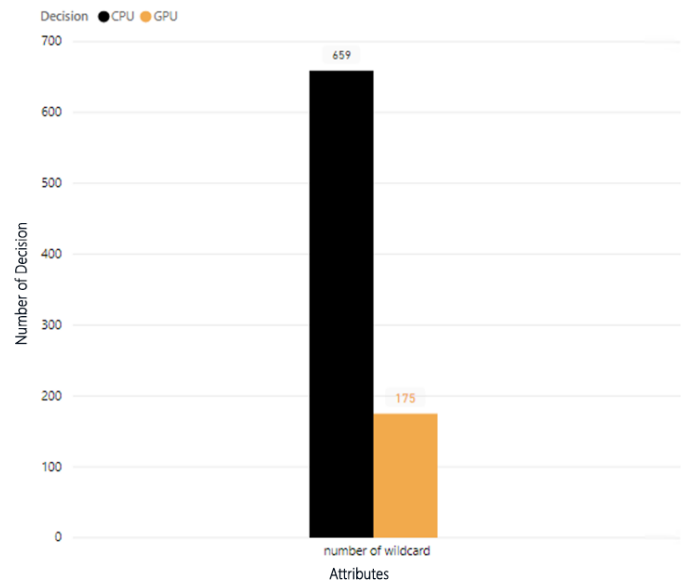


Figure 15: Number of Decision Value Comparison for Query With Wildcard

4.3.7 number_of_having

The chart showed the query type 1, 5, 7, 8, 11, 16, 22 are having a higher number of having attributes, this meaning there might be not the right choice to run a query with having an operator in it. However, this situation also can be changed if the query has another component that strongly optimized on GPU such as a lower number of columns, joins, or a higher number of aggregations.

4.3.8 number_of_avg, number_of_count and number_of_sum

On Figure 11, we can see that query type 1, 5, 7, 8, 11, 16, 22 have a lower number_of_sum, but higher number_of_avg and number_of_count. However, the number_of_avg and number_of_count have a minimal number on it. Hence we reveal the whole number of aggregation, and as we can see on Figure 16, almost all aggregation type query is run faster on GPU, this result confirms the previous research that says the GPU processing is very powerful on an aggregation query type.

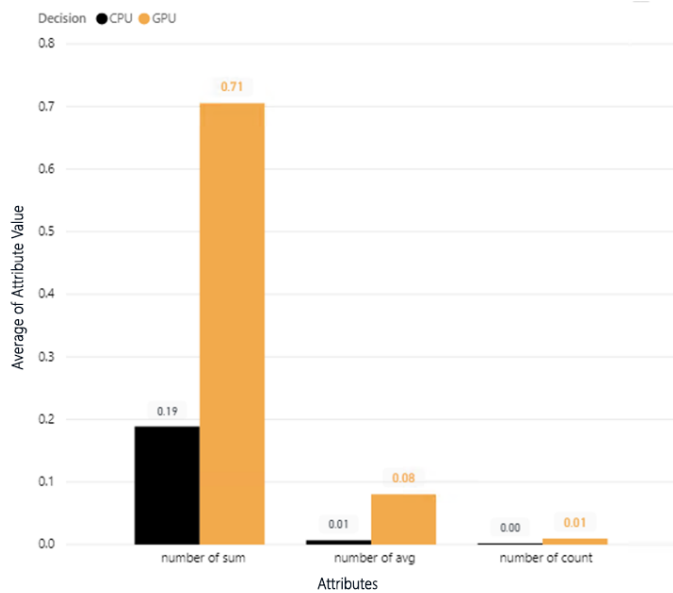


Figure 16: Average Attribute Value for Aggregation

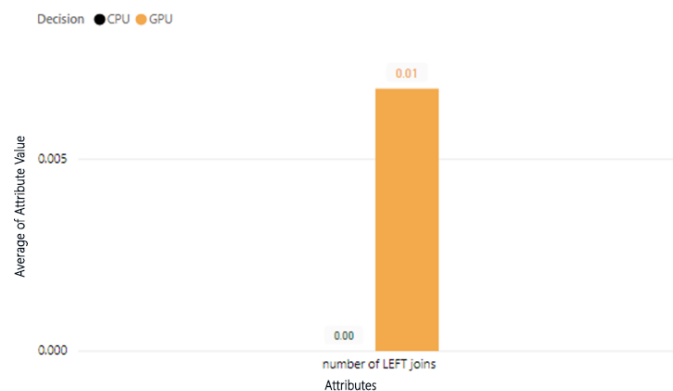


Figure 17: Number of Left Join Value Comparison Between CPU and GPU Decision

4.3.9 number_of_filter

Based on Figure 11, we can see if the GPU can overrun the CPU processor on query with a high amount of filter used on those queries, this proves the GPU SIMD capabilities, as the filtering can be done simultaneously in a parallel way on GPU devices. With this result, we can conclude that if we wanted to run a query with filters on it, GPU co-process could do faster than the CPU only processing. There is a limitation of the above analysis, query Type 1, 5, 7, 8, 11, 16, 22 and query type 6, 10, 15, 19 are do not cover number of left joins; therefore we will separate the analysis for left join. As shown on Figure 17, there is no decision made on CPU when there is left outer join operation query. However, the dataset we use in this research still does not have many outer join operator, which mean this result might be different on another environment where the outer join operator is very commonly used on the real-life ad-hoc queries.

With this result, an appearance of The GPU database is suitable for a simple query with a fewer number of joins and columns selected.

However, the processing speed comparison is significantly higher for query type, which has aggregation in it and a higher amount of filters.

Table 6 is a summarized table for fundamental understanding for an impact of the use of each query function. The Performance impact have High and medium value; this value is determined by comparing the amount of each feature on both GPU and CPU. If the variance is greater than 50% then the value is high, and vice versa, the value will be medium, and if the variance is lower than 20% then the impact grouped as low. As on this research, there is no variance under 20% there is no Low impact grouped.

Table 6: Query Attribute Impact Summary

Feature	Performance Impact	Suitable Platform*
Number Of Joins	High	GPU
Number Of Left Joins	High	CPU
Number Of Inner Joins	High	CPU
Number Of Filter	Medium	CPU
Number Of Count	High	GPU
Number Of Sum	High	GPU
Number Of Min	High	CPU
Number Of Avg	High	GPU
Number Of Column	High	CPU
Number Of Groupby	High	CPU
Number Of Orderby	High	CPU
Number Of Having	High	GPU
Number Of Wildchar	High	CPU
Number Of Squery	Medium	CPU

* The Suitable Platform is valid when feature value has high amount, when the value is low, the suitable Platform result is the opposite.

5 Conclusions & Future Works

In this paper, we have explored the open issues on the GPU database and introduce a machine learning model to enhance the GPU memory usage on the system by eliminating unnecessary data processing on GPU. As on specific queries, CPU processing also still outperform the GPU processing speed, this model also can prevent the system from choosing the processing platform based on the query type wrongly. On the real-world, this approach can be implemented by embed the proposed framework into the upper system. As the OmniSci is a database component which has data related task such as store the data and process the query then pass the processed data into the upper layer, hence to use this database, the upper layer also needed where all the codes and logic happen. In this interface, we can embed the framework and helps the system to choose the right execution platform.

The proposed approach works by automatically parses the input queries, the parsed query then identified by the system to find the parameters that might affect the GPU processing speed, the system then put the query information into the machine learning, and as a result, the system will determine if those input query will be executed on GPU or CPU. To test both CPU and GPU performance on the same system, we use docker as a CPU processing platform, al-

though, to validate the performance result, we conduct a preliminary study to check if there is any speed difference between docker OmniSci vs Linux OmniSci, and as a result, there is no speed difference between them. The machine learning model was validated using cross-validation function within five-folds, and the best result is the Random Forest algorithm with 95% F1-Score value following by AdaBoost on 94% and KNN with 93%. However, from the Random Forest, KNN, and AdaBoost, KNN has the fastest training time, 8.25 times faster than the Random Forest, and 180.5 times faster than Adaboost. Hence with this result, we are choosing KNN as the best algorithm for this framework. From the query side, we also found that usage of filter and subqueries is not affecting the performance difference between CPU only processing vs GPU co-process setup. However, the number of columns on a query is a crucial performance for optimal processing time on a GPU co-processor setup. The more column needs to be shown on a query result, the more challenging for a GPU device to process it.

There are some limitations to this research. First, we are using a middle tier of a graphics card to test the model. Hence if the same method tested on a higher-end tier of graphics cards such as tesla cards, the result might differ. Second, due to OmniSci feature limitation, the GPU data transfer time is not measured on this research. This leads to limited analysis results as we can only see the total query processing time of GPU. Third, due to OmniSci limitations, we unable to identify the data type of each selected column on a query. This may lead to reduced model accuracy as processing time required to process textual data, and numerical data may differ. Forth, the data used for this research is not a real user ad-hoc queries, which mean the query used in this research may not cover all the query scenario used by real users.

As for the future works, this method needs to be tested on a higher-end tier of graphics cards to see if there are still limitations that happened on higher-end tier graphics cards. Second, to get a better insight to do a more detailed analysis, the GPU data transfer time needs to be measured, by measuring the transfer time we can see if the I/O bottleneck happens or not. Third, the query parser may need to be adjusted to calculate the number of numerical columns and non-numerical column selected. This may increase the accuracy of the model. Moreover, this algorithm should be applied to the real-world data and ad-hoc user query where all query types, functions all used on it. The researcher is also planning to extend the work by enhancing the current machine learning model into an unsupervised learning method, where the system will have an ability to learn by itself based on inputted user's query, hence the longer the system work, the smarter it becomes. Unsupervised learning also can reduce the time for initial training in the model.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment We thank Tjeng Wawan Cenggoro from NVIDIA AI R&D Center Bina Nusantara University for helpful feedback and discussions.

References

- [1] KK Yong, Hong Ong, and Vooi Yap. Gpu sql query accelerator. *International Journal of Information Technology*, 22:22, 12 2016.
- [2] Sebastian Bre, Max Heimel, Norbert Siegmund, Ladjel Bellatreche, and Gunter Saake. GPU-Accelerated Database Systems: Survey and Open Challenges. In Abdelkader Hameurlain, Josef Kng, Roland Wagner, Barbara Catania, Giovanna Guerrini, Themis Palpanas, Jaroslav Pokorn, and Athena Vakali, editors, *Transactions on Large-Scale Data- and Knowledge-Centered Systems XV*, volume 8920, pages 1–35. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014. ISBN 978-3-662-45760-3 978-3-662-45761-0.
- [3] Andreas Meister, Sebastian Breß, and Gunter Saake. Toward gpu-accelerated database optimization. *Datenbank-Spektrum*, 15(2):131–140, Jul 2015.
- [4] Peter Bakkum and Kevin Skadron. Accelerating SQL Database Operations on a GPU with CUDA: Extended Results. page 21.
- [5] Iya Arefyeva, David Broneske, Gabriel Campero, Marcus Pinnecke, and Gunter Saake. Memory management strategies in cpu/gpu database systems: A survey. In Stanisław Kozielski, Dariusz Mrozek, Paweł Kasprowski, Bożena Malysiak-Mrozek, and Daniel Kostrzewa, editors, *Beyond Databases, Architectures and Structures. Facing the Challenges of Data Proliferation and Growing Variety*, pages 128–142, Cham, 2018. Springer International Publishing. ISBN 978-3-319-99987-6.
- [6] Esraa Shehab, Alsayed Algergawy, and Amany Sarhan. Accelerating relational database operations using both CPU and GPU co-processor. *Computers & Electrical Engineering*, 57:69–80, January 2017.
- [7] Jaysree Ghorpade. Gpgpu processing in cuda architecture. *Advanced Computing: An International Journal*, 3:105–120, 01 2012.
- [8] Jun Sui, Chang Xu, S.C. Cheung, Wang Xi, Yanyan Jiang, Chun Cao, Xiaoxing Ma, and Jian Lu. Hybrid CPUGPU constraint checking: Towards efficient context consistency. *Information and Software Technology*, 74:230–242, June 2016.
- [9] David B. Kirk and Wen mei W. Hwu. *Programming Massively Parallel Processors: A Hands-on Approach*. Morgan Kaufmann, 2012. ISBN 0124159923.
- [10] Chris Gregg and Kim Hazelwood. Where is the data? Why you cannot debate CPU vs. GPU performance without the answer. In *(IEEE ISPASS) IEEE INTERNATIONAL SYMPOSIUM ON PERFORMANCE ANALYSIS OF SYSTEMS AND SOFTWARE*, pages 134–144, Austin, TX, USA, April 2011. IEEE. ISBN 978-1-61284-367-4.
- [11] *OmniSci Technical White Paper*.
- [12] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct):2825–2830, 2011.
- [13] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.
- [14] Abraham J Wyner, Matthew Olson, Justin Bleich, and David Mease. Explaining the success of adaboost and random forests as interpolating classifiers. *The Journal of Machine Learning Research*, 18(1):1558–1590, 2017.
- [15] Vivek Narayanan, Ishan Arora, and Arjun Bhatia. Fast and Accurate Sentiment Classification Using an Enhanced Naive Bayes Model. In Hujun Yin, Ke Tang, Yang Gao, Frank Klawonn, Minho Lee, Thomas Weise, Bin Li, and Xin Yao, editors, *Intelligent Data Engineering and Automated Learning IDEAL 2013*, pages 194–201. Springer Berlin Heidelberg, 2013. ISBN 978-3-642-41278-3.
- [16] Paraskevas Tsangaratos and Ioanna Iliia. Comparison of a logistic regression and naïve bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size. *Catena*, 145:164–179, 2016.
- [17] David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. *Applied logistic regression*, volume 398. John Wiley & Sons, 2013.
- [18] Jesus Mailló, Sergio Ramírez, Isaac Triguero, and Francisco Herrera. knn-is: An iterative spark-based design of the k-nearest neighbors classifier for big data. *Knowledge-Based Systems*, 117:3–15, 2017.
- [19] Yoav Freund, Robert Schapire, and Naoki Abe. A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence*, 14(771-780):1612, 1999.
- [20] *TPC Benchmark™ H Standard Specification Revision 2.18.0*.
- [21] Naga K. Govindaraju and Dinesh Manocha. Efficient relational database management using graphics processors. In *Proceedings of the 1st international workshop on Data management on new hardware - DAMON '05*, page 1, Baltimore, Maryland, 2005. ACM Press.

- [22] Kevin Angstadt and Ed Harcourt. A virtual machine model for accelerating relational database joins using a general purpose gpu. In *Proceedings of the Symposium on High Performance Computing, HPC '15*, pages 127–134, San Diego, CA, USA, 2015. Society for Computer Simulation International. ISBN 978-1-5108-0101-1.
- [23] Yue-Shan Chang, Ruey-Kai Sheu, Shyan-Ming Yuan, and Jyn-Jie Hsu. Scaling database performance on GPUs. *Information Systems Frontiers*, 14(4):909–924, September 2012.

Evaluation of Type A Uncertainty in a Network Analyzer From 300 kHz to 8.5 GHz

Tan Ming Hui^{*1,2}, Ahmad Yusairi Bani Hashim², Mohd Rizal Salleh²

¹Radio Frequency Calibration Laboratory, National Instruments (M) Sdn Bhd, 11900, Malaysia

²Faculty of Manufacturing Engineering, Universiti Teknikal Malaysia Melaka, 76100, Malaysia

ARTICLE INFO

Article history:

Received: 01 January, 2020

Accepted: 10 May, 2020

Online: 21 May, 2020

Keywords:

Network Analyzer Calibration

Type A uncertainty

Repeatability

ABSTRACT

Network Analyzer is equipment widely used for the execution of radio frequency application scattering parameters. Throughout absolute reading the scattering parameter measured. An absolute reading does not include error, drift, offset, linearity, resolution, coefficient of sensitivity and several other variables that will contribute to the measured measurement dispersion. Type A evaluations of measurement uncertainty clarified according to ISO / IEC Guide 98-1. Type A uncertainty assumption is always done best to characterize an input quantity given in repeated indication values. The assumption was calculated from arithmetic mean, variance of probability distribution and standard deviation respected to the frequency from 300 kHz to 8.5 GHz in a Network Analyzer measurements. Furthermore, ISO/IEC 17025 is the standard for accredited testing and calibration laboratory to calculate the uncertainty of measurement to be declared in the scope of accreditation. Type A uncertainty calculation is one of the mandatory requirements for an accredited testing and calibration laboratory. The Type A uncertainty will be combined with Type B uncertainty to calculate the expanded uncertainty in measurement.

1. Introduction

National Instruments manufacture network analyzer. NI PXIe-5632 Vector Network Analyzer [1] is one of the popular and commonly used models in Radio Frequency (RF) manufacturing. The PXIe-5632 is a Vector Network Analyzer (VNA) able to perform measurements of the scattering parameters which are reflection coefficient and transmission insertion loss. The reflection coefficient measures the voltage reflected from the device under test (DUT). It is known as port 1 VNA reflection coefficient (S11) and port 2 VNA reflection (S22) in a full two ports VNA. The reflected voltage will measure as the mismatch [2] of the DUT. Transmission line from the VNA measured from port 2 to port 1 known as forwards transmission (S21) and reverse transmission port 1 to port 2 (S12) [3]. The transmission voltage will measure as the insertion loss.

The reflected voltage calculated from the prototype or product will help researchers and designers ensure that their design reaches impedance of 50-ohm. It is because RF devices usually use 50-

ohm impedance to apply the principle of maximum power transfer. High reflection voltage will create a high return loss. The return loss will determine by ratio. The best loss of return is 1:1 ratio meaning there is no reflected voltage towards the generator. If metric 4:1 of the return loss is measured as 25% of the reflected power towards the generator [4]. The measured reflected voltage knows as absolute reading. The absolute reading does not contain the errors known as systematic and random errors. ISO / IEC Guide 98-1 [5] clarified the importance of estimated uncertainty would determine the accuracy and performance of the entire measuring system. The measurement uncertainty is calculated based on Type A and Type B contributors.

Type A uncertainty contributor is an evaluation of standard uncertainty based on any valid statistical calculation. The statistical calculation mainly estimated by using the sample of the mean and standard deviation of mean formula. The sample of the mean is based on the sample of repeated measurements under the same condition. The standard deviation of mean is the estimation of Type A uncertainty. The standard deviation of mean will help researchers to perform repeatability [6] and reproducibility [7] statistical analysis. The objective is to set up a stable and accurate test system for manufacturing production. It is mainly to increase

* Tan Ming Hui, National Instruments (M) Sdn Bhd, 8 Lebuh Batu Maung 1, 11900 Bayan Lepas, Penang, Malaysia, Contact No : +604-3776326 & Email: P051510012@student.utem.edu.my ; ming.hui.tan@ni.com

precision and accuracy [8]. Throughout the statistical analysis, the researcher can enhance the system from low accuracy and low precision to a better performance set up. As the results of this paper, it will guide to create a good system with calculating the lower numbers of Type A uncertainty.

This paper is an extension of the work originally presented in the 2016 7th International Conference on Mechanical, Industrial, and Manufacturing Technologies [9], Journal of Engineering Science and Technology Vol. 14, No3 (2019) 1587-1600 [10] and Journal Advances in Science, Technology and Engineering System Journal Vol 5, Issue 3 [11].

2. Network Analyzer PXIe-5632

2.1. Selection of Network Analyzer

Until a manufacturer begins production, engineers are expected to conduct a lot of research until making the decision to invest in any of the production standards. Table 1 contrasts the variations between the network analyzer PXIe-5632, E5080B and ZNB8. Because of the cost-effectiveness and scale, the key reasons PXIe-5632 was chosen. A piece of cost-effective machinery will shorten investment returns. Capital investment risk [12] is low, and profitability is optimised. It is to ensure that the returns on the production product are greater than capital equipment expenditure and very necessary to persuade creditors that a sense of estimating the returns is adequate to cover the debt. Comparison of Network Analyzer between 3 different manufacturers and models was shown in Table 1.

Table 1 : Comparison of Network Analyzer With Other Manufacturers

	Network Analyzer		
	PXIe-5632	E5080B	ZNB8
Cost	Effective	Expensive	Expensive
Frequency Range	300 kHz - 8.5 GHz	9 kHz - 9 GHz	9 kHz - 8.5 GHz
Dynamic Range	110 dB	152 dB	140 dB
Rack Size	4U	8U	8U
Sweep Time	28 mS	3 mS	4 mS
Sweep Points	401	401	401

Every square foot is costing money on the production floor. An 18 slot PXI-Express chassis with 4U (7 Inches) rack size capable of fitting 6 PXIe-5632 Network Analyzer units. Compared to six E5080B and ZNB8 units, they are 36U (63 Inches) and 48U (84 Inches) respectively. As for power consumption, PXIe-5632 Network Analyzer's 6 units consume the same power rate in a single PXIe-1095 chassis with internal clock and trigger synchronization accuracy of timing system working well [13]. E8050B and ZNB8 are, however, going to consume 6 times the rated capacity. The 6 times increase in the rated power would generate a lot of heat on the production floor. Indirectly, the facility team had to raise the flow of air conditioner to sustain at 23 degree Celsius in the production cycle.

As a result, the operating cost increase in production by the volume of E5080B or ZNB8 running.

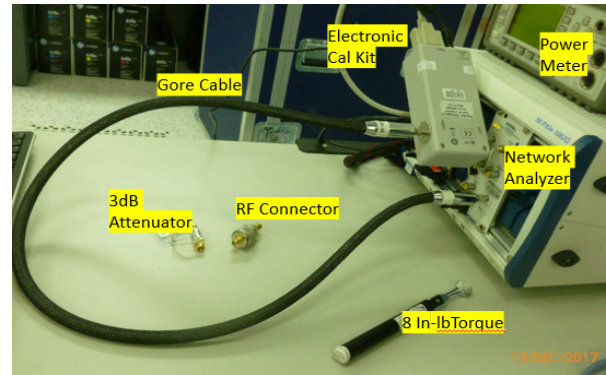


Figure 1: A PXIe-5632 Full 2 Ports Calibration Set Up

Figure 1 demonstrates complete setting up of a PXIe-5632 VNA full 2 ports calibration. It is consisting of a PXIe-1082 chassis and Anritsu electronic calibration module. The PXIe-1082 chassis which can fit with 2 units of PXIe-5632 VNA and work simultaneously. Meanwhile, if the chassis is PXIe-1095, it will fit with 6 units of PXIe-5632 VNA and operate 12 channels at the same time. This configuration is not only cost-effective and required a limited work bench space to run 6 units of PXIe-5632 VNA in the production floor at the same time. The PXIe-5632 owns the advantages of cost effective and space.

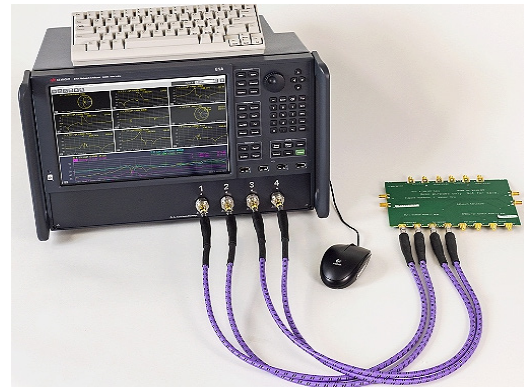


Figure 2 : E5080B 4-Ports Network Analyzer (Adapted from Keysight.com)

Figure 2 shown the Keysight manufactured E5080B 4-ports VNA. It is a highly integrated, complete characterization for radio frequency and microwave impedance measurement solution. The embedded software installed in the E5080B enables complete device characterization for passive device and components for researcher [14]. This VNA package with an extra wide bandwidth of up to 152 dB dynamic range feature improves the RF component and system testing throughput. It is approximate 10 times faster than PXIe-5632 but in terms of price, it is expensive in capital equipment investment.

Figure 3 shown the ZNB8 network analyzer with the mechanical calibration kits used in an electrical laboratory to perform characteristic impedance calibration. ZNB8's key benefit is this VNA capable of extending up to 48 ports by using the matrix switches [15]. The switch matrix benefit is that it allows 48 channels to function simultaneously in output. This feature helps production reduce the waiting and testing period substantially, but the investment and maintenance plan cost will also be high.

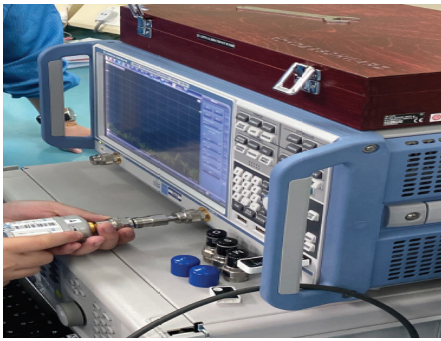


Figure 3 : A ZNB8 Network Analyzer

As a result, the operating cost increase in production by the volume of E5080B or ZNB8 running in a manufacturing production floor compare to PXIe-5632 is not merely cost-effective in terms of capital investment and manufacturing space savings as well.

2.2. Recommended Equipment for PXIe-5632 Calibration

From Table 1, it was found that the PXIe-5632 VNA is low in capital investment and physical small in size are the two main advantages compare to E5080B and ZNB8. The low investment in capital equipment will shorten the return of investment. The PXIe-5632 hardware small in size helps production significantly reduce the space for their test system. Table 2 shown the recommended equipment for PXIe-5632 during calibration.

Table 2 : Recommended Equipment for PXIe-5632 Calibration

Equipment	Model	Where Used	Minimum Requirements
PXI Express Chassis	PXIe-1095	All Tests	3 Adjacent, empty PXI Express slots
PXI Express Controller	PXIe-8880	All Tests	1 System Slot
Precision Coaxial Cable	K-Type (NI Part Number 781611-01)	Scattering Parameter Calibration	Low Loss RF Cables
Automatic VNA Calibration Kit	NI Part Number 782364-01	Automatically Calibrating Full 2 Ports Calibration	K-Type, Male and Female calibration kits up to 9 GHz

To operate the PXIe-5632 a PXI chassis with model PXIe-1095 was required. The PXIe-1095 available with 18 slots, where it can slot and work simultaneously in 6 units of PXIe-5632 VNA. The PXIe-1095 chassis equipped with a PXIe-8880 controller. Only 1 controller needed to operate all the 18 slots in a chassis. An automatic VNA calibration kit used to perform full 2 ports calibration. Figure 1 provided a complete calibration setup for the PXIe-5632 characteristic impedance calibration.

After full 2 ports calibration completed shown in Figure 4, the device under test (DUT) measurement is ready to be measured. In this analysis, a mini-circuit fixed attenuator VAT-3 used to measure the Type A analysis. It is a 3 dB insertion loss fixed attenuator. The

fixed attenuator forwards transmission (S21) and reverse transmission port 1 to port 2 (S12) analysis shown in Figure 5.

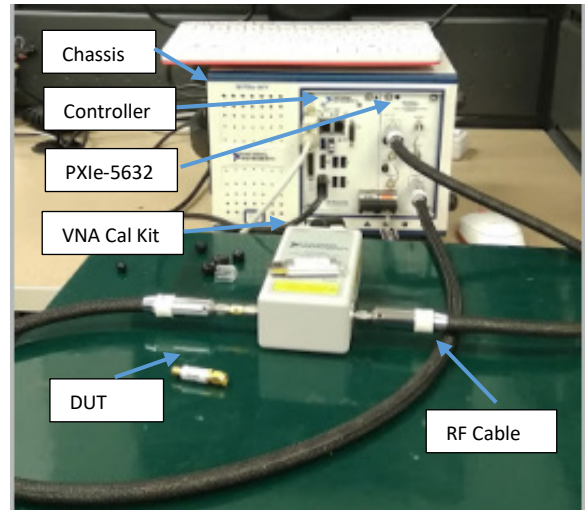


Figure 4 : A set up of PXIe-5632 Calibration

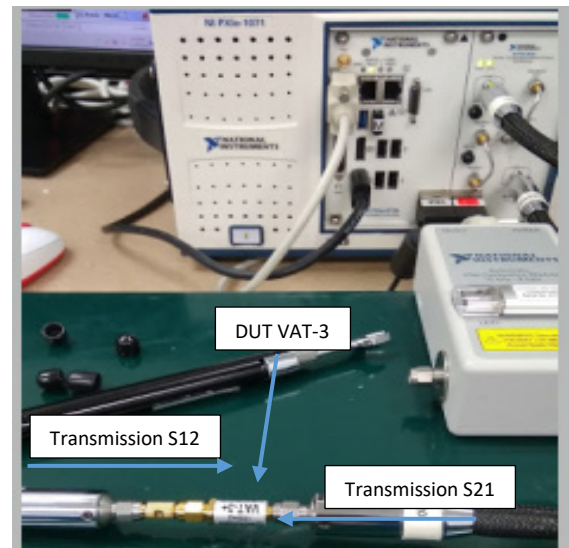


Figure 5 : A Set up of PXIe-5632 Calibration

Type A measurement would acquire a small number of repeated measurements of N number to populate a precision Type A measurement. The minimum of repeated measurement is 2. It is to calculate a small number of standard deviations. The smaller number of standard deviation calculated, the better performance of the network analyzer. In this case study, it is applying 5 repeated measurement under the same condition including equipment, personal and environmental condition. The objective is to minimize the error and standard deviation.

3. Estimation of Type A Uncertainty Calculation

Evaluation of measurement uncertainty classified into two categories Type A and Type B. Type A evaluation evaluates the uncertainty resulted from the statistical study. The statistical analysis may be an undergrowth series of repeated measurements of the same process. The series of repeated measurements will fill in a standard deviation of mean to calculate the Type A uncertainty.

3.1. Arithmetic Mean in Type A Uncertainty Analysis

The estimation of Type A evaluation of uncertainty can be applied when a set of measurement was recorded under the same condition. Minimum repeated measurement is 2. Example a quantity of Y input with N statistically independent ($N > 1$) observed as y_j (where $j = 1, 2, 3, \dots, n$). The estimation of quantity of Y is \bar{y} , arithmetic mean wrote as Equation 1 [16].

$$\bar{y} = \frac{1}{N} \sum_{j=1}^N y_j \quad (1)$$

The system measured in the VNA under test may differ each time it under repeated condition. Even though it is the same set of equipment and the technician performs the measurement in production, it is difficult to determine the true value [17] from 300 kHz to 8.5 GHz frequency swept. This may be caused by various factors such as environmental condition, input line frequency stability, random error effect, impedance mismatch and others. Estimates the mean will help harmonize the cause of effect in the PXIe-5632 VNA towards the measurement. The higher sample size of repeated measurement will be resulting the measurement in lower noise.

3.2. The variance in Type A Uncertainty Analysis

The arithmetic means of \bar{y} is evaluated in Equation 1 will applied in Equation 2 to estimate the variance of probability distribution define as $s^2(y)$ [16]

$$s^2(y) = \frac{1}{N-1} \sum_{j=1}^N (y_j - \bar{y})^2 \quad (2)$$

The size of the arithmetic mean samples determined by PXIe-5632 VNA which corresponds to Equation 2. In this study, it is very important that to keep the probability of variance distribution in a small number. The smaller number of the distribution of variance reflects the capability of PXIe-5632 VNA measured in good repeatability. Although the PXIe-5632 measured in good repeatability condition, it does not determine the system is accurate. The PXIe-5632 VNA framework must be validated against another external reference standard [18], which is specified in discussion.

3.3. Type A Uncertainty Analysis

After estimation of arithmetic mean and variance had been calculated, the experimental standard deviation for Type A uncertainty was calculated in Equation 3 [19].

$$s(\bar{y}) = \frac{s^2(y)}{N} \quad (3)$$

The standard uncertainty $u(\bar{y})$ associated with the initial set of measurement repeated \bar{y} is the experimental standard deviation of mean as shown in Equation 4.

$$u(\bar{y}) = s(\bar{y}) \quad (4)$$

The key purpose of repetitive analysis is to disperse the data into a predictive waveform or pattern. To fill a Type A precision calculation, a high-stability system will acquire several repeated measurements of N (example 5). Indirectly, it calculates a small number of standard deviations. Calculating the smaller number of standard deviations resulting the PXIe-5632 measured precisely. If the standard deviation analysis did not calculate in small number, the higher sample size of repeated measurement is required.

4. PXIe-5632 Type A Uncertainty Calculation

After 5 repetitive measurements were taken by PXIe-5632, a statistical analysis to analyse the data. Figure 3 indicates the unit under test swept at 50 MHz calculated S21 and S12 with a total number of 201 points to approximate the nominal 3 dB error population as guide. The maximum and the minimum were measured at respectively 3.06 dB and 2.94 dB. The error was calculated as 0.02% from 3 dB nominal reference point. A linear graph was plot, and the formula was determined as $y=0.5012x+0.0147$. R square was calculated at 0.3564. R square shown in Figure 3, calculate 35.64% variance distributed along the linear line at $y=0.5012x+0.0147$ as shown in Figure 6.

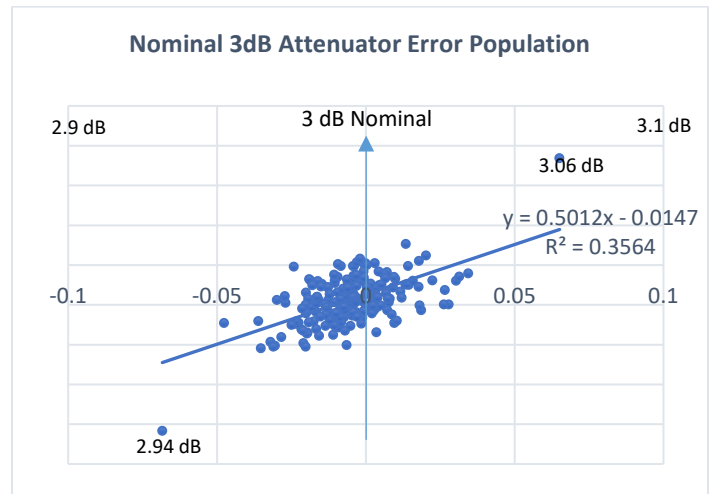


Figure 6 : Nominal 3 dB Fixed Attenuator Error Population

The error calculated 0.02% from 3 dB with R square 35.64% conclude that the measurement variability of the dependent at 3 dB, where approximately at 71 points. The 71 points are distributed from 3.0075 dB (maximum) until 2.9983 dB (minimum). The remaining 130 points are calculated as 64.36% of the variability is still uncounted from total number of 201 points. However, the maximum and minimum error calculated at ± 0.06 dB and it was measure at 3.06 dB and 2.94 dB respectively. The error of ± 0.06 dB was found insignificant respect to the product specification limit at ± 0.15 dB. The population lies within the 3 dB attenuator specification.

A nominal 3 dB fixed attenuator repeatability distribution was plotted as shown in Figure 4. The population distribution close to normal distribution trend with total number of 199 points lies within the 95% confidence interval. There are only 2 points was measures outside the 95% confidence interval as shown in figure 4.

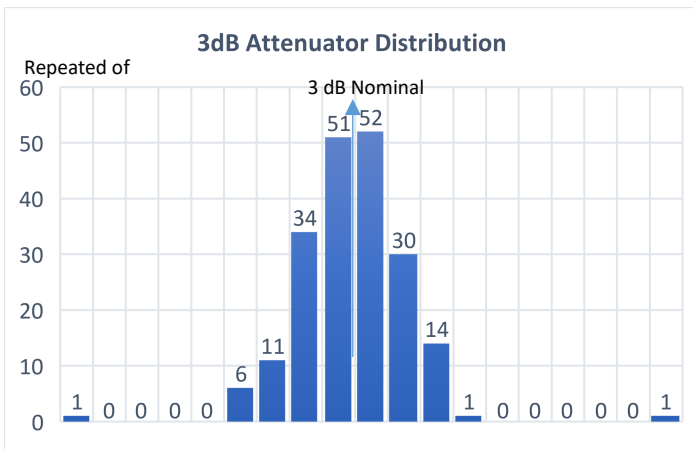


Figure 7 : Nominal 3 dB Fixed Attenuator Distribution Trend

The results from Figure 6 and Figure 7 were found satisfaction, precision and accurate. It can conclude the system repeatability is good and stable.

4.1. PXIe-5632 Experimental Transmission S21 Analysis

Measurement of transmission S21 from port 2 of PXIe-5632 to port 1. It is also known as forward transmission measurement. The configuration shown as Figure 2. The concept is PXIe-5632 VNA power generated from port 2 transmitted to the DUT and received the power at PXIe-5632 VNA port 1. The power is transmitted over the DUT and received at port 1 also known as insertion loss. Equation 1 required minimum 2 repeated measurement to calculate the arithmetic mean. In this experiment, a total of 5 repeated measurements swept from 300 kHz to 8.5 GHz. Frequency swept at 8.5 GHz is the maximum capability for PXIe-5632 VNA. The 5 repeated measurement was proceeded with the same set of equipment used and perform by same technician in 5 consequences day in calibration laboratory. The repeated measurement results obtained in Table 1.

Table 3 : 5 Repeated Measurement S21 Gathered from PXIe-5632

Frequency	Run 1	Run 2	Run 3	Run 4	Run 5
GHz	S21	S21	S21	S21	S21
0.0003	-2.935	-2.941	-2.939	-2.943	-2.936
1.0	-3.087	-3.097	-3.091	-3.090	-3.094
2.0	-3.145	-3.155	-3.156	-3.156	-3.157
3.0	-3.176	-3.188	-3.193	-3.193	-3.193
4.0	-3.273	-3.289	-3.300	-3.301	-3.288
5.0	-3.214	-3.236	-3.257	-3.252	-3.237
6.0	-3.348	-3.374	-3.399	-3.399	-3.377
7.0	-3.381	-3.408	-3.425	-3.421	-3.408
8.0	-3.469	-3.505	-3.510	-3.504	-3.495
8.5	-3.530	-3.559	-3.560	-3.562	-3.557

From Table 3, minimum and maximum insertion loss S21 was measured at -2.935 dB and 3.562 dB at frequency 300 kHz and 8.5 GHz respectively. The frequency swept at 7 GHz , 8 GHz and 8.5 GHz are beyond manufacturer specification. It is expected when the frequency increased will results higher error and drift.

FREQ. RANGE (MHz)	ATTENUATION * (dB)					VSWR (-:1)			MAX. INPUT POWER (W)		
	DC-3 GHz		3-5 GHz	5-6 GHz	DC-6 GHz	DC-3 GHz	3-5 GHz	5-6 GHz			
$f_1 - f_2$	Nom.	Typ.	Typ.	Typ.	Typ.	Typ.	Max.	Typ.	Max.	Typ.	
DC-6000	3±0.3	0.20	0.15	0.15	0.45	1.05	1.20	1.15	1.40	1.40	1.0

* Attenuation varies by 0.3 dB max. over temperature.
 ** Flatness= variation over band divided by 2.

Figure 8 : Mini Circuit VAT-3 Manufacturer Specification

Figure 8 was taken from the manufacturer's manual Mini-Circuit VAT-3 Fixed Attenuator product specification. The manufacturer warranted specification only up to 6 GHz. This experiment was a case study of Type A uncertainty up to 8.5 GHz. The frequency swept from 6 GHz to 8.5 GHz does not adhere to the manufacturer specification. Although the frequency swept beyond the manufacturer warranted specification, this study only required a precise in repeated measurement. However, the repeated measurement study will helps to plot the statistical analysis and the VAT-3 fixed attenuator characteristic. The statistical analysis will calculate the error dispersion from nominal 3 dB attenuation. This error also contribute by drift [20] proportional to the PXIe-5632 VNA frequency increased. The precise repeated measurement will helps to reduce the Type A uncertainty in calculation.

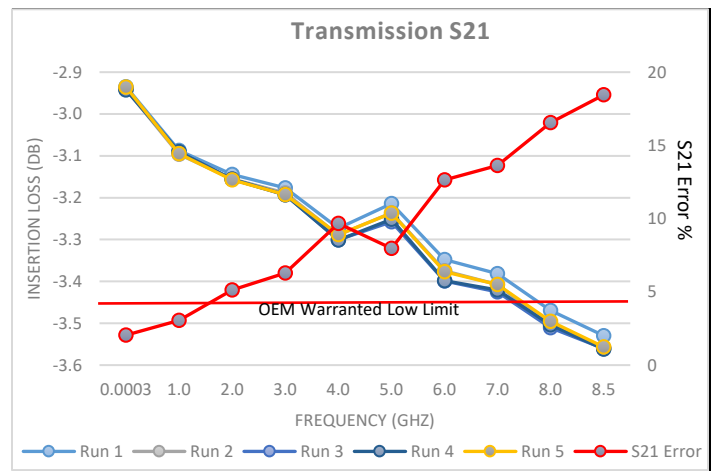


Figure 9: Transmission S21 With Repeated Number of 5

Figure 9 shown the transmission S21 with repeated number of 5 measurement. From 300 kHz found the lowest error calculated as 2.47%. The maximum error was calculated as 18.44% from nominal 3 db reference point. The VAT-3 fixed attenuator warranted specification only up to 6 GHz cut off at -3.45 dB, but this device remain stable up to 8.5 GHz with maximum error less than 20 % from the nominal points. From figure 9, it can conclude that the attenuation loss and error propotional with frequency increased from nominal point at 3 dB.

4.2. PXIe-5632 Experimental Transmission S12 Analysis

Transmission S12 measurement is measurement from PXIe-5632 port 1 to port 2. It is also known as reverse transmission. The set up was shown as Figure 5. The methodology applied from transmission S21 to perform the transmission S12 analysis. The only different the PXIe-5632 VNA measured the transmission S12 opposite direction from transmission S21. The power will be transmitted from Port 1 and received in Port 2 at PXIe-5632. Table

4 shown the 5 repeated measurement gathered under the same calibration condition.

Table 4 : 5 Repeated Measurement S12 Gathered from PXIe-5632

Frequency	Run 1	Run 2	Run 3	Run 4	Run 5
GHz	S12	S12	S12	S12	S12
0.0003	-2.946	-2.952	-2.951	-2.953	-2.951
1.0	-3.108	-3.111	-3.112	-3.114	-3.111
2.0	-3.159	-3.172	-3.169	-3.171	-3.171
3.0	-3.182	-3.197	-3.204	-3.199	-3.194
4.0	-3.278	-3.298	-3.309	-3.304	-3.297
5.0	-3.216	-3.239	-3.261	-3.256	-3.239
6.0	-3.365	-3.394	-3.417	-3.409	-3.390
7.0	-3.395	-3.427	-3.440	-3.439	-3.424
8.0	-3.470	-3.503	-3.510	-3.508	-3.497
8.5	-3.537	-3.569	-3.570	-3.569	-3.562

From Table 2, minimum and maximum measurement was measured at -2.946 dB and 3.570 dB respectively. The minimum and maximum measured were very close to S21 measurement in Table 1.

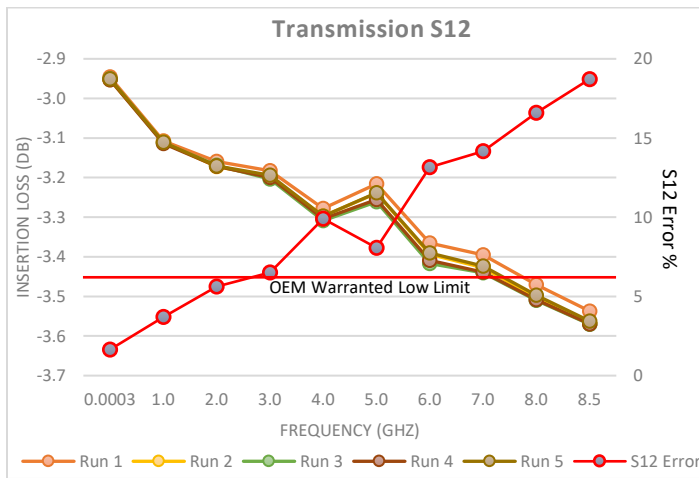


Figure 10: Transmission S12 With Repeated Number of 5

Figure 10 shown the transmission S12 with repeated number of 5 measurement. From 300 kHz found the lowest error calculated as 1.64%. The maximum error was calculated as 18.72% from nominal 3 db reference point.

4.3. Mini-Circuit VAT-3 Transmission S21 and S12 Measurement Error and Limit

Table 5 indicates repeated measurement error on transmission S21 and S12. Transmission S12 error was found from Table 5 to calculate a minimum of 0.217% and a maximum error of 18.926%. The Mini-Circuit VAT-3 was swept beyond its manufacturer specification, so it's important to measure the frequency error maximum. The function of the S21 and S12 limits is used to monitor the repeatability of Type A analysis and identify any deviations or variations in performance by the measurement system as shown in Figure 11. If the S21 or S12 error is measured beyond the control limits, it is required to carry out validation and verification again from chapter 4.1. The main objective to perform

validation and verification again is to detect the characteristic change from the equipment used or the device under test.

Table 5 : S21 and S12 Repeated Measurement Error

Frequency	Measurement Error (%)				
	S21 Error	S12 Error	tolerance	New S21 Limits	New S12 Limits
0.0003	0.461	0.217	0.18926	0.548	0.259
1	5.808	6.228	0.18926	6.907	7.407
2	5.784	6.165	0.18926	6.879	7.332
3	5.598	5.867	0.18926	6.657	6.977
4	10.159	10.521	0.18926	12.082	12.513
5	8.889	9.124	0.18926	10.572	10.851
6	12.694	13.173	0.18926	15.097	15.667
7	11.310	11.946	0.18926	13.451	14.207
8	14.746	15.032	0.18926	17.537	17.877
8.5	18.646	18.926	0.18926	22.175	22.508

The standard used characteristic changed may affect the accuracy in repeated measurement. The inaccuracy of the measured may cause by mishandling of the equipment or the equipment degraded over the time being used in production. The same causes happen to the device under test. The new limit is very important to maintain the stability and repeatability measurement and to ensure the system always running in optimum condition.

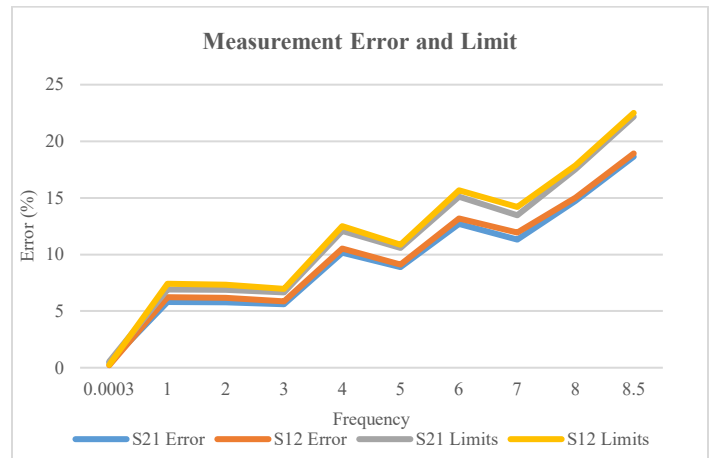


Figure 11: Measurement Error and Limit Control Chart

Figure 11 shown the Transmission S21 and S12 characteristic are close to each other. The device under test, Mini-Circuit VAT-3 fixed attenuator is a known good attenuator and high stability device eventhrought the calibration was proceed approximate 20% beyond its warranted frequency specification.

4.4. The variance of S21 and S12

The variance of S21 and S12 is important to distinguish between the population from the data measured across the frequency. The variance of S21 and S12 were calculated from Equation 2. The variance is calculated from the 5 sample of measurements in S21 and S12. It was found that the observation

does not vary from the group of mean swept from frequency 300 kHz until 8.5 GHz.

The calculated variance concludes that the 5-sample size of mean provide a low number of variances. This experiment does not require more sample of measurement. If the variance calculates a larger number, it represents that the individual frequency results was vary greatly from the group of mean. It could be contributing by any factor such as personal, skills, equipment, environmental condition and other factors.

Table 6: Variance S21 and S12

Frequency GHz	Variance	
	S21	S12
0.0003	1.09857E-05	7.0013E-06
1.0	1.36665E-05	4.8213E-06
2.0	2.58515E-05	2.7513E-05
3.0	5.26069E-05	6.5316E-05
4.0	0.000132802	0.00014041
5.0	0.000286582	0.00031072
6.0	0.000448815	0.00039954
7.0	0.000292401	0.00032702
8.0	0.000262793	0.00026651
8.5	0.000180771	0.00019919

The results from Table 6 is ready to calculate the Type A uncertainty.

4.5. Type A Uncertainty Calculation

The Type A uncertainty calculated by applying the Equation 3. Standard deviation is calculating for each frequency points. It is several variations between each data point relative to the mean. The calculated standard deviation in Table 7 will be the calculated Type A uncertainty as per Equation 4.

Table 7: Standard deviation of S21 and S12

Frequency GHz	Standard Deviation	
	S21	S12
0.0003	0.003314	0.002646
1.0	0.003697	0.002196
2.0	0.005084	0.005245
3.0	0.007253	0.008082
4.0	0.011524	0.01185
5.0	0.016929	0.017627
6.0	0.021185	0.019989
7.0	0.0171	0.018084
8.0	0.016211	0.016325
8.5	0.013445	0.014114

From Table 7, the minimum standard deviation for S21 measure at 0.003314 at 300 kHz meanwhile S12 measure at 0.002196 at 1 GHz. The minimum Type A uncertainty represent the best repeatability measured in the system. The maximum

standard deviation for S21 and S12 were calculated as 0.021 and 0.019 respectively at 6 GHz. The worst case [21] of standard deviation demonstrate the maximum deviation at 6 GHz. From the worst case Type A uncertainty, it can observe that occurs at the maximum frequency point at warranted specification. As the conclusion from Table 7, the minimum and maximum standard deviation did not proportional towards the frequency increased. The minimum and maximum standard deviation could scatter between the frequency range from 300 kHz until 8.5 GHz.

4.6. Type A Uncertainty Associate with Transmission S21 and S12 Measurements

Table 8 and Figure 12 shown the Type A uncertainty evaluation reported into the device under test measurement report. The Type A uncertainty reported respect to its frequency point. The reading of transmission S21 and S12 were the magnitude of absolute measurement from PXIe-5632 VNA and associate with Type A uncertainty. Each of the individual Type A uncertainty will be computed with Type B uncertainty to calculate the total expended uncertainty expended uncertainty which the total expended uncertainty should reported in 2 decimal points [11].

Table 8 : An Example of Calibration Report With Type A Uncertainty

DUT Calibration Report				
Freq. (GHz)	S21	Type A Unc.	S12	Type A Unc.
0.0003	-2.939	0.0033	-2.951	0.0026
1	-3.092	0.0037	-3.111	0.0022
2	-3.154	0.0051	-3.168	0.0052
3	-3.189	0.0073	-3.195	0.0081
4	-3.290	0.012	-3.297	0.012
5	-3.239	0.017	-3.242	0.018
6	-3.379	0.021	-3.395	0.020
7	-3.408	0.017	-3.425	0.018
8	-3.497	0.016	-3.498	0.016
8.5	-3.553	0.013	-3.562	0.014

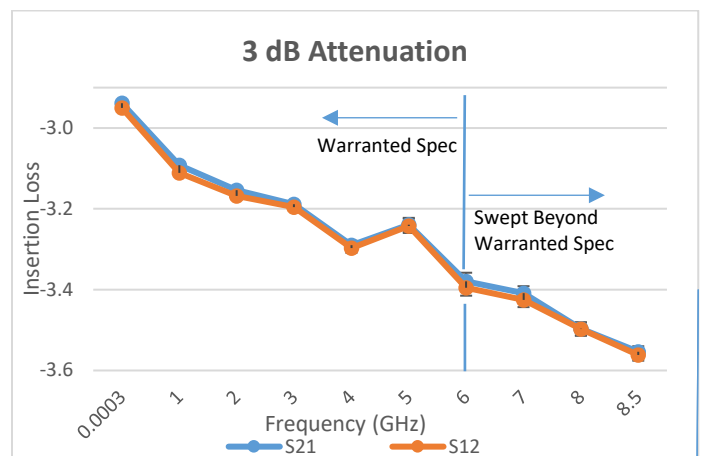


Figure 12: Transmission S21 and S12 associate with Type A Uncertainty

Repeated measurement concept is involved multiple measurement of each frequency. Under the repeated measurements condition, the device under test will exposed the best and worst

conditions from the calibration report. It is a statistical measurement calculate the variation in the response to the variation variable that is cause by random and systematic errors. The random and systematic effect will be analysed when performing the Type A and Type B uncertainty.

5. Discussion

5.1. Validate Type A Uncertainty by Using Correlation S21 and S12 Method

The correlation between S21 and S12 is important to ensure the performance of the Network Analyzer generator and receiver measured the power precise and accurately. The correlation of S21 and S12 calculated as shown in Table 9. The mean of S21 and S12 were calculated from Table 3 and 4 respectively by applying the equation 1 to verify the correlation of the transmission measured by the PXIe-5632 VNA. In ideal case, both transmission S21 and S12 should be the same magnitude in term of forward and reverse insertion loss. It is because the same device under test measured at the same condition. However, this theory does not apply in the experimental. There are gaps identified in the calibration.

Table 9 : Delta of Mean Calculated From S21 and S12

Frequency GHz	Mean		
	S21	S12	Δ Mean
0.0003	-2.9386	-2.9506	0.0120
1.0	-3.0919	-3.1111	0.0192
2.0	-3.1537	-3.1684	0.0146
3.0	-3.1888	-3.1954	0.0066
4.0	-3.2903	-3.2970	0.0067
5.0	-3.2392	-3.2420	0.0028
6.0	-3.3792	-3.3949	0.0157
7.0	-3.4085	-3.4251	0.0166
8.0	-3.4968	-3.4977	0.0008
8.5	-3.5534	-3.5616	0.0082

Figure 13 shown the correlation of S21 and S12 from 300 kHz until 8.5 GHz. From the graph, it was found that the delta of mean is random. It does not proportional to frequency. The maximum delta of mean is 0.0192 dB at frequency 1 GHz and follow by 7 GHz and 6 GHz. Meanwhile, the minimum delta of mean measured at 8 GHz only differentiate with 0.0008 dB.

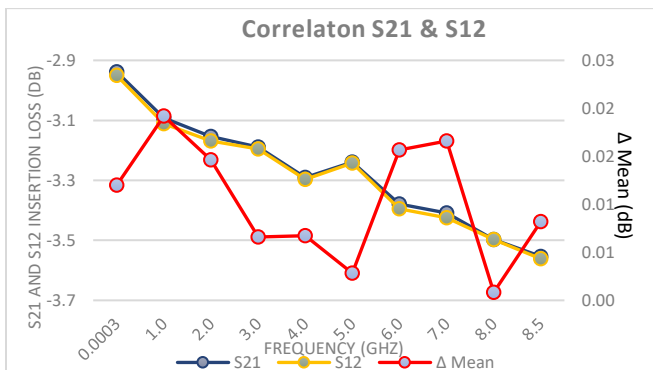


Figure 13 : Correlation S21 and S12

Overall Figure 9 conclude that strong correlation between S21 and S12 trend crossing each other from 300 kHz until 8.5 GHz. The PXIe-5632 VNA port 1 and 2 are two different receivers to measure the signal received. The strong correlation concludes the PXIe-5632 VNA port 1 and port 2 measured insertion loss very close to each other by two different sensors applied in the system.

5.2. Verification of PXIe-5632 VNA Results with Other Original Equipment Manufacturer

Performing the Type A uncertainty analysis only with the PXIe-5632 VNA would not sufficient to determine the accuracy of repeated measurement. With the study had been done is required to perform the gaps analysis with another external system. It is reducing the risk to make mistake in the calibration and to ensure the PXIe-5632 VNA also demonstrate the strong correlation with another system. If the comparison between the PXIe-5632 and external system results are independent, a further detail analysis is required including validate the calibration procedure or hardware being used in this study.

Figure 14 shown the same device under test measured by Rohde & Schwarz ZNB20 VNA from 100 kHz until 8.5 GHz. It was the screenshot from the ZNB20 VNA display. It shown the device under test trend across the frequency swept with markers function turn on. The reading pointed by markers will compare with the PXIe-5632 VNA. From the ZNB20 VNA was found the device under test characteristic does not linear from 1 GHz until 8.5 GHz. It is because the ZNB20 VNA perform the measurement without averaging function turn on. Without the averaging function, the device under test is measured under the “worst case” scenario. If the averaging function turn on, it is expected the device under test measured in a linear characteristic.

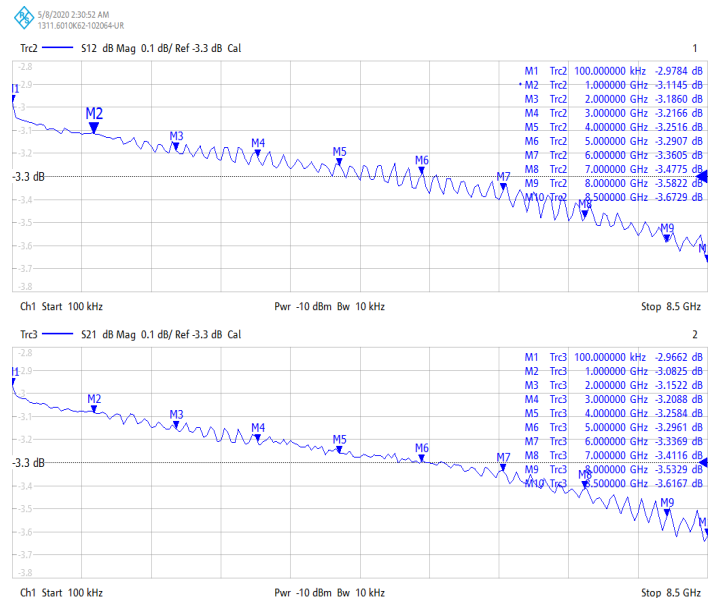


Figure 14: ZNB20 Transmission S21 and S12 Absolute Measurement

Figure 14 observe that the ZNB20 VNA measured the device under test similar to PXIe-5632 VNA. Table 10 and 11 calculate the delta between two independent system in term of precision and accuracy. It is a very useful information to research and development team improve the next generation of VNA design.

Table 10: S21 Comparison Between PXIe-5632 and ZNB20

Frequency GHz	Measurement Error			
	PXIe-5632 S21	ZNB20 S21	Delta S12 (dB)	Delta S12 (%)
0.0003	-2.9862	-2.9662	-0.0200	0.67%
1	-3.1742	-3.0825	-0.0917	2.98%
2	-3.1735	-3.1522	-0.0213	0.68%
3	-3.1679	-3.2088	0.0409	-1.27%
4	-3.3048	-3.2584	-0.0464	1.42%
5	-3.2667	-3.2961	0.0294	-0.89%
6	-3.3808	-3.3369	-0.0439	1.32%
7	-3.3393	-3.4116	0.0723	-2.12%
8	-3.4424	-3.5329	0.0905	-2.56%
8.5	-3.5594	-3.6167	0.0573	-1.59%

From Table 10, the maximum delta between the two system occurs at 1 GHz. It is differentiated by -0.0917 dB or 2.98% above the PXIe-5632 VNA measurement. Overall Table 10 can conclude that the S21 does not measure any significant delta from 300 kHz until 8.5 GHz.

Table 11 : S12 Comparison Between PXIe-5632 and ZNB20

Frequency GHz	Measurement Error			
	PXIe-5632 S12	ZNB20 S12	Delta S12 (dB)	Delta S12 (%)
0.0003	-2.92652	-2.9780	0.0515	-1.76%
1	-2.88306	-3.1145	0.2314	-8.03%
2	-3.10583	-3.1860	0.0802	-2.58%
3	-3.29741	-3.2166	-0.0808	2.45%
4	-3.15151	-3.2516	0.1001	-3.18%
5	-3.36279	-3.2907	-0.0721	2.14%
6	-3.23231	-3.3605	0.1282	-3.97%
7	-3.5807	-3.4775	-0.1032	2.88%
8	-3.75398	-3.5822	-0.1718	4.58%
8.5	-3.76342	-3.6729	-0.0905	2.41%

From Table 11, the maximum delta between the two system occurs at 1 GHz. It is differentiated by 0.2314 dB or -8.98% below the PXIe-5632 VNA measurement. At 4, 6 and 8 GHz were found large deviation measured beyond 3% delta. Overall Table 11 can conclude that the S12 measure multiple significant changes from 300 kHz until 8.5 GHz. Especially at 1 GHz was identify maximum delta in both S21 and S12.

5.3. Summary of Discussion

In the discussion, there are 2 main issues were found. First issue identifies as the frequency swept at 1 GHz and second issue is transmission S12 comparison between PXIe-5632 and ZNB20.

Frequency swept at 1 GHz did not found significant issue when the measurement of S21 and S12 correlation between PXIe-

5632 itself. The changes in insertion loss was low measured at 0.0192 dB. However, the same frequency point resulting the biggest delta compare to another system. This is the first improvement required corrective action in next research.

The second issue was found in transmission S21 comparison between PXIe-5632 and ZNB20. There are 4 out of 10 measurement points are beyond 3% delta. Although the 3% changes are not significant to both systems. If the ± 3% set as the benchmark or pass rate. Overall, only 6 out of 10 measurement points were measured below 3% changes. It is only score 60%. Nevertheless, corrective action required in next research as well.

6. Conclusion

In this study, an evaluation of Type A measurement uncertainty for NI PXIe-5632 Vector Network Analyzer from 300 kHz to 8.5 GHz is calculated. This frequency bandwidth is widely used in telecommunications particularly in the calibration system for the 5 G sub-6 GHz range that required to perform calibration on the telecommunication devices. This paper also carries out several the comparison technique applied in the PXIe-5632 VNA, valued to specific calibration techniques. A complete mathematical review shown the effects in each step involve in the Type A uncertainty analysis. This paper could help improve the scattering of parameters for commercial laboratories to extend the calibration potential in the new 5G telecommunication sector. The calibration analysis by using a 3 dB fixed attenuator also helps the commercial calibration laboratory explore existing measuring capacities and increase the degree of competence in the laboratory. In the discussion, it is explained the limitations are identify in existing PXIe-5632 VNA system. A part of the limitation of this paper would contribute to next improvement research area. As future of this work, the first development is to extend the frequency range beyond 8.5 GHz. The second development is extending to PXIe-5632 VNA Scattering Parameters port 1 (S11) and port 2 (S22). The calculation of the reflections S11 and S22 involved only port 1 or port 2. In addition, a measure of the test uncertainty ratio may also be added in future to satisfy the requirements of ANSI / NCSL Z540. A test uncertainty ratio also can be calculated to meet the ANSI/NCSL Z540 requirement.

Acknowledgements

This work was supported in part by the Malaysian Ministry of Higher Education under the MyBrain15 program and Universiti Teknikal Malaysia Melaka.

References

- [1] National Instruments, Calibration Procedure PXIe-5632 8.5 GHz Vector Network Analyzer, National Instruments USA, 2018.
- [2] Agilent, Fundamentals of RF and Microwave Power Measurements (Part 1), Agilent Technology USA, 2003.
- [3] Williams, A. J. Jeffrey, F. Dylan, "A method for improving high-insertion-loss measurements with a vector network analyzer," in 89th ARFTG Microwave Measurement Conference (ARFTG) 2017, Honolulu, HI. DOI:10.1109/ARFTG.2017.8000839

- [4] Y. S. Meng, A. C. Patel, Y. Shan, H. N. Pandya, Y. Wang, "Performance evaluation of a handheld network analyzer for testing of balanced twisted-pair copper cabling," in 29th Conference on Precision Electromagnetic Measurements (CPEM) 2014, Rio de Janeiro, Brazil. DOI: 10.1109/CPEM.2014.6898402
- [5] Metrology, ISO/IEC Guide 98-1 Uncertainty of Measurement, Switzerland: ISO, 2009.
- [6] G. G. Jing, "How to measure test repeatability when stability and constant variance are not observed" *International Journal of Metrology and Quality Engineering*, vol. 9, Article 10, 2018. <https://doi.org/10.1051/ijmqe/2018007>
- [7] D. Kubatova, M. Melichar, J. Kutlwaser, "Evaluation of Repeatability and reproducibility of CMM equipment," in *Procedia Manufacturing*, Elsevier B.V. Vol 13, Page 558-564. <https://doi.org/10.1016/j.promfg.2017.09.091>
- [8] Agilent, *Fundamentals of RF and Microwave Power Measurements (Part 4)*, Agilent Technologies USA, 2006.
- [9] M.H. Tan, B.H Ahmad Yusairi, S. Mohd Rizal, "An Analysis for 2.4mm-2.4mm RF Connector Insertion Loss Measure From 45MHz Until 50GHz by Using Electronic Calibration Module and Mechanical Calibration kits in a Network Analyzer," in *MIMT 2016*, Cape Town South Africa. <https://doi.org/10.1051/mateconf/20165403006>
- [10] M.H. Tan, B.H Ahmad Yusairi, S. Mohd Rizal, "Radio Frequency Connector Insertion Loss Measured From 300 kHz Until 8.5 GHz By Using Network Analyzer And Mechanical Calibration Kits," in *Journal Of Engineering Science and Technology*, 1587-1600, 2019. <https://doi.org/10.5281/zenodo.3590276>
- [11] M.H. Tan, B.H Ahmad Yusairi, "Evaluation of Uncertainty Measurement Calculation for Vector Network Analyzer From 300 kHz to 8.5 GHz," *Advance in Science, Technology and Engineering System Journal*, Vol 5, Issue 3, Page 01-10 2020. DOI: 10.25046/aj050301
- [12] K. Hitomi, *Manufacturing Systems Engineering: A Unified Approach to Manufacturing Technology, Production Management, and Industrial Economics*, UK: Taylor & Francis Ltd, 1996.
- [13] P.L.Liu, L.L.Ren, J.H.Zhang, C.W.Luo, L.Zhao, "Research Progress of HL-2M Distributed Timing System," *Fusion ENGINEERING and Design* 2020, vol. 153, Article 111491. <https://doi.org/10.1016/j.fusengdes.2020.111491>
- [14] H. Huang, C. Wang, H. Wu, C. Huang, Z. Yang, H. Wang, "Optimization of the two-stage common-emitter transistor amplifier for equalization circuit in visible light communication system," *Optical And Quantum Electronics* 2018, Vol 50, Article 349. <https://doi.org/10.1007/s11082-018-1613-y>
- [15] R. & Schwarz, "Vector Network Analysis with Up to 48 Ports," *Microwave Journal* 2014, Vol 57, Edition 3, <https://doi.org/10.5281/zenodo.3783593>
- [16] D. Nix and A. Weigend, "Estimating the mean and variance of the target probability distribution," in *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN) 1994*, Orlando, FL, USA. IEEE 10.1109/ICNN.1994.374138,
- [17] H. Kirkham, A. Riepnicks, M. Albu, D. Laverty, "The nature of measurement, and the true value of a measured quantity," in *IEEE International Instrumentation and Measurement Technology Conference (I2MTC) 2018*, Houston, TX. DOI: 10.1109/I2MTC.2018.8409771,
- [18] L. K. Pino, B. C. Searle, E. L. Huang, W. S. Noble, A. N. Hoofnagle, M. J. MacCoss, "Calibration Using a Single-Point External Reference Material Harmonizes Quantitative Mass Spectrometry Proteomics Data between Platforms and Laboratories," *Analytical Chemistry* 2018, Vol. 90, Page 13112-13117 DOI: 10.1021/acs.analchem.8b04581
- [19] X. Wan, W. Wang, J. Liu and T. Tong, "Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range," *BMC Medical Research Methodology* 2014, Vol. 14, Page 135. <https://doi.org/10.1186/1471-2288-14-135>
- [20] Agilent, *Fundamentals of RF and Microwave Power Measurements (Part 2)*, Agilent Technologies USA, 2006.
- [21] Agilent, *Fundamentals of RF and Microwave Power Measurements (Part 3)*, Agilent Technologies USA, 2003.

A Review on Autonomous Mobile Robot Path Planning Algorithms

Noraziah Adzhar*, Yuhani Yusof, Muhammad Azrin Ahmad

Intelligent Computing & Optimization Research Group, Centre for Mathematical Sciences, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300, Pahang, Malaysia.

ARTICLE INFO

Article history:

Received: 01 January, 2020

Accepted: 30 April, 2020

Online: 21 May, 2020

Keywords:

Path Planning

Routing Problem

Optimization Method

Mobile Robot

Heuristics Method

Automation

ABSTRACT

The emerging trend of modern industry automation requires intelligence to be embedded into mobile robot for ensuring optimal or near-optimal solutions to execute certain task. This yield to a lot of improvement and suggestions in many areas related to mobile robot such as path planning. The purpose of this paper is to review the mobile robots path planning problem, optimization criteria and various methodologies reported in the literature for global and local mobile robot path planning. In this paper, commonly use classical approaches such as cell decomposition (CD), roadmap approach (RA), artificial potential field (AFP), and heuristics approaches such as genetic algorithm (GA), particle swarm optimization (PSO) approach and ant colony optimization (ACO) method are considered. It is observed that when it comes to dynamic environment where most of the information are unknown to the mobile robots before starting, heuristics approaches are more popular and widely used compared to classical approaches since it can handle uncertainty, interact with objects and making quick decision. Finally, few suggestions for future research work in this field are addressed at the end of this paper.

1. Introduction

Mobile robots are widely used in automated industrial environments such as military, security environments, agriculture, mining and in warehouses [1]. The ultimate goal in designing the path planning for a mobile robot is for the mobile robot to move successfully and able to execute task in the environment. In order to achieve this, a path planning algorithms need to be design first. Such algorithm is meant to provide the free collision path from the starting point to its destination with respect to certain criteria such as distance, time taken, safety level of the path and path smoothness. Thus, determining an optimal path in an environment containing obstacles is one of the challenge in designing path planning algorithm. This problem is categorized as NP-complete or NP-hard, depending on the complexity of the environment [2]. Path planning not only save a lot of time but also reduce the wear and capital investment of mobile robot.

Robot navigation can be further divided into three parts which are global navigation, local navigation and personal navigation. In global navigation, the mobile robot knows the location of obstacles, type of environment and its target point. Local

navigation is more challenging where the location of the obstacles is dynamic. The mobile robot deals with unknown or partially known environment since certain objects might be stationery or moving. In this kind of navigation, the mobile robot need to be able to interact with objects or even carrying out task. Personal navigation is for monitoring of the individual robot and anything in contact with it.

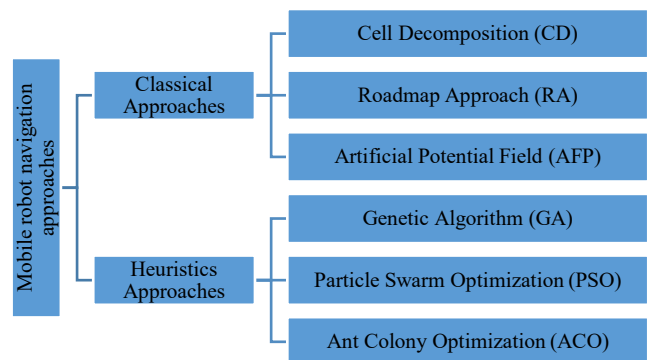


Figure 1: Few Popular Classical and Heuristics Approaches for Mobile Robot Path Planning

*Noraziah Adzhar, Universiti Malaysia Pahang, Malaysia
 Email: noraziahadzhar@ump.edu.my

Path search algorithm for global navigation is based on classical approaches such as cell decomposition, roadmap algorithm and AFP. Local navigational approaches are more intelligence since they need to interact with the dynamic environment and execute plan autonomously. These methods are classified as in Figure 1.

The rest of this paper presents the overview of the latest works done in mobile robot path planning field and is organized as follows. Section 2 discussed the optimization criteria for path planning. Section 3 and 4 presents the latest works of most common used classical and heuristics method respectively. Finally, Section 5 concludes the paper.

2. Optimization Criteria

There are few optimization criteria need to be considered in providing an optimal path whether in global or local navigation category. The criteria include distance, cost, time taken, smoothness and energy level. Not just that the characteristics of the mobile robot and environment also need to be taken into account such as turning radius of the robot, velocity, acceleration, shapes of the obstacles and the occurrence of dynamic environment. Three commonly used optimization objective is as follows:

2.1. Path Length

Through this criteria, the main goal is to have the shortest path as possible. In order to obtain the total path length, all sub length from the source point to destination point will be total up. This calculation is presented as the following formula [3].

$$Path_{length} = \sum_{i=0}^{n-1} \sqrt{(y_{i+1} - y_i)^2 + (x_{i+1} - x_i)^2} \quad (1)$$

In (1), n represent the number of nodes from the source point to its target.

2.2. Smoothness

Smoothness objective tries to have a straight path as possible. This objective will help to reduce energy level as the mobile robot will navigate through the path with minimal turns in a straight way compared to a curvy path that's uses a lot more energy. Path smoothness can be measured through the following equation [4].

$$Path_{smoothness} = \sum_{i=1}^n (180^\circ - \theta_i) \quad (2)$$

In (2), n are number of angles formed from starting position to target destination and θ_i is value of angle for $1 \leq i \leq n$. Figure 2 illustrates the graphical explanation of this objective with $n = 4$.

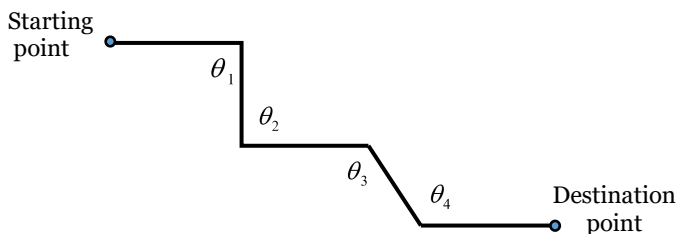


Figure 2: Graphical Explanation to Measure Smoothness Degree of a Path

2.3. Safety Degree

Safety degree of a path is definitely vital to provide free collision paths especially in the environment where robot and human needs to interact to each other. A safe path planning is a must especially for a high degree of freedoms (DOFs) mobile robots in dynamic environment where the obstacles are not necessary stationery or in a crowded environment. The safety degree which is also can represent the collision probability is presented in [3]. It can be defined as

$$Safety_{degree} = \sum_{p=1}^{n-1} S_p = \begin{cases} 0, & d_p \geq \lambda \\ \sum_{p=1}^{n-1} e^{\lambda - d_p}, & d_p < \lambda \end{cases} \quad (3)$$

In (3) d_p is the shortest distance between the i -th segment and its nearest obstacles, and λ is the threshold of the safety degree. Path with a small value of safety degree reflects the small collision probability and will be chosen.

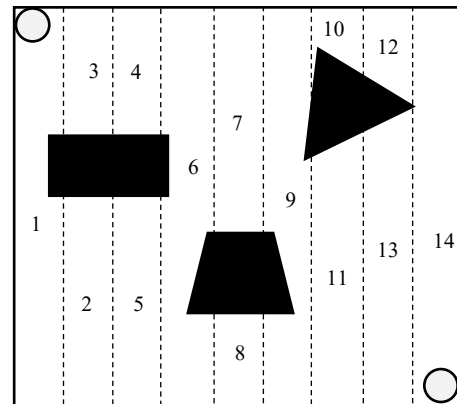
3. Classical Approach

Initially, classical approaches are commonly used in mobile robot path planning before the development of artificial intelligence methods. The major drawback of these methods are high computational cost and unable to make decision prior to dynamic environment and execute new planning [5]. Cell decomposition approach (CD), roadmap approach (RA) and artificial potential field (AFP) are few classical approaches that are being reviewed in this paper.

3.1. Cell decomposition approach (CD)

The idea of this method is to divide the environment into a number of connected regions. The shape of the region or cell can be vertical strip cells, array of rectangular grid, or unequal size rectangular grid. Each cells are connected to each other and not overlapping. If the cell contains any obstacles, it will be considered as 'corrupted cell' and will be decompose further to get free cell. All cell is given its number representation. This idea is illustrated in Figure 3.

Starting position



Destination

Figure 3: Environment is Decomposing into Few Numbered Free Cells

In this decomposition, any two cells are adjacent if they shared the same boundary. This decomposition can be later on convert into graph as in Figure 4 where cell 1 is the starting position of the mobile robot and cell 14 is the destination.

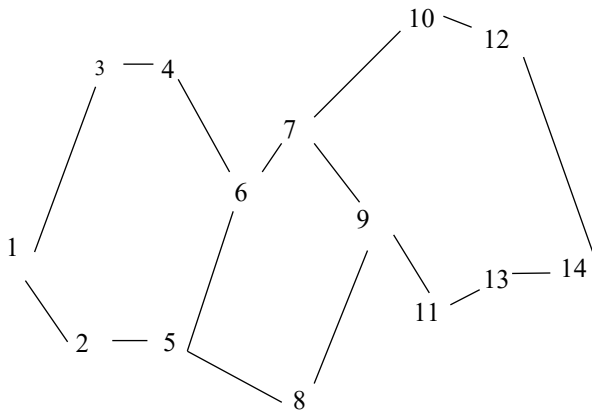


Figure 4: Decomposition is Converted into Graph

This method has been presented in [6] for real-time operation of mobile robot path planning, in [7] to handle multiple activities in three dimensional environments by using greedy depth-first search and GA based method and in [8] to compare the trajectories of various cell decomposition and graph weights.

3.2. Roadmap approach (RA)

Roadmap approach or highway approach uses visibility graphs to develop the roadmap of the environment and further help in finding the shortest path from starting position to the target. In developing visibility graph, all obstacles will first be represented as a polygon. Each of this polygonal obstacles will have nodes correspond to its vertices.

All the nodes will be connected if:

- i. they form an edge of the obstacles.
- ii. the line segment joining them forming a visible connection.

The advantage of this method is if the path found in this roadmap will be the shortest path. However, the performance of this method will be greatly reduced in a complex environment or in higher dimensional-space. The application of roadmap approach using visibility graphs is as presented in [9].

3.3. Artificial potential field (AFP)

Artificial potential field is first developed by Khatib [10] for mobile robot navigation. This method assumes an artificial force field consists of attractive and repulsive force in the environment. The goal will produce attractive force that attract the mobile robot to move towards it. While, the obstacles generate a repulsive force and is pointing out from the obstacles. These imaginary forces attract the robot towards the goal by avoiding all obstacles. See Figure 5. The robot will be reaching the target point by following the negative gradients. The method is applied widely in mobile robot navigation including implementation in a real time robot path planning [11], modification of the method to be use in dynamic environment [12] and hybridization of this method with other techniques such as GA [13], PSO [14] and FL [15].

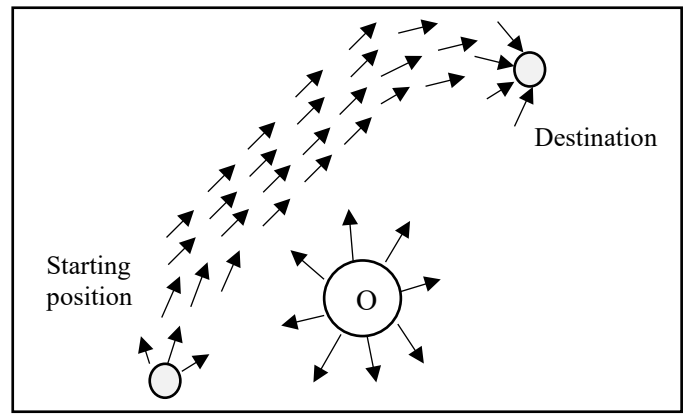


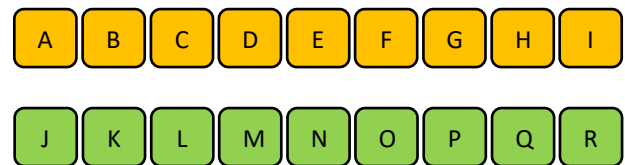
Figure 5: Illustration of the Scheme of Artificial Potential Field

4. Heuristics Approach

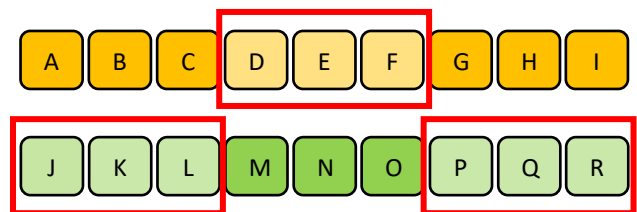
The development of artificial intelligence method helps mobile robots to navigate more successfully in dynamic environment where the obstacles might be moving or stationary. These methods are most popular nowadays compare to conventional approaches since it can handle the uncertainty present in the environment and make decision. Few popular approaches in this field is being reviewed in this section including genetic algorithm (GA), particle swarm optimization approach (PSO) and ant colony optimization method (ACO).

4.1. Genetic algorithm (GA)

Genetic algorithm (GA) is one of the heuristics method that has been used widely to solve constrained and unconstrained optimization problem in science and engineering field. In GA, all candidate solutions (called individuals) of the problem will be encoded into chromosomes which will undergoes few basic operations such as selection operator, crossover operator and mutation operator. The fitness value for each individual will be evaluated based on the objectives. In each generation, multiple individuals will be selected based on their calculated fitness value forming a new population to be used in the next iteration. Figure 6 presents the example of the basic step of this method.



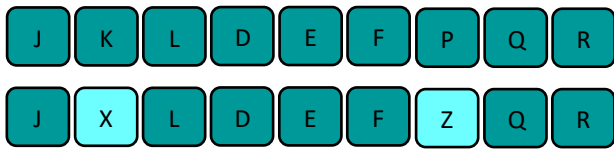
(a) Selection Operator: Two set of individuals encoded into chromosomes



(b) Crossover Operator: multiple of individuals is selected based on their fitness value



(c) New Population Formed.



(d) Mutation Operator: Genes or individuals from other chromosomes are randomly inserted into new population. By doing this, premature convergence can be avoided.

Figure 6: Genetic Algorithm Basic Step

Few stopping conditions that can be used to determine termination include:

- i. Generations – If a maximum number of generations has been produced.
- ii. Time Limit – Algorithm will stop if it has been running for a specific period of time.
- iii. Fitness Value – If satisfactory fitness level has been reached for the population.

The comparison between this method and probabilistic RA method has been reported in [16]. Results show that GA produce a smoother path for robot navigation but more time consuming. In [17], a fitness function that optimizes the number of turns take by the mobile robot is suggested where in [18], a new binary encoding by using matrix is presented for mobile robot navigation in both static and dynamic environments.

4.2. Particle swarm optimization approach (PSO)

Particle swarm optimization (PSO) is derived from the situation of a swarm of birds or a school of fish finding for food. This method mimics the behavior of these animals that doesn't require any leader to lead the group the reach the food source. These swarm of animals do not know where the food is hidden or exactly located, but they have the information of their distance from it. If each single animal tries to reach the food source on their own, it is inefficient since large number of time will be needed and will cause havoc in the situation. Therefore, the best way is to follow the members who is nearest to the food source [5]. In reflect to the algorithm, each single animal in the search space represent the solution. Each of these solution or particle contain two information,

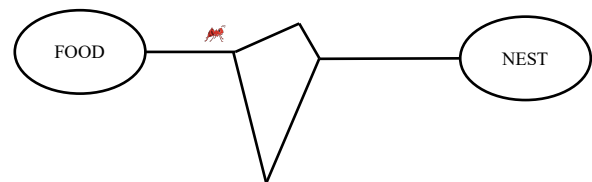
- i. their own fitness value determines by the objective function.
- ii. the velocities which direct the solution to the target.

This algorithm is initialized with a set of solution. In every iteration, each particle will return their fitness value called pbest. The best pbest value in each iteration is recorded and stored as global best value, gbest. After getting these two values, the algorithm will calculate for the velocity and position of particles. This algorithm has few similarities with genetic algorithm such as starting the procedure with a set of randomly generated population and both methods evaluate the population through fitness value.

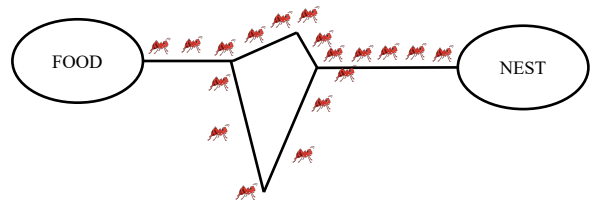
This method has been implemented for navigation of an aerial robot in three dimensional unknown environments [19], humanoid robot [20] and in industrial robot [21].

4.3. Ant colony optimization method (ACO)

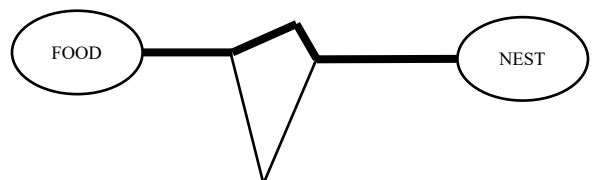
Ant colony optimization method inspired from the biological situation of ants finding the shortest path between their nest and food source using pheromone trails. Since ants hardly use vision, they will lay pheromone trails when they're moving from one place to another as a give signal to other ants. If the next ant decides to follow that path with certain probability, they will lay more pheromone to reinforce the trail. The more ants following the path, the stronger the pheromone. Pheromone of a shorter path builds up faster so more ants will follow it. This behavior is illustrated as in Figure 7.



(a) Ant Making a Trail.



(b) Ant follows and leave pheromone. Shorter trails are run more often as the pheromone on a longer trails evaporate faster and disappear with time.



(c) As time increasing, more ants will follow the shorter trail. The pheromone on the shorter trails will get stronger.

Figure 7: Biological Situation of Ant Colony Optimization Method

In reflect to the algorithm, this algorithm requires some memory to restore data structures. This method has been applied widely in various field including bus routes, garbage collection, telecommunication networks, machine scheduling and composition of products. In [22], ACO is hybridized to be implemented for navigations of humanoids in cluttered environment where in [23], a new fuzzy approach is presented for diversity control of ACO.

5. Conclusion

This study discuss few commonly used classical approaches such as cell decomposition (CD), roadmap approach (RA),

artificial potential field (AFP), and heuristics approaches such as genetic algorithm (GA), particle swarm optimization (PSO) approach and ant colony optimization (ACO) for mobile robot navigation. From the discussion, heuristics approaches are more popular and widely used compared to classical approaches due to its robustness and perform well in various environmental condition. This field is in high demand can be extended to several situations. Mobile robot path planning can be inoculated with multi sensor. The use of multi sensor improve mobile robot performance by giving extra information on the surrounding and able to reduced uncertainty given by single sensor. Few topics can be studied in depth related to multi mobile robots for example their task assignment, communication and cooperation between them and group path planning. Mobile robot path planning in high dimensional environment cooperated with static or dynamic obstacles can be another new focus area in this field.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to thank Universiti Malaysia Pahang for supporting part of this research through research grant RDU1703188.

References

[1] H. Y. Zhang, W. M. Lin, A. X. Chen, "Path planning for the mobile robot: A review" *Symmetry*, **10**(10), 450, 2018. <https://doi.org/10.3390/sym10100450>

[2] J. Han, Y. Seo, "Mobile robot path planning with surrounding point set and path improvement" *Applied Soft Computing*, **57**, 35-47, 2017. <https://doi.org/10.1016/j.asoc.2017.03.035>

[3] X. Y. Wang, G. X. Zhang, J. B. Zhao, H. Rong, F. Ipaté, R. Lefticaru, "A modified membrane-inspired algorithm based on particle swarm optimization for mobile robot path planning" *Int. J. Comput. Commun. Control*, **10**, 732-745, 2015. <https://doi.org/10.15837/ijccc.2015.5.2030>

[4] A. Hidalgo-Paniagua, M. A. Vega-Rodríguez, J. Ferruz, N. Pavón, "MOSFLA-MRPP, multi-objective shuffled frog-leaping algorithm applied to mobile robot path planning" *Eng. Appl. Artif. Intell.*, **44**, 123-136, 2015. <https://doi.org/10.1016/j.engappai.2015.05.0111>

[5] B. K. Patle et al., "A review: On path planning strategies for navigation of mobile robot" *Defence Technology*, **15**(4), 582-606, 2019. <https://doi.org/10.1016/j.dt.2019.04.0111>

[6] P. T. Tunggal et al., "Pursuit algorithm for robot trash can based on fuzzy-cell decomposition" *Int J Electr Comput Eng*, **6**(6), 2863-2869, 2016. <https://doi.org/10.11591/ijeece.v6i6.10766>

[7] A. Mark, C. Gill, Y. Z. Albert, "A cell decomposition-based collision avoidance algorithm FOR robot manipulators" *Cybern Syst*, **29**(2), 113-135, 2010. <https://doi.org/10.1080/019697298125759>

[8] R. Gonzalez, M. Kloetzer, C. Mahulea, "Comparative study of Trajectories Resulted from Cell Decomposition Path Planning Approaches" in 2017 21st International Conference on System Theory, Control and Computing (ICSTCC), Sinaia, 2017. <https://doi.org/10.1109/ICSTCC.2017.8107010>

[9] L. Lulu, A. Elnagar, "A Comparative Study Between Visibility-Based Roadmap Path Planning Algorithms" in 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, Alta., 2005. <https://doi.org/10.1109/IROS.2005.1545545>

[10] O. Khatib, "Real Time Obstacle Avoidance for Manipulators and Mobile Robot" in 1985 IEEE International Conference on Robotics and Automation, St. Louis, MO, USA, 1985. <https://doi.org/10.1109/ROBOT.1985.1087247>

[11] P. Vadakkepat, K. C. Tan, M. L. Wang, "Evolutionary Artificial Potential Fields and Their Application in Real Time Robot Path Planning" in 2000 Congress on Evolutionary Computation. CEC00 (Cat. No.00TH8512), La Jolla, CA, USA, 2000. <https://doi.org/10.1109/CEC.2000.870304>

[12] M. G. Park, M. C. Lee, "Artificial Potential Field Based Path Planning for Mobile Robots using Virtual Obstacle Concept" in 2003 IEEE/ASME

International Conference on Advanced Intelligent Mechatronics (AIM 2003), Kobe, Japan, 2003. <https://doi.org/10.1109/AIM.2003.1225434>

[13] R. Raja, A. Dutta, K. S. Venkatesh, "New potential field method for rough terrain path planning using genetic algorithm for a 6-wheel rover" *Robotics and Autonomous Systems*, **72**, 295-306, 2015. <https://doi.org/10.1016/j.robot.2015.06.002>

[14] P. H. Kuo, T. H. Li, G. Y. Chen, Y. F. Ho, C. J. Lina, "Migrant-inspired path planning algorithm for obstacle run using particle swarm optimization, potential field navigation, and fuzzy logic controller" *The Knowledge Engineering Review*, **32**(5), 2017. <https://doi.org/10.1017/S0269888916000151>

[15] J. M. Abdel Kareem, M. H. Garibeh, E. A. Feilat, "Autonomous mobile robot dynamic motion planning using hybrid fuzzy potential field" *Soft Computing*, **16**(1), 153-164, 2012. <https://doi.org/10.1007/s00500-011-0742-z>

[16] C. S. Robert Martin, D. O. Anton Louise, T. U. Aristotle, A. B. Argel, P. D. Elmer, "Path Planning for Mobile Robots using Genetic Algorithm and Probabilistic Roadmap" in 2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Manila, 2017. <https://doi.org/10.1109/HNICEM.2017.8269498>

[17] L. Chaymaa, B. Said, E. Ali, "Genetic algorithm based approach for autonomous mobile robot path planning" *Procedia Computer Science*, **127**, 180-189, 2018. <https://doi.org/10.1016/j.procs.2018.01.113>

[18] B. K. Patle, D. R. K. Parhi, A. Jagadeesh, K. Sunil Kumar, "Matrix-binary codes based genetic algorithm for path planning of mobile robot" *Computers and Electrical Engineering*, **67**, 708-728, 2018. <https://doi.org/10.1016/j.compeleceng.2017.12.011>

[19] M. A. Rendon, F. F. Martins, "Path following control tuning for an autonomous unmanned quadrotor using particle swarm optimization" *IFAC-Papers on Line*, **50**(1), 325-330, 2017. <https://doi.org/10.1016/j.ifacol.2017.08.054>

[20] P. B. Kumar, K. K. Pandey, C. Sahu, A. Chhotray, D. R. Parhi, "A Hybridized RA-APSO Approach for Humanoid Navigation" in 2017 Nirma University International Conference on Engineering (NUICONE), Ahmedabad, 2017. <https://doi.org/10.1109/NUICONE.2017.8325611>

[21] M. Gao, P. Ding and Y. Yang, "Time-Optimal Trajectory Planning of Industrial Robots Based on Particle Swarm Optimization," in *2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*, Qinhuangdao, 2015. <https://doi.org/10.1109/IMCCC.2015.410>

[22] B. K. Priyadarshi, S. Chinmaya, R. P. Dayal, "A hybridized regression-adaptive ant colony optimization approach for navigation of humanoids in a cluttered environment" *Applied Soft Computing*, **68**, 565-585, 2018. <https://doi.org/10.1016/j.asoc.2018.04.023>

[23] C. Oscar, N. Héctor, S. José, M. Patricia, V. Fevrier, "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. <https://doi.org/10.1016/j.asoc.2014.12.002>

Study of Wrinkling and Thinning Behavior in the Stamping Process of Top Outer Hatchback Part on the SCGA and SPCC Materials

Sri Wahyanti, Agus Dwi Anggono*, Waluyo Adi Siswanto

Department of Mechanical Engineering, Universitas Muhammadiyah Surakarta, Jl.Ahmad Yani, PO.BOX 1 Pabelan, Surakarta 57162, Indonesia.

ARTICLE INFO

Article history:

Received: 14 April, 2020

Accepted: 12 May, 2020

Online: 21 May, 2020

Keywords:

Deep Drawing

Thinning

Wrinkling

Friction Coefficient

Forming Limit Diagram

ABSTRACT

The objective of the research is to determine the changes in wrinkling and thinning behavior in the stamping process of the top outer hatchback. The study was conducted using two types of materials, i.e., SCGA (Steel Cold rolled Galvanized Annealed) and SPCC (Steel Plate Cold rolled Coiled) with a thickness of 0.80 mm. During the stamping process, the coefficient of friction values varied of 0.0, 0.05, 0.10, and 0.15 for each material. The stamping process was carried out by using the simulation method to investigate the wrinkling and thinning behavior. The forming limit diagram (FLD) provides information on material changes and the amount of safe and unsafe areas on the blank. The results showed that in the fifth step, the parts that experienced high values of major-minor stress and strains, thinning, and wrinkling decreased in the last step. That caused by removing unused areas after the trimming process. Therefore, the safe zone of SCGA was increased from 9.52% to 18.19%. For the SPCC material, the safe area increased from 8.63% to 16.31%. The coefficient of friction affects the thinning and wrinkling defects. The greater of the friction coefficient will increase the value of thinning and wrinkling. Based on the Non-Linear FLD analysis, both materials SCGA and SPCC are still in a safe condition.

1. Introduction

The development of the automotive industry today, requires the manufacturing industry to make products effectively and efficiently [1]. Automotive panel components are made by stamping [2], [3]. In the stamping process, product quality is strongly influenced by several factors, such as material properties, die geometry, friction characteristics, and punch forces [4]-[6]. In the stamping process, there are some defect problems that are usually present, namely wrinkle, fracture, springback, and cracking [7]-[9]. To solve these defects, stamping companies require more time and cost. An accurate sheet metal forming simulation is needed to develop an effective production process so that it can reduce the cost of producing automotive components. The use of stamping simulations to reduce defects can shorten the time of making molds, increasing material utilization, and reducing component damage. As such, stamping simulations make a significant contribution to the automotive industry to reduce product defects [10], [11].

In previous studies, researchers have used sheet metal forming simulations to predict defect problems during the stamping process. One of the efforts made is to analyze the influence of drawbead geometry to reduce failures and defects in the deep drawing process [12]. Researchers make changes in material flow by using drawbead to optimize cavity pressure. Thus, wrinkling can be reduced [13]. Analyzing blank holding forces related to draw-bead depth [14]. The drawbead retaining force is evaluated using the finite element method (FEM) for a variety of blank size variations. Recently, a design sensitivity analysis was also carried out to optimize the blank holder force (BHF) and draw-bead strength in the stamping process [15]. Reducing and eliminating cracking in the deep drawing process is done by analyzing the blank holder force and punch stroke variants [16]. Researchers suggest the factors that need to be considered while making a stamping simulation are the holding force, the shape of the draw-bead, the distance between the tool and the blank, and the boundary conditions [17]. In the stamping simulation, it uses not only the holding force but also the lift force because the draw-bead and strain become the boundary conditions in the stamping process [18].

* Agus Dwi Anggono, Email: adal26@ums.ac.id

Researchers have used FEM to analyze thinning and wrinkling behavior [19]-[21] and the factors that influence wrinkling [22]. BHF optimization and friction coefficient studies have been conducted to reduce the occurrence of wrinkling [23]. The combination of displacement adjustment (DA) and spring forward (SF) methods into a hybrid method (HM) has been carried out to reduce springback and optimize die in the stamping process [24]. Case studies for analyzing springback and compensation in automotive components have been carried out using FEM [25].

FLD is an empirical diagram used to determine the safe area of the plate during the forming process, which is applied to avoid failure [26]. FLD can be used to observe the behavior of isotropic hardening and strain in the deep drawing process [27][28]. That is also used to analyze the damage and thinning due to sheet metal forming processes such as fracture, necking, wrinkling, local bonding, and wrinkling [29]. FLD uses surface strain measurements by calculating major and minor strains [30][31].

This study aims to investigate wrinkling and thinning behavior. The effect of the coefficient of friction on the formation of wrinkling and thinning was also investigated. The automotive component studied is a top outer hatchback with material using two different materials, namely SCGA and SPCC.

2. Research Method

The study was conducted by using numerical simulations of Autoform software. The design was created by using CATIA V5 software. The CAD (Computer-Aided Design) data in the form of surface models, then developed into upper and lower dies. Other CAD data were blank holder, outer trimming, and inner trimming. All models in the form of surfaces have been prepared from CATIA V5, as shown in Figure 1. The models were then saved in the type of an IGS file for input files in Autoform.

The material used in this study was a commercial plate for the stamping industry, namely SCGA and SPCC. The material properties was described in Table 1. Plate thickness is 0.80 mm with a length of 1,610 mm and a width of 1,120 mm. Variations in the friction coefficient used were 0.0, 0.05, 0.10 and 0.15 for each material. It was done to determine the effect of the coefficient of friction on the formation of wrinkling and thinning defects. The component studied was the back part of the Micro Bus, the top outer hatchback.

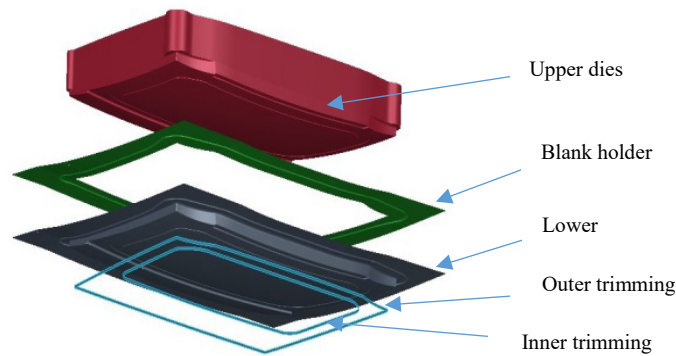


Figure 1. Surface definition for sheet metal forming.

Table 1. Material properties of SCGA and SPCC [32][33].

Mechanical properties	Material type	
	SCGA	SPCC
Young's Modulus (GPa)	160	210
Poisson's ratio	0.3	0.3
Yield Stress (MPa)	157.1	157.1
Yield Strain (ϵ_{0}^{TM})	0.00869	0.00869
Strain Rate Exponents (m)	0.228	0.225
Strength Coefficient C (MPa)	551.4	551.4

Strain hardening for hardening curves use the Swift formula:

$$\sigma = C * (\epsilon_{pl} + \epsilon_0)^m \tag{1}$$

Hardening curves are defined using a combination of Swift and Hockett-Sherby approaches. The combination factor of α determines the composition of the equation.

$$\sigma = (1 - \alpha)C * (\epsilon_{pl} + \epsilon_0)^m + \alpha \{ \sigma_{sat} - (\sigma_{sat} - \sigma_i) e^{-\alpha \epsilon_{pl}^p} \} \tag{2}$$

where σ is stress equivalent, ϵ_{pl}^{TM} is equivalent plastic strain and the others are material parameters.

Tools setting parameters were described in the Table 2. Analyzes were performed on both SCGA and SPCC materials with a thickness of 0.8 mm each. The cushion stroke was 90 mm, and the blank holder force 60 kN. The investigation was carried out using FLD in 6 steps. The step is made based on the distance between punch and die or distance to bottom, as shown in Table 3. FLD is a curve that represents the boundaries of the process of material formation based on the major and minor strain stresses of an element. That means that material that experiences a strain above the curve's limit will experience thinning, wrinkling, springback or cracking. Material changes in the process of forming the top outer hatchback underwent several changes as indicated by the difference in the color of the formation, which consists of areas of thickening, compressing, stretch insuffitun, safe, risk of split, excess thinning and split as shown in Figure 2.

Table 2. Tools setting parameters for stamping simulation.

Setting Tool	Punch	Die	Binder	Cutting Tool
Tool Contact	Manual Upper Side of Blank	Manual Lower Side Blank	Manual Upper Side of Blank	Inner Cutting Outer Cutting
Support Type	Force Control	Rigid	Force Control	-
Displacing Tool	Die	-	Die	-
Cushion Stroke	90. mm	-	90. mm	-
Tool Stiffness	50. MPa/mm	-	50. MPa/mm	-
Force/ Pressure	60.0 kN	-	60.0 kN	-
Maximum Cutting Depth	-	-	-	70. mm

Table 3. Definition of each step for investigation.

Step	Distance to Bottom (mm)	Process
1	-125	Drawing
2	-115	Drawing
3	-105	Drawing
4	-95	Drawing
5	0.00	Drawing
6	0.00	Trimming

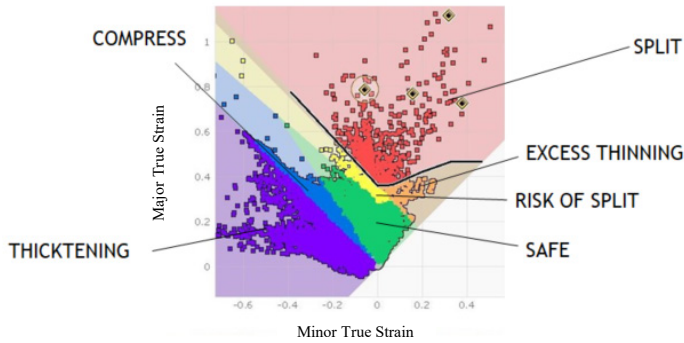


Figure 2. FLD description for forming investigation.

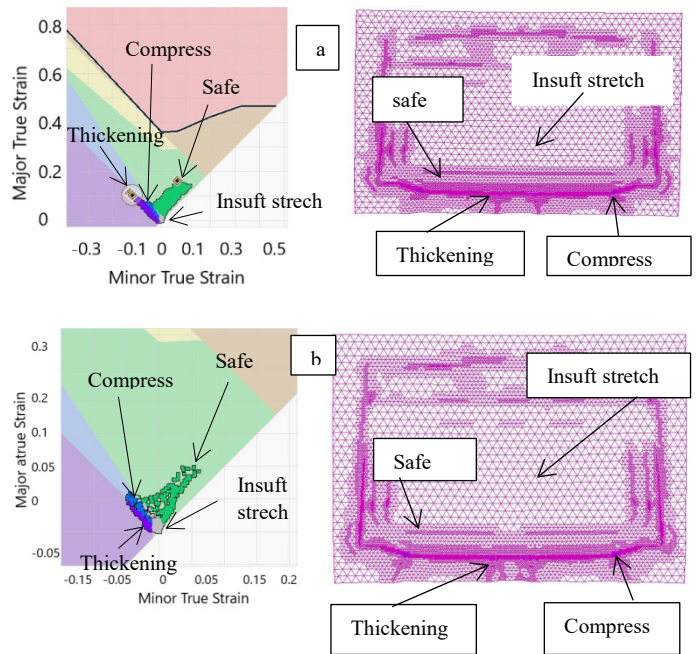


Figure 3. FLD results at step 1 for (a) SCGA, and (b) SPCC.

3. Results and Discussion

3.1. FLD based Investigation

The FLD results of the SCGA and SPCC material in the first step with a distance of -125 mm from the die was shown in Figure 3. Plate material flow behavior has been seen, and material changes occur from several conditions in both materials. Plate material changes that occur are the thickening area for SCGA material 0.12% and SPCC of 0.13%. The compressed area for SCGA material was 0.03% and SPCC 0.04%. Insuft stretch area for SCGA material was 99.84% and SPCC material 99.80%. Both material SCGA and SPCC have a safe area of <0.01%. The risk of split area, excess thinning, and split of all materials remained unchanged at 0.00%.

In the second step, the distance was -115 mm from the die. The thickening area of the SCGA and SPCC material was 0.32% and 0.39%. The compressed area was 0.11% and 0.13%, respectively. The area of insuft stretch on SCGA material was 99.56% and SPCC of 99.46%. The safe zone for SCGA material was 0.01% and 0.02% for SPCC material.

The investigation FLD based results for all steps were described in Table 4 for SCGA material and Table 5 for SPCC. The thickening area continues to increase until the fifth step for both SCGA and SPCC materials. Likewise, for compressed areas, insuft stretch, and safe regions. The forming simulation parameters used have shown that the stamping process can work safely. It can be seen from the area of risk of a split, exes thinning, and split, giving zero results. But in step six, the split area arose. The areas of risk of split and exes thinning was shown in Figure 4. (a) and (c). Figure 4. (b) and (d) show the FLD results after trimming. The thickening area has given zero results in the sixth step, and the compressed area as well. The area has decreased after the trimming process. The results of the analysis based on FLD indicate that the thickening area occurs in the regions that are not used. The results can be seen the sixth step or after trimming.

Table 4. FLD based results of SCGA material for all steps.

Parameters (%)	Step for SCGA					
	1	2	3	4	5	6
Thickening	0.12	0.32	1.16	6.21	14.6	0
Compress	0.03	0.12	0.43	1.43	3.76	0.01
Insuft stretch	99.8	99.6	91.9	91.9	72.1	81.6
Safe	0.01	0.01	0.1	7.2	9.52	18.2
Risk of split	-	-	-	-	-	-
Exes thinning	-	-	-	-	-	-
Split	-	-	-	-	-	-

Table 5. FLD based results of SPCC material for all steps.

Parameters (%)	Step for SPCC					
	1	2	3	4	5	6
Thickening	0.13	0.39	1.33	6.27	15.04	0
Compress	0.04	0.13	0.56	1.29	3.66	0.19
Insuft stretch	99.8	99.5	97.9	92.0	72.6	81.6
Safe	0.01	0.02	0.45	8.2	8.63	16.3
Risk of split	-	-	-	-	0.01	0
Exes thinning	-	-	-	-	0.01	0
Split	-	-	-	-	0.01	0

3.2. Non-Linear FLD based Investigation

Nonlinear FLD analysis was performed to determine material behavior based on strain changes. Material failure behavior depends on the strain path. Figure 5 shows the results of the nonlinear FLD analysis after the trimming process for both materials SCGA and SPCC. Based on these results, there were no

cracks or splits occur in the component. Material failures have occurred in unused areas. That has been explained before. Local necking was generated due to unstable deformation. This phenomenon occurs before the appearance of a crack or split.

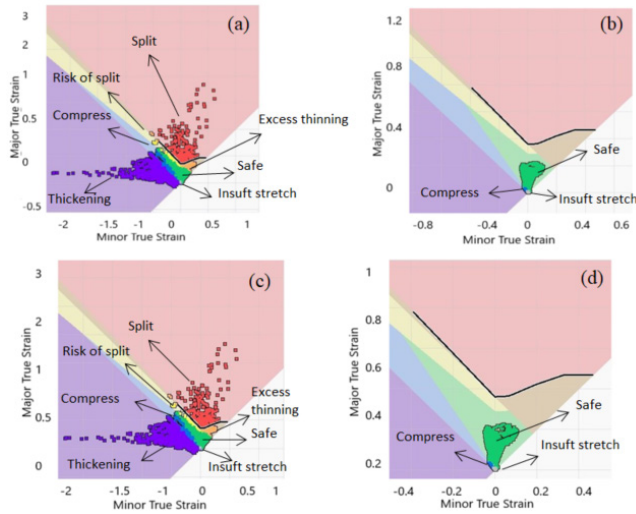


Figure 4. FLD results of SCGA before and after trimming (a-b), SPCC before and after trimming (c-d).

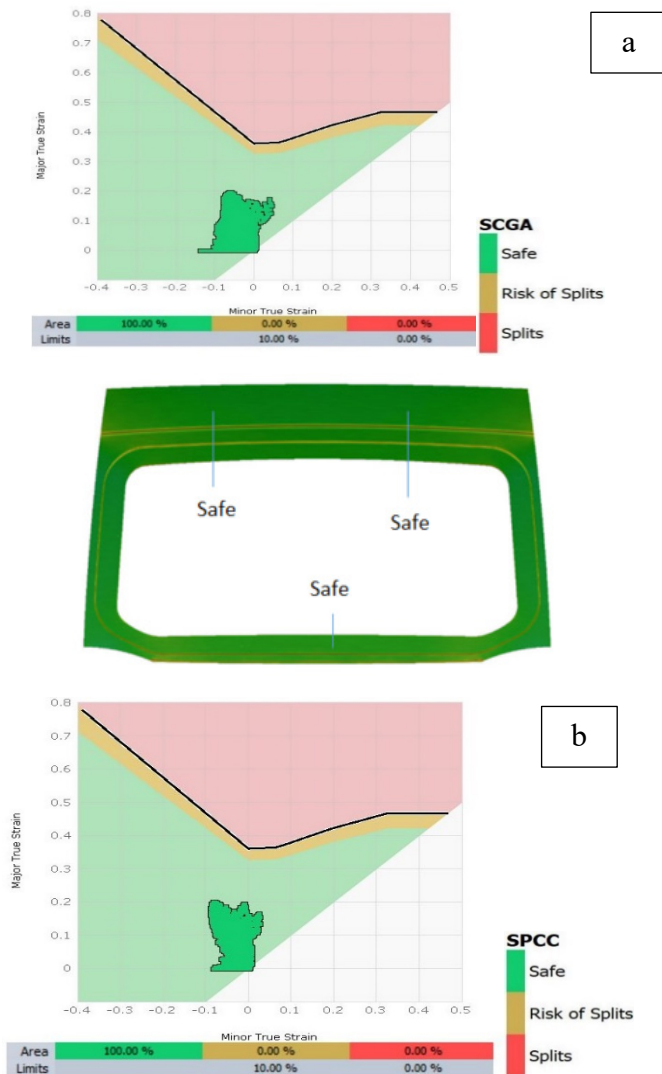
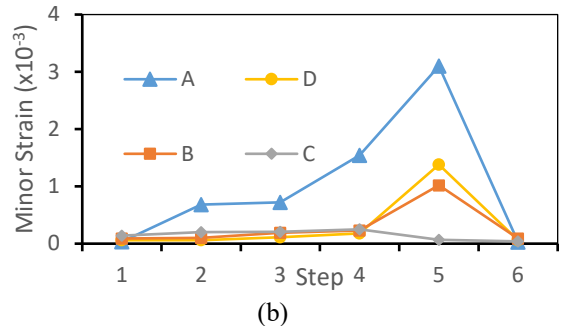
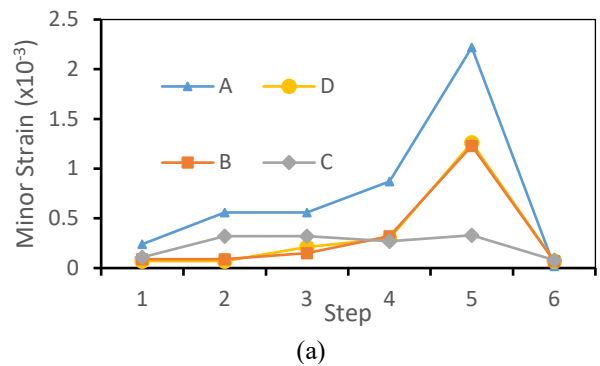


Figure 5. Nonlinear FLD results of SCGA (a) and SPCC (b) material.

3.3. Strain Based Investigation

The material behavior has been observed in six predetermined steps. Observations have been made for minor strains, as shown in Figure 6 (a-b) for both SCGA and SPCC material. The materials have changed the value of minor strains in step five. Observations have been made in four areas with potential defects, namely in areas A, B, C, D, as shown in Figure 6 (c-d). In area A, the highest minor strain values were 2.2×10^{-3} for SCGA material and 3.10×10^{-3} for SPCC material. It has happened in the drawing process. In the sixth step or the trimming process, the minor strain has decreased. It was 0.03×10^{-3} and 0.02×10^{-3} for SCGA and SPCC respectively.

The minor strain is a smaller value of principal logarithmic strain. Minor strain values indicate that the material has a stretching. The value of this strain affects the wrinkling phenomenon. On observing the process of changing minor strains for SCGA material from steps 3, 4, and 5, the value of minor strains was increased. That was resulted in wrinkling in part, shown in Figure 6 (c-d). The same thing has happened to the SPCC material. The minor strain and wrinkling values also increase in steps 4 and 5. The wrinkling area for both materials occurs at the edge of the trimming area.



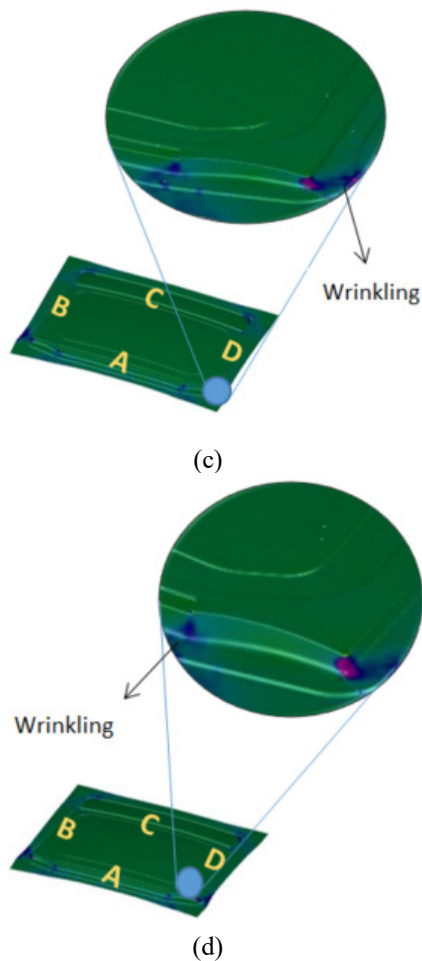


Figure 6. Minor strain results and wrinkling.

Figure 7 shows the results of major strains for both SCGA and SPCC materials. Observations have been made in four areas that have the potential to experience defects, namely in areas A, B, C, D, the same as previous observations. Major strains show the greater value of principal logarithmic strains. It affects the behavior of thinning and wrinkling during the forming process.

In the C area, the SPCC material was experienced a more significant major strain than SCGA. The most considerable major strain was occurred at the end of the fifth drawing step. In this step, the dominant strain values are SCGA 8.0×10^{-3} material and SPCC 7.9×10^{-3} material. Major strains in SCGA material have increased significantly from step 4 to step 5. It has resulted in wrinkling in several areas of the blank. In the SPCC material, the significant stress begins to increase from stage 4 to stage 5, generating more wrinkling on the edge of the blank. In step six, the major strain has dropped to zero because the process has been completed for all materials. Based on the analysis of major and minor strains, the forming process has been seen to be safe because it did not exceed the yield strain of material.

3.4. Stress Based Investigation

The results of investigations based on the stress experienced by the blank during the forming process were shown in Figure 8. Major stress and minor stress appear to have similar patterns. These results were greater and smaller values of principal stress. Major stress was closely related to the appearance of surface

defects and springback. While minor stress was affected by wrinkling defect. The observations have been made on areas A, B, C, and D for all steps from 1-6 as well as strain observations (see Figure 6).

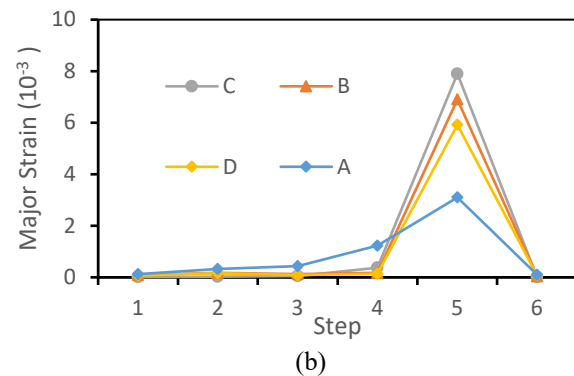
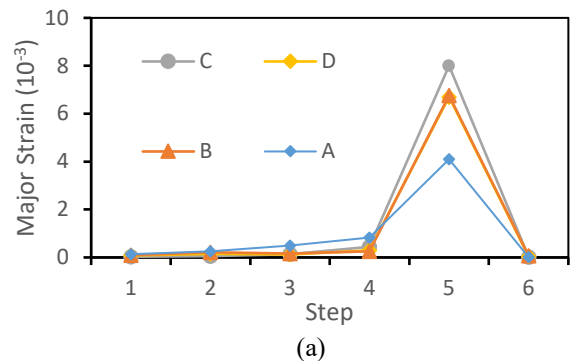


Figure 7. Major strain results for SCGA (a) and SPCC (b)

The value of major stress was always greater than minor stress. SCGA material was delivered major stress of 299.4 MPa. While at the end of the forming process, the major stress becomes 29 MPa (see Figure 8 (a-b)) because it has not suffered pressure from the dies. The most significant minor stress has occurred in area A, which was 199 MPa in the fifth step. After the trimming process, minor stress has decreased to 19 MPa because the process has been completed. Almost the same results were obtained for SPCC material for both major and minor stress, as shown in Figure 8 (c-d).

Figure 9 shows the results of wrinkling defect analysis for SCGA and SPCC materials. Wrinkling defects were corrugated surfaces at the end of the sheet metal forming process. That happened because of a non-uniform strain on the surface of the plate. These strains occur as a result of compressive stress in the direction of the plate. Other factors affecting wrinkling were material properties, friction coefficient, plate thickness, the shape of dies, blank holder force, and draw bead patterns.

Wrinkling analysis was carried out in areas A, B, C, and D as before (see Figure 10). The highest wrinkling was recorded at every step from 1-6. Based on the evaluation of the geometry, the results were obtained during the forming process. Wrinkle geometry was calculated by its radius (r). It was used to calculate the wrinkle constant, which was then called "C". The c value is one per radius (1/r). So wrinkle was calculated by multiplying C by half of the plate thickness (t/2). The results of this wrinkle are dimensionless.

In the fifth step, the highest wrinkle values were seen in all parts A, B, C, and D (Figure 10) for both materials (see Figure 9). It was related to the results of stress and strain analysis, where the highest value has occurred in the fifth step. The highest wrinkles occur in section A. It was 1.71 and 1.48 for both SCGA and SPCC materials, respectively. While at the end of the process, the wrinkle value was 0.5 and 0.24 after trimming. The wrinkle value was decreased after the trimming process because the highest wrinkling occurs outside the used area.

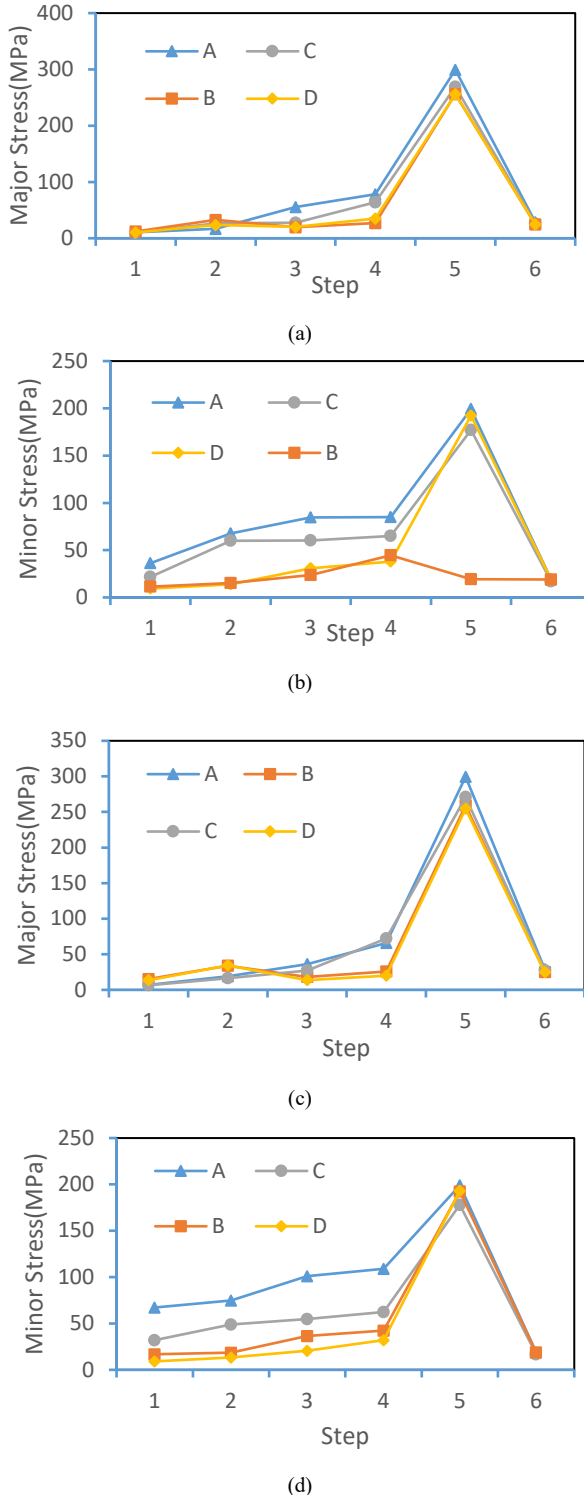


Figure 8. Major and minor stress for SCGA (a-b) and SPCC (b-c).

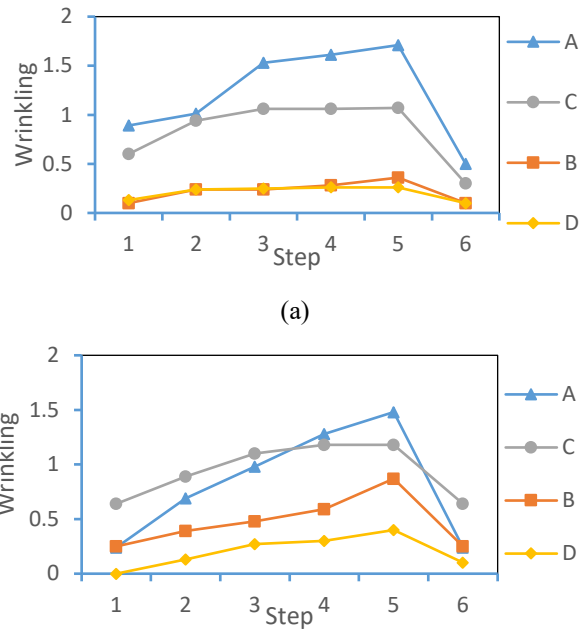


Figure 9. Results of wrinkling for SCGA (a) and SPCC (b).

The effect of the coefficient of friction on wrinkling was done by making variations of the coefficient of friction of 0, 0.05, 0.1, and 0.15. Based on the observed wrinkle value, the coefficient of friction was affected the wrinkling behavior during the forming process. A high coefficient of friction gives high wrinkle value. The same results were obtained from thinning observations. An increase in the coefficient of friction gives an increased amount of thinning, as shown in Figure 11. Thinning calculation is the difference in thickness divided by the initial thickness. The thickness difference was obtained from the final thickness minus the initial thickness. Observations were made in areas B and D, which have a higher chance of thinning compared to other areas.

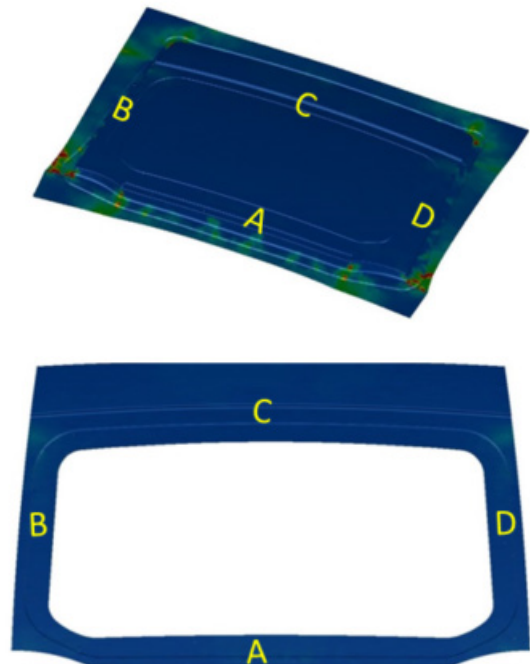


Figure 10. Wrinkling and thinning area investigation at step 5-6.

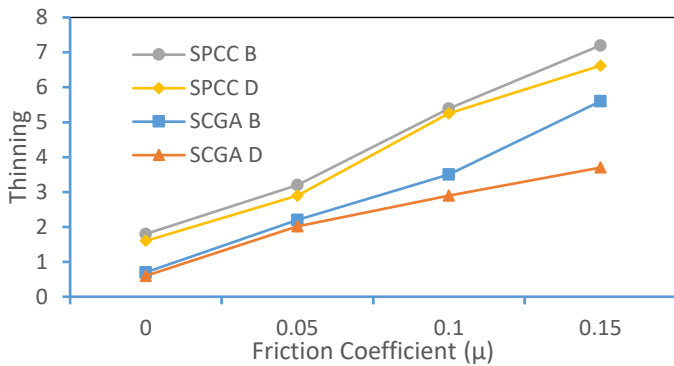


Figure 11. Influence of friction coefficient to the thinning.

4. Conclusions

The changes in wrinkling and thinning behavior on SCGA and SPCC material were successfully observed in six steps. The observations has shown that the plate received the higher stress-strain. The higher chance of wrinkling and thinning was delivered. It cannot be avoided because the deep drawing process always has high compressive stress on the plate. Major stress was increased from 78.46 MPa to 299.3 MPa in step 5. The coefficient of friction was affected the wrinkling and thinning behavior during the sheet metal forming process. The efforts to reduce friction was significant to reduce the phenomenon of wrinkling and thinning. At the end of the investigation, the forming results of the top outer hatchback part have yielded good results with minimal wrinkling and thinning defects.

Conflict of Interest

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Acknowledgment

The authors would like to acknowledge the Ministry of Research Technology and Higher Education, the Republic of Indonesia for the funding support of the research project under the schema of Thesis Magister Research No. 133.52/A.3-III/LPPM/IV/2020, and Master Program of Mechanical Engineering Department of Universitas Muhammadiyah Surakarta for the support of the project.

References

- [1] A. Zoesch, T. Wiener, and M. Kuhl, "Zero Defect Manufacturing: Detection of Cracks and Thinning of Material during Deep Drawing Processes," *Procedia CIRP*, vol. 33, pp. 179–184, 2015. <https://doi.org/10.1016/j.procir.2015.06.033>.
- [2] M. Sigvant et Sigvant, Mats Pilthammar, Johan Hol, Johan Wiebenga, Jan Harmen Chezan, ToniCarleer, Bartvan den Boogaard, Ton, "Friction in sheet metal forming: influence of surface roughness and strain rate on sheet metal forming simulation results," *Procedia Manuf.*, vol. 29, pp. 512–519, 2019. <https://doi.org/10.1016/j.promfg.2019.02.169>
- [3] M. El Sherbiny, H. Zein, M. Abd-Rabou, and M. El shazly, "Thinning and residual stresses of sheet metal in the deep drawing process," *Mater. Des.*, vol. 55, pp. 869–879, 2014. <https://doi.org/10.1016/j.matdes.2013.10.055>
- [4] X. L. Geng, B. Wang, Y. J. Zhang, J. X. Huang, M. M. Duan, and K. S. Zhang, "Effect of crystalline anisotropy and forming conditions on thinning and rupturing in deep drawing of copper single crystal," *J. Mater. Process. Technol.*, vol. 213, no. 4, pp. 574–580, 2013. DOI: 10.1016/j.jmatprotec.2012.11.009

- [5] M. A. Hassan, N. Takakura, and K. Yamaguchi, "Friction aided deep drawing of sheet metals using polyurethane ring and auxiliary metal punch. Part 1: experimental observations on the deep drawing of aluminum thin sheets and foils," *Int. J. Mach. Tools Manuf.*, vol. 42, no. 5, pp. 625–631, 2002. DOI: 10.1016/S0890-6955(01)00146-8
- [6] P. Ramanjaneyulu, P. Venkataramaiah, and K. D. Reddy, "Multi parameter optimization of deep drawing for cylindrical cup formation on brass sheets using Grey Relational Analysis," *Mater. Today Proc.*, vol. 18, pp. 2772–2778, 2019. <https://doi.org/10.1016/j.matpr.2019.07.142>
- [7] A. Atrian and H. Panahi, "Experimental and finite element investigation on wrinkling behaviour in deep drawing process of Al3105/Polypropylene/Steel304 sandwich sheets," *Procedia Manuf.*, vol. 15, pp. 984–991, 2018. <https://doi.org/10.1016/j.promfg.2018.07.396>
- [8] W. Liu, Y. Xu, and S. Yuan, "Effect of Pre-bulging on Wrinkling of Curved Surface Part by Hydromechanical Deep Drawing," *Procedia Eng.*, vol. 81, pp. 914–920, 2014. <https://doi.org/10.1016/j.proeng.2014.10.117>
- [9] E. Doege, T. El-Dsoki, and D. Seibert, "The prediction of necking and wrinkles in deep drawing processes using the FEM," in *Materials Processing Defects*, vol. 43, S. K. Ghosh, Ed. Elsevier, 1995, pp. 91–105. [https://doi.org/10.1016/S0922-5382\(05\)80007-8](https://doi.org/10.1016/S0922-5382(05)80007-8)
- [10] A. K. Choubey, G. Agnihotri, C. Sasikumar, and M. Singh, "Analysis of Die Angle in Deep Drawing Process Using FEM," *Mater. Today Proc.*, vol. 4, no. 2, pp. 2511–2515, 2017. <https://doi.org/10.1016/j.matpr.2017.02.104>
- [11] I. A. Choudhury, O. H. Lai, and L. T. Wong, "PAM-STAMP in the simulation of stamping process of an automotive component," *Simul. Model. Pract. Theory*, vol. 14, no. 1, pp. 71–81, 2006.
- [12] N. Triantafyllidis, B. Maker, and S. K. Samanta, "An Analysis of Drawbeads in Sheet Metal Forming: Part I — Problem Formulation," 2018. <https://doi.org/10.1115/1.3225889>
- [13] B. Meng, M. Wan, X. Wu, S. Yuan, X. Xu, and J. Liu, "Inner wrinkling control in hydrodynamic deep drawing of an irregular surface part using drawbeads," *Chinese J. Aeronaut.*, vol. 27, no. 3, pp. 697–707, 2014. <https://doi.org/10.1016/j.cja.2014.04.015>
- [14] J. Cao and M. C. Boyce, "Sae Technical Draw Bead Penetration as a Control Element of Material Flow," no. 412, 1993. DOI: <https://doi.org/10.4271/930517>
- [15] T. H. Choi and H. Huh, "Draw-bead Simulation by an Elasto-plastic Finite Element Method with Directional Reduced Integration," vol. 0136, no. 1, 1997. [https://doi.org/10.1016/S0924-0136\(96\)02704-5](https://doi.org/10.1016/S0924-0136(96)02704-5)
- [16] S. Candra, I. M. L. Batan, W. Berata, and A. S. Pramono, "Analytical study and FEM simulation of the maximum varying blank holder force to prevent cracking on cylindrical cup deep drawing," *Procedia CIRP*, vol. 26, pp. 548–553, 2015. <https://doi.org/10.1016/j.procir.2014.08.018>
- [17] M. Nozic, "Numerical Simulation Of Deep Drawing Process With Faculty of Mechanical Engineering," no. September, pp. 10–11, 2013. <https://doi.org/10.5545/sv-jme.2010.258>
- [18] T. Meinders, H.J.M. Geijselaers, J. Huétink "Equivalent Drawbead Performance In Deep Drawing Simulations," 1994.
- [19] B. R. Billade and P. S. K. Dahake, "Optimization of Forming Process Parameters in Sheet Metal Forming Of Reinf-Rr End Upr-Lh / Rh for Safe Thinning," vol. 8, no. 8, pp. 1–7, 2018. DOI: 10.9790/9622-0808010107. 1
- [20] F. E. M. Study, "Plastic Wrinkling Investigation of Sheet Metal Product Made by Deep Forming Process : A FEM Study," vol. 3, no. 10, pp. 186–191, 2014.
- [21] V. Laxman and S. R. Srivatsa, "Sheet Metal Forming Processes – Recent Technological Advances," *Mater. Today Proc.*, vol. 5, no. 1, pp. 2564–2574, 2018. <https://doi.org/10.1016/j.matpr.2017.11.040>
- [22] J. P. De Magalhães Correia and G. Ferron, "Wrinkling predictions in the deep-drawing process of anisotropic metal sheets," *J. Mater. Process. Technol.*, vol. 128, no. 1–3, pp. 178–190, 2002. DOI: 10.1016/S0924-0136(02)00448-X
- [23] A. D. Anggono, W. A. Sharif, A. Trianto, and M. Y. Darmawan, "Influence of lubrication and blank holder force in dome wrinkling defect on cup drawing process," *ARNP J. Eng. Appl. Sci.*, vol. 11, no. 16, pp. 9985–9991, 2016. http://www.arnpjournals.org/jeas/research_papers/rp_2016/jeas_0816_4870.pdf
- [24] W. A. Siswanto, A. D. Anggono, B. Omar, and K. Jusoff, "An alternate method to springback compensation for sheet metal forming," *Sci. World J.*, vol. 2014, 2014. <https://doi.org/10.1155/2014/301271>
- [25] S. Jadhav, M. Schoiswohl, and B. Buchmayr, "Applications of Finite Element Simulation in the Development of Advanced Sheet Metal Forming ProcessesAnwendungen der Finite-Elemente-Simulation für die Entwicklung hochwertiger Blechumformprozesse," *BHM Berg- und Hüttenmännische Monatshefte*, vol. 163, no. 3, pp. 109–118, 2018.

- [26] R. Dwivedi and G. Agnihotri, "Study of Deep Drawing Process Parameters," *Mater. Today Proc.*, vol. 4, no. 2, pp. 820–826, 2017. <https://doi.org/10.1016/j.matpr.2017.01.091>
- [27] L. Zhang, H. Liu, and W. Wang, "Numerical Simulation and Analysis of Hydromechanical Deep Drawing Process for Half-three-way Tube," *Procedia Eng.*, vol. 174, pp. 524–529, 2017. <https://doi.org/10.1016/j.proeng.2017.01.181>
- [28] G. Sun, W. Zhang, Z. Wang, H. Yin, G. Zheng, and Q. Li, "A novel specimen design to establish the forming limit diagram (FLD) for GFRP through stamping test," *Compos. Part A Appl. Sci. Manuf.*, vol. 130, p. 105737, 2020. DOI: 10.1016/j.compositesa.2019.105737
- [29] S. Basak, S. K. Panda, and M. G. Lee, "Formability and fracture in deep drawing sheet metals: Extended studies for pre-strained anisotropic thin sheets," *Int. J. Mech. Sci.*, vol. 170, no. November 2019, p. 105346, 2020. DOI: 10.1016/j.ijmecsci.2019.105346
- [30] V. R. Shinge and U. A. Dabade, "Experimental Investigation on Forming Limit Diagram of Mild Carbon Steel Sheet," *Procedia Manuf.*, vol. 20, pp. 141–146, 2018. <https://doi.org/10.1016/j.promfg.2018.02.020>
- [31] H. J. Bong, F. Barlat, M. G. Lee, and D. C. Ahn, "The forming limit diagram of ferritic stainless steel sheets: Experiments and modeling," *Int. J. Mech. Sci.*, vol. 64, no. 1, pp. 1–10, 2012. <https://doi.org/10.1016/j.ijmecsci.2012.08.009>
- [32] C. D. Schwindt, M. Stout, L. Iurman, and J. W. Signorelli, "Forming Limit Curve Determination of a DP-780 Steel Sheet," *Procedia Mater. Sci.*, vol. 8, pp. 978–985, 2015. <https://doi.org/10.1016/j.mspro.2015.04.159>
- [33] H. Zein, M. El Sherbiny, M. Abd-Rabou, and M. El shazly, "Thinning and spring back prediction of sheet metal in the deep drawing process," *Mater. Des.*, volume 53, pp. 797–808, 2014. <https://doi.org/10.1016/j.matdes.2013.07.078>

Improved Nonlinear Fuzzy Robust PCA for Anomaly-based Intrusion Detection

Amal Hadri^{*1}, Khalid Chougali², Raja Touahni¹

¹LASTID Laboratory, Faculty of Science, Ibn tofail University, Kenitra, Morocco

²GREST Research Group, National School of Applied Sciences (ENSA), Kenitra, Morocco

ARTICLE INFO

Article history:

Received: 05 February, 2020

Accepted: 10 May, 2020

Online: 28 May, 2020

Keywords:

KDDCup99

Principal Component Analysis
PCA

NFRPCA

IDS

Feature extraction methods

NSL-KDD

ABSTRACT

Among the most popular tools in security field is the anomaly based Intrusion Detection System (IDS), it detects intrusions by learning to classify the normal activities of the network. Thus if any abnormal activity or behaviour is recognized it raises an alarm to inform the users of a given network. Nevertheless, IDS is generally susceptible to high false positive rate and low detection rate as a result of the huge useless information contained in the network traffic employed to build the IDS. To deal with this issue, many researchers tried to use a feature extraction methods as a pre-processing phase. Principal Component Analysis (PCA) is the excessively popular method used in detection intrusions area. Nonetheless, classical PCA is prone to outliers, very sensitive to noise and also restricted to linear principal components. In the current paper, to overcome that we propose a new variants of the Nonlinear Fuzzy Robust PCA (NFRPCA) utilizing the popular data sets KDDcup99 and NSL-KDD. The results of the conducted experiments demonstrated that the proposed approaches is more effective and gives a promising efficiency in comparison to NFRPCA and PCA.

1 Background

Thanks to the major shift in technology tools in the twenty first century, the complexity of network security has greatly increased, which gave birth to highly developed attacks. There are several traditional security methods like firewalls, data encryption and user authentication. Those techniques are insufficient to protect the network systems against all the existing threats. As a consequence, they may be less effective in detecting several dangerous attacks. Therefore, we need to strengthen our systems by adding more powerful systems such as intrusion detection system (IDS). The IDS protects the network systems by preventing the eventual damages that could be caused by an intrusion. Commonly, there is 2 principal categories of IDS, misuse-based and anomaly-based. The misuse-based method aim to classify an attack via comparing its signature with the attacks currently existing in a database of signatures of attacks and produce an alarm if any malicious activity is detected. The most two well-known misuse detection methods are STAT [1] and Snort [2]. This technique has proved its effectiveness in detecting the attacks stored in the datasets but it can not detects new intrusions or attacks and maintaining the databases is very expensive. Hence

anomaly-based detection was initiated by Anderson [3] & Denning [4], the fundamental idea behind this concept is to specify the normal behaviour or model and generate an alarm if the difference between an observation and the defined model surpasses a threshold already defined. The uniqueness of this concept is its capability to categorize new and unusual intrusions.

Nonetheless, the current network traffic data, which are usually tremendous, are in fact a big challenge to anomaly based IDS. This type of traffic may decrease the whole detection mechanism and lead frequently to a falsified classification accuracies. This kind of huge dataset often have redundant and noisy data, that can be very difficult to model.

To address that, many feature extraction techniques have been used to increase network IDS efficiency. For instance, the paper [5] used a Discrete Differential Evolution to recognize the most important features. The detection accuracies were enhanced significantly. Likewise, the work [6] presented an IDS that can detect several attacks by exploring just a small number of features, the algorithm utilized is called Ant Colony Optimization algorithm. In [7] the authors used a feature selection method called cuttlefish

*Corresponding Author Amal Hadri, LASTID Laboratory, Faculty of Science, Ibn tofail University, Kenitra, Morocco, amal.hadri@uit.ac.ma

which suppress the noisy and redundant data and simultaneously guarantee the quality of data. The authors of [8] proposed to utilize the Principal Component Analysis (PCA) and Fuzzy Principal Component Analysis (FPCA) as a pre-processing step, before applying the k nearest neighbor (KNN) classifier, the same authors suggest in another publication [9] an improved variant method called Robust Fuzzy PCA. The acquired results show the promising performance of the technique proposed with regard to network attacks detection, as well as false alarms reduction.

Nevertheless, PCA [10]-[12] and its linear variants are known to be sensitive to noise and outliers, which can impact on the deriving principal component (PCs) [13], therefore they effect as well the results of classification. In addition to that, PCA allows uniquely a linear dimensionality reduction [14]. Hence, in the case of complicated structures like nonlinear structures, the data will not be correctly expressed in a linear space, linear variants of PCA will not be the optimal solution. To deal with this issue, NFRPCA (Non-linear Fuzzy Robust PCA) [15] was suggested to calculate PCs by utilizing a non-linear technique.

Nevertheless, this method is based on the L_2 -norm that is highly sensitive to outliers and it also squares the error, and so the model can have a much bigger error. So as to tackle this issue, we introduce a new variant of NFRPCA called L_p -norm NFRPCA. Note that this paper is an extension of work originally published in 2018 IEEE 5th International Congress on Information Science and Technology (CiSt) [16], in that work we suggested a variant of NFRPCA employing L_1 -norm rather than the classical Euclidian norm. In this paper, we propose another variant of NFRPCA using L_p -norm, we conducted new experiments besides the ones previously proposed in [16].

The remainder of this paper is structured as follows: Section 2 deliver a brief presentation of PCA, Section 3 will present an overview of NFRPCA. We present the suggested techniques in Section 4, Section 5 is dedicated to give shortly an overview of the two popular datasets namely KDDcup99 and NSL-KDD. Section 6 presents the conducted experiments and discuss the results and conclusions are summarized in section 7.

2 Principal Component Analysis Method

Principal Component Analysis (PCA) [17] is as an exploratory data analysis tool that involves a dataset with observations on variables, it was employed extensively in several research areas. The principal concept of PCA is to change data into a downsized form and preserve most of the initial variance from the original data at the same time. In other terms, PCA major role is to change variables n that were correlated into uncorrelated state d , the uncorrelated variables d are usually called the principal components (PCs) [14] [18].

Consider we already have a data set of M vectors $v_1, v_2, v_3, \dots, v_M$ where each vector is represented by N features. To obtain the PCs we comply to the strategy explained through the steps underneath:

- Compute the mean μ of the data set

$$\mu = \frac{1}{M} \sum_{i=1}^M v_i \quad (1)$$

- Determine the deviation from the mean

$$\theta_i = v_i - \mu \quad (2)$$

- Compute the covariance matrix of the corresponding data set:

$$C_{n*n} = \frac{1}{M} \sum_{i=1}^M \theta_i \theta_i^T = \frac{1}{M} A A^T \quad (3)$$

where $A = [\theta_1, \theta_2, \theta_3, \dots, \theta_n]$

- Assume that U_k is the k^{th} eigenvector of C , λ_k the related eigenvalue and let's say that the $U_{n*d} = [U_1, U_2, \dots, U_d]$ is the matrix of these eigenvectors. Hence:

$$C U_k = \lambda_k U_k \quad (4)$$

- Sort the Eigenvalues in decreasing order, hence pick the eigenvectors (known as principal components PC_i) that have the largest Eigenvalues. We can compute the number of PCs that we could keep as follow:

$$\tau = \frac{\sum_{i=1}^d \lambda_i}{\sum_{i=1}^n \lambda_i} \quad (5)$$

- Consider t as a new sample column vector, t is projected on the reduced subspace covered by the PC_i :

$$y_i = U_i^T t \quad (6)$$

3 Nonlinear Fuzzy Robust pca (nfrpca) Method

The Nonlinear fuzzy robust principal component analysis employed in the current paper was suggested initially by Luukka in [15]. It was inspired from the robust principal component techniques that Yang & Wang proposed in [19] that fundamentally introduced by Xu & Yuilles techniques [20] where PCA learning rules are associated to energy functions and they proposed a cost function in regard to outliers. In Yang & Wang's proposed methods the cost function was modified to be fuzzy and it involved Xu & Yuilles techniques as a particular case. We introduce briefly these methods in this section. Xu and Yuille [20] suggested an objective function, based on $u_i \in \{0, 1\}$:

$$E(U, w) = \sum_{i=1}^n u_i e(x_i) + \eta \sum_{i=1}^n 1 - u_i \quad (7)$$

The variables are defined as follows: η is the threshold, $U = \{u_i | i = 1, \dots, n\}$ is the membership set, & $X = \{x_1, x_2, \dots, x_n\}$ is the dataset. The principal objective is to minimize $E(U, w)$ with regard of u_i & w . It should be noted that w is a continuous variable and u_i is a binary

variable that engender optimization challenging to resolve with a gradient descent technique. To solve the problem employing the gradient descent technique the minimization problem is simplified into maximization of Gibbs distribution as underneath:

$$P(U, w) = \frac{\exp(-vE(U, w))}{Z} \quad (8)$$

Where Z is the partition function confirming $\sum_U \int_w P(U, w) = 1$. The measure $e(x_i)$ might be, e.g. one of the below functions

$$e_1(x_i) = \|x_i - w^T x_i w\|^2 \quad (9)$$

$$e_2(x_i) = \|x_i\|^2 - \frac{\|w^T x_i\|^2}{\|w\|^2} = x_i^T x_i - \frac{w^T x_i x_i^T w}{w^T w} \quad (10)$$

To minimize $E_1 = \sum_{i=1}^n e_1(x_i)$ and $E_2 = \sum_{i=1}^n e_2(x_i)$, the gradient descent rules are

$$w^{new} = w^{old} + \alpha_t [y(x_i - u) + (y - v)x_i] \quad (11)$$

$$w^{new} = w^{old} + \alpha_t (x_i y - \frac{w}{w^T w} y^2) \quad (12)$$

Where $y = w^T x_i$, $u = yw$, $v = w^T u$ and α_t is the learning rate. The nonlinear case of PCA was presented as underneath :

$$e_3(x_i) = \|x_i - w^T g(y)\| \quad (13)$$

And $y = x_i w$ and g could be selected as nonlinear functions. The weight updating in this situation is

$$w^{new} = w^{old} + \alpha_t (x_i e^T w^{old} F + e_3(x_i) g(y)) \quad (14)$$

$$\text{And: } F = \frac{d}{dy}(g(y)).$$

The cost function proposed by Yang and Wang:

$$E = \sum_{i=1}^n u_i^{m_1} e(x_i) + \eta \sum_{i=1}^n (1 - u_i)^{m_1} \quad (15)$$

subject to $u_i \in [0, 1]$ and $m_1 \in [1, \infty)$. Now u_i is the membership value of x_i associated to the data cluster and $(1 - u_i)$ is the membership value of x_i associated to the noise cluster and m_1 is the fuzziness variable. And $e(x_i)$ is the error between the class center and x_i .

As u_i is a continuous variable now, the complexity of an amalgamation of continuous and discrete optimization could be obviated and the gradient descent technique can be employed.

The gradient of $E(15)$ is calculated with regard to u_i .

By choosing $\frac{\partial E}{\partial u_i} = 0$, we obtain

$$u_i = \frac{1}{1 + \left(\frac{e(x_i)}{\eta}\right)^{\frac{1}{m_1-1}}} \quad (16)$$

Replacing this membership back, we get

$$E = \sum_{i=1}^n \left(\frac{1}{1 + \left(\frac{e(x_i)}{\eta}\right)^{\frac{1}{m_1-1}}} \right)^{m_1-1} e(x_i) \quad (17)$$

The gradient with regard to w would be

$$\frac{\partial E}{\partial w} = \left(\frac{1}{1 + \left(\frac{e(x_i)}{\eta}\right)^{\frac{1}{m_1-1}}} \right)^{m_1} \left(\frac{\partial e(x_i)}{\partial w} \right) \quad (18)$$

Consider

$$\beta(x_i) = \left(\frac{1}{1 + \left(\frac{e(x_i)}{\eta}\right)^{\frac{1}{m_1-1}}} \right)^{m_1} \quad (19)$$

where m_1 is the fuzziness variable. If $m_1 = 1$, the fuzzy membership downsizes into hard membership and could be picked following the concept:

$$u_i = \begin{cases} 1 & \text{if } e(x_i) < \eta \\ 0 & \text{otherwise} \end{cases}$$

right now η is a though threshold in this situation. The setting of m_1 has no rule. We sum up NFRPCA steps in Algorithm1.

Algorithm 1 NFRPCA algorithm

Step 1: At first set the count of iteration $t = 1$, bound of iteration T , learning coefficient $\alpha_0 \in (0, 1]$ soft threshold η to a very small positive value then initialize in random way the weight w .

Step 2: As long as t is smaller than T , do steps 3-9.

Step 3: Calculate $\alpha_t = \alpha_0(1 - t/T)$, set $i = 1$ and $\sigma = 0$.

Step 4: As long as i is smaller than n , do steps 5-8

Step 5: Calculate $y = w^T x_i$, $u = yw$ and $v = w^T u$.

Step 6: Calculate $g(y)$, $F = \frac{d}{dy}(g(y))$, $e_3(x_i) = x_i - w^{old} g(y)$, then the weight is updated:

$$w^{new} = w^{old} + \alpha_t (x_i e^T w^{old} F + e_3(x_i) g(y))$$

In [15] $g(y)$ was selected to be a sigmoid function such as $g(y) = \tanh(10y)$, F is the first derivative of $g(y)$.

4 The Proposed Methods

4.1 Nonlinear Fuzzy Robust L_1 -norm PCA method

We can clearly remark that, NFRPCA technique utilize an Euclidian norm for computing the reconstruction error in (13), and it is widely recognized that the Euclidian norm usually squares the reconstruction error, thus the approach have a much bigger error. Consequently, this could falsify the results and deteriorate the quality of solutions. For the sake of addressing this issue, the paper [16] suggest utilizing L_1 -norm to compute the reconstruction error.

The reconstruction error equation could be re-written as below:

$$e'_3(x_i) = \|x_i - w'^T g(y')\|_1 \quad (20)$$

$y' = x_i w'$ and g could be picked as nonlinear functions. Here the weight updating is

$$w^{new'} = w^{old'} + \alpha_t (x_i e'^T w^{old'} F' + e'_3(x_i) g(y')) \quad (21)$$

Where $F' = \frac{d}{dy'}(g(y'))$.

Equally as in algorithm 1, we utilize updating weight to compute the PCs.

Algorithm 2 L_1 -NFRPCA algorithm

Step 1: At first set the count of iteration $t = 1$, bound of iteration T , learning coefficient $\alpha_0 \in (0, 1]$ soft threshold η to a very small positive value then initialize in random way the weigh w .

Step 2: As long as t is smaller than T , do steps 3-9.

Step 3: Calculate $\alpha_t = \alpha_0(1 - t/T)$, set $i = 1$ and $\sigma = 0$.

Step 4: As long as i is smaller than n , do steps 5-8

Step 5: Calculate $y' = w'^T x_i$, $u' = y'w'$ and $v' = w'^T u'$.

Step 6: Calculate $g(y')$, $F' = \frac{d}{dy'}(g(y'))$, $e'_3(x_i) = x_i - w'^{old}g(y')$, then the weight is updated: $w'^{new} = w'^{old} + \alpha_t(x_i e'^T w'^{old} F' + e'_3(x_i)g(y'))$

In the proposed algorithm $g(y')$ was picked to be sigmoid like the function $g(y') = \tanh(10y')$, & F' is the first derivative of $g(y')$. We use the term L_1 -norm NFRPCA to refer to this technique in the rest of this paper.

4.2 Nonlinear Fuzzy Robust L_p -norm PCA method

The technique that we suggest is L_p -norm NFRPCA where we propose to substitute the L_2 -norm with generalized L_p -norm in the computation of the reconstruction error, in order to minimize the large error that L_2 -norm can cause. The reconstruction error equation would be re-written as follow:

$$e''_3(x_i) = \|x_i - w''^T g(y'')\|_p \quad (22)$$

Where $0 > p \geq 2$.

And $y'' = x_i w''$ and g could be selected as nonlinear function. Here the weight updating is

$$w''^{new} = w''^{old} + \alpha_t(x_i e''^T w''^{old} F'' + e''_3(x_i)g(y'')) \quad (23)$$

Where: $F'' = \frac{d}{dy''}(g(y''))$.

In the same way as in algorithm 1, we use the updating weight to compute the PCs. Finally the main steps of the proposed method is summarized in Algorithm 3.

Algorithm 3 L_p -norm NFRPCA algorithm

Step 1: At first set the count of iteration $t = 1$, bound of iteration T , learning coefficient $\alpha_0 \in (0, 1]$ soft threshold η to a very small positive value then initialize in random way the weight w'' .

Step 2: As long as t is smaller than T , do steps 3-9.

Step 3: Calculate $\alpha_t = \alpha_0(1 - t/T)$, set $i = 1$ and $\sigma = 0$.

Step 4: As long as i is smaller than n , do steps 5-8

Step 5: Calculate $y'' = w''^T x_i$, $u'' = y''w''$ and $v'' = w''^T u''$.

Step 6: Calculate $g(y'')$, $F'' = \frac{d}{dy''}(g(y''))$, $e''_3(x_i) = x_i - w''^{old}g(y'')$, then the weight is updated: $w''^{new} = w''^{old} + \alpha_t(x_i e''^T w''^{old} F'' + e''_3(x_i)g(y''))$

The function $g(y'')$ was picked to be sigmoid function, such as $g(y'') = \tanh(10y'')$, and F'' is the first derivative of $g(y'')$. We use the term L_p -norm NFRPCA to refer to this technique in the rest of this paper.

5 The Simulated Datasets and its Preprocessings

5.1 Datasets

In the experiments we conducted, the popular public intrusion data sets were used, they are called KDDcup99 and NSL-KDD. We present them shortly in the next subsections.

5.1.1 KDDCup99 Dataset

KDDCup99 [21, 22] is a dataset that contains many TCPdump raws, that was captured during 9 weeks. This dataset was prepared and still conducted by Intrusion Detection Evaluation Program called DARPA. The main aim of KDDCup99 is to establish a generalized data set to examine researchs in intrusion detection field.

The training data set represents 4 gigabytes of data compressed (most of it is binary TCP dump), it contains 4,898,431 records, while the test dataset contains around 311,029 connection records and each record in dataset contains exactly 41 features. The attacks existing in KDDCup99 are categorized as underneath:

- Denial of Service (DoS): cyber-attack where the perpetrator attempt to consume network or machine resources to make it unavailable or limited to its legitimate users.
- Remote to Local (R2L): Where the attacker control the remote machine pretending that he is a user of the system, by exploiting the system vulnerabilities.
- User to Root (U2R): by exploiting the vulnerabilities and the flaws existing in a machine or on a network, the hacker tries to start accessing from a normal user account as to get the root access privilege to the system.
- Probing: The intruder tries to collect useful information of all services and machines existing in the same network to exploit it later.

5.1.2 NSL-KDD Dataset

The NSL-KDD [23] data set has been created to alleviate a couple of the major shortcoming of the KDDCup99 data set. This new version has some improvement compared to KDDCup99 data set and has solved a few of its fundamental issues. The main advantages of NSL-KDD are as follow:

- It suppressed redundant instances in the training set.
- The test dat set does not contain replicate records.

- The current version allow us to use the whole data set. Correspondingly, the random selection method will be needless because of the reduction of the number of instances in both train set and test set. Consequently, the accuracy and consistency of the evaluation and review of research works will increase.
- Every complexity level group involve a couple of instances that is oppositely corresponding to the percentage of instances in the KDDCup99 dataset. Therefore, we obtain more realistic examination of various machine learning techniques.

6 Experiments and Discussion

In the current part of this paper, several experiments were performed to examine the efficiency of the suggested method utilizing KDDcup99 and NSL-KDD databases, for the sake of evaluating the effectiveness of the suggested technique. We compute the measures: detection rate (DR), F-measure and false positive rate (FPR) described underneath:

$$DR = \frac{TP}{TP + FN} * 100 \tag{24}$$

$$FPR = \frac{FP}{FP + TN} * 100 \tag{25}$$

$$FMeasure = \frac{2 * TP}{2 * TP + FP + FN} * 100 \tag{26}$$

When :

- An intrusion is successfully predicted we call it a True positives (TP).
- An intrusion is wrongly predicted we call it a False negatives (FN).
- A normal connection is wrongly predicted we call it a False positive (FP).

5.2 Normalization Phase

The normalization phase is more than important, because it allow us to apply the techniques cited above on the data set in a correct manner. To perform effectively this phase, we replaced the discrete values with continuous values for all the discrete attributes existing in the data sets through the same process used in [10], the process is briefly clarified as follow: for every attribute y that accepts x variant values. The attribute y is illustrated by x coordinates contains zeros and ones. E.g., the attribute of the protocol type, that accepts three, values i.e. tcp, udp or icmp. Utilizing this logic, these values are modified to the following coordinates (1,0,0), (0,1,0) or (0,0,1).

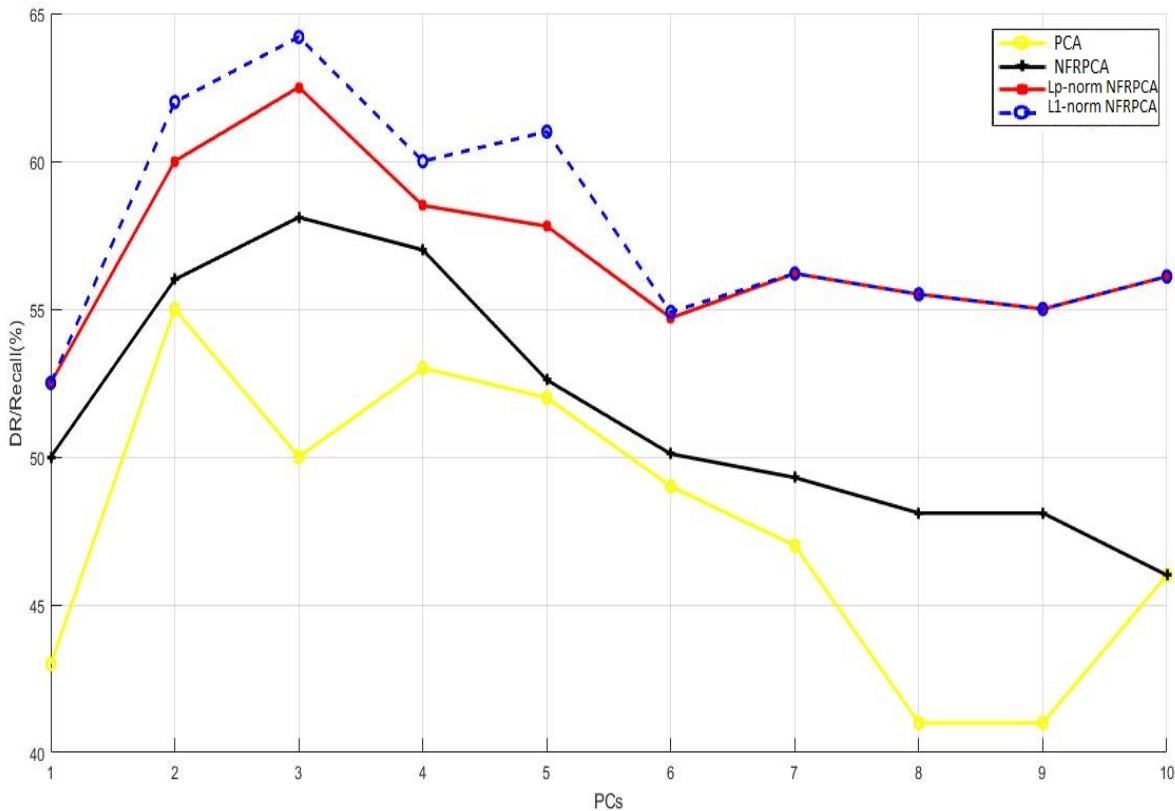


Figure 1: detection rate vs. PCs using KDDcup99 dataset

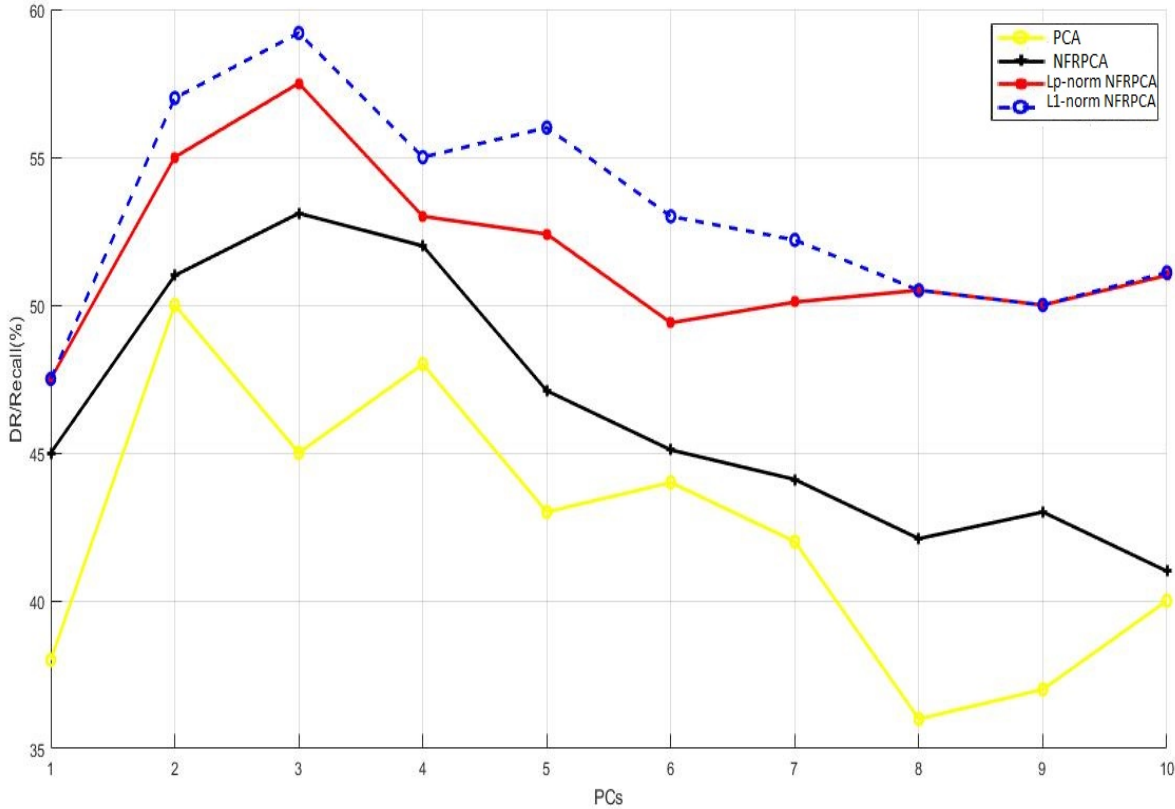


Figure 2: detection rate vs. PCs using NSL-KDD dataset

- A normal connection is correctly predicted we call it a True negatives (TN).

The classifier that was utilized in all our conducted experiments is the nearest neighbor, and for the sake of obtaining more trustworthy results we computed the average of twenty runs. The highest robust feature extraction technique must have the highest DR and F-measure rates and as much as possible the lowest FPR rate.

The simulation settings used in our first experiments are as follow: for the training sample we randomly selected 1000 normal, 100 DOS, 50 U2R, 100 R2L and 100 PROBE as for the test sample we have this structure: 100 normal data, 100 DOS data, 50 U2R data, 100 R2L data and 100 PROBE also chosen randomly using the test database. So as to both KDDcup99 and NSL-KDD data sets the settings of the simulation are similar. Also the value of p is set to 0.5 in all our experiments.

During our first experiments and to choose the ideal number of principal components(PCs) for all the feature extraction techniques which helps drastically to raise F-measure and detection rate (DR) and decrease FPR, examine and make a comparison between PCA, NFRPCA, L_1 -NFRPCA and L_p -norm NFRPCA. To achieve that, we have performed PCA, NFRPCA, L_1 -NFRPCA and L_p -norm NFRPCA to train our model first. Consequently, we obtained the PCs. The number of principal components PCs represent the dimension of the new downsized samples. Furthermore, the test sample is projected on the new downsized subspace established via the PCs.

The aim of our first experiment is to compute the measures detection rate, F-measure and false positive rate for each single principal component through choosing 10 of 41 principal component and changing their number iteratively throughout the test. We can observe clearly in the Figure.1 and Figure.2 in both datasets KDDCup99 and NSLKDD the L_1 -NFRPCA and L_p -norm NFRPCA takes the lead over the original NFRPCA and the linear PCA which

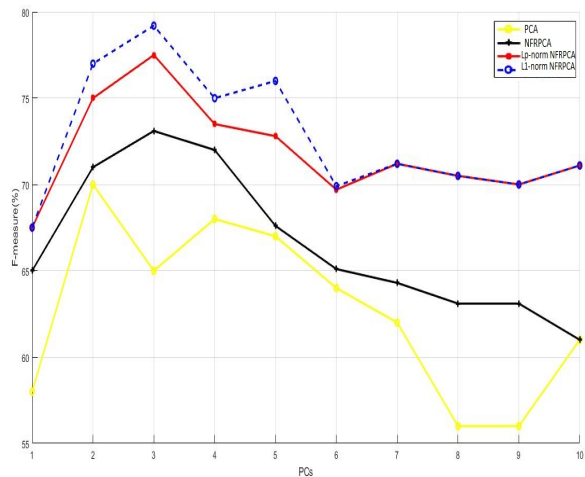


Figure 3: F-measure vs. PCs using KDDcup99 dataset

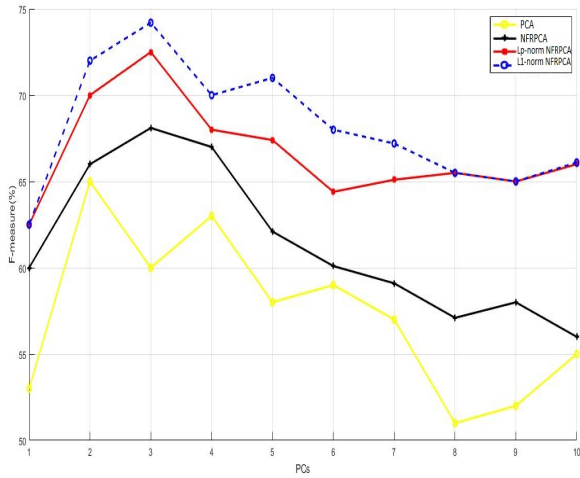


Figure 4: F-measure vs. PCs using NSL-KDD dataset

Following the previous concept, we computed the F-measure by increasing the number of PCs. As we can see from the Figure.3 and Figure.4, for both datasets the L_1 -NFRPCA and L_p -norm NFRPCA got the highest values for the F-measure comparing to the original NFRPCA and the classical PCA which support our previous results.

And so as to calculate the false positive rate (FPR), all the algorithms cited above were implemented, but only the proposed algorithms (L_1 -NFRPCA and L_p -norm NFRPCA) who has the lowest FPR for both datasets, as we see clearly in the figures (Figure 5 and Figure 6).

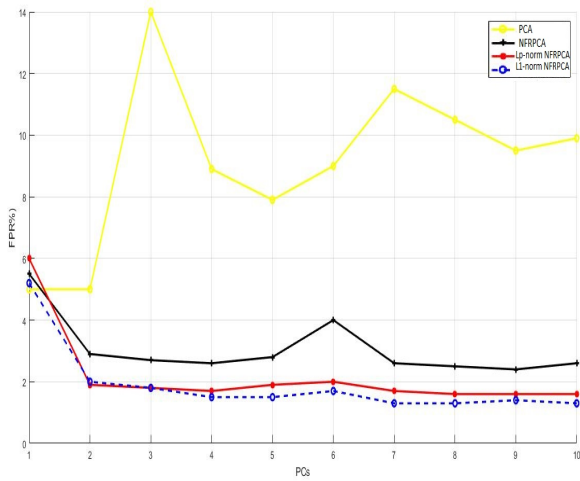


Figure 5: FPR vs. PCs using KDDcup99 dataset

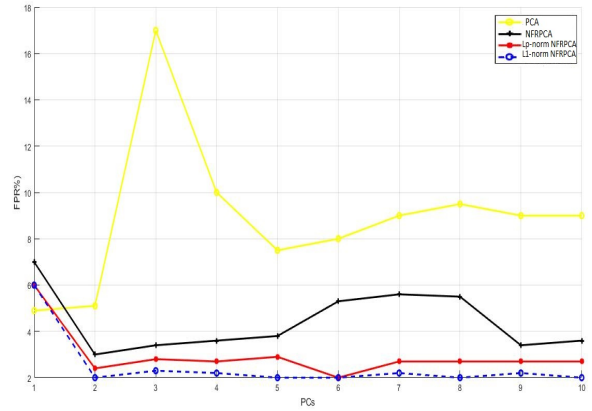


Figure 6: FPR vs. PCs using NSL-KDD dataset

In the second stage of our experiments, we intend to examine all the techniques cited above under a wide range of different training dimensionnality, and examine their impact on the DR, FPR, and F-measure. To achieve that, the structure of the test data set was kept intact by fixing it at 100 normal connections, 100 DOS, 50 U2R, 100 R2L, and 100 PROBE.

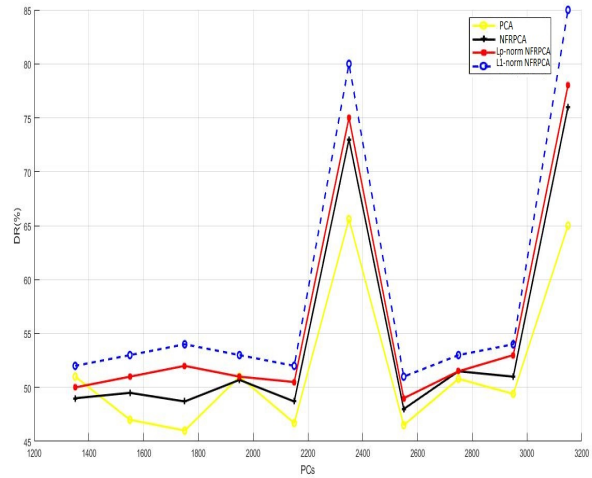


Figure 7: detection rate vs. Training data using KDDcup99 dataset

Table 1: DR of Attacks for PCA, NFRPCA, L_1 -NFRPCA and L_p -norm NFRPCA using KDDCup99 dataset

	Method	DOS	U2R	R2L	Probing
DR (%)	PCA	68,6656	8,6923	4,7734	92,1121
	NFRPCA	72,1123	14,4615	4,1165	90,2345
	L_1 -NFRPCA	73,1993	15,9815	4,1775	91,8325
	L_p -norm NFRPCA	74,2314	16,1111	4,5556	92,1211

Concerning DR, Figure.7 and Figure.8 assert that the proposed methods produce a detection rate higher than the original ones. It proves that the methods are very powerful in differentiating between normal connections and attacks.

In Figure.9, Figure.10, we can clearly see that the L_1 -NFRPCA achieve at least 5% improvement over L_p -norm NFRPCA, 10% over original NFRPCA and the classical PCA, the new approaches surpasses permanently the original techniques. In terms of FPR, the Figure.11 and Figure.12 show that the L_1 -NFRPCA still gives the lowest FPR even under different dimensionnality. These results support the great capability of the new approaches to classify the connections autonomously of the training samples size.

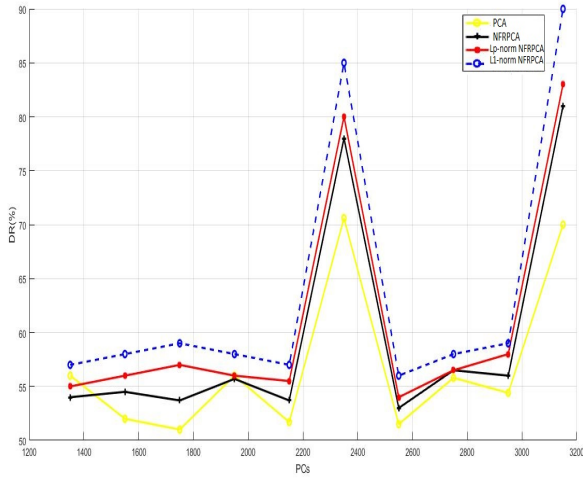


Figure 8: detection rate vs. Training data using NSL-KDD dataset

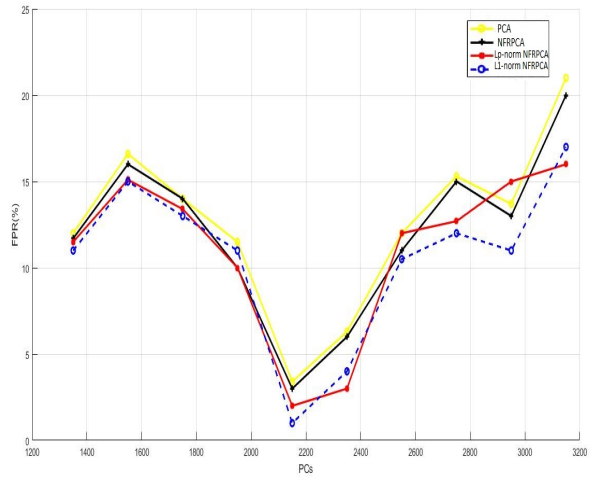


Figure 11: FPR vs. Training data using KDDcup99 dataset

Table 2: DR of Attacks for PCA, NRFPCA, L_1 -NRFPCA and L_p -norm NRFPCA using NSL-KDD dataset

	Method	DOS	U2R	R2L	Probing
DR (%)	PCA	67,5546	7,9723	4,8872	93,1121
	NRFPCA	71,1211	13,6415	4,1435	90,3478
	L_1 -NRFPCA	72,1341	13,9995	4,1435	91,3698
	L_p -norm NRFPCA	73,2215	15,1123	4,7656	93,1128

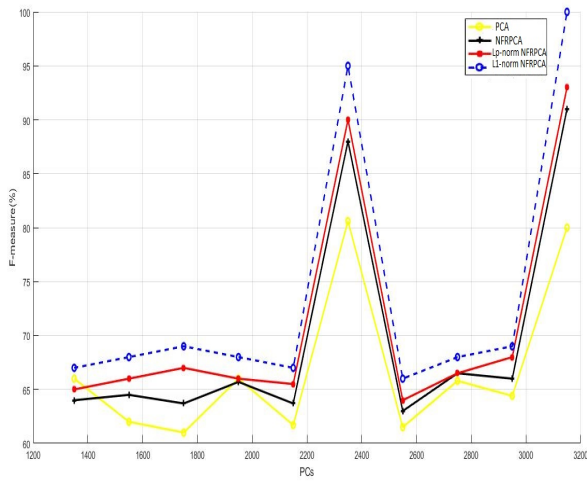


Figure 9: F-measure vs. Training data using KDDcup99 dataset

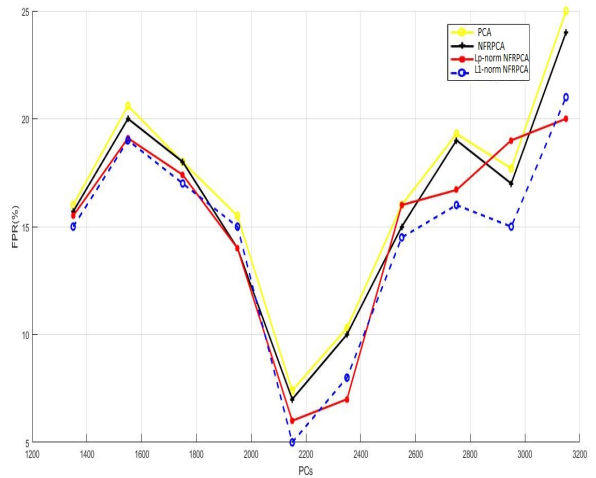


Figure 12: FPR vs. Training data using NSL-KDD dataset

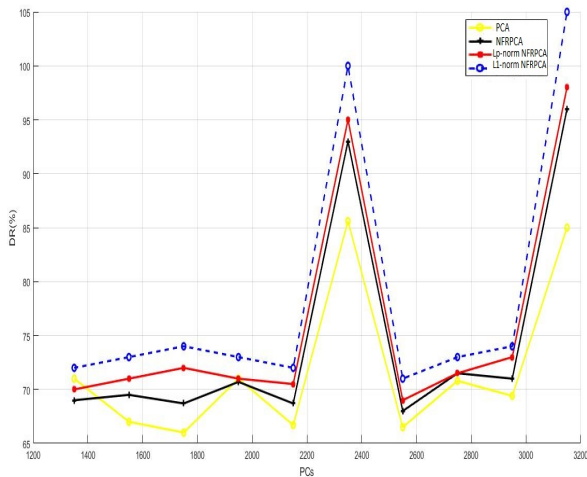


Figure 10: F-measure vs. Training data using NSL-KDD dataset

The last figures (Figure.13 and Figure.14) exhibit the correlation between CPU time and the number of principal components. As it is indicated clearly we observe that increasing the number of principal components PCs engender a huge consuming time. In addition to that, we can observe clearly in the figures that the suggested techniques are computationally speedy than the original algorithms.

To obtain higher precise results, we did an experiment in which we compared side by side the DR of every single category of attacks for PCA, NRFPCA, L_1 -NRFPCA and L_p -norm NRFPCA. According to Table I and Table II. It is obviously clear that the DR of the

L_1 -NRFPCA and L_p -norm NRFPCA for U2R and DOS attacks are often the highest compared to U2R and DOS attacks of NRFPCA and PCA.

in the detection of the most categories of attacks and in reducing the false positive alarms.

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] E. Spafford and S. Kumar, *A software architecture to support misuse intrusion detection*, in Proceedings of the 18th National Information Security Conference, 1995, pp. 194-204.
- [2] B. Caswell and J. Beale, *Snort 2.1 intrusion detection*. Syngress, 2004.
- [3] J. P. A. Co, *Computer Security Threat Monitoring and Surveillance*, 1980.
- [4] D. E. Denning, *Intrusion-Detection Model*, *IEEE Trans. Softw. Eng.*, vol. SE-13, no. 2, pp. 222-232, 1987.
- [5] E. Popoola and A. O. Adewumi, *Efficient feature selection technique for network intrusion detection system using discrete differential evolution and decision.*, International Journal Network Security, vol. 19, no. 5, pp. 660-669, 2017.
- [6] M. H. Aghdam and P. Kabiri, *Feature selection for intrusion detection system using ant colony optimization.*, International Journal Network Security, vol. 18, no. 3, pp. 420-432, 2016.
- [7] A. S. Eesa, Z. Orman, and A. M. A. Brifcani, *A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems.*, Expert Systems with Applications, vol. 42, no. 5, pp. 2670-2679, 2015.
- [8] A.Hadri, K.Choug dali and R.Touahni, *Intrusion detection system using PCA and Fuzzy PCA techniques*, in the proceeding of the International Conference on Advanced Communication Systems and Information Security (ACOSIS), 17-19 October 2016, Marrakesh, Morocco.
- [9] A.Hadri, K.Choug dali and R.Touahni, *Identifying intrusions in computer networks using Robust Fuzzy PCA*, in the proceeding of the IEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), 30 October-3 November 2017, Hammamet , Tunisia.
- [10] Y. Bouzida and N. Cuppens-boulahia, *Efficient Intrusion Detection Using Principal Component Analysis*. pp. 381-395. 2004
- [11] M.-L. Shyu, S.-C. Chen, K. Sarinnapakorn, and L. Chang, *A novel anomaly detection scheme based on principal component classifier*, Miami Univ. Dept Electr. Comput. Eng. Tech. Rep, 2003.
- [12] W. Wang and R. Batitti, *Identifying intrusions in computer networks with principal component analysis*, 2006, p. 8.
- [13] L. Xu and A. L. Yuille, *Robust principal component analysis by self-organizing rules based on statistical physics approach*, *IEEE Trans. on Neural Net.*, vol. 6, no. 1, pp. 131-143, 1995.
- [14] R. L. Kashyap, M. J. Paulik, N. Loh, A. Automation, A. K. Jain, G. M. Jenkins, S. Francisco, and L. Sirovich, *Application of the Karhunen-Lokve Procedure for the Characterization of Human Faces*, vol. 12, no. 4, 1990.
- [15] P. Luukka, *A New Nonlinear Fuzzy Robust PCA Algorithm and Similarity Classifier in Classification of Medical Data Sets*, *Int. J. Fuzzy Syst.*, vol. 13, no. 3, pp. 153-162, 2011.
- [16] Amal HADRI, Khalid Choug dali, and Raja Touahni, "A Network Intrusion Detection based on Improved Nonlinear Fuzzy Robust PCA." 2018 IEEE 5th International Congress on Information Science and Technology (Cist).IEEE, 2018.
- [17] M. Ringnr, *What is principal component analysis?*, vol. 26, no. 3, pp. 303-304, 2008.
- [18] J. Shlens, M. View, and I. Introduction, *A Tutorial on Principal Component Analysis*, 2014.
- [19] T. N. Yang and S. D. Wang, *Robust algorithms for principal component analysis*, *Pattern Recognit. Lett.*, vol. 20, pp. 927-933, 1999.
- [20] L. Xu and A. L. Yuille, *Robust principal component analysis by self-organizing rules based on statistical physics approach*, *IEEE Trans. Neural Net.*, vol. 6, no. 1, pp. 131-143, 1995.

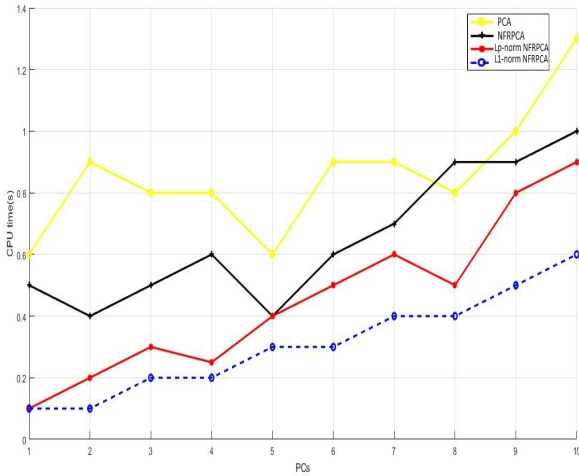


Figure 13: CPU time(s) vs. PCs using KDDCup99 dataset

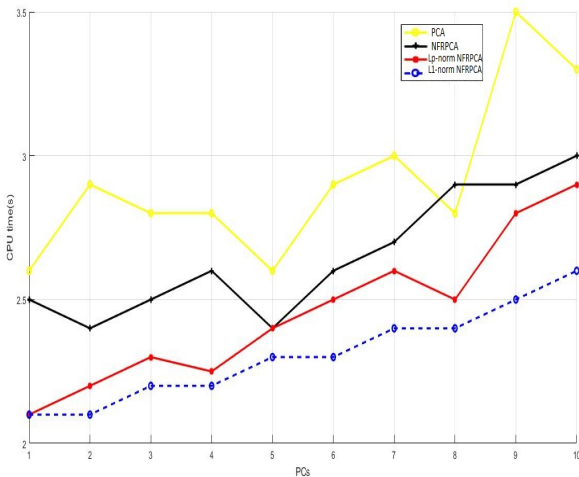


Figure 14: CPU time(s) vs. PCs using NSL-KDD dataset

7 Conclusion

As several linear statistical techniques Principal Component Analysis (PCA) has many shortcoming, it is limited just to Gaussian distribution and it has basically a high sensitivity to noise. In addition to that, principal components are frequently damaged by outliers, therefore feature extraction utilizing PCA are not credible if outliers exists in data. To tackle this issue, we proposed an effective new variants of nonlinear feature extraction techniques called Nonlinear Fuzzy Robust PCA for anomaly-based intrusion detection. The experiments performed on the popular databases (KDDcup99 and NSL-KDD), approved the effectiveness of the suggested approaches, the New variants outperform NRFPCA and PCA

- [21] M. Tavallae, E. Bagheri, W. Lu, and A. A. Ghorbani, *A Detailed Analysis of the KDD CUP 99 Data Set*, no. Cisd, pp. 1-6, 2009.
- [22] [Online] Available: <http://kdd.ics.uci.edu/databases/kddcup99/task.html>.
- [23] [Online] Available: <http://www.unb.ca/cic/research/datasets/nsl.html>.

Enhanced Collaborative Constellation for Visible Light Communication System

Manh Le Tran, Sunghwan Kim*

School of Electrical Electronics, University of Ulsan, 93, Daehak-ro, Nam-gu, Ulsan, 44610, Korea

ARTICLE INFO

Article history:

Received: 03 April, 2020

Accepted: 05 May, 2020

Online: 28 May, 2020

Keywords:

Maximum-likelihood detection

Constellation design

Wireless optical communication

ABSTRACT

Visible light communication (VLC) that simultaneously gives illumination and information transmission abilities, is recognized as a hopeful competitor for prospective wireless networks. Furthermore, a channel adaptive collaborative constellation (CASCC) with the capacity of adapting according to the channel condition to enhance the bit error rate (BER) while concurrently improving the versatility of the receiver mobility was regarded to be further effective than the contemporary CC. Nonetheless, the early CASCC barely presents modest performance enhancement in a strong correlation channel. Hence, by this study, we provide a design to form a channel-adaptation CC, namely the enhanced CC (ECC) for VLC systems. More specifically, from the basic constellation points, we form the optimization problem of efficient size that can be solved by any convex optimization solving technique. Also, the computational simulation outcomes confirm that the ECC is more beneficial than preceding constellations in term of BER for different channels. Moreover, we also provide the result comparison of the proposed scheme with other schemes in the imperfect channel condition. Overall, by effectively reducing the distance among the constellation points, a significant signal-to-noise ratio gain can be achieved.

1 Introduction

As a substitute competitor for radio frequency (RF) communications, visible light communication (VLC) has lately brought significant attention of researchers, accompanied by the growing popularity of the emerging solid-state lighting like the laser diode (LD) and then light-emitting diode (LED) to realize the conventional illumination sources thanks to their extraordinary efficiency and massive bandwidth, exceptional security, inexpensive cost, independence from the spectral licensing problem, simple implementation into existing infrastructure [1, 2]. In conventional VLC systems that convey information based on LEDs, because of the low hardware-cost and simplicity of implementation, the intensity-modulation (IM) with direct-detection (DD) designs are employed. Furthermore, the communications systems with optical multiple-input multiple-output (MIMO) techniques [3–6] can be engaged to obtain various data transmission rates [7]. Though, MIMO-VLC systems, which heavily depend on channel correlations, severely influence the achievable performance [8]. On the other hand, various modulation schemes have been proposed for MIMO-VLC, such as repetition code (RC), spatial modulation (SM), and spatial multiplexing (SMP) [9]. In RC the same signal is emitted simultaneously from all transmitters.

In SM, the conventional signal constellation diagram is extended to the spatial dimension, in which only one transmitter is activated at any symbol duration. In SMP, independent data streams are simultaneously emitted from all transmitters.

Recently, an appealing design named collaborative constellation (CC) concerning MIMO-VLC systems with various numbers of PDs and LEDs was introduced [10]. To outlining the transmitted constellation of varied LED spaces, the Euclidean distances (EDs) were simultaneously maximized following the constrains in average transmit power. Later, to considerably enhanced the design of CC, the channel adaptive space CC (CASCC) was proposed in [11] by taking the channel influences on the system model into consideration. More particularly, in expressions including four primary points, CASCC design intended to a 2×2 MIMO systems and later prolonged to higher-order constellations. Noticeably, the idea about CASCC mentioned in [11] can accommodate to different channels among variations in positions of both LEDs and PDs.

In this study, we offer a practical algorithm to outline a multi-layer CC (ECC) for MIMO-VLC system utilizing simultaneously the space-collaboration and channel-adaptation ideas [3]. Furthermore, the introduced approach attends to produce the constellations of each size based on particular primary points. Besides, with the

*Corresponding Author Sunghwan Kim, School of Electrical Electronics, University of Ulsan, 93, Daehak-ro, Nam-gu, Ulsan, 44610, Korea, sungkim@ulsan.ac.kr

help of the multilayer constellation scheme, the offered design can minimize the amount of constraint presented in optimization problems. More precisely, our offered design analyzes the distance of every pair of points; also, the form of the constellation in the PDs does not have to be distinct as in CASCC. Our proposed scheme considers the distance of any two constellation points and the shape of constellation in receiver space does not have to be a diamond shape. Hence, effectively design the adaptive collaborative constellation and achieve SNR gain in comparison with CASCC, both under perfect and imperfect CSI. Consequently, ECC efficiently produces the adaptive constellation and gains a notable SNR enhancement in comparing to CASCC including different designs.

2 Transmission model

In this study, the introduced principle of the ECC for MIMO-VLC systems, comparable to [10, 11], have the numbers of LEDs and PDs are $n = 2$, respectively. Furthermore, we denote C that contains $m = 2^r$ symbols as the constellation at the transmitting side, that is $C = \{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ with r bits as the transmission rate. Moreover, with $x_\alpha^{(1)}$ denotes the signal value transmitted from 1th-LED while $x_\alpha^{(2)}$ denotes the signal value transmitted from 2th-LED, the overall transmitted signal \mathbf{x}_α is represented as $\mathbf{x}_\alpha = (x_\alpha^{(1)}, x_\alpha^{(2)})^T \in C$. Since we have the number of constellation points in C should be a power of two, hence we at the beginning present detail procedure to design a constellation with several layers, so-called \mathcal{X} . After the constellation \mathcal{X} is initially formed, by picking a subset of constellation point in \mathcal{X} , the constellation set C is optimally obtained. This is clear since the proposed constellation \mathcal{X} with the number of layers is l , will have $l(l+1)/2$ as the cardinality.

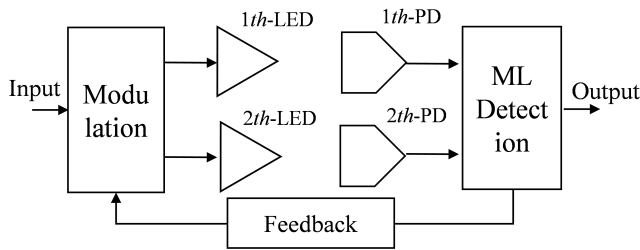


Figure 1: System model.

Following the constellation introduced in [10], a set of constellation points that contained in a number of g layers is first described. It can be seen that, in the g -th layer of constellation ($g=1, \dots, l$), as shown in Fig.1 there are g points. Moreover, begin by the primary points $\mathbf{x}_{11} = (0, 0)^T$, $\mathbf{x}_{21} = (x_{21}^{(1)}, x_{21}^{(2)})^T$, and $\mathbf{x}_{22} = (x_{22}^{(1)}, x_{22}^{(2)})^T$, the g layers constellation set can be obtained with $\mathbf{x}_{\alpha p} = (x_{\alpha p}^{(1)}, x_{\alpha p}^{(2)})^T$ denotes the p -th constellation point in the α -th layer of the proposed constellation set, and $x_{\alpha p}^{(1)}, x_{\alpha p}^{(2)}$ denote the signal values that can be transmitted simultaneously by both 1th-LED and 2th-LED, respectively. Moreover, the at the transmitter side, the constellation set \mathcal{X} will be expressed by

$$\mathcal{X} = \left\{ \mathbf{x}_{\alpha p} | \mathbf{x}_{\alpha p} = (x_{\alpha p}^{(1)}, x_{\alpha p}^{(2)})^T = \left[(\alpha - p)x_{21}^{(1)} + (p - 1)x_{22}^{(1)}, (\alpha - p)x_{21}^{(2)} + (p - 1)x_{22}^{(2)} \right]^T \right\} \quad (1)$$

with $1 \leq p \leq \alpha, 1 \leq \alpha \leq l$.

The line-of-sight (LOS) VLC systems, which comprise two LEDs and two PDs is analyzed in this study. Moreover, the received signal \mathbf{y} at the receiver, with $\mathbf{x} = (x^{(1)}, x^{(2)})^T$ are the conveyed signals and the i -th LED signal is denoted as $x^{(i)}$, similar to, [11], can be rewritten as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \quad (2)$$

Besides, we assume that \mathbf{n} is the $n_i \times 1$ additive white Gaussian noise with zero mean and σ^2 is the variance. Moreover, the assumption that the LEDs' sufficient knowledge about the channel matrix \mathbf{H} of size 2×2 through feedback channel is acceptable. Also, the channel element h_{ji} between the j -th PD and the i -th LED is described by [1]

$$h_{ji} = \begin{cases} \frac{(k+1)A_r}{2\pi\rho_{ji}^2} \cos^k(\phi) \cos(\psi) & , 0 \leq \psi \leq \psi_{1/2} \\ 0 & , \psi > \psi_{1/2} \end{cases} \quad (3)$$

where the angle-of-emergence with the LED axes and the angle-of-incidence with its normal axes can be denoted as ϕ and ψ , respectively. Also, the receiving area of the PDs can be generally expressed by A_r . $\psi_{1/2}$ indicates the field-of-view semi angle of the PDs. The distance from the j -th PD to the p -th LED is denoted as ρ_{ji} . The Lambertian order k will be described by $k = \frac{-\ln 2}{\ln(\cos \Phi_{1/2})}$ with the half-power semi angle of the LEDs is $\Phi_{1/2}$.

On the other hand, with $y_{\alpha p}^{(1)}, y_{\alpha p}^{(2)}$ denoting signals received by PD₁ and PD₂, the received symbol at both PDs $\mathbf{y}_{\alpha p}$ can be expressed as follows

$$\mathbf{y}_{\alpha p} = \mathbf{H}\mathbf{x}_{\alpha p} \Leftrightarrow \begin{bmatrix} y_{\alpha p}^{(1)} \\ y_{\alpha p}^{(2)} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_{\alpha p}^{(1)} \\ x_{\alpha p}^{(2)} \end{bmatrix} \quad (4)$$

where the corresponding transmitted constellation $\mathbf{x}_{\alpha p} \in \mathcal{X}$, $\mathcal{X} = \{\mathbf{x}_{11}, \mathbf{x}_{21}, \mathbf{x}_{22}, \mathbf{x}_{31}, \mathbf{x}_{32}, \mathbf{x}_{33}, \dots, \mathbf{x}_{l1}, \mathbf{x}_{l2}, \dots, \mathbf{x}_{ll}\}$ and the channel matrix \mathbf{H} of size 2×2 . Therefore, the constellation in the PDs will be expressed as $\mathcal{Y} = \{\mathbf{y}_{11}, \mathbf{y}_{21}, \mathbf{y}_{22}, \mathbf{y}_{31}, \mathbf{y}_{32}, \mathbf{y}_{33}, \dots, \mathbf{y}_{l1}, \mathbf{y}_{l2}, \dots, \mathbf{y}_{ll}\}$. Additionally, we have $|\mathcal{Y}| = |\mathcal{X}|$ with $\mathcal{Y} = \mathbf{H}\mathcal{X}$. Any symbol belong to \mathcal{X} , as previously remarked in [11], will be exclusively assigned to \mathcal{Y} . As a result, \mathcal{Y} is

$$\mathcal{Y} = \left\{ \mathbf{y}_{\alpha p} | \mathbf{y}_{\alpha p} = (y_{\alpha p}^{(1)}, y_{\alpha p}^{(2)})^T = \left[(\alpha - p)y_{21}^{(1)} + (p - 1)y_{22}^{(1)}, (\alpha - p)y_{21}^{(2)} + (p - 1)y_{22}^{(2)} \right]^T \right\}. \quad (5)$$

With a l layers-constellation set, the cardinality of it is calculated as $l(l+1)/2$. Furthermore, there is the need to bound the number of layers l by using the constrain that $l(l+1)/2 \geq M$ with a $M = 2^y$ points constellation C . Therefore, we arrive at the condition that $l = \left\lceil \frac{\sqrt{8M+1}-1}{2} \right\rceil$. Besides, to maintain a r bits constellation C from the primary constellation-points $\mathbf{x}_{11}, \mathbf{x}_{21}, \mathbf{x}_{22}$, following steps are performed to obtain the full constellation:

Input: matrix $\{\mathbf{x}_{11}, \mathbf{x}_{21}, \mathbf{x}_{22}\}$ where $\mathbf{x}_{11} = (0, 0)^T$, size of constellation.

Output: the optimal constellation set C

- Compute g
- Assign $\mathcal{X} = \{\mathbf{x}_{11}, \mathbf{x}_{21}, \mathbf{x}_{22}\}$
- form the full $l(l+1)/2$ symbols
- For $\alpha = 3$ to l
- For $p = 1$ to α

$$- \text{Set } \mathcal{X} \leftarrow \mathcal{X} + \left\{ \mathbf{x}_{\alpha p} | \mathbf{x}_{\alpha p} = \left(x_{\alpha p}^{(1)}, x_{\alpha p}^{(2)} \right)^T = \left[(\alpha - p)x_{21}^{(1)} + (\alpha - p)x_{22}^{(1)}, (\alpha - p)x_{21}^{(2)} + (\alpha - p)x_{22}^{(2)} \right]^T \right\}$$

- End

- End

Choose the set \mathcal{C} from the set \mathcal{X} to be the optimal constellation set.

- For $\mathbf{x}_{\alpha p} \in \mathcal{X}$

- Choose $M = 2^y$ constellation symbols that are nearest from $(0,0)$ to make the set \mathcal{C} .

- End

- With the power level requirement, normalize the set \mathcal{C} .

- Output the set \mathcal{C} as the final optimal constellation set.

3 Proposed design

In this section, the set of primary constellation points is adequately decided to maximize the minimum Euclidean distance between any two constellations on the receiver side. The optimization problem which can be manipulated to create a l layers constellation \mathcal{X} that enhance the distance between each pair of points from the obtained constellation under the total average transmitting power P_{\max} . More specifically, the optimization problem purpose is to determine optimal primary points $\mathbf{x}_{11}, \mathbf{x}_{21}, \mathbf{x}_{22}$. Also, at the receiver, the constellation is $\mathcal{Y} = \mathbf{H}\mathcal{X}$. Then, the resultant primary points will then be used to acquire l layers constellation. Moreover, we set $\mu = \alpha - \beta; \gamma = p - q$ with the assumption that $\alpha \geq \beta \geq 1$ without any loss of generality. We also define in this study the minimum Euclidean distance $d_{\mathcal{Y}}$ among any two constellation points $\mathbf{y}_{\alpha p}, \mathbf{y}_{\beta q}$. In here, $\mathbf{y}_{\alpha p}, \mathbf{y}_{\beta q}$ are the elements of a particular \mathcal{Y} . Consequently, the primary points $\mathbf{x}_{11}, \mathbf{x}_{21}, \mathbf{x}_{22}$, based on the aforementioned criterion, would be optimally designed. From equation (5), The Euclidean distance among any two points, $\mathbf{y}_{\alpha p}, \mathbf{y}_{\beta q} \in \mathcal{Y}$ will be expressed as

$$\begin{aligned} d^2(\mathbf{y}_{\alpha p}, \mathbf{y}_{\beta q}) &= \|\mathbf{y}_{\alpha p} - \mathbf{y}_{\beta q}\|^2 = \left\{ \left[(\alpha - \beta) - (p - q) \right] y_{21}^{(1)} + (p - q)y_{22}^{(1)} \right\}^2 + \left\{ \left[(\alpha - \beta) - (p - q) \right] y_{21}^{(2)} + (p - q)y_{22}^{(2)} \right\}^2 \\ &= \left[(\mu - \gamma)y_{21}^{(1)} + \gamma y_{22}^{(1)} \right]^2 + \left[(\mu - \gamma)y_{21}^{(2)} + \gamma y_{22}^{(2)} \right]^2 \\ &= \left\| \begin{bmatrix} \mu - \gamma & \gamma & 0 & 0 \\ 0 & 0 & \mu - \gamma & \gamma \end{bmatrix} \begin{bmatrix} y_{21}^{(1)} \\ y_{22}^{(1)} \\ y_{21}^{(2)} \\ y_{22}^{(2)} \end{bmatrix} \right\|_2^2 \\ &= \|\mathbf{v}_{(\mu-\gamma)\gamma} \mathbf{H} \mathbf{t}\|_2^2 = \mathbf{u}^T \mathbf{H}^T \mathbf{v}_{(\mu-\gamma)\gamma}^T \mathbf{v}_{(\mu-\gamma)\gamma} \mathbf{H} \mathbf{u} \\ &= \mathbf{u}^T \mathbf{v}_{(\mu-\gamma)\gamma} \mathbf{u} \quad (6) \end{aligned}$$

$$\text{where } \mathbf{u} = \begin{bmatrix} x_{21}^{(1)} & x_{22}^{(1)} & x_{21}^{(2)} & x_{22}^{(2)} \end{bmatrix}^T, \quad \mathbf{v}_{(\mu-\gamma)\gamma} = \begin{bmatrix} \mu - \gamma & \gamma & 0 & 0 \\ 0 & 0 & \mu - \gamma & \gamma \end{bmatrix},$$

$$\text{and } \mathbf{v}_{(\mu-\gamma)\gamma} = \mathbf{H}^T \mathbf{v}_{(\mu-\gamma)\gamma}^T \mathbf{v}_{(\mu-\gamma)\gamma} \mathbf{H}.$$

In this paper, we consider the set $\Phi = \{(\mu - \gamma, \gamma) | \gamma = p - q; \mu = \alpha - \beta; 1 \leq p \leq \alpha; 1 \leq \beta \leq \alpha \leq l; 1 \leq q \leq \beta; \{\alpha, p\} \neq \{\beta, q\}\}$.

Consequently, the optimization problem can be formed as

$$\mathcal{P}-1: \quad \max \min \mathbf{u}^T \mathbf{v}_{(\mu-\gamma)\gamma} \mathbf{u}; (\mu - \gamma, \gamma) \in \Phi$$

$$\text{s.t.:} \quad P_{\mathcal{X}} \leq P_{\max}$$

and can be reformed as follows

$$\mathcal{P}-2: \quad \max p$$

$$\text{s.t.:} \quad \mathbf{u}^T \mathbf{v}_{(\mu-\gamma)\gamma} \mathbf{u} \geq p; (\mu - \gamma, \gamma) \in \Phi$$

$$P_{\mathcal{X}} \leq P_{\max}$$

Still, to determine the global solutions of $\mathcal{P}-2$ is challenging because $\mathcal{P}-2$ is a non-convex optimization problem. Accordingly, the convex relaxation strategy can be engaged in order to approximate $\mathcal{P}-2$. More specifically, by linearizing $\mathbf{u}^T \mathbf{v}_{(\mu-\gamma)\gamma} \mathbf{u}$ in \mathbf{u}_k , the relaxed optimization problem can be presented as

$$\mathcal{P}-3: \quad \max p$$

$$\text{s.t.:} \quad 2\mathbf{u}_k^T \mathbf{v}_{(\mu-\gamma)\gamma} \mathbf{u} - \mathbf{u}_k^T \mathbf{v}_{(\mu-\gamma)\gamma} \mathbf{u}_k \geq p; (\mu - \gamma, \gamma) \in \Phi$$

$$P_{\mathcal{X}} \leq P_{\max}$$

Furthermore, because $\mathcal{P}-3$ surely is a convex optimization problem, the optimization package helpful in solving convex optimization problems can be engaged to solve (P3). Besides, the initial primary constellation points are assumed as $\mathbf{x}_{11} = (0, 0)^T, \mathbf{x}_{21} = (0.1, 0.1)^T, \mathbf{x}_{22} = (0.2, 0.2)^T$. Moreover, to efficiently overcome and lessen the complexity cost of solving the design optimization problem, we try to reduce the total number of constraints by defining the set Φ as $\Phi = \{(0, 1); (1, 0); (a, -b) | 1 \leq a, b \leq (l-1); \gcd(a, b) = 1\}$. Moreover, to calculate the greatest common divisor between a and b , we denote a function as $\gcd(a, b)$. With the M'obius function $\mu(k)$, the cardinality of the constrain set is $|\Phi| = 2 + \sum_{k=1}^{l-1} \mu(k) \left\lfloor \frac{l-1}{k} \right\rfloor^2$.

4 Generalized Constellation Design for MIMO-VLC systems

Because of the restrictions of this study, the constellations designing problem is chiefly centered on 2×2 MIMO VLC systems. This limitation was mentioned in [10, 11] because of the mathematical complexity and large volume of numerical contents. Therefore, we shortly illustrate the idea in designing the ECC for larger numbers of LEDs and PDs. Perceive that optimization can be acquired, and the equivalent design rule would be achieved without any problem. Nevertheless, the main restriction in getting the optimal points for a larger numbers of LEDs and PDs rests in huge complexities of the optimization-problem solving when high quantities of variables and constraints are examined.

Consequently, in here a MIMO-VLC system with transmitter constellation \mathcal{C} of $M = 2^r$ symbols is considered and can be denoted as $\mathcal{C} = [\mathbf{a}_1, \dots, \mathbf{a}_M]$ where $\mathbf{a}_i = [a_i^{(1)}, a_i^{(2)}, \dots, a_i^{(N_T)}]^T \in \mathcal{C}$ is a signal vector that can be conveyed through multiple LEDs. With n_t transmitters, we define a multidimensional layered constellation that each N_i -th-dimension (N_i -D) layer comprises all $(N_i - 1)$ -D layers; for example, with $n_t = n_r = 2$, an 1D-layer can belong to a line, as in Fig.1; with $n_t = n_r = 3$, a 2D-layer consists of all 1D-layers that lie in it, and every 1D-layers is a line that can be similarly constructed as in $n_t = n_r = 2$; and so on the proposed constellation for any

$n_t = n_r$ can be built by generating all constellation-points belong to $(n_t - 1)D$ -layers. Similar to the 2×2 case, a constellation-point can be represented as $\mathbf{x}_{\alpha_1 \alpha_2 \dots \alpha_{n_t}} = (x_{\alpha_1 \alpha_2 \dots \alpha_{n_t}}^{(1)}, x_{\alpha_1 \alpha_2 \dots \alpha_{n_t}}^{(2)}, \dots, x_{\alpha_1 \alpha_2 \dots \alpha_{n_t}}^{(n_t)})^T$ where the signal transmitted by the p -th LED will be rewritten as

$$x_{\alpha_1 \alpha_2 \dots \alpha_{n_t}}^{(i)} = (\alpha_1 - \alpha_2)x_{211\dots 1}^{(i)} + \dots + (\alpha_{n_t-1} - \alpha_{n_t})x_{222\dots 2}^{(i)} + (\alpha_2 - 1)x_{211\dots 1}^{(i)} + \dots + (\alpha_{n_t} - 1)x_{222\dots 2}^{(i)} \quad (7)$$

where the primary points of the constellation here are $\mathbf{x}_{11\dots 1} = (0, 0, \dots, 0)^T$, $\mathbf{x}_{211\dots 1} = (x_{211\dots 1}^{(1)}, x_{211\dots 1}^{(2)}, \dots, x_{211\dots 1}^{(n_t)})^T, \dots$, $\mathbf{x}_{222\dots 2} = (x_{222\dots 2}^{(1)}, x_{222\dots 2}^{(2)}, \dots, x_{222\dots 2}^{(n_t)})^T$ and $1 \leq \alpha_{n_t} \leq \alpha_{n_t-1} \leq \dots \leq \alpha_2 \leq \alpha_1 \leq L$. Moreover, there will be $l!/n_t!$ points in the constellation. Then, with any transmitted constellation $\mathbf{x}_{\alpha_1 \alpha_2 \dots \alpha_{n_t}} \in \mathcal{X}$, $\mathcal{X} = [\mathbf{x}_{11\dots 1}, \mathbf{x}_{211\dots 1}, \dots, \mathbf{x}_{n_t n_t \dots n_t}]$, we define the corresponding received constellation \mathcal{Y} at the n_r PDs by vector $\mathbf{y}_{\alpha_1 \alpha_2 \dots \alpha_{n_t}}$ using equation (4) and $|\mathcal{Y}| = |\mathcal{X}|$. It is observed that the constellation when $n_t = n_r = 2$ can be consider as a particular case. Even though not detailedly considered in this article, the constellation \mathcal{C} of higher dimensions will be produced from a collection of primary points $\{\mathbf{x}_{211\dots 1}, \mathbf{x}_{221\dots 1}, \dots, \mathbf{x}_{222\dots 2}\}$ by a generalization of the algorithm in previous section. Following that, the optimization problem of the constellation forming can be readily determined in an equivalent way similar to $\mathcal{P}-0$. The method of simplifying and solving the optimization problem is conceivable by orderly extending the method outlined in this article without any mathematical challenge. Nonetheless, this article focuses on the 2×2 VLC-MIMO scheme as earlier studies [10, 11] and the expansion to the system with a higher numbers of PDs and LEDs will be presented in later study.

5 Simulation results

In this sector, the performances of the proposed ECC are presented in comparison with various conventional constellation design techniques in MIMO-VLC. Moreover, the CC and also the CASCC designs are utilized in systems with two LEDs and two PDs. Furthermore, $A_r = 1\text{cm}^2$, $\Phi_{1/2} = 60^\circ$, and $\psi_{1/2} = 60^\circ$. Besides, the full transmitting power value is the same between the constellations to ensure between the compared schemes the fairness. As mentioned in [10, 11], SNR value can be determined with the normalized channel matrix \mathbf{H} and \mathbf{x} . We generate a channel realization that

$$\mathbf{H} = \begin{bmatrix} 1.0000 & 0.8000 \\ 0.2000 & 0.3000 \end{bmatrix}$$

is the channel matrix.

5.1 BER Performance With Perfect CSI

Initially, the BER performances of CC, ECC, and CASCC are shown in Fig. 2 for a low bit rate and in Fig. 3 for a high bit rate. More particularly, in the simulations, we fixed the values of the transmission rate r to 4, 8, 10, and 12 bits. It is perceived that when related to different designs such as CC and CASCC, ECC presents more beneficial BER performances. An SNR gain of around 2dB in relation to CASCC can be achieved in ECC. Nonetheless, ECC and CASCC produce remarkable performance gains through employing

the spatial source. Besides, by considering the influence of channel coefficients in the optimization problem, the offered ECC and CASCC exceptionally relieve the association among LEDs. Moreover, ECC seeks to lessen the influence of channel correlation by further efficiently enhancing the Euclidean distances among symbols in the constellation under the equivalent power constraint in corresponding with CASCC. Hence, the offered ECC in overall can obtain more excellent performance for MIMO-VLC in comparison among different constellations.

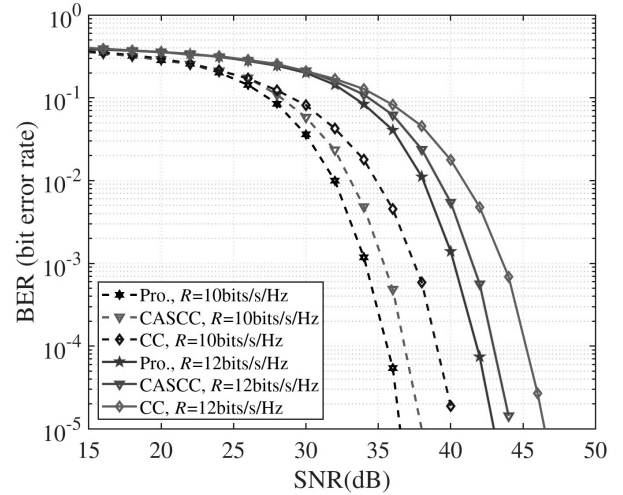


Figure 2: BER performances comparison where $r = 4, 8$ bits.

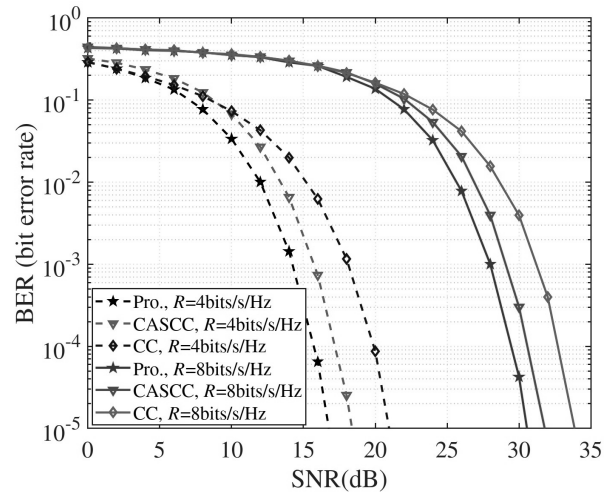


Figure 3: BER performances comparison where $r = 10, 12$ bits.

5.2 BER Performance With Imperfect CSI

Meanwhile, the channels in VLC for any particular transmitter and receiver fixed positions or specifications are deemed deterministic. In fact, the hypothesis of ideal CSI at the transmitter side is generally not vigorously practical for the indoor VLC environment [12]. To favorably obtain the conveyed symbol with a necessary degree of certainty, the information of each channel coefficient at received PDs would be of significant concern. Consequently, to assess and

produce insights on the impairment of incomplete CSI on MIMO-VLC system performance by the introduced ECC, in this section, suitable methods to imitate the channel estimating imperfection required to be carried out. Without the loss of generalization, the channel matrix evaluated at the receiver can be expressed as

$$\mathbf{H}_\epsilon = \mathbf{H} + \epsilon \quad (8)$$

where \mathbf{H}_ϵ is the estimated version of \mathbf{H} with the estimation error values expressed in a matrix ϵ of dimensions $n_r \times n_t$. The channel estimation error ϵ which is independent of \mathbf{H} and elements follow $\mathcal{N}(0, \sigma_\epsilon^2)$.

In later simulations, we illustrate the situation at $r=4$ bits in channel \mathbf{H}_2 . To set the overall impact of the estimating error on the deploying of systems, the power of the estimating error σ_ϵ^2 was set to 0.05 value for every SNR values, i.e., approximating 6.25% of the channel gain power. Particularly with ECC and CASCC that take the incomplete CSI to the design approach. Furthermore, regard that the estimation error will likewise worsen the performance of the MIMO-VLC system that engages all modulation as discussed designs such as CC, ECC, and CASCC; because ML detector is employed. Fig.4 presents the performance comparison of ECC with different designs under profoundly CSI error where $\sigma_\epsilon^2 = 0.05$.

It is clear that because of the ineffective utilization of energy that causes such small Euclidean distances in constellation-points, the performances of conventional schemes become worse. Consequently, while channel estimation errors are significant, both of them operate inadequately. In contrast, ECC notwithstanding outperforms various schemes.

Moreover, since the inadequate CSI can directly influence the design approach of both ECC and CASCC, we can easily obtain a thorough glimpse of those two designs in some incomplete CSI conditions in Fig.4. According to Fig.4, ECC can obtain at least 2dB of SNR gain in comparison with CASCC in moderate to great CSI imperfection. It is also deserving of noticing that the performance of both designs quickly deteriorates when CSI error rising.

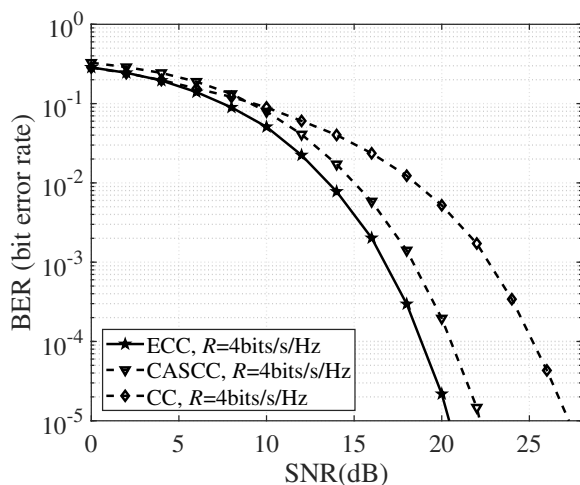


Figure 4: BER comparisons at $r = 4$ bits for incomplete CSI where $\sigma_\epsilon^2 = 0.05$.

6 Conclusion

In this study, a designing procedure concerning a constellation, particularly ECC, was offered for MIMO-VLC schemes. More specifically, the proposed scheme effectively exploits the layered structure of collaborative constellation and overcomes the aforementioned drawbacks of CASCC. By using simulations validation, the introduced method can give constellation, which achieves better BER results in comparison with several well-known constellation designs for VLC systems.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by the Research Program through the National Research Foundation of Korea (NRF-2019R1A2C1005920).

References

- [1] A. Jovicic, J. Li, T. Richardson, "Visible light communication: Opportunities, challenges and the path to market", *IEEE Commun. Mag.*, 51(12), 26-32, Dec. 2013. <https://doi.org/10/gfxf7x>
- [2] S. Dimitrov, H. Haas, Principles of LED Light Communications, en. Cambridge University Press, Mar. 2015. ISBN:978-1-107-04942-0
- [3] M. L. Tran, S. Kim, "Novel bit mapping for generalized spatial modulation in VLC systems", *IEEE Photonics Technol. Lett.*, 31(15) 1257-1260, Aug. 2019. <https://doi.org/10/gf4gp4>
- [4] M. Le Tran, S. Kim, T. Ketsoglou, E. Ayanoglu, "LED Selection and MAP Detection for Generalized LED Index Modulation", *IEEE Photonics Technol. Lett.*, 12(8), 1254-1260, Aug. 2018. <https://doi.org/10/gfi979>
- [5] M. Simon, V. Vlnrotter, "Alamouti-type space-time coding for freespace optical communication with direct detection", *IEEE Trans. Wirel. Commun.*, 4(1), 35-39, Jan. 2005. <https://doi.org/10/d6hvvk>
- [6] M. L. Tran, S. Kim, "Joint power allocation and orientation for uniform illuminance in indoor visible light communication", *Opt. Express*, OE, 27(20), 28575-28587, Sep. 2019. <https://doi.org/10.1364/OE.27.028575>
- [7] L. An, H. Shen, J. Wang, Y. Zeng, R. Ran, "Energy Efficiency Optimization for MIMO Visible Light Communication Systems", *IEEE Wirel. Commun. Lett.*, 12(10) 1322-1325, 2019. <https://doi.org/10.1109/LWC.2019.2958802>
- [8] Q. Wang, Z. Wang, L. Dai, "Multiuser MIMO-OFDM for Visible Light Communications", *IEEE Photonics J.*, 7(6), 1-11, Dec. 2015. <https://doi.org/10/gfxdm6>
- [9] M. L. Tran, S. Kim, "Layered adaptive collaborative constellation for MIMO visible light communication", *IEEE Access*, 6, 74895-74907, 2018. <https://doi.org/10.1109/access.2018.2883346>
- [10] Yi-Jun Zhu, Wang-Feng Liang, Jian-Kang Zhang, Yan-Yu Zhang, "Space-collaborative constellation designs for MIMO indoor visible light communications", *IEEE Photonics Technol Lett*, 27(15), 1667-1670, Aug. 2015. <https://doi.org/10/gfxdjb>
- [11] K. Xu, H.-Y. Yu, Y.-J. Zhu, H.-B. Cai, "Channel-adaptive space-collaborative constellation design for MIMO VLC with fast maximum likelihood detection", *IEEE Access*, 5, 842-852, Jan. 2017. <https://doi.org/10/gfxdj9>
- [12] C. Zhang, H. Du, Z. Wu, "Robust Signal Recovery for MIMO VLC System with Incomplete Channel", *IEEE Commun. Lett.*, 16(8), 3327-3331, Jan. 2019. <https://doi.org/10.1109/LCOMM.2019.2957115>

Trajectory Tracking Control of a DC Motor Exposed to a Replay-Attack

Reda El Abbadi*, Hicham Jamouli

The Laboratory of System Engineering and Decision Law, National School of Applied Sciences Ibn Zohr University Agadir, Morocco.

ARTICLE INFO

Article history:

Received: 24 April, 2020

Accepted: 18 May, 2020

Online: 28 May, 2020

Keywords:

Networked control system

Replay-attack

Data packet dropout

Delay

Linear matrix inequalities

ABSTRACT

This paper investigates the trajectory tracking control (TTC) problem of a networked control system (NCS) against a replay-attack. The impact of data packet dropout and communication delay on the wireless network are taken into account. A new mathematical representation of the NCS under network issues (packet dropout, delay, and replay-attack) is proposed, the resulting closed-loop system is written in the form of an asynchronous dynamical system. Linear matrix inequalities (LMIs) formulation and a cone complementary linearization (CCL) approach are used to calculate the controller gain F_1 and the trajectory tracking gain F_2 . Finally, a DC motor simulation with MATLAB is carried out to demonstrate the effectiveness of our approach.

1 Introduction

This article is an extended version of a conference paper presented in 2019 at the International Conference on Control and Fault-Tolerant Systems [1].

Networked Control System (NCS) is a new generation of systems where the control loop is closed through a communication network. The defining feature of an NCS is that system states and control law exchange between the components of the system (sensors, actuators, and controllers) in the form of information packages via a wireless network. The use of wireless network diminishes system wiring, simplifies maintenance and diagnosis, and improves the system agility [2, 3]. Nonetheless, the introduction of the wireless network to control the physical process brought some challenges, such as induced delay, and packet dropout, which harm the stability and reduce the system's performances. The model of the NCS with data packet dropout and network induced delay has been treated in [4–6]. There are also other challenges of using the network as a mean to impart the information between the system components, like malicious intrusions, viruses, and cyber-attackers who always find a way to access the system network and make critical damages. Consequently, reinforcing system security attracts the attention of the specialists. Remarkable papers have studied the security of the network and the cyber-attacks [7].

This article addresses a particular cyber-attack, termed as replay-attack, studied first in [8]. At our best knowledge, the trajectory

tracking control (TTC) problem of NCS under a replay-attack based on an accurate mathematical model of the replay-attack is not fully investigated, and this will be the subject of this paper. In our previous work [1, 9, 10], we modeled the NCS under a replay-attack in three different ways. In [10], we used the Markovian jump linear system to model the replay-attack. In [1], we neglected the communication delay, and we assumed that the adversary attacks just the sensor reading, and in [9], we supposed that the adversary attacks the sensor reading and the actuator's inputs simultaneously. This article involves a new contribution compared to the studies mentioned above. In this study we will give a new model of the NCS against a replay-attack, in which we take into account the effect of the delay that exists in the communication channel (sensor-controller).

This extended version deals with the TTC of a DC motor controlled through a wireless network and exposed to a replay-attack. The adversary seeks to destabilize the system by recording secretly the sensor reading and subsequently replayed it to the controller. To protect its anonymity on the network and stay undetectable for the most prolonged period, the adversary appears at different times (randomly), that sounds similar to the Stuxnet cyber-attack [11]. Moreover, we will take into account the packet dropout and the delay in the communication channel (sensor-controller). Nevertheless, the communication channel (controller-actuator) will be supposed perfect, that means 100% of packets have successfully arrived at the actuator from the controller without any delay.

The study will contain four fundamental parts. We will start by

*Corresponding Author Email: reda.elabbadi@edu.uiz.ac.ma

defining the structure of the global system. Then, we will give the model of the network. After that, we will provide our model of the replay-attack. Finally, by using the linear matrix inequalities (LMIs) formulation and the cone complementary linearization (CCL) approach we will calculate the controller gain which helps us to find the TTC gain.

2 Structure of the Global System

The structure of the NCS against a replay-attack is given in Figure 1. The sensor measurement $z(k_i)$, which has successfully arrived at the controller (the switch S1 is closed), will be first saved in a buffer. If a measurement $z(k_i)$ is lost during the transmission (the switch S1 is opened), the controller will utilize the measurements that are already stocked in the buffer to calculate the new control law $v(k_i)$. To avoid being detected by the classical detectors, the adversary will apply the attack at various times (randomly). The switch S2 models the replay-attack, if S2 is opened this means there is no attack in the buffer, whereas if S2 is closed this means there is an attack; this switching will harm the stability and the system's performances, which will reflect negatively on the trajectory tracking.

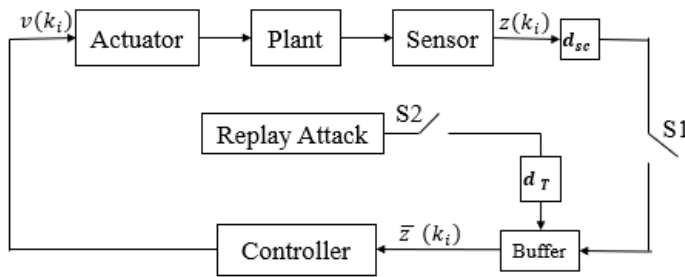


Figure 1: System architecture.

The plant of the system is defined as follows:

$$z(k_{i+1}) = Gz(k_i) + Hv(k_i), \tag{1}$$

where $z(k_i) \in \mathbb{R}^n$ is the system state and $v(k_i) \in \mathbb{R}^m$ is the control law. $G \in \mathbb{R}^{n \times n}$ is the state matrix and $H \in \mathbb{R}^{n \times m}$ is the input matrix. The state feedback controller is:

$$v(k_i) = F_1 \bar{z}(k_i), \tag{2}$$

where F_1 is the controller gain, and $\bar{z}(k_i) \in \mathbb{R}^n$ is the controller input which will be defined in the next subsections.

2.1 Wireless network model

The iterative approach described in [12, 13] was adjusted to give a model of the wireless network with the packet dropout and the communication delay d_{sc} in the channel (sensor-controller).

The Figure 2 shows the packets sent from the sensor to the controller. The green packets represent the received packets, whereas the red ones represent the dropouts packets. The notations k_i and $k_i + m$ represent respectively the green packets and the red packets, where $i \in \mathbb{Z}$ and $m \in \mathbb{N}^*$.

At first, we will ignore the effect of the replay-attack and we will concentrate on the data packet dropout and the communication delay. Hence, the controller input can be written as:

$$\bar{z}(k_i) = z(k_{i-d_{sc}}). \tag{3}$$

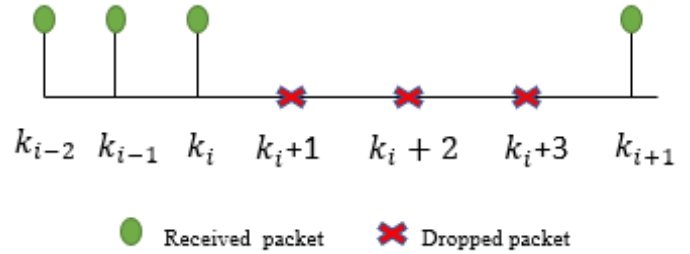


Figure 2: Received and lost packets [12].

As reported by Figure 2, the iterative approach can be defined as:

$$\begin{aligned} z(k_i + 1) &= Gz(k_i) + HF_1 z(k_{i-d_{sc}}), \\ z(k_i + 2) &= Gz(k_i + 1) + HF_1 z(k_{i-d_{sc}}), \\ &= G^2 z(k_i) + GHF_1 z(k_{i-d_{sc}}) + HF_1 z(k_{i-d_{sc}}), \\ &= G^2 z(k_i) + (GHF_1 + HF_1) z(k_{i-d_{sc}}), \\ z(k_i + 3) &= Gz(k_i + 2) + HF_1 z(k_{i-d_{sc}}), \\ &= G^3 z(k_i) + (G^2 HF_1 + GHF_1 + HF_1) z(k_{i-d_{sc}}), \end{aligned}$$

For time instant k_{i+1} , the mathematical model of the system with packet dropout and delay is:

$$z(k_{i+1}) = G^N z(k_i) + \sum_{j=0}^{N-1} G^j HF_1 z(k_{i-d_{sc}}), \tag{4}$$

with N is the number of successive dropped packets.

2.2 Replay-attack model

We assume that an attacker has connected to the buffer can replace the received packet $z(k_i)$ by the previous one $z(k_{i-d_T})$, with d_T is the replay-delay. For example, in the Figure 3 the third packet (102) was exposed to an attack with $d_T=2T_e$, so it was replaced by the first packet (100). The same for the packet (200) which was replaced by the packet (198).

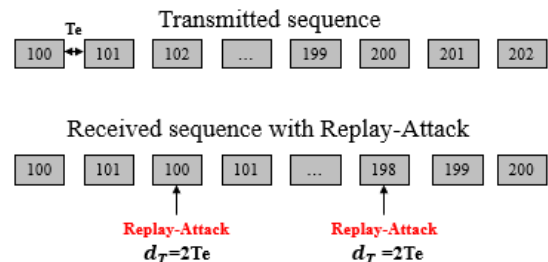


Figure 3: Replay-attack.

As reported by Figure 3, the system under a replay-delay can be defined as:

$$\begin{aligned} z(k_i + 1) &= Gz(k_i) + HF_1z(k_{i-d'_T}). \\ z(k_i + 2) &= Gz(k_i + 1) + HF_1z(k_{i-d'_T}), \\ &= G^2z(k_i) + GHF_1z(k_{i-d'_T}) + HF_1z(k_{i-d'_T}), \\ &= G^2z(k_i) + (GHF_1 + HF_1)z(k_{i-d'_T}). \\ z(k_i + 3) &= Gz(k_i + 2) + HF_1z(k_{i-d'_T}), \\ &= G^3z(k_i) + (G^2HF_1 + GHF_1 + HF_1)z(k_{i-d'_T}). \end{aligned}$$

$$\Phi_1 = \begin{bmatrix} G^N & 0 & \dots & \sum_{j=0}^{N-1} G^j HF_1 \\ I & 0 & \dots & 0 \\ 0 & I & 0 & \vdots \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & 0 & 0 \\ \vdots & \dots & \vdots & \vdots \\ \vdots & \dots & \vdots & \vdots \\ 0 & I & 0 & 0 \end{bmatrix}, \quad (10)$$

For time instant k_{i+1} , the mathematical model of the system against a replay-attack is:

$$z(k_{i+1}) = G^N z(k_i) + \sum_{j=0}^{N-1} G^j HF_1 z(k_{i-d'_T}), \quad (5)$$

where $d'_T = d_T + d_{sc}$.

The overall system will switch between two subsystems. Subsystem 1 if S2 is "off", and subsystem 2 if S2 is "on".

From (4) and (5) the global system becomes:

$$z(k_{i+1}) = G^N z(k_i) + \gamma \sum_{j=0}^{N-1} G^j HF_1 z(k_{i-d_{sc}}) + (1-\gamma) \sum_{j=0}^{N-1} G^j HF_1 z(k_{i-d'_T}), \quad (6)$$

where the variable γ equals one if S2 is "off", and equals zero if S2 is "on".

The augmented state can be written as:

$$\hat{z}(k_i) = [z^T(k_i) \ z^T(k_{i-1}) \ \dots \ z^T(k_{i-d_{sc}}) \ \dots \ z^T(k_{i-d'_T})]^T. \quad (7)$$

The overall system (6) can be written as:

$$\hat{z}(k_{i+1}) = \Phi_\sigma \hat{z}(k_i), \quad (8)$$

in which $\sigma = 1, 2$, and

$$\Phi_\sigma = \begin{bmatrix} G^N & 0 & \dots & \gamma \sum_{j=0}^{N-1} G^j HF_1 \\ I & 0 & \dots & 0 \\ 0 & I & 0 & \vdots \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & (1-\gamma) \sum_{j=0}^{N-1} G^j HF_1 \\ \vdots & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & I & 0 & 0 \end{bmatrix}. \quad (9)$$

Therefore, the overall system can be equivalent to an asynchronous dynamical system expressed in (8),

and

$$\Phi_2 = \begin{bmatrix} G^N & 0 & \dots & 0 \\ I & 0 & \dots & 0 \\ 0 & I & 0 & \vdots \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & 0 & 0 \\ \vdots & \dots & \vdots & \vdots \\ \vdots & \dots & \vdots & \vdots \\ 0 & I & 0 & 0 \end{bmatrix}. \quad (11)$$

3 Stability and Control Design

Lemma 1 [14] *The asynchronous dynamical system $z_{k+1} = f_s(z_k)$, $s = 1 \dots N'$, is exponential stable if there exists a Lyapunov function where $\beta_1 \|z\|^2 \leq V(z) \leq \beta_2 \|z\|^2$, $\beta_{1,2} > 0$, and given positive scalars α_s satisfying:*

$$V(z_{k+1}) - V(z_k) < (\alpha_s^{-2} - 1)V(z_k), \quad (12)$$

$$\alpha_1^{r_1} \alpha_2^{r_2} \dots \alpha_s^{r_s} > 0, \quad (13)$$

with r_s , ($s \in \mathbb{N}$) is the occur rate of discrete event satisfying this two conditions, $r_s > 0$ and $\sum_{s=1}^{N'} r_s = 1$.

Theorem 1 *If there exist symmetric matrices $P_1 > 0$, $P_2 > 0$ and given scalars $\alpha_\sigma > 0$, $\sigma = 1, 2$ satisfying:*

$$\alpha_1^r \alpha_2^{1-r} > 0, \quad (14)$$

$$\begin{bmatrix} -P_1 & \Phi_\sigma^T \\ \Phi_\sigma & -\alpha_\sigma^{-2} P_2 \end{bmatrix} < 0, \quad (15)$$

with minimizing the trace (P_1, P_2) subject to:

$$\begin{bmatrix} P_1 & I \\ I & P_2 \end{bmatrix} \geq 0. \quad (16)$$

Then, the system (8) is exponential stable.

Proof: applying (12) to (8), we have

$$V(\hat{z}_{k_{i+1}}) - V(\hat{z}_{k_i}) < (\alpha_\sigma^{-2} - 1)V(\hat{z}_{k_i}).$$

Hence,

$$V(\hat{z}_{k_{i+1}}) < \alpha_\sigma^{-2}V(\hat{z}_{k_i}).$$

Since $V(\hat{z}_{k_i}) = \hat{z}^T(k_i)P_1^{-1}\hat{z}(k_i)$,

$$\hat{z}^T(k_i)\Phi_\sigma^T P_1^{-1}\Phi_\sigma \hat{z}(k_i) < \alpha_\sigma^{-2}\hat{z}^T(k_i)P_1^{-1}\hat{z}(k_i).$$

Therefore,

$$\Phi_\sigma^T P_1^{-1}\Phi_\sigma - \alpha_\sigma^{-2}P_1^{-1} < 0. \quad (17)$$

To rewrite (17) in a matrix form. We will utilize the Schur complement. Then, (17) becomes:

$$\begin{bmatrix} -P_1 & \Phi_\sigma^T \\ \Phi_\sigma & -\alpha_\sigma^{-2}P_1^{-1} \end{bmatrix} < 0. \quad (18)$$

Remark 1 It is clear that (18) is not linear because of the existence of P_1 and its inverse P_1^{-1} in the same matrix. To fix this problem we will make a change of variable ($P_2 = P_1^{-1}$), and to guarantee the convergence of P_2 to P_1^{-1} , we will use the CCL approach [15]. The inequality matrices (18) becomes:

$$\begin{bmatrix} -P_1 & \Phi_\sigma^T \\ \Phi_\sigma & -\alpha_\sigma^{-2}P_2 \end{bmatrix} < 0, \quad (19)$$

with

$$P_2 = P_1^{-1}. \quad (20)$$

The CCL approach is an algorithm used to guarantee that, P_2 equals P_1^{-1} , by minimizing the trace ($P_1.P_2$) subject to:

$$\begin{bmatrix} P_1 & I \\ I & P_2 \end{bmatrix} \geq 0. \quad (21)$$

Solving the LMIs using Yalmip Toolbox, we can calculate the gain F_1 which will help us to find the trajectory tracking gain F_2 .

$$F_2^{-1} = C(I - (G + HF_1))^{-1}H. \quad (22)$$

4 Application

4.1 DC motor Model

A DC motor is a machine which converts direct current electrical power into mechanical power. The DC motor has vast applications in many fields including NCS. Owing to this importance, we chose the DC motor as an application system where we will apply our approach.

In this paragraph, the mathematical model of a DC motor will be studied.

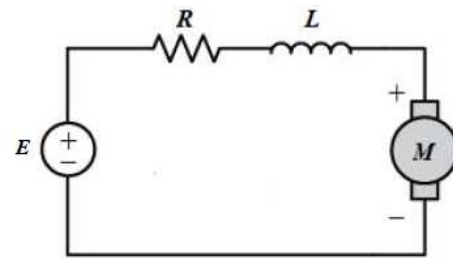


Figure 4: DC motor equivalent circuit model.

As shown in Figure 4, the system input is the voltage source (E), while the system output is the rotational speed $\dot{\theta}$. The physical parameters of the DC motor are given in Table 1:

Table 1: Physical parameters of the Dc motor.

Symbol	Description	Value
J	Moment of inertia	0.02 Kg.m ²
b	Motor viscous friction constant	0.2 N.m.s
K_e	Constant of emf	0.02V.s/rad
K_t	Motor torque constant	0.02N.m/A
R	Resistance	1.5 Ω
L	Inductance	0.5H

The motor torque and the back emf (e) are given in (23) and (24):

$$T_q = K_t.i, \quad (23)$$

$$e = K_e.\dot{\theta}. \quad (24)$$

Let us consider the constant K such that $K = K_t = K_e$. From the Figure 4, and employing the Kirchoff's voltage law, the electrical equation of the DC motor is described as:

$$J\ddot{\theta} + b\dot{\theta} = Ki, \quad (25)$$

$$L\frac{di}{dt} + Ri = E - K\dot{\theta}. \quad (26)$$

If we choose $[\dot{\theta}, i]^T$ as a state variables, the state space representation will be written as:

$$A = \begin{bmatrix} -\frac{b}{J} & \frac{K}{J} \\ -\frac{K}{L} & -\frac{R}{L} \end{bmatrix}, B = \begin{bmatrix} 0 \\ \frac{1}{L} \end{bmatrix}, C = [1 \quad 0], D = 0.$$

Replacing the parameters by their values, the state space representation becomes:

$$A = \begin{bmatrix} -10 & 1 \\ -0.04 & -3 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, C = [1 \quad 0], D = 0.$$

The command "c2d" in Matlab is used to passe from the continuous-time to discrete-time, where the sampling time is $T_e = 0.1s$. The discrete-system can be written as follow:

$$G = \begin{bmatrix} 0.3678 & 0.0563 \\ -0.0021 & 0.7407 \end{bmatrix}, H = \begin{bmatrix} 0.0066 \\ 0.1728 \end{bmatrix}, C = [1 \quad 0], D = 0.$$

4.2 Simulation and results

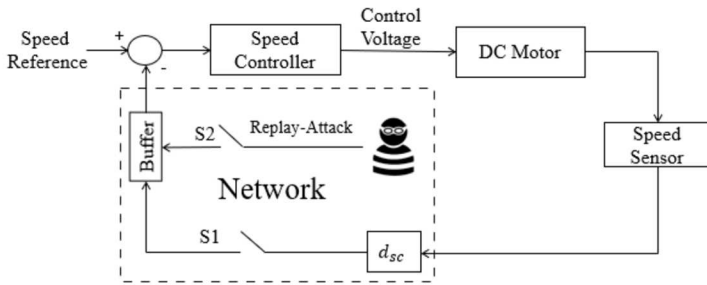


Figure 5: DC motor control under network issues.

The Figure 5 shown the structure of the DC motor under wireless network issues. The initial condition is $z(0) = [0 \ 0]^T$, $v_i = 0$, for $i \leq 0$, the maximum number of the successive packets losses during the transmission is $N=3$ packets, the communication delay equals to 0.1s, and the replay-delay equals to 0.3s, that means $d_{sc} = 1$, and $d_T=3$. We will study three different situations. In the first situation, the event rate of the switch S2 equals 0.1, in the second situation, the event rate equals 0.5, in the third, the event rate equals 0.9.

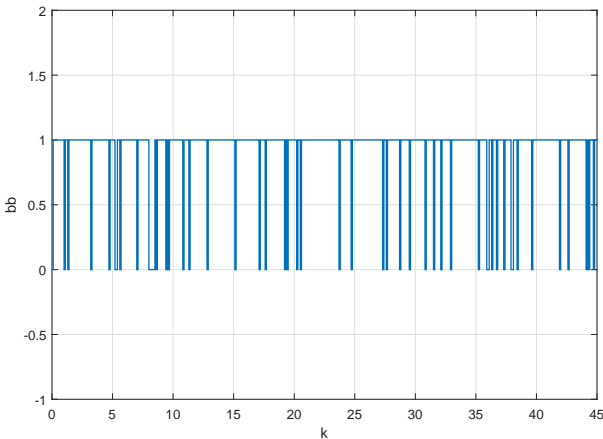


Figure 6: Event rate of S2 is 0.1.

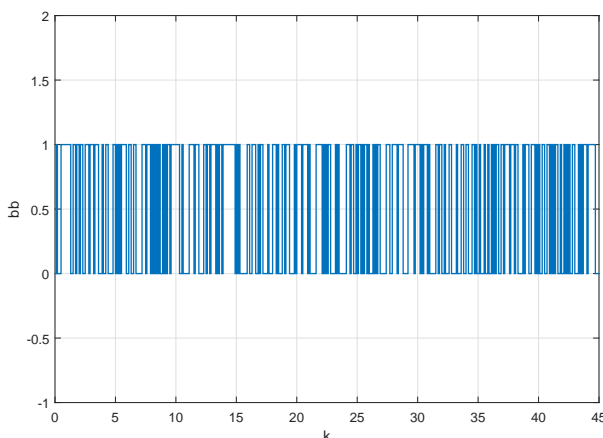


Figure 7: Event rate of S2 is 0.5.

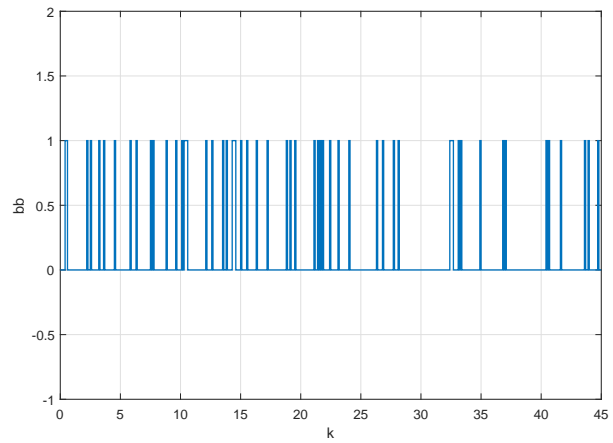


Figure 8: Event rate of S2 is 0.9.

The figures (Figure 6, Figure 7, Figure 8) show the different event rate of the switch S2. 0.1, 0.5 and 0.9 respectively.

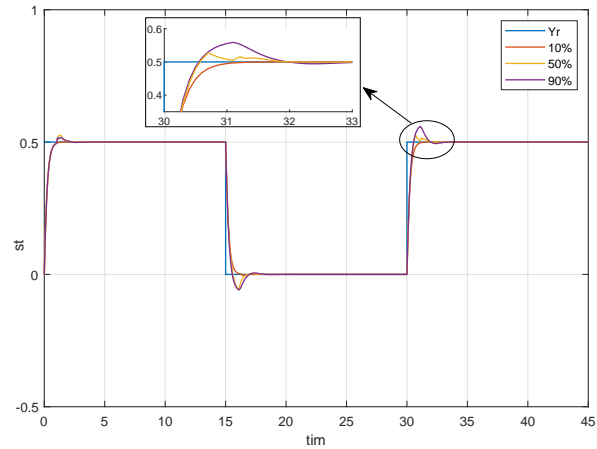


Figure 9: Trajectory tracking.

According to Theorem 1, the LMIs are feasible for $\alpha_1 = 0.1$ and $\alpha_2 = 0.45$. Solving the LMIs in Theorem 1 utilizing the Yalmip toolbox, we can find the controller gain and the trajectory tracking gain of the system as follow:

$$F_1 = [0 \ 0.0012] \text{ and } F_2 = 20.9085.$$

From Figure 9, we can see that if the event rate of the switch S2 equals 0.1 (the percentage to have an attack is 10%) the output can tracks perfectly the trajectory Y_r , the same thing happens if the chance to have an attack rises to 50% or 90%. But, in these two cases an overshoot appears. However, the results stay acceptable, which reflects the potency of our approach.

5 Conclusion

This paper dealt with the TTC issue of a DC motor controlled through a wireless network. In this extended version, we took into account the packet dropout and the delay in the communication channel (sensor-controller). On the other hand, the communication channel (controller-actuator) was assumed perfect, that means all

data have successfully transferred from the controller to the actuator without any delay. We also considered that the DC motor exposed to a replay-attack, where a cyber-adversary sought to destabilize the system and diminished its performances. A new mathematical model of the NCS under replay-attack was proposed. A sufficient condition for the stability of the resulting asynchronous dynamical system was given in the form of LMIs. The controller gain F_1 and the trajectory tracking gain F_2 were obtained by solving these LMIs employing the CCL approach. Finally, the simulation results proved the effectiveness of our approach. As a perspective of this study, our attention will be oriented towards studying the same problem with communication delay and packets losses in both communication channels.

References

- [1] R. El Abbadi and H. Jamouli, "Stabilization of cyber physical system with data packet dropout and replay attack via switching system approach," in *2019 4th Conference on Control and Fault Tolerant Systems (SysTol)*. IEEE, 2019, pp. 325–329, doi: 10.1109/SYSTOL.2019.8864787.
- [2] T. C. Yang, "Networked control system: a brief survey," *IEE Proceedings-Control Theory and Applications*, vol. 153, no. 4, pp. 403–412, 2006, doi: 10.1049/ip-cta:20050178.
- [3] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138–162, 2007, doi: 10.1109/JPROC.2006.887288.
- [4] J. Xiong and J. Lam, "Stabilization of linear systems over networks with bounded packet loss," *Automatica*, vol. 43, no. 1, pp. 80–87, 2007, doi: 10.1016/j.automatica.2006.07.017.
- [5] L. Qiu, Q. Luo, S. Li, and B. Xu, "Modeling and output feedback control of networked control systems with both time delays; and packet dropouts," *Mathematical Problems in Engineering*, vol. 2013, 2013, doi: 10.1155/2013/609236.
- [6] L. Qiu, Q. Luo, F. Gong, S. Li, and B. Xu, "Stability and stabilization of networked control systems with random time delays and packet dropouts," *Journal of the Franklin Institute*, vol. 350, no. 7, pp. 1886–1907, 2013, doi: 10.1016/j.jfranklin.2013.05.013.
- [7] J. M. R. Hernan, "Detection of attacks against cyber-physical industrial systems," Ph.D. dissertation, 2017.
- [8] Y. Mo and B. Sinopoli, "Secure control against replay attacks," in *2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2009, pp. 911–918, doi: 10.1109/ALLERTON.2009.5394956.
- [9] R. EL Abbadi and H. Jamouli, "Stabilization of a cyber physical system with network issues," in *2019 8th International Conference on Systems and Control (ICSC)*. IEEE, 2019, pp. 508–512, doi: 10.1109/ICSC47195.2019.8950544.
- [10] R. El Abbadi and H. Jamouli, "Stabilization of cyber physical system exposed to a random replay attack modeled by markov chains," in *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*. IEEE, 2019, pp. 528–533, doi: 10.1109/CoDIT.2019.8820311.
- [11] P. Shakarian, J. Shakarian, and A. Ruef, "Attacking iranian nuclear facilities: Stuxnet," *Introduction to cyber-warfare: A multidisciplinary approach*, pp. 223–239, 2013, doi: 10.1016/B978-0-12-407814-7.00013-0.
- [12] A. Routh, S. Das, and I. Pan, "Stabilization based networked predictive controller design for switched plants," in *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)*. IEEE, 2012, pp. 1–6, doi: 10.1109/ICCCNT.2012.6396001.
- [13] M. Yu, L. Wang, T. Chu, and G. Xie, "Stabilization of networked control systems with data packet dropout and network delays via switching system approach," in *2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601)*, vol. 4. IEEE, 2004, pp. 3539–3544, doi: 10.1109/CDC.2004.1429261.
- [14] A. Rabello and A. Bhaya, "Stability of asynchronous dynamical systems with rate constraints and applications," *IEE Proceedings-Control Theory and Applications*, vol. 150, no. 5, p. 546, 2003, doi: 10.1049/ip-cta:20030704.
- [15] L. El Ghaoui, F. Oustry, and M. AitRami, "A cone complementarity linearization algorithm for static output-feedback and related problems," *IEEE transactions on automatic control*, vol. 42, no. 8, pp. 1171–1176, 1997, doi: 10.1109/9.618250.

Digital Sovereignty Between “Accountability” and the Value of Personal Data

Nicola Fabiano^{*1,2}

¹Studio Legale Fabiano, 00179, Italy

²International Institute of Informatics and Systemics (IIIS), 34787, Florida, USA

ARTICLE INFO

Article history:

Received: 03 April, 2020

Accepted: 18 May, 2020

Online: 28 May, 2020

Keywords:

Digital sovereignty

Data Protection

Privacy

Ethics

ABSTRACT

In the last year, especially in Europe, the expression “digital sovereignty” has been used very frequently to describe, above all, the primacy of a State. Indeed, the “digital sovereignty” is a complex concept, which entails cross-reference with several sectors and contexts. We believe that the concept of “digital sovereignty” can be two sides of a coin. On the one hand, we can use the expression “digital sovereignty” to describe the supremacy and full control of a State on the digital area. On the other hand, we can use the same expression “digital sovereignty” to refer to the power on the digital domain - as we will explain in our contribution - that anyone is potentially able to use in the private or public sector. Our contribution aims to demonstrate that where someone, public or private, can have the control on the digital domain, there is “digital sovereignty”.

1 Premise

Digital sovereignty has multidisciplinary connotations, and it can assume different meaning or describe several aspects depending on the contest in which we refer to it. We would demonstrate how it is possible to find other “digital sovereignty” scenarios different from the traditional description of the digital power that a State uses to protect its cyberspace borders.

We aim not to deepen here on the entire digital sovereignty topic but, starting from the definition of both the terms “digital” and “sovereignty”, we demonstrate how it is possible to realise a “digital sovereignty” also by a private organisation and not only by States.

Furthermore, we analyse what are “digital sovereignty” impacts on data protection and privacy, highlighting the consequent effects and possible approaches.

We think that there is undoubtedly existing “digital sovereignty” also in the private sector, which is expressed mainly through the adopted internal approach on the digital by some organisations that have relevant effects outside them indeed.

Indeed, starting from the analysis of “digital sovereignty” traditional concept, we would highlight how is preeminent nowadays, the digital aspect in any contest and how private organisations carry it out.

2 The meaning of “sovereignty”

The purpose of this contribution, as we said, is to investigate looking for an adequate definition, starting from the terms “digital” and “sovereignty” and analysing the meaning of both single words, till the expression “digital sovereignty” and so to have a proposal of complete definition.

The sovereignty topic is not recent and anyway related to the description of nature and characteristics of a State. Indeed, we can find many references to the sovereignty in the juridical literature about the study of a State. We bypass to deepen the traditional sovereignty concept because it is well-known.

3 The meaning of “digital”

For some time, there have been casual use of the word “digital” (in a sense opposed to “analogic”), with which commonly reference made to the possibility of representing information in the form of numbers.

The word “digital” derives from the Latin “digitus” (finger) because the ancient Romans used fingers to count. Nowadays, the term “digital” - among other definitions - refers in general to the number system. And specifically to the binary number system (0 or 1, off or on) on which a computer is based. With the spread of new technologies and, hence, with the use of techniques or algorithms

*Corresponding Author Nicola Fabiano - info@fabiano.law

based on numbers or binary system, became common to refer to the term “digital”.

In the most common and widespread language, it is customary to intercept the word “digital” when it is used to describe, with a non-technical approach, only the use of a device (smartphone, tablet, computer) and/or the Internet. In general, the digital term represents, in the collective imagination, a general sense of innovation in its most heterogeneous manifestations and applications.

Hence, what is digital?

We reach a conventional definition, even if not nearly deep, broad or basic enough [1], according to which digital is synonymous with a set of electronic computing techniques¹.

4 What do we mean by “Digital sovereignty”?

In light of the synthetically described panorama, it would emerge a definition of “digital sovereignty” as the power expressed in an innovative context.

In summary, therefore, we would affirm that with the expression “digital sovereignty” we intend to refer to the power attributed to the State in the sphere that concerns any activity classifiable as “digital”, that is connected to the use of the technologies or derived from them.

The outcome of a brief survey on international scientific production related to the topic, aimed at considering whether there is a theoretical convergence in the qualification of digital sovereignty, has produced thought-provoking results.

Among the most recent publications, Couture [2] claim that the expression “digital sovereignty” is characterized by five different perspectives (“*Cyberspace Sovereignty*”, “*Digital Sovereignty, Governments and States*”, “*Indigenous Digital Sovereignty*”, “*Social Movements and Digital Sovereignty*” and “*Personal Digital Sovereignty*”). Not wanting to go into detail, the constant reference to the digital term proposed by Peters [1] emerges, namely a generic and conventional definition based on the calculation.

It would seem that only in the nineties was the term digital super-gravity introduced [3, 4], used to envisage the internet as an opportunity to exercise independence from state control.

However, especially in recent times and more increasingly, we assist in the spreading of the use of the expression “digital sovereignty” to refer to the extraordinary power of a State, particularly in the digital domain.

This approach has drawn some attention limited to describing a phenomenon related to cyberspace, and specifically to the power of a State regarding its digital borders.

Digital sovereignty, moreover, has aroused the interest of Data Protection Authorities by aiming to investigate what kind of impact it would have on the protection of personal data. In fact, “digital sovereignty” was the topic of the event organised by the Italian Privacy Data Protection Authority (DPA) on the occasion of the Data Protection Day (Rome - 29/1/2019)².

It is well-known that with the term “sovereignty,” we generally refer to a power (of State, of people, of economy, etc.), original and independent from any other, and expressed by the manifestation of a will.

The different definitions of “digital sovereignty” have in common only the meaning to express primacy on something but not on the digital domain in a broad sense; the technology’s primary role, whose development or diffusion involves the manifestation of power anyway, might be “digital sovereignty.”

The primary reference is to the technological scenario which sees a current fierce competition between the USA, China, and Europe, hoping from this latter an effective intervention [5] to improve technological development and counter the supremacy of other countries.

We have the impression that with these positions, there is the aim at soliciting European development policies that are adequate to support confrontation with other states rather than aimed at the expression of power over a domain. In essence, increasing competitiveness in Europe, it implies an improvement both in the internal and in the global market: sovereignty, therefore, would express as supremacy on the market. Some people have doubts about whether this is a case of protectionism [6].

Furthermore, there is a widespread fear of losing control over technologies, both at the national and European or international level; there are different resources on the Internet. Moreover, Ursula von der Leyen, President of the European Commission, in the document entitled “**My agenda for Europe**”, states, “*It may be too late to replicate hyperscalers, but it is not too late to achieve technological sovereignty in some critical technology areas*”.

We can find a lot of news, already published, that express the same concern. Among several contributions, we highlight the article entitled “Digital sovereignty does not need EU champions” published on 14 November 2019, in the Financial Times where we read: “Building an ecosystem of services which protect user data would fill a neglected niche between the corporate wild west of the US and the state panopticon of China. Its appeal would not be restricted to Europe, either.” The positions highlighted, in summary, can be considered concurrent, since the common denominator is constituted by a widespread desire not to allow the big five tech companies - GAFAM - to process the personal data of European citizens. Fear, market, and technological supremacy converge towards the need for greater security for personal data.

4.1 Digital sovereignty and cyberspace

The reference to the power of the State over the digital domain has led to limiting the scope of this power to cyberspace, so that, for example, in Italy the recent Legislative Decree no. 105/2019, converted with modifications by the Law 18 November 2019, n. 133, on “*Urgent provisions on cybernetic national security perimeter and discipline of special powers in sectors of strategic importance*”, with article 1, paragraph 1, institutes the “cybernetic national security perimeter”. This recent legislative innovation, which undoubtedly deserves further study, has led to the affirmation of digital

¹The author says “That conventional sense in which digital is synonymous with discrete electronic computing techniques is not nearly deep, broad, or basic enough.”

²Here: <https://www.garanteprivacy.it/documents/10160/0/I+confini+del+digitale.+Nuovi+scenari+per+la+protezione+dei+dati+-+Atti+del+Convegno.pdf/89efdb61-c0c3-cc6f-8037-f0b283bad2b4?version=1.0>, last access May 2020

sovereignty understood as the power of the State over cyberspace.

However, recently, Roguski [7] affirmed that we are facing “layered sovereignty in cyberspace” approach. The author identified logical and social layers of cyberspace that “may be open to the exercise of State authority based on a criterion of proximity, i.e. whenever the State can establish a genuine link with the digital objects or online personae over which authority is to be asserted”.

In our opinion, in relation also to what we referred to, it is possible to identify further profiles of the exercise of digital sovereignty that need not necessarily be taken over or dominated by the State.

Moreover, Couture [2] states that the notion of sovereignty in the world of digital “is increasingly used to describe various forms of independence, control, and autonomy over digital infrastructures, technologies, and data”. not necessarily state and “meanings, and definitions of sovereignty can significantly differ from one group to another.”

These authors, having registered as a common denominator of technological sovereignty (of which the digital one is a part) autonomy, independence and control, conclude their research with the following question: “unsettling digital sovereignty?” [2].

This statement should make people think.

4.2 *Digital sovereignty: proposal for a definition*

In any case, in light of this, it is possible to affirm that “digital sovereignty” - in general terms - is not exclusively identified with the power exercised by the State.

In fact, “digital sovereignty” can be expressed in any model adopted by the private sector through which the power over one’s digital domain is exercised (in autonomy and with full control). This power may correspond to actions undertaken, to choices of particular work technologies, and hence, to the intention of preserving the digital heritage.

Thus, we can define “digital sovereignty” as the power over one’s digital domain exercised by a State’s or even a private organisation one. The key-point is related to the “power over ones’ digital domain”. In the case of a State, that power will consist of any activities aimed to protect its cyberspace. A private organisation may exercise that power carrying out any activity focused on the own digital domain (protect, develop, spread, propose, sell, etc.). Ultimately, we can have different “digital sovereignty” approaches depending on the (private or public) bodies. It is not a matter of subjective profile, but the main point is the power and how it is exercised.

We agree with Roguski [7] - although he refers to a different field - regarding a concept of layered “digital sovereignty”, depending on the specific area (public, private, etc.). It may be the likelihood of being in front of different kinds of “digital sovereignty”.

This approach, undoubtedly, also significantly affects the aspects related to the protection of personal data in the exercise of digital sovereignty.

³lay down by Article 25 of the EU Regulation n. 2016/679 - GDPR

⁴(16) This Regulation does not apply to issues of protection of fundamental rights and freedoms or the free flow of personal data related to activities which fall outside the scope of Union law, such as activities concerning national security. This Regulation does not apply to the processing of personal data by the Member States when carrying out activities in relation to the common foreign and security policy of the Union.

5 Digital Sovereignty and Inclusion

Digital sovereignty, besides, should also be characterized by an inclusion process of individuals fundamental rights in their domain, and thus avoid to be confined outside the protection of personal data.

So far, the phenomenon of digital sovereignty has been described as power over a domain.

Nevertheless, apart from the definition as described before, digital sovereignty is so versatile that it cannot be ruled out that it may also constitute the opportunity for one or more individuals to acquire digital autonomy and sovereignty. In this direction, to increase the knowledge of individuals, sovereignties could be enriched by awareness campaigns and in this way, obtain added value.

In fact, according to Nitot [8], awareness is an integral part of what means technological sovereignty. The (perfect) awareness of the current digital condition of the user will favourably affect his choices also regarding technologies and his personal data.

The data subject, that be aware, will be able to decide, by exercising his power of self-determination, even in the context of digital sovereignty. That decision is not by merely refusing the technologies, but by implementing appropriate choices aimed at avoiding the expropriation (in part or all) of his data personal, losing control over them.

According to Nitot [8], the “privacy by design” principle³ is precisely in this sense, as it is the tool to induce the user to increase his awareness to acquire the necessary tools to defend himself.

6 Digital Sovereignty: the Limits

The most crucial matter is if “digital sovereignty” can be a limit for privacy and data protection.

Digital sovereignty in its layers or different perspectives, qualifying as a power over the digital domain, cannot, however, constitute a limit, intended as pre-eminence on the individual and his rights, especially those related to the protection of personal data.

Indeed, the only limits are those provided for by the law and, concerning the digital sovereignty and specifically to the sovereignty over cyberspace, it is evident that the law on the protection of personal data does not apply in cases of national security.

Moreover, this is expressly envisaged by the recital nr. 16⁴, as well as Articles 2 and 23 of the EU Regulation 2016/679 (GDPR) [9].

7 Digital Sovereignty and the Rules on the Protection of Personal Data

The EU Regulation 2016/679, General Data Protection Regulation (GDPR) regulates the protection of natural persons and places the data subject, the person who has the power over their data, at the centre of the entire system, of the processing. Technological evolution does not mean abuse his (its or her) power on the individual but

ensuring a necessary balance between innovation and protection of humans. In the current globalised system that leads to the acquisition of an overall and non-analytical view, it is needed to refer to a general legal framework⁵ [10] of principles regarding privacy and protection of personal data that is widely valid. An instrument is already available today: the Convention 108+ and one can proceed from the principles outlined in it.

Personal data is an absolute value because it belongs to any natural person and it is inevitably and ontologically linked to it. Furthermore, personal data contribute to characterising the primacy of human dignity from which one cannot ignore and even clumsily try to disqualify by treating such information as if it were a secondary aspect of the person.

As already stated in another contribution [11], the protection of personal data and privacy are discussed solely and exclusively as there are ad hoc regulations; otherwise, there would be no problem of dwelling on the essence and relevance of personal data.

We cannot overlook, however, that personal data must be considered as an absolute value, also through an ethical approach and in any case, regardless of any norm [12].

The preventive criterion based on the principle according to which the personal data is an absolute value and requires awareness and ethics must be considered as a “prerequisite”: this constitutes the true and real starting point, not codified, which stands as an ultra-legal element [13].

The “level zero”, the right starting point, is also the ethical consideration of the high value attributable to personal data; without this assumption, it is difficult to have a suitable approach to the law. The “level one” will be that of legal rules.

8 Digital Sovereignty and Accountability: a Possible Challenge

The data controller must comply with the principle of “accountability” as required by art. 5, paragraph 2, of the GDPR. We should not attribute to this concept merely a legal meaning, because it is laid down by the GDPR, but also a programmatic nature. In fact, in qualifying the accountability and, therefore, considering the data controller as accountable, it should be necessary an assessment of the organisational measures to be implemented. In this way, the controller, respecting of every available instrument (good practice, guidelines, standards, etc.), minimises risks and protects the personal information belonging to the individual.

We should add to this not only the respect for ethics but also, equally, the development of a genuinely ethical conscience; if we apply ethics together with the juridical norms, we could connote the principles enunciated by the Convention 108+ and the GDPR in concrete.

It is no coincidence that the 41st International Conference of Data Protection and Privacy Commissioners, held in Tirana in October 2019, has adopted the “*International resolution on privacy as a fundamental human right and a precondition for the exercise*

⁵We proposed this approach in the contribution entitled “Privacy and Security in the Internet of Things” published by Cutter IT Journal8 (Vol. 26, No. 8 August 2013); see references.

⁶<http://globalprivacyassembly.org/wp-content/uploads/2019/10/Resolution-on-privacy-as-a-fundamental-human-right-2019-FINAL-EN.pdf> [retrieved: May, 2020].

of other fundamental rights”⁶ where we read the following, explicit statement: “**Reaffirm a strong commitment to privacy as well as to right and value in itself, given various international obligations**”.

In conclusion, the principle of accountability appears to be compatible with public or private digital sovereignty, where the primary reference value remains the natural person and human dignity. Digital sovereignty that is in contrast with respect for human dignity is not acceptable.

9 Conclusions

Digital sovereignty is a broad concept which can refer to the national security but also to the (digital) power expressed by someone (company or organisation or Public Body). Thus, we believe that nowadays, it is possible to discuss in terms of “sovereignty” related to anyone public or private ones. By this approach, it is evident that any case or situation deserves appropriate evaluation to verify whether we are dealing with a hypothesis of “digital sovereignty”.

References

- [1] B. Peters, Digital Keywords - A Vocabulary of Information Society and Culture, Princeton University Press Princeton and Oxford, 2016, 94, <http://culturedigitally.org/wp-content/uploads/2016/07/Peters-2016-Digital-Digital-Keywords-Peters-ed.pdf> [retrieved: May, 2020]
- [2] S. Couture - S. Toupin, What does the notion of “sovereignty” mean when referring to the digital?, 2019, New Media & Society, 21(10), pp. 23052322. doi: 10.1177/1461444819865984.
- [3] J. Perry Barlow, A Declaration of the Independence of Cyberspace, 1996, Electronic Frontier Foundation - <https://www.eff.org/fr/cyberspace-independence> [retrieved: May, 2020];
- [4] F. Turner, From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism, 2006, Chicago, University of Chicago Press.
- [5] ENISA - Consultation paper - EU ICT Industrial Policy: breaking the cycle of failure, July 2019, <https://www.enisa.europa.eu/publications/enisa-position-papers-and-opinions/eu-ict-industry-consultation-paper> [retrieved: May, 2020]
- [6] European ‘tech sovereignty’ or ‘tech protectionism’?, 30/10/2019, <http://www.project-disco.org/european-union/103019-european-tech-sovereignty-or-tech-protectionism/> [retrieved: May, 2020]
- [7] P. Roguski, Layered Sovereignty: Adjusting Traditional Notions of Sovereignty to a Digital Environment, [in:] T. Minrik, S. Alatalu, S. Biondi, M. Signoretti, I. Tolga, G. Visky (eds.), 11th International Conference on Cyber Conflict: Silent Battle, 2019, p. 347-359.
- [8] T. Nitot, Numrique: reprendre le contrle, Paris: Framasoft, 2016, 15. <https://framabook.org/docs/NRC/NumeriqueReprendreLeControleCC-Byimpress.pdf> [retrieved: May, 2020].
- [9] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN> [retrieved: May, 2020]
- [10] N. Fabiano, Privacy and Security in the Internet of Things, in Cutter IT Journal, Vol. 26, No. 8, August 2013.

- [11] N. Fabiano, Protezione dei dati personali e privacy: qual lo starting-point?, 2019, <https://www.nicfab.it/protezione-dei-dati-personali-privacy-qual-lo-starting-point/> [retrieved: May, 2020].
- [12] Charter of Fundamental Rights of the European Union, 2016 https://www.ecb.europa.eu/ecb/legal/pdf/oj_c_2016_202_full_en_txt.pdf
- [13] The Treaty on the functioning of the European Union (2016/C 202/01), 2016. https://www.ecb.europa.eu/ecb/legal/pdf/oj_c_2016_202_full_en_txt.pdf [retrieved: May, 2020]

Angular Orientation of Steering Wheel for Differential Drive

Rajesh Kannan Megalingam*, Deepak Nagalla, Ravi Kiran Pasumarthi, Vamsi Gontu, Phanindra Kumar Allada

Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, 690525, India

ARTICLE INFO

Article history:

Received: 15 January, 2020

Accepted: 29 April, 2020

Online: 29 May, 2020

Keywords:

Differential drive

Robot Operating System (ROS)

Wheeled Robot

Independent Control

Angular Motion

Gazebo

Universal Robot Description

Format (URDF)

ROS_Node

Trajectory

ABSTRACT

Several drive mechanisms for different robots are at hand in current days. Bicycle steering, Ackerman steering, differential drive are some principal drive mechanisms that are being deployed in robots these days. The differential drive needs the wheel rotations to be updated very frequently. But it is most commonly deployed on the robots with two wheels and casters, as discussed in this work. It also can be used to have an independent control for each of the wheels with independent control signals. This work deals with the modeling of the differential drive mechanism for a robot with two main drive wheels and two casters, which takes the angular orientation of the steering wheel as input. For simplicity, this work considers that casters have no influence on any aspect of the differential drive. An adaptive model, whose output depends on the real-time input from the gamers steering wheel and produces required output has been formulated. This work differs from the other differential drives in the context that the steering wheel and the robot wheels have no physical connection. The proposed model has been implemented in python and integrated with the Robot Operating System (ROS). The steering wheel, which is used to generate commands using, is attached to the controller at the control station and the ROS_Node thus created is used to read the values from the steering and generate commands for each of the left and right wheels. These commands are transferred to the controller on the mobile platform, which in turn generates control signals for actuators. This work also deals with the deployment of the proposed model using the Universal Robot Description Format (URDF) of the robot in the Gazebo simulation and evaluating it using the Nitho Drive Pro steering wheel. To prove that the differential drive mechanism can be used to control the robot efficiently in any type of terrain, a ROS python node is used to control and maneuver the robot.

1. Introduction

Mobile robots are nowadays being deployed in varied scenarios. Different designs are needed in dissimilar scenarios for effective deployment and efficient use. Each design will be based on a unique drive mechanism [1], which best suits the arena for the implementation of the robot. Some of the most commonly used steering mechanisms are Mecanum drive, Synchronous drive, XR4000, Differential drive and Ackerman steering. Each mechanism has its own not only pros but also cons, which can sometimes be advantageous during installation in the area of interest. Differential drive, which is of interest in this work, works

based on the difference in angular velocities of the two wheels of the robot as in work [2]. It has its own disadvantage in requiring the velocities to be updated very frequently. But this mechanism best suits the requirement when the wheels should be operated independently as stated in paper [3]. Especially when the wheels or their alignment is not identical, they should be controlled independently for stability and better control. The robots, though designed based on a drive mechanism, need the control signals to maneuver. These signals can be either autonomously generated or from an HMI device as discussed in paper [4]. Gamer's steering wheel is one such device which consumes low power and gives an interactive control experience to the user. This can be used to generate the required commands for the differential drive. Unlike implementing this mechanism in cars, most of the robots are wireless and the user cannot be on the robot in all instances. The

*Rajesh Kannan Megalingam, Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India, 09496120900 & rakeshkannan@ieee.org

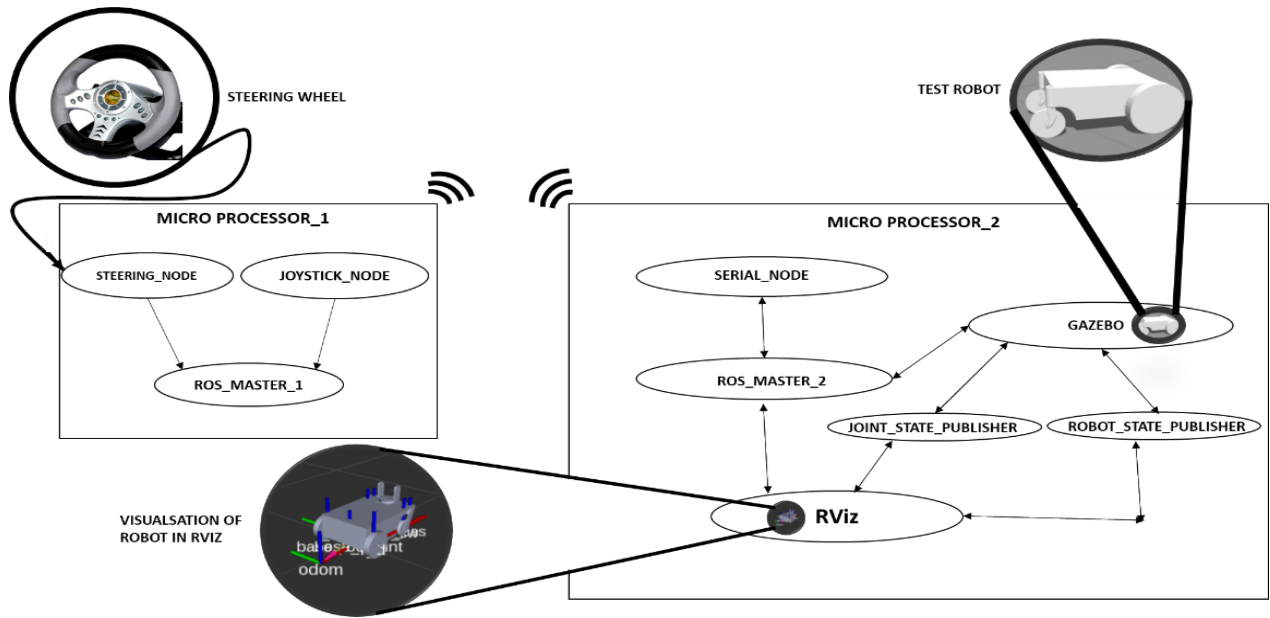


Figure 1(a): Internal ROS communication

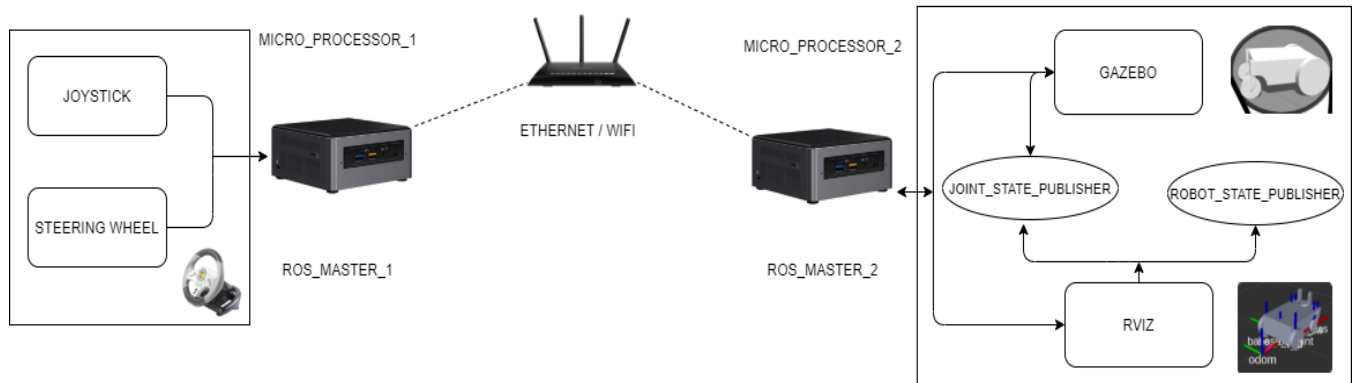


Figure 1(b): Architectural diagram

robot might be at one place and the HMI can be at another. ROS comes handy in such scenarios, especially when data should be exchanged between different nodes, which might be on different systems. ROS acts as a strong platform with a variety of tools and tremendous support for any robotics application. Gazebo, a simulation tool, which is interfaced with ROS as described in the work [5] and can be used to simulate mostly all kinds of robots with simulated arenas, which resemble the real-time environments. And the simulation is plain sailing, a robot or an environment can be built directly in the simulation tool or a URDF file generated from solid works can be used for simulation.

2. Motivation and Problem Statement

Considering the best example for this case is cars with power steering gives an appropriate turn without using much power on the steering. This theory of power steering helps to avoid accidents in vehicles by reducing the effort on the driver and with exact turnings. Inaccurate performance of the vehicle highlights neither the model of the robot nor the working. Many gamers might have used this steering wheel in different PlayStation games in order to get the exact sensation of the game. Comparing

this robot with the cars in the games like F1 race, blur and many. Considering the same case in the wheeled robots, implementing this idea may give an appropriate solution. To overcome this problem the following concepts have been proposed. The concept and the modeling [6] in this paper mainly concerned the differential drive with two castor wheels (freewheels) [7], and two main actuators [8]. This model helps the user to control from a remote area where there is no point in direct contact with the robot. The ROS-Gazebo [9] interface with this robot gives the simulation measures that help in achieving accurate results.

3. Related Works

Paper [10] mainly deals with the Varying speeds of differential drive motors and for no of stages. This approach of considering the varying speeds, which is a basic case in any differential drive approach has been adapted. This work deals with a better way of handling highly varying speeds in different scenarios. Kinematic algorithms and mobility control algorithms are explained in [11]. High reliability and precise working mechanisms with powerful dynamic equations can be taken from [12]. The system modeling of a four-wheeled differential drive

robot which used to achieve accuracy in Control by using sliding mode control is proposed in [13]. A hierarchical coordination control approach by defining various layers has been taken from [14]. Designing and simulating the robot, using the solid works, nonlinear dynamics to control the simulated robot is presented in [15]. The algorithm of the simulation control of robots and the behavior of the robot during simulation is presented in [16]. Working principle of controllers and the new form of feedback to controllers, an easy understanding of the control system is presented in [17]. Designing our own form of the controller and building the interface between the controller and the robot platform is presented in [18]. Wireless communication between the master controller and the slave robots, the algorithm for a wheeled robot is presented in [19]. The differential drive motor control technique has been taken from learning robotics using python as presented in [20]. Paper [21] explains the motion planning of the robot using the differential drive and trajectory smoothing using optimization techniques. Mathematical modeling and behavior of the robot are explained using the skid steering model, the tracking of the robot is also mentioned in paper [22].

4. Architectural Diagram

Figure 1(a) shows the ROS connections and communications of the entire system and how it controls the robot from a control station to the remote position of the robot. The left block in the diagram i.e. with the steering wheel explains the internal process happening at the control station. Similarly, the right block explains the internal connections and the processes happening within the robot. Figure 1(b) shows the architectural block diagram of this robot.

4.1. ROS_Master_1

The master controller of the MICRO_PROCESSOR_1 is entirely dependent on the ROS_MASTER_1 node which is independent of the entire package once the roscore is launched it automatically starts. Once the ROS_MASTER_1 is activated it acts as a host and starts connecting to different nodes (or vice versa) with different topics as a bridge and communicating among themselves. This node runs in the control station and communicates within the nodes i.e.; STEERING_NODE and JOYSTICK_NODE.

4.2. Steering_Node

The input to the ROS_MASTER_1 is given from either keyboard or steering wheel, which is connected to the processor through I/O peripherals. Keys like 'i', ',', 'j', 'l' in the keyboard are used for the movement of the robot in forward, backward, right, left respectively. Speed control can also be done using the keyboard. The steering wheel can be used for the same and the control resembles driving a car. These devices' steering wheel, joystick should be integrated and accessed with the ROS_MASTER_1 through STEERING_NODE and JOYSTICK_NODE.

4.3. Joystick_Node:

These nodes form a bridge between I/O devices and ROS. Each key of these devices has a unique ID in these nodes. So, when a key is pressed, the respective ID gets active and this, in

turn, can be assigned with a function to be performed. The values or commands generated in this function are published over a topic that can be accessed by the Gazebo node, where the robot is being simulated in different environments.

4.4. ROS_Master_2

Microcontroller_2 has another ROSmaster named ROS_MASTER_2 which is in the mobile robot that communicates with the nodes at the remote area i.e. within the robot. Since the entire communication of the system is done through wireless connectivity, both the ROS_MASTER's are controlled using the master-slave protocol.

4.5. Gazebo

The gazebo is open-source for the simulation of different environments. Gazebo interfaces with ROS and communication are done using the GAZEBO node. It can act as a publisher as well as a subscriber. GAZEBO node subscribes to the commands published by the STEERING_NODE and JOYSTICK_NODE to move the joints of the robot accordingly. It also publishes the transforms of the URDF node which in turn helps to move the robot. The Joint_states of the robot's movement is published (JOINT_STATE_PUBLISHER) over the Odom topic which is used by the differential drive plugin to correct the robot's path from the intended path. This data is also used by RViz to mimic the robot in Gazebo.

4.6. Joint_State_Publisher

This is used to read data from various origins. This has special features such as inputting the values through the graphical user interface, sending those values to the other joint and can set the default values even. It includes various parameters such as robot description, velocities, position, dependent joints...etc.

4.7. Robot_State_Publisher

ROBOT_STATE_PUBLISHER is used to publish the state of the robot to a tool where it can keep track of multiple coordinate frames per overtime. For a package, the position of the joint is taken as the initial values. This package subscribes to the joint states of the robot and publishes a three-dimensional representation of each and every link using the kinematic involved in the URDF model. Internally it has a kinematic model of the robot. General parameters used by this node are robot description, tf_prefix, publish frequency and it subscribes to the node joint states.

4.8. RViz

RViz is a platform for visualization and displaying different sensor values and information using ROS. It helps us to analyze different aspects of the robot-like transforms.

Transforms play a dominant role in avoiding collisions and in establishing coordination amongst the different frames and regulating the movement of all the joints. Not just the transforms of physical systems but also with the odometry frame can be calculated. This is not a simulation tool but a visualization interface where we can visualize both simulated robots and a real-time robot as shown in Figure 2.

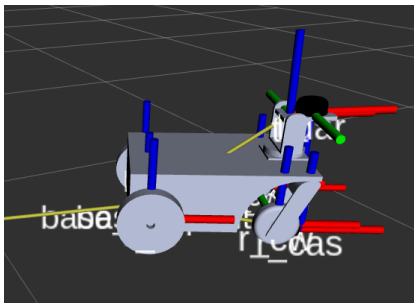


Figure 2: Robot model in RViz with TF's

4.9. Serial_Node

SERIAL_NODE introduces serial communication which is a bit to bit transmission of data. Rosserial helps this type of communication in different hardware parts of microcontrollers such as Arduino, teensy and also acts as an interface between the hardware devices and software tools.

5. Design and implementation

5.1. Implementation in a software platform

Nodes are a part of ROS in which they communicate with the other nodes in order to transfer the data or information. These nodes are generally connected by a bridge called a topic. The topic is a unidirectional data channel. The communication between these nodes and topics can be viewed through graphs(Rqt_graph) or frames.

The simulator consists of launch and configuration files. A virtual robot can be built in Gazebo which produces a xacro file that has all the links connected to the parent link. Similar to the Gazebo tool we have many other platforms (v-rep, webots, solid works) in which Solid works are most prominent for designing.

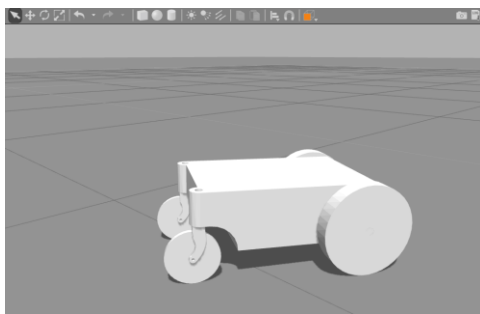


Figure 3: Gazebo Bot

Solid works to URDF exporter has been used to generate a URDF file from the solid works model, which includes inertia on the body, torque on the wheel links, parent links, velocity as shown in Figure 3. For the obtained URDF file adding the differential-drive [23] motor controls the velocity of the wheels independently. Forward kinematics equations are used in differential-drive to solve the problems obtained during the slippery surfaces controlling velocity.

$$V = r\omega \quad (1)$$

where V is the velocity of the wheel, r is the radius of the wheel, ω is the angular velocity of the wheel.

This above Equation (1) is the basic fundamental for this differential-drive. This URDF file is added into a ROS package and connected to the master and we can navigate the robot in any environment. Once the body is computationally ready with all the connections within the code, the controller launches all the required files using ROS commands. These launch files communicate through the nodes to publish or access the required data. The topic of Joint_states published by the Gazebo node and subscribed to the keyboard node using the command velocity(/cmd_vel) topic.

5.2. Steering wheel control

Nitho Drive Pro is a Steering wheel compatible with PC2, PC3 and PC. It helps us to control the Wheeled robot using teleoperation. In the steering wheel, the Joint_state values are published in the terminal, which is in turn subscribed by the robot, which helps in teleoperation. This robot simulation helps us to get the virtual feedback in Gazebo, so the operator knows what is happening just by looking in the simulation. By using the steering wheel, we can even adjust the speed and also the direction we want to turn using axis values. We can assign different values for buttons ranging from 0 to 1. We can change the values of the axes to turn based on our needs. The Steering Wheel system consists of two different controls for the forward and backward motion. One is pedal motion and the other is by using the buttons on the steering wheel.

This Steering wheel system can be implemented in an actual car driving for handicapped people who cannot apply much force to turn the wheel. Nitho Drive Pro does not require much force to rotate its wheel and it can be programmed in such a way that we can increase or decrease the speed by pressing the designated buttons. We can also find out which buttons are assigned for what number and its functions. By giving the command we can access the control system of the steering wheel and find out the mapping of the steering wheel and be able to change the calibration of the steering wheel. In the mapping option of the steering wheel, we can find out which button is assigned to which number and helps to change the controls accordingly. In calibration default values are assigned to the buttons and axes for movement, we can change the calibration values manually or by altering the values. We can even change the mode of controlling the axes. If the PS/MODE button is on it means that the steering wheel gives the continuous values of the axes, else if it is in the steady mode it means that the axes are in button state, it behaves just like a button and gives a single value. If it is off that means that the steering wheel is off.

5.3. Four-wheeled Differential drive

For the design of a Two-wheeled differential drive model initially, a Four-Wheeled drive model is discussed. This drive system is mainly composed of control and drive systems. The mechanical structure is the basis for the entire robot. It turns using all the four wheels, unlike normal vehicles. Each wheel has a separate motor mechanism to drive the steering of the wheel. Therefore, the motion of all the wheels of the robot are independent of each other.

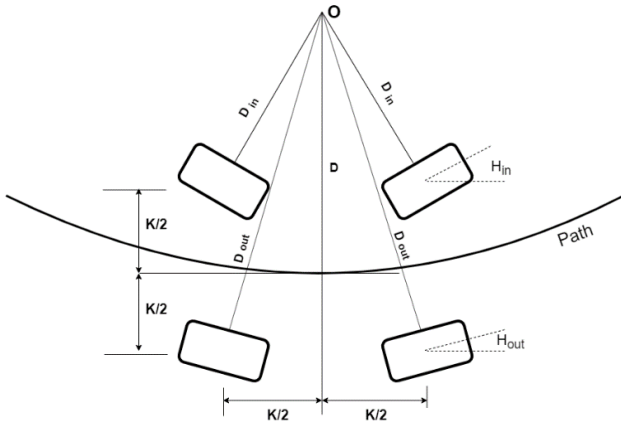


Figure 4: four-wheeled differential drive robot.

From geometric relation, from (Figure 4) it can be considered as follows

$$D = \frac{K(1+\tan(H_{in}))}{2} \tag{2}$$

$$D_{in} = ((D - \frac{K}{2})^2 + (\frac{K}{2})^2)^{\frac{1}{2}} \tag{3}$$

$$D_{out} = ((D + \frac{K}{2})^2 + (\frac{K}{2})^2)^{\frac{1}{2}} \tag{4}$$

$$H_{in} = \arctan\left(\frac{K}{2D+K}\right) \tag{5}$$

$$H_{out} = \arctan\left(\frac{K}{2D-K}\right) \tag{6}$$

Where D is the radius of the body in which it rotates, D_{in} is the radius of inner wheels in which it rotates, D_{out} is the radius of outer wheels in which it rotates, K is the separation between left and right wheels/separation between front and back wheel, H_{out} is the slip angle of outer wheels, H_{in} is the slip angle of inner wheels. The speed relationship is as follows:

$$s_{in} = \frac{D_{in} s_n}{D} \tag{7}$$

$$s_{out} = \frac{D_{out} s_n}{D} \tag{8}$$

Where s_{in} is the speed of the inner wheels, s_{out} is the speed of the outer wheels, s_n is the speed of the robot

$$\psi_{if} = \kappa_{if} \beta_{if} \tag{9}$$

$$\psi_{ir} = \kappa_{ir} \beta_{ir} \tag{10}$$

where $i=L$ refers left, $i=R$ refers right. ψ_{if} and ψ_{ir} denote the lateral force of front wheels and rear wheels respectively. κ_{if} and κ_{ir} are the lateral stiffness of the front wheels and rear wheels respectively. β_{if} is the slip angle for front wheels and β_{ir} is the slip angle for rear wheels.

5.4. Implementation of a two-wheeled robot

Implementing a robot with a simple 2 wheeled mechanism which has a common axis of rotation. And two caster wheels in the front so there will be no slip angle for the wheels. There will be only one inner wheel and one outer wheel. The moment of the body in circular turns are modeled by the varying speeds of the two wheels. from (Figure 6) the geometric relations of the two-wheeled robots are as follows.

$$D_{in} = D - \frac{K}{2} \tag{11}$$

$$D_{out} = D + \frac{K}{2} \tag{12}$$

$$s_{in} = \frac{D_{in} s_n}{D} \tag{13}$$

$$s_{out} = \frac{D_{out} s_n}{D} \tag{14}$$

$$s_n = \frac{s_{in} + s_{out}}{2} \tag{15}$$

where D is the radius of the body in which it rotates, D_{in} is the radius of inner wheels in which it rotates, D_{out} is the radius of outer wheels in which it rotates, K is the separation between left and right wheels/separation between front and back wheels, s_{in} is the speed of the inner wheels, s_{out} is the speed of the outer wheels, s_n is the speed of the robot.

Varying Speeds of front and back motors in terms of rotations per minute.

$$R_{in} = \frac{D_{in} \rho}{D} \tag{16}$$

$$R_{out} = \frac{D_{out} \rho}{D} \tag{17}$$

$$\rho = \frac{[s_{in} - s_{out}]}{k} \tag{18}$$

where R_{in} is the speed of the inner wheel (Rpm), ρ is the angular velocity of the robot.

The D (Radius of robot rotation) value will be defined by the steering control as shown in the following graph.



Figure 5: The varying values of D (Radius of robot rotation) according to the steering control (steering readings).

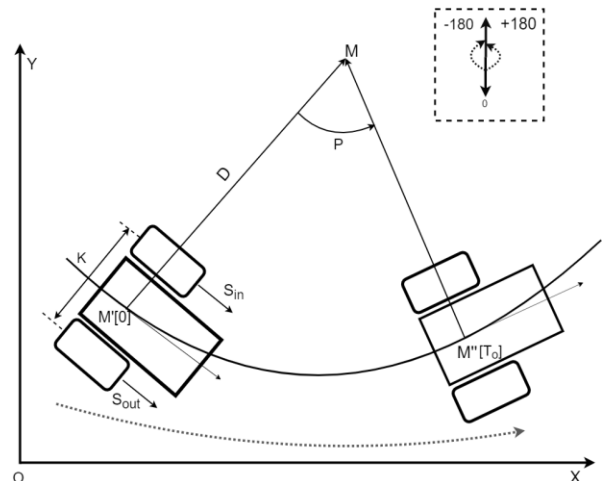


Figure 6: The transformation of the body in a 2D frame.

The robot started at M' at time $t=0$, the coordinates of the robot at time $t=T_o$ is M'' the transformation equations are:

$$M = [x - D\sin(M'_\theta), y + D\cos(M'_\theta)] \quad (19)$$

$$M''_x = (\cos(\rho T_o))(M'_x - M_x) - (\sin(\rho T_o))(M'_y - M_y) + M_x \quad (20)$$

$$M''_y = (\sin(\rho T_o))(M'_x - M_x) + (\cos(\rho T_o))(M'_y - M_y) + M_y \quad (21)$$

$$M''_\theta = M'_\theta + \rho T_o \quad (22)$$

where $[M'_x, M'_y, M'_\theta]$ are the coordinates of robot and angle of orientation at M' , $[M''_x, M''_y, M''_\theta]$ are the coordinates of robot and angle of orientation at M'' , $[M'_x, M'_y, 0]$ are the coordinates of the virtual center of rotation M

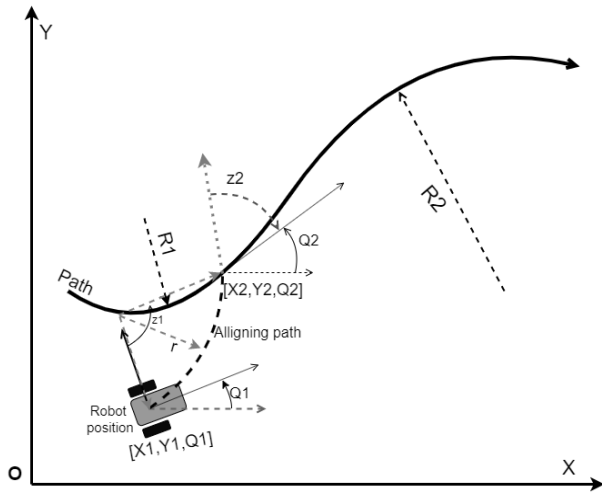


Figure 7: Tracking of the robot according to the steering control values.

The robot is at a random position $[X1, Y1, Q1]$, in order to follow the path on which it should follow it should first align with the path along the aligning path which is mentioned in Figure 7. The aligning path of radius 'r' which meets at a point on the path to be followed after $z1$ degrees of rotation. The robot now should rotate the $z2$ angle for making it tangential to the path and follow the path with the radius $R2$. This path can be tracked in the simulation when controlled using the steering wheel. The following equations give the position of the robot at a time T_o , obtained from the transformation equations.

$$M''_x = \left(\frac{1}{2}\right) \int_0^{T_o} [s_{out}(T_o) + s_{in}(T_o)] \cos(M'_\theta T_o) dt \quad (23)$$

$$M''_y = \left(\frac{1}{2}\right) \int_0^{T_o} [s_{out}(T_o) + s_{in}(T_o)] \sin(M'_\theta T_o) dt \quad (24)$$

$$\int M''_\theta = \left(\frac{1}{k}\right) \int_0^{T_o} [s_{out}(T_o) - s_{in}(T_o)] dt \quad (25)$$

6. Experiments

To get the real experience of controlling a robot, the Nitho Pro drive steering wheel has been used as shown in Figure 8, to control the robot which is compatible with any type of system. The robot is controlled in Gazebo to get a clear idea of real-life scenarios. Using the differential drive plugin, the angle of rotation of the robot is adjusted to the turning ratio of the steering wheel to get a better idea of the testing scenario and for easy control of

the robot as mentioned in [6]. For the controlling of the robot in multiple scenarios using Gazebo, a directory was created. This directory was used to control the robot using a keyboard. After successfully controlling using the keyboard a directory to control using the Steering wheel has been created.



Figure 8: Steering Wheel Control

Control of the robot using the Steering wheel in gazebo has been successfully implemented using ROS. Differential drive plugin has been added to the URDF file for the control of the robot. The robot has been tested on multiple terrains to test the control and manoeuvrability in different conditions. It was designed in such a way that the speed and angle of rotation of the robot can be monitored from (1), (2) and (3). If the power is kept constant the forces acting on the robot will differ due to the terrain conditions, stiffness, slipping and other forces acting in the body as shown in (9) and (10).

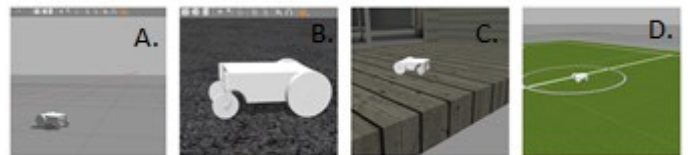


Figure 9: A. Plane surface, B. Asphalt Plane, C. Rough surface, D. Grass field

Multiple testing scenarios were considered for the implementation of the robot model. For the case of flat surfaces in Figure 9.A-D, is considered in many daily life scenarios. Figure A. is a Plane smooth surface where frictional value is zero which cannot be found in our daily life. When the robot is given even a small amount of speed the body moves freely as there is no opposing force but for cases such as Figure 9.B-D, which can be found in our daily a certain amount of force needs to be applied to overcome the frictional force and start moving. In the case of grass fields, the body moves freely due to less frictional force and covers a certain distance in less amount of time as shown in Table.

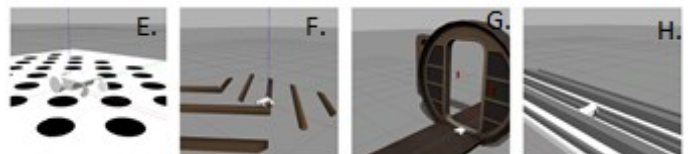


Figure 10: E. Crater surface, F. Uneven path, G. Doorway, H. Moat plane

In the case of scenarios such as rough surfaces in Figure 10.E-H., more speed is required to move over the obstacles but as the power is kept constant the robot moves at a slower speed compared to the plane surface and as a result, it takes more time. For the crater surface, maneuverability and angle of rotation are important to avoid the craters and advance smoothly. Whereas for the case of other rough surfaces in Figure 10.F-H., speed is required to avoid the obstacles or move over them.

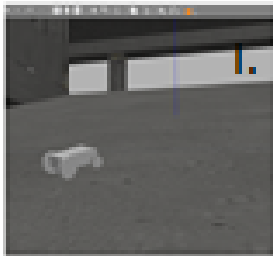


Figure 11: I. Slope

In the case of the plane slope in Figure 11.I, more speed should be provided than in the case of a plane surface as the robot needs to go to a higher place carrying its entire weight as the frictional force is acting in the opposite direction while moving up and in the direction of motion while coming down. But if the speed is given above a certain value the robot topples and if we rotate the body at high speed it topples.



Figure 12: J. Collapsed roof, K. Slippery Slope, L. Rough slope

In the case of the collapsed roof in Figure 12.J, and the rough slope in Figure 12.L, the body does not move easily as there is a frictional force between the surface and the robot compared to the case of plane slope and as a result, the body does not topple easily. But, for the case of driving terrain in Figure 12.K, the surface is damp and the robot slips and drifts while moving with more than the required speed but when given an appropriate speed the robot moves smoothly without sliding as the inclination of the plane is not constant.

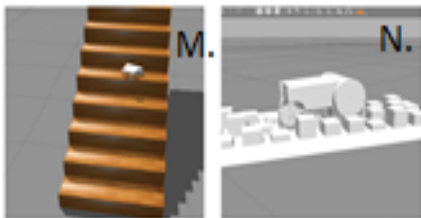


Figure 13: M. Steps, N. Uneven Path

Stairs are also a type of terrain but with both inclination and rough surface as shown in Figure 13.M, where control and mobility play an important part in the manoeuvrability of the robot. The uneven path in Figure 13.N is like rough paths in Figure 10.F-H, which requires more speed and mobility to avoid getting stuck in between the obstacles.

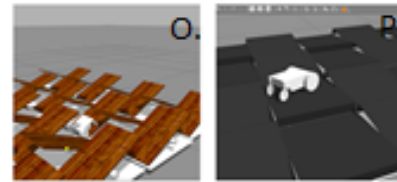


Figure 14: O. Teeter Ramps, P. Ramps

For testing the mobility and stability of the robot teeter ramps are chosen as shown in Figure 14.O The robot should move from one ramp to another while balancing the weight and move steadily over the ramps. Similarly, ramps were placed in Figure 15.P, to test the maneuverability of the robot to avoid obstacles and move smoothly over the ramps.

7. Results

The below Table provides the details of the frequencies at which the data is being published on different topics in different terrains and time taken for the robot to complete the first half of the distance as T1(sec) and the other half as T2(sec). From which we can infer that even in different scenarios the rate at which the data is being published is the same and the change in the behavior is mainly due to the speed and control, which affects the time taken by the robot to travel a certain distance.

Table 1: Frequencies of different Topics are published in multiple terrains

Test surface	Joy_topic	Cmd_vel	Joint_states	Tf	Time (T1)	Time (T2)
A	24.42	24.86	9.97	54.87	10	10
B	24.14	24.94	9.98	55.0	11.8	11.8
C	24.33	24.92	9.92	55.04	12.4	12.4
D	24.01	24.66	9.96	54.99	10.7	10.7
E	24.23	24.53	10.0	54.88	11.2	11.2
F	24.46	24.79	9.98	55.09	13.6	13.6
G	24.39	24.35	9.96	54.89	14	14
H	24.19	24.95	9.99	54.93	19.6	19.5
I	24.53	24.92	10.0	54.91	12.2	7.93
J	24.36	24.61	9.97	54.99	14.7	11.9
K	24.29	24.42	9.99	54.89	13.2	9.1
L	24.12	24.84	10.0	55.02	15.6	12.3
M	24.24	24.18	9.92	54.9	-	10.8
N	24.2	24.4	10.0	55.01	16.2	17.5
O	24.18	24.49	9.76	54.96	15.3	15.7
P	24.09	24.26	10.0	54.88	17.4	16.9

While rotating the body the amount of speed required and the angle of rotation differs for different surfaces. When the body is moving in a straight line, the center point of rotation lies at an infinite distance from the robot. Similarly, when the angle of rotation of the robot increases the center point of rotation decreases until it approaches the center of the robot which works on the same principle of a differential drive plugin. When the angle of rotation for the robot is maximum the robot rotates at its own position as a center. When the angle of rotation increases the center point of rotation moves away from the robot and it occupies more space to rotate. As the outer wheel is at a larger distance from the center point of rotation the outer wheel covers more distance compared to the inner rotating wheel. This shows that angular orientation also plays a major role in the movement of the robot. Under normal circumstances, it might take a while to turn or control the robot in different conditions but using differential drive mechanism we can maneuver the robot easily through any type of terrains.

8. Conclusion

The major focus of this work is to model the differential drive for a wheeled robot. The mathematical equations have been derived for a robot with two caster wheels and two main drive wheels, by assuming that the casters are ideal and will not affect the drive in any form. The proposed model has been implemented using python and the file has been successfully converted to a ROS node for integrating it with ROS. The model has been evaluated using the simulation tools. A URDF file of the robot model has been generated from SolidWorks (SW) using SW to URDF exporter plugin and this has been triumphantly launched in the Gazebo platform and integrated with the node generated previously. The hardware integration of the steering wheel with the Gazebo environment has also been achieved. All the works referred to in this work are either dealing with the modeling of different aspects of different drive mechanisms or the ROS aspect of the robots, controllers, control systems. All the previous works referred to in this work discuss the usage of joysticks or have steering on the robot. As stated earlier, the major difference of this work is that the steering wheel and the robot are different entities. They do not have any physical contact. But virtually it is similar to having the steering physically on the robot. The ROS node takes care of the exchange of control information between these two entities. The robot has been tested in different environments and different readings from ROS have been collected to evaluate the performance, which shows a steady performance of the published velocity values and stable performance of the robot.

8.1. Future work

Gazebo simulation and RViz visualization with two-wheeled differential drive vehicle integrating with mapping modules and some sensors make the robot autonomous and increase the performance and efficiency of the robot. To implement and evaluate these applications on the real-time heavy-duty robot in the future plan. Improvising the performance of the robot by optimizing the algorithms which can reduce the delay in the response. Evaluating the performance of the heavy-duty may lead to further improvement of the robot. Optimizing the robot with

high profile sensor components will get accurate and precise results.

Acknowledgment

Firstly, we would like to thank our institute Amrita Vishwa Vidyapeetham for providing the laboratory and components. The editors of HuTLabs and its staff have supported this work in each and every point of time and the people from the department of electronics and communication engineering have been constant supporters for the project at any point. Finally, we'd like to thank all the authors, maintainers and all the ROS users who worked and helped for the project with great positivity, composure, and determination.

References

- [1] R. K. Megalingam, D. Nagalla, R. K. Pasumarthi, V. Gontu, and P. K. Allada, "ROS Based, Simulation and Control of a Wheeled Robot using Gamer's Steering Wheel," 2018 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2018 DOI:10.1109/CCAA.2018.8777569
- [2] Wei Yu, Oscar Ylaja Chuy, Jr., Emmanuel G. Collins, Jr., Patrick Hollis, "Analysis and Experimental Verification for Dynamic Modeling of a Skid-Steered Wheeled Vehicle", IEEE Transactions on Robotics., vol. 26, pp. 340-353, Mar. 2010. DOI: 10.1109/TRO.2010.2042540
- [3] Li Wang, Xinhui Liu, Xin Wang, Beibei Fu, Ran Xu, "Research on differential performance of four-wheel independent steering of a hydraulic wheel-driving off-road vehicle", The Journal of Engineering., vol. 2019, pp. 68-73, May. 2019. DOI:10.1049/joe.2018.8956
- [4] Nobutake Hiraoka, Katsuhiko Inagaki, "A study of a new controller interface for omnidirectional robots", 10th Asian Control Conference (ASCC), 2015. DOI: 10.1109/ASCC.2015.7244592
- [5] Xudong Ma Fang Fang; Kun Qian; Can Liang, "Networked robot systems for indoor service enhanced via ROS middleware", 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2018. DOI: 10.1109/ICIEA.2018.8397832
- [6] Rajesh Kannan Megalingam; Sricharan Boddupalli; K. G. S. Apuroop, "Robotic arm control through mimicking of miniature robotic arm", 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), 2017. DOI:10.1109/ICACCS.2017.8014622
- [7] Jianfeng Liao, Zheng Chen, Bin Yao, "Performance-Oriented Coordinated Adaptive Robust Control for Four-Wheel Independently Driven Skid Steer Mobile Robot", IEEE Access., vol. 5, pp. 19048-19057, Sept. 2017. DOI: 10.1109/ACCESS.2017.2754647
- [8] Jianfeng Liao, Zheng Chen, Bin Yao, "Model-Based Coordinated Control of Four-Wheel Independently Driven Skid Steer Mobile Robot with Wheel-Ground Interaction and Wheel Dynamics.", IEEE Transactions on Industrial Informatics., vol. 15, pp. 1742-1752, Sept. 2018. DOI:10.1109/TII.2018.2869573
- [9] Masayoshi Wada, "A Joystick Steering Control System with Variable Sensitivity For Stable High Speed Driving", IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society, 2013. DOI: 10.1109/IECON.2013.6699781
- [10] F.U. Rehman, M.M. Ahmed, "Steering Control Algorithm for a class of wheeled mobile robots", IET Control Theory and Applications., vol. 1, pp. 915-924, Jul. 2007. DOI: 10.1049/iet-cta:20060189
- [11] M.A. Minor, B.W. Albiston, C.L. Schwensen, "Simplified Control of a two-axle compliant framed wheeled mobile robot", IEEE Transactions on Robotics, vol. 22, pp. 491-506, Jun. 2006. DOI: 10.1109/TRO.2006.875503
- [12] Jingang Yi, Hongpeng Wang, Junjie Zhang, Dezhen Song, Suhada Jayasuriya, Jingtai Liu, "Kinematic Modeling and Analysis of Skid-Steered Mobile Robots with Applications to Low-Cost Inertial-Measurement -Unit-Based Motion Estimation", IEEE Transactions and Robotics., vol 25, pp. 1087-1097, Jul. 2009. DOI: 10.1109/ROBOT.2009.5152342
- [13] Wei Shen, Zhichun Pan, Min Li, Hui Peng, "A Lateral control method for Wheeled-Footed Robot based on Sliding mode control and Steering Prediction", IEEE Access., vol. 6, pp. 58086-58095, Oct. 2018. DOI: 10.1109/ACCESS.2018.2873020

- [14] Junnian Wang, Zheng Luo, Yan Wang, Bin Yang, Francis Assadian, “Coordinated Control of Differential Drive Assist Steering and Vehicle Stability Control for Four-Wheel-Independent-Drive EV”, IEEE Transactions on Vehicular Technology., vol. 67, pp. 11453-11467, Oct. 2018. DOI 10.1109/TVT.2018.2872857
- [15] Yousuf, L.S.Email Author,” Experimental and simulation investigation of nonlinear dynamic behavior of a polydyne cam and roller follower mechanism”, Mechanical Systems and Signal Processing 116, pp. 293-309, 2018. DOI: 10.1016/j.ymsp.2018.06.028
- [16] Altarazi, S.A. Ammouri, M.M.,” Concurrent manual-order-picking warehouse design: a simulation-based design of experiments approach”, International Journal of Production Research pp. 1-19, 2017. DOI: 10.1080/00207543.2017.1421780
- [17] Ashley L. Guinan; Nathaniel A. Caswell; Frank A. Drews, “William R. ProvancherA video game controller with skin stretch haptic feedback,” 2013 IEEE International Conference on Consumer Electronics (ICCE), 2013. DOI: 10.1109/ICCE.2013.6486973
- [18] Rajesh Kannan Megalingam; Sarath Sreekanth; Govardhan; Chinta Ravi Teja; Akhil Raj, “Wireless gesture controlled wheelchair”, 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), 2017. DOI: 10.1109/ICACCS.2017.8014621
- [19] Rajesh Kannan Megalingam; Deepak Nagalla; Pasumarthi Ravi Kiran; Ravi Teja Geesala; Katta Nigam, “Swarm based autonomous landmine detecting robots”, 2017 International Conference on Inventive Computing and Informatics (ICICI), 2017. DOI: 10.1109/ICICI.2017.8365205
- [20] Lentin Joseph, “learning Robotics using Python, Packt Publishing, 2015.
- [21] W. Huang, X. Wu, Q. Zhang, N. Wu, and Z. Song, "Trajectory optimization of autonomous driving by differential dynamic programming," 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), Singapore, 2014 DOI: 10.1109/ICARCV.2014.7064582
- [22] Y. Wu et al., "Experimental kinematics modeling estimation for wheeled skid-steering mobile robots," 2013 IEEE International Conference on Robotics and Biomimetics (ROBIO), Shenzhen, 2013 DOI: 10.1109/ROBIO.2013.6739470
- [23] El-Dessoky, M.M, Alzahrani, E.O, Almohammadi, N.A, “Chaos control and function projective synchronization of novel chaotic dynamical system”, Journal of Computational Analysis and Applications 27(1), pp. 162-172, 2018. DOI: 10.1155/2011/452671

Controller Design Using Backstepping Algorithm for Fixed-Wing UAV with Thrust Vectoring System

Shogo Hirano*, Kenji Uchiyama, Kai Masuda

Nihon University, Department of Aerospace Engineering, Chiba 274-8501, Japan

ARTICLE INFO

Article history:

Received: 15 January, 2020

Accepted: 02 May, 2020

Online: 29 May, 2020

Keywords:

Backstepping control

Thrust vectoring system

Fixed-wing UAV

ABSTRACT

This paper describes the design method of a nonlinear flight controller for a fixed-wing UAV with a thrust vectoring system (TVS) using the backstepping method. The flight dynamics of the UAV exhibits strong nonlinear coupling behavior between its translational and rotational motion. The backstepping algorithm has been successfully applied to controller design for such a nonlinear system. However, the main idea of the method is to use some of the state variables as virtual control inputs that need ungeneratable forces by the UAV. To overcome this problem, we use the TVS that can generate thrust in an arbitrary direction. Numerical simulation is performed to confirm the effectiveness of the proposed control method for a fixed-wing UAV with the TVS.

1. Introduction

Unmanned aerial vehicles (UAVs) have been used for national projects, and multirotor are already introduced in various scenes. The fixed-wing UAVs have been used in a specific field [1]. The UAV has some advantages such as high cruising ability or large capacity of payloads. In particular, a small UAV with a wingspan of several meters is superior to a UAV with a wingspan of dozens of meters in maneuverability and maintainability. However, a small UAV has the problem that it is easily affected by the wind disturbance during a flight because of its small moment of inertia. As a result, it is difficult to achieve stable flight of the small UAV when a strong wind blows in a real mission. However, the small UAV cannot avoid an unexpected situation such as a strong wind during a flight because a high-performance device that measures the flight environment cannot be used due to its payload capacity.

One of the ways to solve the problem is to apply a nonlinear control method. The backstepping method [2]–[3], the nonlinear control methods, has also been studied in the fixed-wing UAV flying under the wind disturbance [4]–[7]. The advantage of this method is that it makes separated subsystems stabilized by putting the input successively. Therefore, in this study, we attempt to control translational and rotational motion individually to implement the flight at the constant attitude angle under wind disturbance. However, there is a problem that the conventional fixed-wing UAV has a structural restriction that the altitude and the pitch angle cannot be controlled at the same time.

Therefore, we introduce the thrust vectoring system (TVS) for the UAV to solve the problem. The TVS has been used to operate

the engine nozzle on aircraft that require maneuverability such as fighter aircraft but also a study on the UAV mounting a propeller-type TVS has been done at late years [8]–[13]. However, in those studies, rather than getting maneuverability as a fighter aircraft, the TVS is mainly used to control the attitude of a tail-sitter UAV during transition flight or the fixed-wing UAV with a broken dynamic surface. In this research, we apply the TVS to the generation of translational force to fly stably under wind disturbance. Finally, we confirm the effectiveness of the proposed method by the numerical simulation.

2. Nonlinear model of UAV

The nonlinear model of UAV can be described separately in translational motion and rotational motion. Control inputs are the thrust vector and the steering angles of an aileron, a rudder, and an elevator. The body coordinate and the inertial coordinate system of UAV are defined as shown in Fig.1.

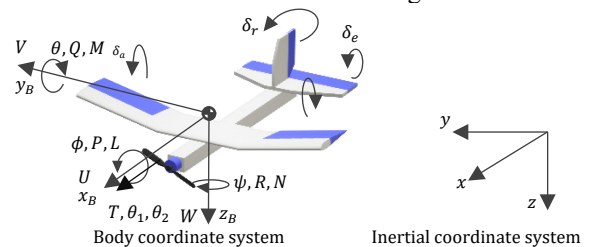


Figure 1: Definition of the coordinate system

2.1. Translational motion

The navigation equation of the UAV is given by

* Shogo Hirano, 7-24-1 Narashinodai, Funabashi, Chiba 274-8501, +81-47-469-5448, cssh19023@g.nihon-u.ac.jp

www.astesj.com

<https://dx.doi.org/10.25046/aj050337>

$$\dot{\mathbf{x}}_1 = \mathbf{g}_1(\mathbf{x}_3)\mathbf{x}_2 \quad (1)$$

where $\mathbf{x}_1 = [x \ y \ z]^T$, $\mathbf{x}_2 = [U \ V \ W]^T$, and $\mathbf{x}_3 = [\phi \ \theta \ \psi]^T$ are the position of the UAV in the inertial coordinate system, its velocity in the body coordinate system, and the Euler angle, respectively. $\mathbf{g}_1(\mathbf{x}_3)$ is defined as

$$\mathbf{g}_1(\mathbf{x}_3) = \begin{bmatrix} c(\psi)c(\theta) & c(\psi)s(\theta)s(\phi) - s(\psi)c(\phi) & c(\psi)s(\theta)c(\phi) + s(\psi)s(\phi) \\ s(\psi)c(\theta) & s(\psi)s(\theta)s(\phi) + c(\psi)c(\phi) & s(\psi)s(\theta)c(\phi) - c(\psi)s(\phi) \\ -s(\theta) & c(\theta)s(\phi) & c(\theta)c(\phi) \end{bmatrix}$$

Here, $c(\cdot)$ and $s(\cdot)$ are abbreviations for the function $\cos(\cdot)$ and the function $\sin(\cdot)$, respectively.

Next, the translational motion equation of UAV is expressed by

$$\dot{\mathbf{x}}_2 = \mathbf{f}_2(\mathbf{x}_4)\mathbf{x}_2 + \mathbf{g}_2(\mathbf{x}_3) + \mathbf{F}_a + \mathbf{F}_u \quad (2)$$

where $\mathbf{x}_4 = [P \ Q \ R]^T$ is the angular velocity of the body coordinate system, and nonlinear matrices $\mathbf{f}_2(\mathbf{x}_4)$, nonlinear vector $\mathbf{g}_2(\mathbf{x}_3)$ are defined as

$$\mathbf{f}_2(\mathbf{x}_4) = \begin{bmatrix} 0 & R & -Q \\ -R & 0 & P \\ Q & -P & 0 \end{bmatrix}, \mathbf{g}_2(\mathbf{x}_3) = g \begin{bmatrix} -s\theta \\ s\phi c\theta \\ c\phi c\theta \end{bmatrix}$$

$\mathbf{F}_a = [X_a \ Y_a \ Z_a]^T/m$ is the aerodynamic force vector, $\mathbf{F}_u = [X_u \ Y_u \ Z_u]^T/m$ is the control input vector, and m is the mass of the UAV. Each element of \mathbf{F}_a and \mathbf{F}_u is the force along each axis.

2.2. Rotational motion

The following equation expressed the kinematics equation of the UAV.

$$\dot{\mathbf{x}}_3 = \mathbf{g}_3(\mathbf{x}_3)\mathbf{x}_4 \quad (3)$$

Nonlinear matrix $\mathbf{g}_3(\mathbf{x}_3)$ is defined as

$$\mathbf{g}_3(\mathbf{x}_3) = \begin{bmatrix} 1 & s\phi t\theta & c\phi t\theta \\ 0 & c\phi & -s\phi \\ 0 & s\phi/c\theta & c\phi/c\theta \end{bmatrix}$$

where $t(\cdot)$ denotes the function $\tan(\cdot)$.

The rotational motion equation of the UAV is given as

$$\dot{\mathbf{x}}_4 = \mathbf{f}_4(\mathbf{x}_4)\mathbf{x}_4 + \mathbf{u}_a + \mathbf{u}_{st} \quad (4)$$

Nonlinear matrix $\mathbf{f}_4(\mathbf{x}_4)$ is defined as

$$\mathbf{f}_4(\mathbf{x}_4) = \mathbf{J}^{-1} \begin{bmatrix} 0 & R & -Q \\ -R & 0 & P \\ Q & -P & 0 \end{bmatrix} \mathbf{J}$$

\mathbf{J} is the moment of inertia. $\mathbf{u}_a = \mathbf{J}^{-1}[L_a \ M_a \ N_a]^T$ is the aerodynamic moment vector, and $\mathbf{u}_{st} = \mathbf{J}^{-1}[L_u \ M_u \ N_u]^T$ is the control input vector. Each element of \mathbf{u}_a and \mathbf{u}_{st} is the moment of the force around each axis.

3. Controller design

In the backstepping method, the control system is designed considering the stability of each subsystem. In this paper, a control

system is designed by putting (1)-(4) as subsystems S1 to S4, and its construction is as shown in Fig.2.

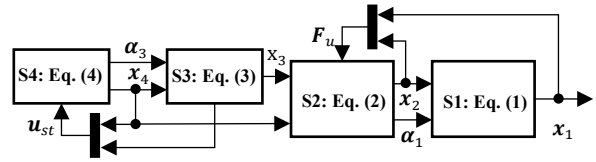


Figure 2: The structure of subsystem

In the above figure, control systems are constructed with two extensive subsystems that consist of the translational subsystems (S1, S2) and the rotational subsystems (S3, S4).

3.1. Translational control system

First, in order to give a target value \mathbf{x}_{1c} , the position, and velocity for UAV, we define error equation \mathbf{x}_{1e} as follows:

$$\mathbf{x}_{1e} = \mathbf{x}_1 - \mathbf{x}_{1c} \quad (5)$$

When giving a target value of the velocity, the target value of the position that is the integrated value of the velocity is given instead, assuming the target value of velocity as steady. Then, to introduce a candidate of control Lyapunov function for the subsystem S1, V_1 is expressed as

$$V_1(\mathbf{x}_{1e}) = \frac{1}{2} \mathbf{x}_{1e}^T \mathbf{x}_{1e} \quad (6)$$

The derivative of V_1 is

$$\begin{aligned} \dot{V}_1(\mathbf{x}_{1e}) &= \mathbf{x}_{1e}^T \dot{\mathbf{x}}_{1e} \\ &= \mathbf{x}_{1e}^T (\dot{\mathbf{x}}_1 - \dot{\mathbf{x}}_{1c}) \\ &= \mathbf{x}_{1e}^T (\mathbf{g}_1(\mathbf{x}_3)\mathbf{x}_2 - \dot{\mathbf{x}}_{1c}) \end{aligned} \quad (7)$$

Here, considering state variable \mathbf{x}_2 as virtual input of S1, and virtual input is defined as $\alpha_1(\mathbf{x}_1, \mathbf{x}_3, t)$. When $\dot{V}_1 < 0$, the Lyapunov stability theory is satisfied, and V_1 becomes a Lyapunov function. i.e., the subsystem S1 gets stable. The following equation can determine the virtual input $\alpha_1(\mathbf{x}_1, \mathbf{x}_3, t)$ that satisfies the condition for the system stability because $\mathbf{g}_1(\mathbf{x}_3)$ is a regular matrix.

$$\alpha_1(\mathbf{x}_{1e}, \mathbf{x}_3, t) = \mathbf{g}_1(\mathbf{x}_3)^{-1} \{ \dot{\mathbf{x}}_{1c} - \mathbf{K}_1 \mathbf{x}_{1e} \} \quad (8)$$

where \mathbf{K}_1 is a positive definite and diagonal matrix.

The error equation between state variable \mathbf{x}_2 and virtual input α_1 is as (9).

$$\mathbf{z}_1 = \mathbf{x}_2 - \alpha_1 \quad (9)$$

Then, we define V_2 to introduce a candidate as a control Lyapunov function for the whole of the translational systems.

$$V_2(\mathbf{x}_{1e}, \mathbf{z}_1) = V_1 + \frac{1}{2} \mathbf{z}_1^T \mathbf{z}_1 \quad (10)$$

The derivative of V_2 is

$$\begin{aligned} \dot{V}_2(\mathbf{x}_{1e}, \mathbf{z}_1) &= \mathbf{x}_{1e}^T \{ \mathbf{g}_1(\mathbf{x}_3)(\mathbf{z}_1 + \boldsymbol{\alpha}_1) - \dot{\mathbf{x}}_{1c} \} + \mathbf{z}_1^T \dot{\mathbf{z}}_1 \\ &= \mathbf{x}_{1e}^T (\mathbf{g}_1(\mathbf{x}_3)\boldsymbol{\alpha}_1 - \dot{\mathbf{x}}_{1c}) \\ &\quad + \mathbf{z}_1^T \{ \dot{\mathbf{z}}_1 + (\mathbf{g}_1(\mathbf{x}_3))^T \mathbf{x}_{1e} \} \\ &= \mathbf{x}_{1e}^T (\mathbf{g}_1(\mathbf{x}_3)\boldsymbol{\alpha}_1 - \dot{\mathbf{x}}_{1c}) \\ &\quad + \mathbf{z}_1^T \{ \mathbf{f}_2(\mathbf{x}_4)\mathbf{x}_2 + \mathbf{g}_2(\mathbf{x}_3) + \mathbf{F}_a + \mathbf{F}_u \\ &\quad - \dot{\boldsymbol{\alpha}}_1 + (\mathbf{g}_1(\mathbf{x}_3))^T \mathbf{x}_{1e} \} \end{aligned} \quad (11)$$

The first term on the right side of (11) is negative-definite according to (8). Therefore, putting the input \mathbf{F}_{uc} as written in (12), the second line of the summation of (11) becomes negative-definite that $\dot{V}_2 < 0$ is satisfied, and V_2 is determined as Lyapunov function.

$$\mathbf{F}_{uc} = -\mathbf{K}_2\mathbf{z}_1 - \mathbf{f}_2(\mathbf{x}_4)\mathbf{x}_2 - \mathbf{g}_2(\mathbf{x}_3) - \mathbf{F}_a + \dot{\boldsymbol{\alpha}}_1 - (\mathbf{g}_1(\mathbf{x}_3))^T \mathbf{x}_{1e} \quad (12)$$

where \mathbf{K}_2 is a positive definite and diagonal matrix.

3.2. Rotational control system

Similar to the previous section, the error equation \mathbf{x}_{3e} is defined by using the target value \mathbf{x}_{3c} of the Euler angle.

$$\mathbf{x}_{3e} = \mathbf{x}_3 - \mathbf{x}_{3c} \quad (13)$$

The candidate of control Lyapunov function, V_3 for S3 is defined as

$$V_3(\mathbf{x}_{3e}) = \frac{1}{2} \mathbf{x}_{3e}^T \mathbf{x}_{3e} \quad (14)$$

The time derivative of V_3 is

$$\begin{aligned} \dot{V}_3(\mathbf{x}_{3e}) &= \mathbf{x}_{1e}^T \dot{\mathbf{x}}_{1e} = \mathbf{x}_{3e}^T (\dot{\mathbf{x}}_3 - \dot{\mathbf{x}}_{3c}) \\ &= \mathbf{x}_{3e}^T (\mathbf{g}_3(\mathbf{x}_3)\mathbf{x}_4 - \dot{\mathbf{x}}_{3c}) \end{aligned} \quad (15)$$

In (15), considering the state variable \mathbf{x}_4 as a virtual input of S3, we replace it by a virtual input $\boldsymbol{\alpha}_3(\mathbf{x}_3, \mathbf{x}_4, t)$. The following equation gives a virtual input that satisfies the Lyapunov stability theory.

$$\boldsymbol{\alpha}_3(\mathbf{x}_{3e}, \mathbf{x}_4, t) = \mathbf{g}_3(\mathbf{x}_3)^{-1} (\dot{\mathbf{x}}_{3c} - \mathbf{K}_3\mathbf{x}_{3e}) \quad (16)$$

where \mathbf{K}_3 is a positive definite and diagonal matrix.

The error equation between state variable \mathbf{x}_4 and virtual input $\boldsymbol{\alpha}_3$ is defined as

$$\mathbf{z}_3 = \mathbf{x}_4 - \boldsymbol{\alpha}_3 \quad (17)$$

Then, the candidate of control Lyapunov function for the rotational motion is defined as

$$V_4(\mathbf{x}_{3e}, \mathbf{z}_3) = V_3 + \frac{1}{2} \mathbf{z}_3^T \mathbf{z}_3 \quad (18)$$

The time derivative of V_4 is obtained as follows:

$$\begin{aligned} \dot{V}_4(\mathbf{x}_{3e}, \mathbf{z}_3) &= \mathbf{x}_{3e}^T \{ \mathbf{g}_3(\mathbf{x}_3)(\mathbf{z}_3 + \boldsymbol{\alpha}_3) - \dot{\mathbf{x}}_{3c} \} + \mathbf{z}_3^T \dot{\mathbf{z}}_3 \\ &= \mathbf{x}_{3e}^T (\mathbf{g}_3(\mathbf{x}_3)\boldsymbol{\alpha}_3 - \dot{\mathbf{x}}_{3c}) \\ &\quad + \mathbf{z}_3^T \{ \dot{\mathbf{z}}_3 + (\mathbf{g}_3(\mathbf{x}_3))^T \mathbf{x}_{3e} \} \\ &= \mathbf{x}_{3e}^T (\mathbf{g}_3(\mathbf{x}_3)\boldsymbol{\alpha}_3 - \dot{\mathbf{x}}_{3c}) \\ &\quad + \mathbf{z}_3^T \{ \mathbf{f}_4(\mathbf{x}_4)\mathbf{x}_4 + \mathbf{u}_a + \mathbf{u}_{st} - \dot{\boldsymbol{\alpha}}_3 \\ &\quad + (\mathbf{g}_3(\mathbf{x}_3))^T \mathbf{x}_{3e} \} \end{aligned} \quad (19)$$

By calculating the input command value \mathbf{u}_{stc} to be as the following equation, the second and third lines of the last summation of (19) becomes negative-definite, and it makes the rotational system stable.

$$\mathbf{u}_{stc} = -\mathbf{K}_4\mathbf{z}_3 - \mathbf{f}_4(\mathbf{x}_4)\mathbf{x}_4 - \mathbf{u}_a + \dot{\boldsymbol{\alpha}}_3 - (\mathbf{g}_3(\mathbf{x}_3))^T \mathbf{x}_{3e} \quad (20)$$

where \mathbf{K}_4 is a positive definite and diagonal matrix.

Fig.3 Shows the block diagram that includes the controller obtained in sections 3.1 and 3.2.

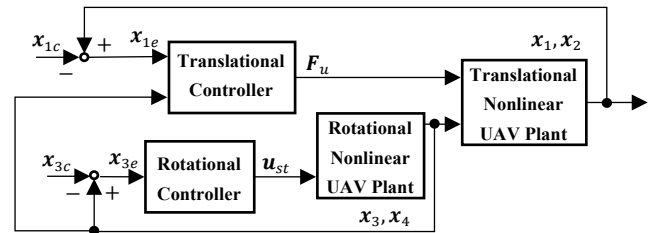


Figure 3: Block diagram of the proposed controller

3.3. Stability analysis

In this section, we analyze the stability of the entire control system that we designed. The candidate of control Lyapunov function of the entire control system is defined.

$$\begin{aligned} V &= V_1 + V_2 + V_3 + V_4 \\ &= 2V_1 + \frac{1}{2} \mathbf{z}_1^T \mathbf{z}_1 + 2V_3 + \frac{1}{2} \mathbf{z}_3^T \mathbf{z}_3 \end{aligned} \quad (21)$$

The time derivative of V is obtained as

$$\begin{aligned} \dot{V} &= 2\dot{V}_1 + \mathbf{z}_1^T \dot{\mathbf{z}}_1 + 2\dot{V}_3 + \mathbf{z}_3^T \dot{\mathbf{z}}_3 \\ &= -2\mathbf{K}_1\mathbf{x}_{1e}^T \mathbf{x}_{1e} - \mathbf{K}_2\mathbf{z}_1^T \mathbf{z}_1 - 2\mathbf{K}_3\mathbf{x}_{3e}^T \mathbf{x}_{3e} - \mathbf{K}_4\mathbf{z}_3^T \mathbf{z}_3 \end{aligned} \quad (22)$$

It is clear that the right side in (22) is negative definite because of the coefficient matrix \mathbf{K}_i is positive, i.e., the entire system satisfies the Lyapunov stability theory.

3.4. Realization of control input

In section 3.1, we calculated the input command for translational motion. The absolute value of thrust and the target

value of the deflection angle are obtained from those commands. The thrust coordinate system is defined as shown in Fig.4.

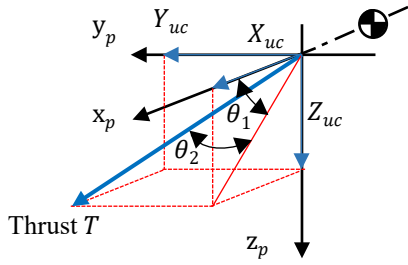


Figure 4: Definition of thrust coordinate system

The absolute value of thrust is obtained by (23).

$$T = \sqrt{X_{uc}^2 + Y_{uc}^2 + Z_{uc}^2} \quad (23)$$

X_{uc}, Y_{uc}, Z_{uc} are the commands of the input that is obtained to satisfy the ideal input of F_u . The following inequalities denote the range of angles θ_1 and θ_2 .

$$\begin{cases} -\pi \leq \theta_1 \leq \pi \\ -\pi/2 \leq \theta_2 \leq \pi/2 \end{cases}$$

where θ_1 and θ_2 are defined as

$$\begin{aligned} \tan \theta_1 &= \frac{Z_{uc}}{X_{uc}} \\ \theta_1 &= \tan^{-1} \left(\frac{Z_{uc}}{X_{uc}} \right) \\ \sin \theta_2 &= \frac{Y_{uc}}{\sqrt{X_{uc}^2 + Z_{uc}^2}} \\ \theta_2 &= \tan^{-1} \left(\frac{Y_{uc}}{\sqrt{X_{uc}^2 + Z_{uc}^2}} \right) \end{aligned} \quad (24)$$

$$(25)$$

4. Numerical simulation

Numerical simulations are conducted to confirm the effect of the TVS on control performance in two cases: the level flight (Case 1) and the improvement of maneuverability by the TVS under the follow-up flight to the target altitude (Case 2) by numerical simulations. The conventional UAV without the TVS, its thrust command is obtained by using only X_{uc} from the translational controller. Moreover, regardless of the presence of the TVS, we introduce the first-order lag filter as written in (26) and (27) to the thrust, the TVS, and dynamic surface. T_m and T_s are the time constants.

$$\frac{1}{T_m s + 1} \quad (26)$$

$$\frac{1}{T_s s + 1} \quad (27)$$

Conditions of numerical simulation are shown in Tables 1 to 4.

Table 1: Specification of UAV

Full length L , m	0.823
Total height H , m	0.322
Wingspan b , m	1.40
MAC \bar{c}_a , m	0.200
Mass m , kg	0.220

Table 2: Initial condition and target value

Position \mathbf{x}_{1i} , m	$[0 \ 0 \ -10]^T$
Velocity \mathbf{x}_{2i} , m/s	$[5 \ 0 \ 0]^T$
Euler angle \mathbf{x}_{3i} , rad	$[0 \ 0 \ 0]^T$
Angular rate \mathbf{x}_{4i} , rad/s	$[0 \ 0 \ 0]^T$
Target position in Case 1 \mathbf{x}_{1c} , m	$[0 \ 0 \ -10]^T$
Target position in Case 2 \mathbf{x}_{1c} , m	$[0 \ 0 \ -12]^T$
Target velocity \mathbf{x}_{2c} , m/s	$[6 \ 0 \ 0]^T$

Table 3: Design control parameters

Translational controller gain \mathbf{K}_1	diag[1 1 1]
Translational controller gain \mathbf{K}_2	diag[0.5 0.5 0.5]
Rotational controller gain \mathbf{K}_3	diag[8 8 8]
Rotational controller gain \mathbf{K}_4	diag[10 10 10]

Table 4: Constraints on control input

Aileron angle δ_a , deg	± 45
Elevator angle δ_e , deg	± 45
Rudder angle δ_r , deg	± 30
Deflection angle θ_1 , deg	± 135
Deflection angle θ_2 , deg	± 90
Thrust T , N	0 or more
Time constant on thruster T_m , s	0.2
Time constant on dynamic surface and thrust vectoring system T_s , s	0.0076

In addition, the Dryden model [14] is used as a wind disturbance on numerical simulations. The wind disturbance profile is shown in Fig.5.

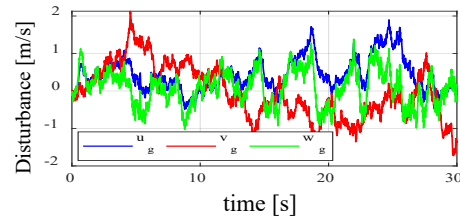


Figure 5: Wind disturbance

Fig.6 to Fig.9 show the numerical results in Case 1. It is clear that the UAV without the TVS cannot keep the altitude as shown in Figs.6(a), 6(b), and 6(d).

Furthermore, it should be noted that the TVS is also effective to control the UAV along y axis as shown in Fig.6(c). One of the main reasons for these flight is that the TVS can generate thrust in arbitrary direction as shown in Fig.8(a) under mechanical constraint. In fact, the inputs for translational motion in Fig.7(a) is constant because the UAV without the TVS cannot generate the thrust for the negative command as shown in Fig.7(b). Fig.7(c) and Fig.8(c) show the attitude angles of the UAV without and with

the TVS, respectively. It can be seen from these figures that the TVS suppressed the high-frequency vibration of pitch angle of the UAV. Fig.8(d) shows that the thrust change in Fig.8(a) leads to the sudden change of the elevator deflection to keep the altitude.

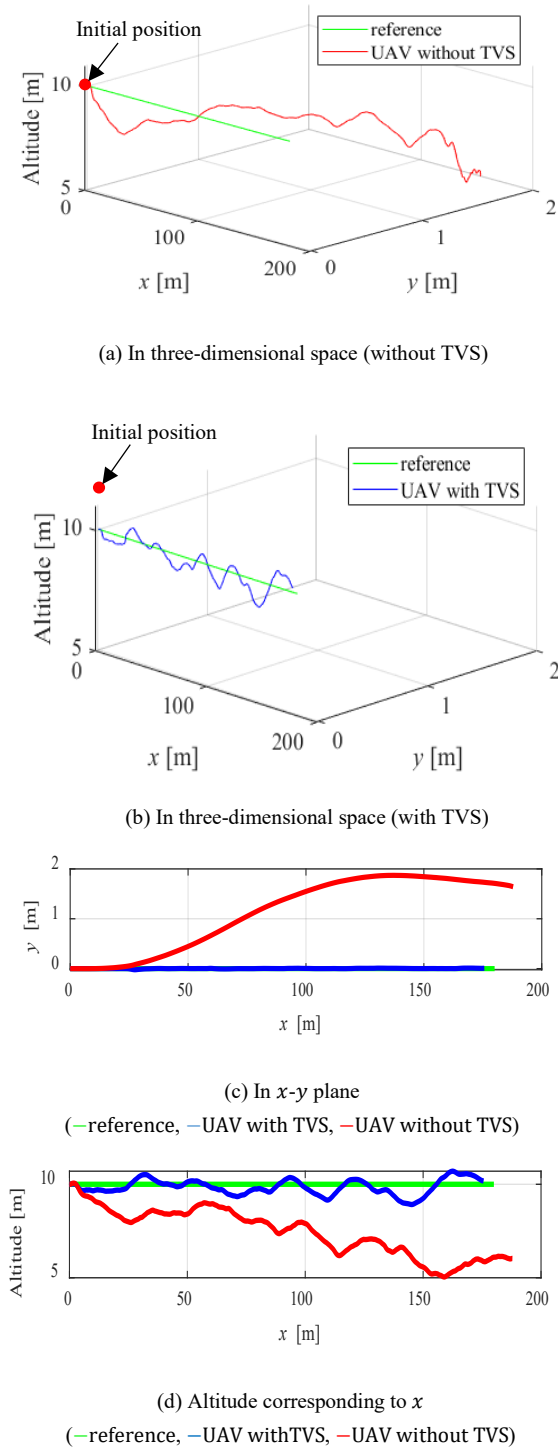


Figure 6: Trajectories of UAV in Case1

Fig.9 shows the phase portrait during the control of the UAV in Case 1. Using the proposed controller, the variables in the phase space follow those references, unlike the controller without the TVS.

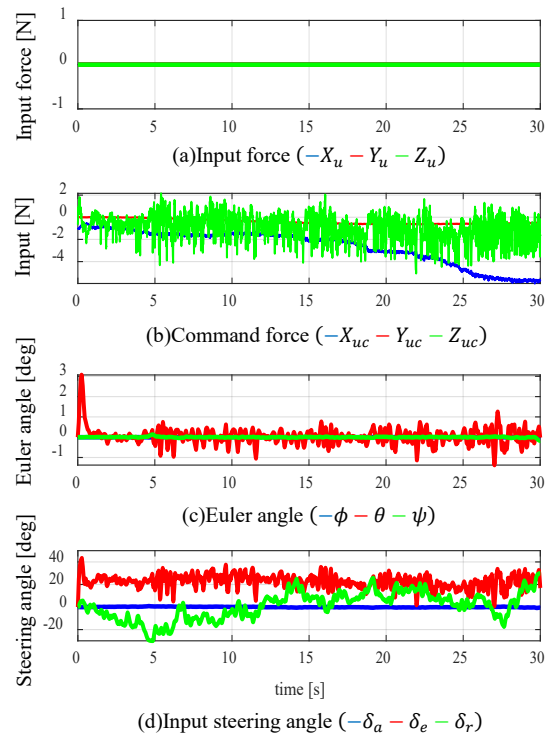


Figure 7: Results in case of UAV without TVS (Case 1)

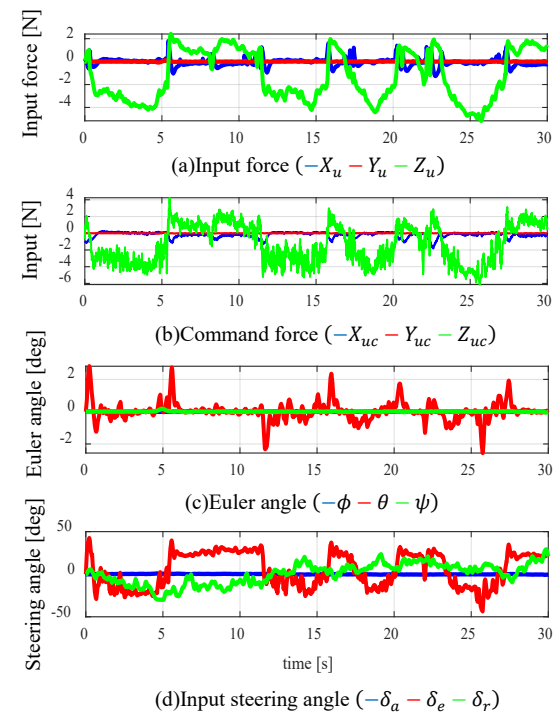


Figure 8: Results in case of UAV with TVS (Case 1)

Fig.10 and Fig.11 show the numerical results in Case 2. The altitude of the UAV is controlled from the initial value 10 m to the final value 12 m as shown in Fig. 10(a). In this paper, the final value was decided by 20% higher than the initial value. However, there is no limitation to altitude in the proposed method, but it is necessary to be careful so that the altitude changes slightly. The attitude of the UAV with the TVS in Fig.10(c) is also controlled by control inputs in Fig.10(b) and Fig.10(d) similar to the Case 1.

Moreover, it can be seen in Fig.11 that the UAV follows the reference in the phase space.

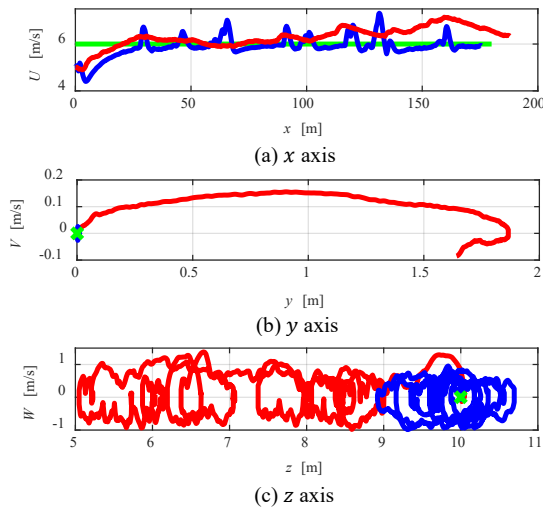
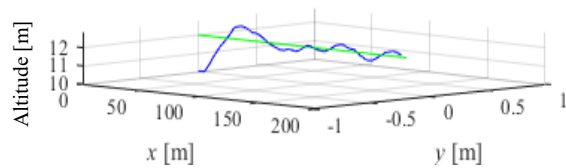
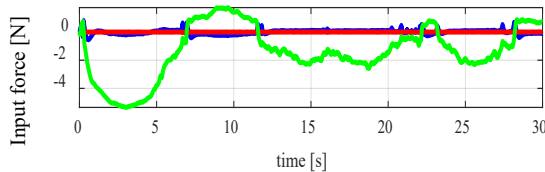


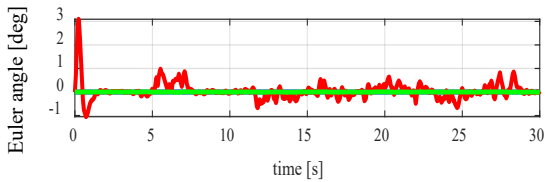
Figure 9: Phase portrait in Case1
(— reference, — UAV with TVS, — UAV without TVS)



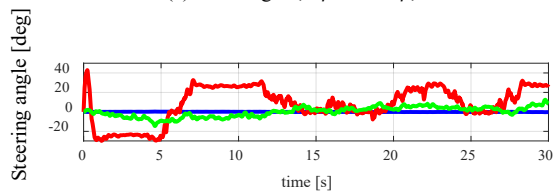
(a) Trajectory of UAV



(b) Input force ($-X_u - Y_u - Z_u$)



(c) Euler angle ($-\phi - \theta - \psi$)



(d) Input steering angle ($-\delta_a - \delta_e - \delta_r$)

Figure 10: Results of the UAV with TVS in Case2

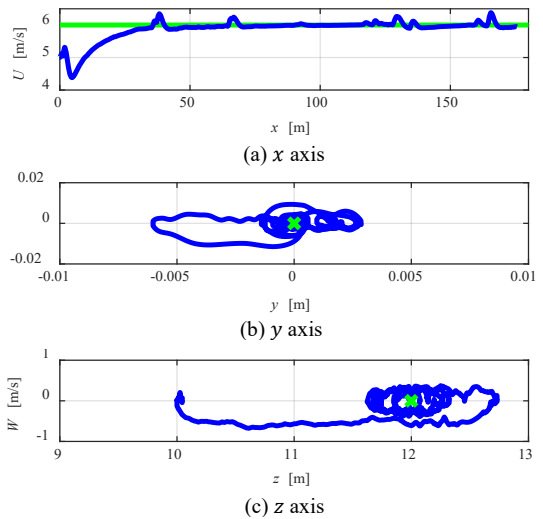


Figure 11: Phase portrait in Case2
(— reference, — UAV with TVS.)

5. Conclusion

This paper proposed the nonlinear flight control system under disturbance for the UAV with a thrust vectoring system (TVS) using the backstepping method. The results of numerical simulation show that the TVS enables the stable flight even under disturbances for the two cases proposed.

The accurate estimation of the aerodynamic forces would afford the key to succeed the stable flight of the UAV with the TVS. In future work, we will verify the effectiveness of the proposed method experimentally.

Reference

- [1] I V Kovalev, A A Voroshilova, and M V Karaseva, "Analysis of the current situation and development trend of the international cargo UAVs market," *J. Phys. Conf. Ser.*, 1399(5), 055095, 2019. <https://doi.org/10.1088/1742-6596/1399/5/055095>
- [2] K. Shimizu, *Feedback Control Theory*. Corona Publishing Co., 2013.
- [3] M. Krstic, I. Kanellakopoulos, and P. V. Kokotovic, *Nonlinear and Adaptive Control Design*. Wiley-Interscience, 1995.
- [4] T. Espinoza, A. Dzul, R. Lozano, and P. Parada, "Backstepping — Sliding mode controllers applied to a fixed-wing UAV," in *2013 International Conference on Unmanned Aircraft Systems (ICUAS)*, Atlanta, GA, USA, 2013. <https://doi.org/10.1109/ICUAS.2013.6564678>
- [5] A. Brezoescu, R. Lozano, and P. Castillo, "Bank to turn approach for airplane translational motion in unknown wind," in *2013 International Conference on Unmanned Aircraft Systems (ICUAS)*, Atlanta, GA, USA, 2013. <https://doi.org/10.1109/ICUAS.2013.6564790>
- [6] K. Wu, B. Fan, and X. Zhang, "Trajectory following control of UAVs with wind disturbance," in *2017 36th Chinese Control Conference (CCC)*, Dalian, China, 2017. <https://doi.org/10.23919/ChiCC.2017.8028144>
- [7] A. T. Espinoza Fraire, R. P. Parada Morado, A. E. Dzul López, and R. Lozano Leal, "Design and implementation of fixed-wing MAV controllers," in *2015 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED-UAS)*, Cancun, Mexico, 2015. <https://doi.org/10.1109/RED-UAS.2015.7441004>
- [8] J. W. Roberts, R. Cory, and R. Tedrake, "On the controllability of fixed-wing perching," in *2009 American Control Conference*, St. Louis, MO, USA, 2009. <https://doi.org/10.1109/ACC.2009.5160526>
- [9] K. Z. Y. Ang et al., "Development of an unmanned tail-sitter with reconfigurable wings: U-Lion," in *11th IEEE International Conference on Control Automation (ICCA)*, Taichung, Taiwan, 2014. <https://doi.org/10.1109/ICCA.2014.6871015>

- [10] C. Liang and C. Cai, "Modeling of a rotor/fixed-wing hybrid unmanned aerial vehicle," in *2017 36th Chinese Control Conference (CCC)*, Dalian, China, 2017. <https://doi.org/10.23919/ChiCC.2017.8029181>
- [11] W. Wang, J. Zhu, M. Kuang, and X. Zhu, "Adaptive Attitude Control for a Tail-Sitter UAV with Single Thrust-Vectored Propeller," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, QLD, Australia, 2018. <https://doi.org/10.1109/ICRA.2018.8463158>
- [12] M. Kuang, J. Zhu, W. Wang, and Y. Tang, "Flight controller design and demonstration of a thrust-vectoring tailsitter," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, Singapore, 2017. <https://doi.org/10.1109/ICRA.2017.7989605>
- [13] H. Kikkawa and K. Uchiyama, "Attitude control of a fixed-wing UAV using thrust vectoring system," in *2017 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED-UAS)*, Linkoping, Sweden, 2017. <https://doi.org/10.1109/RED-UAS.2017.8101677>
- [14] *Flying Qualities of Piloted Aircraft*. Department of Defense Handbook. MIL-HDBK-1797. Washington, DC: U.S. Department of Defense, 1997.

Based on Reconfiguring the Supercomputers Runtime Environment New Security Methods

Andrey Molyakov*

Institute of information technologies and cybersecurity, Russian State University for the Humanities, Moscow, 117534, Russia

ARTICLE INFO

Article history:

Received: 09 April, 2020

Accepted: 18 May, 2020

Online: 29 May, 2020

Keywords:

Kripke structures

Sandbox

Stationary and on-board

Supercomputer systems

CTL logic

ABSTRACT

This paper is an extension of work originally presented in 2019 Third World Conference on Smart Trends in Systems Security and Sustainability (WorldS4) [1]. Author describes two new methods: reactive protection method (without delay after detecting an attack), which consists in virtualizing the execution environment of supercomputers processes if the calculated state descriptor falls into the "risk" zone and based on monitoring requests for allocation of resources in accordance with the rules of the security policy in the form of temporal modal structures CTL logic and method for reconfiguring the runtime environment of the supercomputers taking into account the mobility requirements (built-in computations) based on the application of the trajectories of computing state security descriptors on Kripke structures. The methods develop a number of provisions of the theory of information security, based on the new concept of Information Security of stationary and onboard supercomputer computing systems as a calculated convolution of the states of the execution environment (hardware or virtual) and system software.

1. Introduction

Widely used computers of various classes are essentially systems with von Neumann architecture, and the program execution model is a universal Turing machine with a tape of calculations, left and right shifts.

The performance criterion in the classical Popek - Goldberg theory for Turing machines is ambiguous. In the original work it was formulated twice [1]:

- a) "a statistically prevailing subset of virtual processor instructions must be executed directly by the physical processor, without the intervention of a virtual machine monitor";
- b) "all harmless instructions are executed directly by the physical processor, without the intervention of a virtual machine monitor".

These conditions are not identical. They are identical only if most of the instructions executed by the virtual machine are harmless, which cannot be guaranteed in the general case. In proving sufficient conditions for constructing a virtual machine monitor, Popek and Goldberg use the second definition.

Since the 70-80s, electronic technology has undergone significant changes. With the advent and development of hardware virtualization successfully implemented [2, 3]:

- *SMEP (Supervisor Mode Execution Prevention)*. Prevention of code execution in supervisor mode) is a technology developed by Intel to protect your computer from hacker attacks and other threats that use the so-called "supervisor mode".

- *SMAP (Supervisor Mode Access Prevention)*. It prevents writing to memory and reading code from it that unauthorized uses supervisor mode, while SMEP only prevents the execution of this code.

Supervisor mode is the preferred processor mode used by the kernel of the operating system. This mode is also called kernel mode. The opposite is the user mode in which user applications work [4, 5].

For SMEP to work, in addition to the corresponding processor, a suitable operating system is required. However, SMEP, although significantly complicating the task of hacking the system, still does not guarantee its complete protection. Therefore, later (in Broadwell architecture processors), in order to increase security and protect against vulnerabilities that were not resolved by SMEP, SMAP technology was additionally introduced [6, 7].

Direct execution of guest instructions whose encodings correspond to the privileged instructions of the physical processor on the host is impossible, since they certainly cause a trap to be thrown when executed in the user mode in which virtual machines are running. For example, the encodings of new secure instructions for Intel-64 architecture usually correspond to invalid encodings on processors

*Andrey Molyakov, Moscow, 117534, Kirovogradskaya street, 25/2, Russia. Tel: 8-495-388-0888 & andreimolyakov@mail.ru

www.astesj.com

<https://dx.doi.org/10.25046/aj050338>

of previous generations of this architecture. That is, new instructions are usually privileged on processors of previous generations, because they cause exceptions [8, 9].

Many hardware processors and various research groups are interested in supporting hardware transactional memory [10]. IBM was the first to develop a system with transactional memory hardware support in IBM Blue Gene [11] supercomputers and z-Enterprise EC12 servers [12]. AMD announced the development of the Advanced Synchronization Facility (ASF) extension [13], which is a variant of the hardware transactional memory for the Intel 64 architecture. Sun Microsystems has developed a processor code-named ROCK [14], but this project was closed. The Intel® Transactional Synchronization Extensions (Intel® TSX) instruction set [15], consisting of Intel® Hardware Lock Elision (Intel® HLE) and Intel® Restricted Transactional Memory (Intel® RTM) extensions, has been added to the sixth generation Intel Core processors [16].

There are also many software implementations of transactional memory ideas that do not require special support from the hardware. However, the study of existing solutions shows their extremely low productivity [17].

Moreover, with the advent of hardware transactional memory, the Intel RTM extension allows you to perform nested transactions, but at the same time, all internal XBEGIN and XEND pairs should not create new save points, that is, rollback of a nested transaction leads to a rollback of the entire transaction chain. The introduction of transactional memory extensions implies that the semantics of all instructions working with memory, as well as instructions affecting the state of the processor that are not stored in savepoints, must be changed in order to support the new execution mode.

In order to avoid critical errors, it is necessary to implement a multi-level “sandbox” in which the agents of the hypervisor function and monitor the execution of processor processing instructions at all levels of the hierarchy.

2. New actual requirements for information security

The Based on the Theorem on a protected hypervisor OS, formulated and proved by Zegzhda, two requirements can be formulated [18]:

- a) *Requirement 1 for virtualization of protected resources:* in a supercomputer computing system, all operations of application OS applications on protected resources must be carried out by virtualizing the resource in a virtual application environment;
- b) *Access control requirement 2:* in a supercomputer computing system, all operations of application OS applications on protected resources must be controlled by hypervisor protection tools.

Further developing the idea, one can formulate theoretical premises for the solvability of the supervenience problem: all possible operating modes of the processor should be considered - the real hardware mode, the protected hardware mode, the real virtual machine mode, the protected virtual machine mode. This means that the multi-domain protected hypervisor must ensure the correct execution (software using the interpreter or hardware) of all the instructions of the guest system, as well as guarantee isolation of virtual machines. At the same time, all safe instructions

common to the guest and host systems are executed directly by the physical processor (privileged modes).

The virtual machine monitor operates in mode 0 (in a separate domain), therefore, any interpretation errors, memory leaks, incorrect pointers, addressing errors of data code segments, program codes, and page addressing errors do not lead to an abnormal termination.

The solution to the problem of providing a modified instruction mnemonics for processors of different generations in order to support hardware transactional memory is to prevent the direct execution of these instructions in a virtual environment - they must be simulated programmatically on virtual processors (unprivileged modes).

Taking into account the development of the element design base and hardware technologies, we can formulate **new requirements for information protection** [18]:

- a) Protective equipment should control all informational interactions without exception;
- b) Security features should be developed independently of application programs and rely on an abstract representation of information interactions;
- c) Protective equipment should control information interactions based on clearly defined rules that make up the formal model;
- d) A mechanism should be provided to assess the safety of both the current state of the system and predict the safety of future states.

The principle of homogeneity of memory is the main postulate of von Neumann architecture. This means that for supercomputer systems in which record high performance is important, a fundamentally new architecture is required. Sharing the bus for program memory and data memory leads to a bottleneck in von Neumann's architecture, namely limiting the bandwidth between the processor and memory compared to the amount of memory. Due to the fact that program memory and data memory cannot be accessed at the same time, the processor-memory channel bandwidth and memory speed significantly limit the processor speed. Storage of data and commands in different places solves the problem of the “memory wall” on highly loaded computational problems [1].

3. Fundamental scientific problem

Over the past 25 years, OS security tools (for example, SMEP, SMAP, PaX, ExecShield, ASLR, DEP, Flusk, Patchguard and etc.) have come a long way, but failed to provide protection against current threats. The fundamental problem for this field of knowledge (computer science and computer technology) is known as the **problem of supervenience**: taking into account the high asynchrony and multi-connectedness of processes performed with supercomputer systems, the huge amount of processed data (Petabytes of data) there is no logical correspondence between changes in the programs of the runtime environment of supercomputers and changes hardware components in terms of implementing an isolated environment and monitoring security policy rules. For the chosen class of speakers, the solution to the fundamental problem is the development of the theoretical, scientific, practical and organizational and technical foundations

of reactive protection of stationary and onboard supercomputers, which consists in virtualizing the execution environment of the processes of applied and system software, in the interest of increasing their security and reliability, as well as substantiation of the principles for constructing mechanisms for implementing such protection for next generation automated systems.

The main contradiction for this class of computing systems is the ratio of energy intensity, speed and security. With the introduction of additional functional elements using the example of a hypervisor and a transactional memory controller, the characteristics of performance and energy consumption change. Concurrently executed queries conflict when they read and modify a certain database element, and the resulting conflict can lead to an erroneous result that could not be obtained if these queries were executed sequentially. Transactional memory provides a lightweight transaction mechanism for control flows running in a shared address space. It guarantees atomicity and isolation of parallel tasks [10, 11].

Atomicity ensures that changes in the state of a program made by code that is executed in a transaction are invisible from the point of view of other transactions executed in parallel [12].

Isolation ensures that concurrent tasks do not affect the outcome of the transaction, so that the transaction produces the same result as if no other task was being performed. Transactions provide the basis for constructing parallel abstractions, which are building blocks that can be combined without knowing their internal details, much like procedures and objects provide suitable abstractions for composing sequential code [13].

Integration with hardware transactional memory requires solving 5 main problems of computer science: limited application, debugging complexity, process synchronization and exclusion of access, resource control in conditions of parallelism and high asynchronous processes, emulation of different types of processors [14]. For multi-domain protected systems, trust is ensured by the following factors: “*Transparency*” – invisibility for the application OS, working directly with equipment, the amount of hypervisor code is small compared to the OS and applications, therefore it is easier to ensure that there are no vulnerabilities, all actions and events inside the virtualized systems are reversible - attacks can be quickly neutralized by rollback or reset.

Monitoring the exchange of a virtualized system with the external environment allows you to abandon a thorough study of the virtualized system itself.

Limitations associated with loss of productivity are extremely low, due to the record high performance characteristics of supercomputers.

4. Research methodology: new concepts and definitions

Information security is based on access control to objects of managing and guest OS, these objects can be attributed to different levels of protection. Examples of objects are directories, files, network sockets, registers, and interrupt handlers and some special-use memory areas by operating systems. The traditional access control approach involves the use of access attributes (rights) in requests to these objects for performing certain operations on them. If the verification of such attributes is successful, then access to the object at its security level is allowed, then the requested operation is performed on it. With this approach,

it is technically possible to intercept a request and use its access rights in a substitute request aimed at malicious impact.

New concepts and definitions are introduced in the work: supersecurity, information security of stationary and onboard supercomputer computing systems, a descriptor for assessing the security of system states.

Supersecurity is a property of a supercomputer system based on self-configuration and self-control. By using the high-performance reserves of these supercomputers, you can create a hypervisor that provides reliable protection against attacks associated with high asynchronous processes of the supercomputers and the ability to perform false transactions as a result of destabilizing external influences and accidental hardware failures.

The information security of stationary and onboard supercomputer computing systems is formalized as “a calculated convolution of the states of the execution environment (hardware or virtual) and system software”.

The descriptor for assessing the security of system states is a convolution that is “calculated” based on the attributes of ongoing processes that are implemented at the level of microprocessor cores and the interworking environment. All processes have their own descriptors calculated and methods for their processing — get, set, delete and etc. are specified.

5. New Safe Operations Model

The safe operations model includes 3 components: commands, data and timestamps. The space of operations in calculating the descriptors is a variety of Kripke's “worlds”. Each world of Kripke is assigned its own digital double and form a knowledge base for machine learning. The class of operations that are characteristic of supercomputers is the operations of multiplication and division in the form of successive shifts in the tagged structures of packets of requests from clients, guest and control OSs.

From the point of view of category theory, we work with the Group object - the category group. Objects are a group in the form of a residue ring, morphisms are mappings preserving the group structure. Category theory studies concepts through how these concepts interact with each other. We forget how these concepts are implemented, and we only look at the properties of connections, abstracting from the type of processor architecture (scalar, vector, MIPS, classic x86_64, tile architecture like Tiler, mass-multithread, hybrid and etc.), processor capacity (32-bit, 64-bit, 128-bit and etc.).

A type refers to a class type when a type provides certain operations with a specific expected behavior. For example, the tau type may belong to the Functor class if it has a specific behavior similar to a collection:

- a) The type **tau** is parameterized over another type, which you should consider as the type of the collection element. The type of the complete collection is then similar to `Scheme_type = {Int, String, Bool and etc.}`;
- b) If you contain integers, strings, or Booleans, respectively. If the element type is unknown, it is written as a parameter of type **a**. Examples include lists (zero or more elements),

type “Unknown” (zero or one element of type a), sets of elements of type a, arrays of elements of type a, all kinds of search trees containing values of type a, etc.

2. Another property that tau must satisfy is that if you have a function like $a \rightarrow b$ (function on elements), then you should be able to use this function and product for a related function over collections. You do this with the `fmap` operator, which is shared by each type of Functor class. The operator is actually overloaded, so if you have an even function with type $\text{Int} \rightarrow \text{Bool}$, then

Definition 1. A functor is a kind of collection (a set of configurations for recursively computing a hash), for which, if you are provided with a function on elements, `fmap` will return the function in the collections.

Definition 2. The context of the operation is a set of tuples consisting of logical variables. A collection of collections is a conjunction of atomic predicates defined on many tuples.

I appeal to the theory of functional programming, we will define a functor for comparing `is_equal` of two hash values with support for different class type templates :

```
class is_equal
{ private: scheme_type v; public:
is_equal(scheme_type value) : v(value) {}
bool operator () (scheme_type x)
{ return x == this->v; } };
scheme_type count_zero(const std::vector< scheme_type>& data)
{ return std::count_if(data.begin(), data.end(), is_equal(0)); },
scheme_type – data type (int, uint, float, double and etc.).
```

It is not possible to implement a complete hash function that is valid for all types. You cannot just convert an object to raw memory and hash bytes.

In addition, this idea fails due to padding technology when creating an index for each record. Because of this, it is necessary to take into account the context of a particular operation. To implement a universal convolution calculation algorithm, it is necessary to take into account different processor operating modes and the principle of type conversion and alignment of orders. We need to normalize the presentation of metadata operations in terms of aligning the boundaries of digital structures.

The basic principles of the algebraic structure on which the whole theory and methodology are based: The principle of Soft power, Self-organization, Supersecurity. The descriptors of the state safety assessment function are calculated on the set, which is a multiplicative group of the residue ring modulo 8.

Classical propositional logic is a “black and white” model; utterances are static, unchanged in time. In the ordinary propositional logic, sentences that do not explicitly or implicitly contain properties whose truth changes with time do not

adequately formalize. We want to study and verify systems that evolve over time.

The proposed approach to the description of operations is based on the classification of risks of information security breaches and analysis of the context of the implementation of outgoing directives, through which data can be transmitted bypassing the requirements of the adopted security policy, which leads to a violation of the security of computing nodes resources.

The carriers of the analyzed operations are the sets of objects and access subjects to which various security levels (labels) are assigned. To control the security level of operations generating new entities, for example, operations, we will use the sign of immutability of the object generating the access entity.

6. Theoretical calculations of the reactive protection method

6.1. Algebra of operations with objects processed on supercomputers

For supercomputers the sign of immutability cannot be a constant indefinitely: simultaneously, a huge number of processes are launched. Temporal logic is needed to describe the states of subjects and access objects. We need a model of threats to the integrity of the execution environment of IC processes, in which instead of the classic link “subject, object, predicate” a new paradigm of writing security policy rules is implemented and new entities are defined - “subject”, “object” and “descriptor for assessing state security”: for each *i*-th threat, the value of the state security assessment function is calculated, the arguments of which are given in the form of a conjunction of predicates of eight logical variables, the subjects are the guest and control OS, and the objects are the components of the hype Sizer and transactional memory controller.

Temporal logic of branching time consider possible calculations (paths on a tree) - trajectories on a scan of the Kripke structures.

The Kripke structure is a transition system with labeled states and unlabeled transitions. Sweep defines infinite chains of states - possible calculations. Each state can have not one, but many chains - continuations, and is the root of its tree of stories (calculations).

The structure of Kripke M is the five $M = (S, S_0, R, L, AF)$.

In our case [19, 20], $AF = \{\text{context_id} \mid \text{Dom_id} \mid S \mid \text{Ord} \mid \text{Context_type} \mid \text{TCU} \mid \text{TR}\}$. Let an arbitrary formula F_i of CTL logic and a Kripke structure M. be given. For each subformula ψ_i of formula F_i , the marking algorithm performs the following steps:

- a) We proceed to the construction of the numbering of the programs. Each operator is uniquely characterized by a pair - type (name) and a list of parameters, including operator labels, variables, functional and logical expressions. We only need to encode each of the possible values of the operator parameters. Labels of operators do not need special coding, since they are natural numbers;
- b) Since each program calculates a certain function, the introduced Gödel numbering generates some numbering

of the functions. On the one hand, this numbering is not one-to-one, since we code the program syntax, and any syntactically different programs have different codes. On the other hand, every computable function is computed by an infinite class of programs.

However, such numbering has fundamental and practically useful properties. Let A be the countable class of functions for assessing the safety of supercomputers states: $A = \{f_1, f_2 \dots, f_n\}$.

Even though each of the f_i functions is computable by some algorithm, this does not guarantee the existence of a single algorithm for computing all functions. We call a class A uniformly enumerable if there exists a two-place computable function F such that the class A consists exactly of functions of the form $F(n, x)$ for some n from the set N .

We call the function a universal function of class A . $nF(n, x)$ is the uniform numbering of class A .

Our main goal is to show that the numbering of the computable functions that we have determined is uniform with respect to other numbers and that by the number of the computable function in the given uniform numbering we can effectively find its number. Using the function for assessing the safety of SC states, we can subsequently accurately calculate the numbering states on the whole variety of Kripke structures, and then identify the classes of safe and dangerous operations.

Lemma 1. If A is an effective set, then for any effective set B : AB is effective, as well as any Cartesian product $A_1, A_2 \dots$. An of effective sets is effective, the set A^* , $A^* = A_n$, of all finite sequences of elements of A is effectively countable.

The proof of the statement is trivial and follows from the main theorem of arithmetic. We formulate and prove a theorem on the enumerability of the class of operations.

Theorem 1. Class A operations on the residue ring modulo m are uniformly enumerable, and the hash value calculation function F is a universal function of class A only if the number of hierarchy (nesting) levels is 8.

Evidence. In fact, operations are performed on data tuples, which are a set of i -th elements of an allergic structure (bit, atomic predicate, predicate conjunction, any set of predicative and functional symbols, vector variable, scalar, real number) of any dimension m (regardless of encodings and positional number system). We work with matrices, where the number of rows is equal to the number of nesting (hierarchies) of the computation space N , and the i -th column is an element of a structure of dimension m in the residue ring $Z(m)$, $i = 1 \dots n$. Using the Kornfeld formula to assess the confidence probability of security, each factor $P\Delta_i$ is a geometric decreasing progression [21]. The most suitable values are 0.1 and 0. (1). Criteria of temporal logic (lack of new patterns over time) satisfies only the solution $N = 8$.

Moreover, $nF(n, x)$ is an effective numbering of class A , since with a change in the dimension of the parameter n the regular relations between the elements of the set do not change, new properties of objects of our algebraic structure do not appear.

The main result of theoretical calculations is that we have proved universality and uniformity for given classes of operations of functional transformations. Regardless of the dimension of the input parameters, the nature of the relationships between the types, the type of operands, the type of processor architecture, the operation scheme and implementation of the algorithms, we have found such a number of hierarchy levels ($N = 8$) that there is an invariant in the algebraic system. If we evaluate the time parameters, this regularity property is preserved indefinitely (from 0 to $+\infty$).

6.2. Integration of the algebraic structure with architectural modifications of supercomputers

We proceed directly to the construction of a weighted multigraph-model for performing operations of supercomputers, which is a tree. The starting point is the top of the tree structure, describes a transactional memory controller that interacts directly with the hypervisor verifier module. The controller in conjunction with the verifier implements security mechanisms: an isolated multi-domain address space is represented as non-overlapping memory areas, each of which corresponds to a vertex of the graph lying in the second level of the root structure. The vertices of the third level correspond to the subsets of the components that make up the hypervisor. Since we have 8 levels of the request processing hierarchy, the block is divided into eight subsets of level details of possible states indicated by S_8, \dots, S_1 .

The last level of the root structure of a multigraph is represented by agents of the hypervisor (graph leaves) through which communication with the external environment and all possible attacks on the supercomputers occur. Advances to each next level are a higher level of abstraction of the description of request processing, a transition from the lower level of the specification to the upper. Each state can have not one, but many chains - continuations, and is the root of its tree of stories (calculations).

The reactive protection method should include not only deductive algorithms for calculating descriptors for assessing the security of states and identifying threats based on marking the states of fulfillment of supercomputers requests, but also implementing inductive learning based on self-diagnostics and explanatory decision-making mechanisms based on the concept of machine learning. We calculate such impacts that ensure the safe functioning of supercomputers, and block dangerous and suspicious ones.

7. Method for reconfiguring the runtime environment of the supercomputers

The hash is calculated recursively for all processes (subprocesses) and an effective value is obtained for the regular measurement of the query path (based on past experience, pre-training). In the process, the system is being trained. Each new configuration is classified and recognized in the future based on the characteristic set of markers. Coding of identification features - a set or conjunction of simple predicates on which the formula F_i is given [20]. The subformula ψ is a recursive call of the same algorithmic procedure, only with a different set of characteristic features and properties.

6 configurations of “colors” of the tree of computation histories. We select two colors: Blue - the predicate of the presence of the transient process, Red - the predicate of changing the type of the context of the operations performed.

- a) configuration allows you to track changes in local states at the same hierarchy level and identify the launch of child processes as part of the base process with the markers of interest to us;
- b) configuration analyzes the entire state tree based on the selected marker (the mode of a full run of calculations in case of restarting the task). The story tree is being rebuilt;
- c) analyzes the change in two parameters in the near branches of the tree (local scale, in one domain);
- d) configuration to track a single (irregular) parameter change in the far branches of the computation tree;
- e) configuration regularly repeated changes of the selected identification parameter along the entire trajectory of movement;
- f) configuration studies the branches of the tree; over time, in the future, changes in two parameters may appear on different trajectories. At the same time, processes can be started in different domains.

Sets of configurations adapted to different contexts of operations execution, in the form of matrices of access rules, are stored by agents of the hypervisor monitoring information security events in the memory of the transactional memory controller. In the process of training, the rules of safety rules are updated and adjusted. The system dynamically evolves and modifies sets of modal rules based on CTL temporal logic grammars for responding to signaling events. Identification criteria for threats in the semantic interpretation of a set of atomic predicates.

Thus, the hypervisor in conjunction with the transactional memory controller is a multi-agent system with machine learning.

There are deductive checks that correspond to the principles of classical logic, but we are using inductive algorithms with metadata about the operation of system components at different levels of the hierarchy, where, along with checking the integrity in the form of changing the values of the calculated hash functions for each process, access control is implemented in the form of Labeling allowed transients when switching between security domains and thread migrations. The history of various system configurations is kept, new events are recognized due to a variant new set of predicates and the values of their trajectories on Kripke structures.

8. Discussion

8.1. Optimization, normalization of descriptor calculation

For optimization and normalization, it is necessary to specify an exactly safe data volume. The *axiomatic of our algebraic structure* are as follows:

- a) Multiplication and division operations are specified (cyclic shifts to the right or left);
- b) A lot of descriptor calculations is a residue ring modulo $m = 8$;

- c) The ring must be symmetrical with respect to the performance of the operations of multiplication and division;
- d) A request for any tap of a processor with a classic von Neumann architecture is a combined tagged structure in which a data segment and a program code segment are stored in the same RAM sections. In this case, the L-operand for the recursive calculation of ψ_i is stored in the data section, the R-operand for the recursive calculation of ψ_i is stored in the code section.

The principle of supervenience is the absence of differences of one kind in the absence of differences of another type: the absence of differences in the set of process descriptors in the absence of changes in the configuration of software and hardware agents. There is a one-to-one logical correspondence between a change in the set of a vector variable stored in the hypervisor generative tables, a map of transactional memory states (hardware level), and a hash function value (software level). In the process of passing the request, the key line does not change normally. If it changes, then it means that the route for passing the request by a third-party agent is being modified.

A-property – the identity of the values of the security assessment descriptors for the i -th process at the n -level of the hierarchy (running programs at the OS level, each process is identified by a hash value, formalized as a hash function value).

B-property – configuration of software and hardware components (formalized as a set of matrices in the form of 8 16-bit key lines, based on which the value of the function is calculated).

8.2. The solvability condition for the supervenience problem for supercomputer systems

In this case, one must take into account the upper or lower case of addresses (a segment of a data code or a command code, a combined format for representing the structure in RAM). Then the amount of precisely safe calculations doubles and you need to introduce an additional $2 * 4 = 8$ register fields in the form of significant bits of the n -dimensional vector $\Psi(n) = 8$ to control the addressing and offset boundaries in the pipeline of processed requests and prevent the execution of “false transactions”. Thus, the dimension of the control parameter k is 8. In this case, the bits (upper and lower registers) of the pointer to the RAM memory cells are not mixed.

If this condition is not fulfilled for the comparison functor, then the boundaries of the data segment and code are violated, the descriptor is recursively calculated: for $N < 8$, empty bits remain (you can write zeros or junk data in them), and normalize the processed data for $N > 8$ (BDC - format and type conversion of operands) is not possible.

To get this type of representation of a data set in the form of tuples of n -logical variables, the minimum set (basis) is 8. When the number of iterations is a multiple of 8, we increase the amount of data of recursive calculations, but the effects and patterns do not change. With a decrease in the number of iterations (less than 8 levels of the query processing hierarchy), automatic alignment of the processed data in the BDC format leads to the fact that

individual bits or groups of bits remain uninitialized, which allows dangerous shifts in the supercomputers.

The optimization of descriptor calculations is due to eight interval constraints in the processing cycle to obtain a productive sample, using only eight significant bits (rather than infinitely large numbers) for storing and processing the results. Based on the foregoing, we formulated the solvability condition for the fundamental problem of establishing logical correspondence (supervenience) of the program execution process on the supercomputer, by changing the state of the hypervisor components and supercomputers hardware: the number of levels of the request processing hierarchy by the hypervisor N in interaction with the transactional memory controller should be eight.

Solvability condition. The number of levels of the request processing hierarchy by the hypervisor N in interaction with the transactional memory controller should be eight.

A change in the states of processes at the upper level leads to a coordinated change in states at the lower level; there is a similarity relation. When using two rings of protection ($N = 2$) and in the absence of physically separated storage of programs and data, the process can gain access to the "alien" segment and increase privilege levels.

9. Results

Development, taking into accounts the specifics of supercomputers, of a fundamentally new technological solution creating an isolated program execution environment in the form of an 8-level sandbox with the implementation of control mechanisms both at the level of hypervisors and at the level of transactional memory controllers. Thus, an important scientific and technical problem has been solved in the field of creating information security tools for a new class of systems - stationary and on-board supercomputers. All the means of protecting information that existed today were the hardware and software components of the protected system itself, without integration with hardware transactional memory and the introduction of multi-level control, it was impossible to solve the problem of detecting and identifying various types of threats at all levels of the hierarchy of query execution. In the course of experimental studies, new scientific results were obtained that confirm the efficiency and minimal performance loss of applying hardware virtualization technology in the form of a multi-level "sandbox" for promising supercomputers compared to using traditional clusters.

Since it is impossible to control the operation of all equipment, but only to monitor the execution of requests at the level of the components of the hypervisor and the controller of transactional memory, during the research, the maximum level of functioning of information protection agents was found - S8. With this configuration, when the number of hierarchy levels is $N = 8$, the execution of context-sensitive operations becomes quasi-determined with a confidence probability of approximately 0.9.

The threshold for performance loss when using the verification complex is less than 6-7%. The effectiveness of the developed security system was assessed based on the use of various security

tools (used to protect clusters and mainframes) and analysis of the number of successful recognitions and errors.

The following results of experimental studies are obtained:

- a) with the number of training samples $n \geq 86$, stable detection with critically minimal errors of the 1st and 2nd kind is ensured;
- b) only 8 clusters are enough for effective detection of malicious code, which significantly reduces the cost of implementing a software and hardware solution. Each cluster is bound to a security domain. Their number is also 8.

Thus, an unambiguous correspondence has been achieved between the number of hierarchy levels and the number of protection domains and the number of countable clusters tied to a certain "sandbox" level. To ensure the requirements of the international standard ISO 5725 for the convergence and reproducibility of measurement data, the experiments were performed in a series of 5 repeated experiments for 14 days until the observed data were stabilized.

10. Conclusion

- a) The theorem on the uniform enumerability of functions for assessing the safety of states is proved. In this case, the means of protection of stationary and onboard supercomputers are considered as an object of research for the first time;
- b) The solvability condition is formulated, as a consequence of the theorem on the uniform enumerability of functions, the problems of establishing the logical correspondence (supervenience) of the program execution process on a supercomputer, by changing the states of the components of the hypervisor and equipment;
- c) A reactive protection method has been developed (without delay after detecting an attack), which consists in virtualizing the execution environment of supercomputers processes if the calculated state descriptor falls into the "risk" zone and based on monitoring requests for allocation of resources in accordance with the rules of the security policy in the form of temporal modal structures CTL logic;
- d) A method has been developed for reconfiguring the runtime environment of supercomputers taking into account the mobility requirements (built-in computations) based on the application of the trajectories of computing state security descriptors on Kripke structures;
- e) It is proposed to use a model of threats to the integrity of the execution environment of supercomputers processes, in which instead of the classic link "subject, object, predicate" a new paradigm of writing security policy rules is implemented and new entities are defined - "subject", "object" and "descriptor for assessing state security" : for each i -th threat the value of the state security assessment function is calculated, the arguments of which are given in the form of a conjunction of predicates of eight logical variables, the subjects are the guest and operating systems, the objects are components of the hypervisor and transactional memory controller;

- f) A safe operation model has been developed that considers, from the standpoint of “integrity” of the runtime environment, both supercomputers hardware at the microprocessor level and system software, describing its processes in the form of decomposition into an 8-level hierarchical structure: each process privilege level uniquely corresponds to a domain number protection and operation mode of the microprocessor.

- [20] Molyakov A. S., “Token Scanning as a New Scientific Approach in the Creation of Protected Systems: A New Generation OS MICROTEK” / Automatic Control and Computer Sciences, 2016, 50(8), 687-692.
- [21] Molyakov A. S., “Threat model and theoretical foundations of the reactive protection method of supercomputers” / Natural and technical sciences, Company Sputnik +, 2019, 7, 197–201.

References

- [1] Molyakov A. S., “New security descriptor computing algorithm of Supercomputers” / 2019 Third World Conference on Smart Trends in Systems Security and Sustainability (WorldS4), IEEE Xplorer Digital Library. <https://doi.org/10.1109/WorldS4.2019.8903965>
- [2] Popek G. J., Goldberg R. P., “Formal Requirements for Virtualizable Third Generation Architectures” / Communications of the ACM, 1974.
- [3] F. Leung, G. Neiger, D. Rodgers et al., “Intel® Virtualization Technology” : Intel Technology Journal., 2006, 10.
- [4] AMD Corporation. — AMD64 Architecture Programmer’s Manual Volume 2: System Programming, 2013.
- [5] Stephan Diestelhorst, Martin Pohlack, Michael Hohmuth et al., “Implementing AMD’s Advanced Synchronization Facility in an out-of-order x86 core” / 5th ACM SIGPLAN Workshop on Transactional Computing, 2010.
- [6] Intel® Software Development Emulator, 2012. <https://software.intel.com/en-us/articles/intel-software-development-emulator>.
- [7] Rechistov Grigory, Plotkin Arnold, “Implementation of Intel Restricted Transactional Memory ISA Extension in Simics” / Procedia Computer Science, 2013, 18, 1804 – 1813.
- [8] Herlihy Maurice, Moss J. Eliot B., “Transactional memory: architectural support for lock-free data structures” / Proceedings of the 20th annual international symposium on computer architecture, ISCA '93, New York, USA : ACM, 1993, 289–300.
- [9] Amy Wang, Matthew Gaudet, Peng Wu et al., “Evaluation of Blue Gene/Q Hardware Support for Transactional Memories” / Proceedings of the 21st International Conference on Parallel Architectures and Compilation Techniques, PACT '12, 2012, 127–136.
- [10] Jacobi Christian, Slegel Timothy, Greiner Dan, “Transactional Memory Architecture and Implementation for IBM System Z” / Proceedings of the 2012 45th Annual IEEE/ACM International Symposium on Microarchitecture, 2012., 12, 25–36.
- [11] Jaewoong Chung, Stephan Diestelhorst, Martin Pohlack et al., “ASF: AMD64 Extension for Lock-free Data Structures and Transactional Memory” / Proceedings of the 2010 43rd Annual IEEE/ACM International Symposium on Microarchitecture, 2010.
- [12] Moir Mark, Moore Kevin, Nussbaum Dan, “The adaptive transactional memory test platform: a tool for experimenting with transactional code for ROCK” / Proceedings of the twentieth annual symposium on Parallelism in algorithms and architectures, SPAA '08, New York, USA : ACM, 2008, 362–362.
- [13] Maurice Herlihy, Victor Luchangco, Mark Moir, William N. Scherer , “Software Transactional Memory for Dynamic-sized Data Structures” / Proceedings of the Twenty-second Annual Symposium on Principles of Distributed Computing, PODC '03, New York, USA : ACM, 2003, 92–101.
- [14] Dice Dave, Shalev Ori, Shavit Nir, “Transactional Locking II” / Proceedings of the 20th International Conference on Distributed Computing, DISC'06, Berlin, Heidelberg : Springer-Verlag, 2006, 194–208.
- [15] Milo M. K. Martin, Daniel J. Sorin, Bradford M. Beckmann et al., “Multifacet’s general execution-driven multiprocessor simulator (GEMS) toolset” / SIGARCH Comput. Archit. News, 2005, Vol. 33, no. 4, 92–99.
- [16] Yi Liu, Yangming Su, Cui Zhang et al., Efficient Transaction Nesting in Hardware Transactional Memory / Architecture of Computing Systems - ARCS 2010, Berlin, Heidelberg : Springer Berlin Heidelberg, 2010, 138–149.
- [17] Moore Kevin E, “LogTM: Log-Based Transactional Memory” / In Proceedings of the Twelfth IEEE Symposium on High-Performance Computer Architecture, 2006, 258–269.
- [18] Zegzhda D.P., “Building secure operating systems based on virtualization technology” [presentation], St. Petersburg, 2018, 33 p. <https://docplayer.ru/storage/77/76383820/1589898846/qi9gw3tca9toeeJG7VgqWQ/76383820.pdf>
- [19] Molyakov A. S., “A Prototype Computer with Non-von Neumann Architecture Based on Strategic Domestic J7 Microprocessor” / Automatic Control and Computer Sciences., 2016, 50(8), 682 -686.

Analysis of Local Rainfall Characteristics as a Mitigation Strategy for Hydrometeorology Disaster in Rain-fed Reservoirs Area

Kartono Kartono^{1,2,*}, Purwanto Purwanto^{1,3}, Suripin Suripin⁴

¹*Doctorate Program of Environmental Science, School of Postgraduate Studies, Universitas Diponegoro, Semarang, 50241, Indonesia*

²*Department of Mathematics, Faculty of Sciences and Mathematics, Universitas Diponegoro, Semarang, 50275, Indonesia*

³*Department of Chemical Engineering, Faculty of Engineering, Universitas Diponegoro, Semarang, 50275, Indonesia.*

⁴*Department of Civil Engineering, Faculty of Engineering, Universitas Diponegoro, Semarang, 50275, Indonesia*

ARTICLE INFO

Article history:

Received: 09 April, 2020

Accepted: 13 May, 2020

Online: 29 May, 2020

Keywords:

Hydrometeorology Disasters

Rain-fed Reservoir

Mann-Kendall Test

Time series model

Mitigation

ABSTRACT

The Gembong reservoir in Pati Regency, Java, Indonesia is a rain-fed reservoir, which experiences a depletion of its carrying capacity. The characteristic of local rainfall is one of the important factors in assessing the potential of hydrometeorology disasters in its area. Sedimentation in watersheds and reservoirs has covered water sources, so local rainfall determines the dynamics of water availability. This research is needed in the development of mitigation strategies. This article contains an analysis of the characteristics of local rainfall, and forecasting based on local daily rainfall data. This data was obtained from the rainfall gauge station in the Gembong reservoir area in 2007-2019. Variation coefficient, anomaly index, rainfall concentration index, and Mann-Kendall test were used to identify its characteristics. The time series model is used as modeling for forecasting. The results of empirical analysis show that rainfall volatility with irregular changes in high variability, meteorological drought in moderate category, rainfall trends follow fluctuating patterns and do not follow monotonic trend patterns but high concentrations of rainfall. Forecasting results with the Autoregressive Integrated Moving Average model for the wet months show an increasing in total rainfall by 17% in the next year. So, the potential for flooding is greater than the potential for drought. Based on the analysis of the local rainfall characteristics, then mitigation on flood is preferred.

1. Introduction

Water is the most important natural resources for sustainable life in the world. Rainfall is one of the environmental variables that is often used to track the level and variability of climate change and is influencing the availability of water in the future [1]. Long-term climate change associated with changing patterns and intensity of rainfall has a potency to increase the frequency of drought or flooding.

Trend analysis based on historical rainfall data is needed to detect trends in climate variables [2], and in general, patterned sinusoidal periodic waves [3]. Climate change with volatile changes between water abundance and water scarcity make it difficult to manage its water availability [4]. Volatility of rainfall

shows the size of fluctuations in rainfall data that is the size of the deviation to the average rainfall needs to be observed properly [5].

Rainfall volatility explains changes in rainfall intensity, which is very important to consider in the preparation of hydrometeorology disaster mitigation strategies [3-5]. The number of losses incurred due to natural disasters of climate change is necessary to mitigate disasters.

Observations at the location confirmed that the Gembong reservoir in Pati Regency, Indonesia is a rain-fed reservoir. The water collected in the Gembong Reservoir comes from direct rainwater or water through watersheds that inflow to the reservoir during the rainy season. The rivers are not able to drain their water into the reservoir in the dry season. The availability of water in the reservoir depends only on local rainfall. Climate change and its

*Corresponding Author: Kartono, Email: kartono@live.undip.ac.id

effect on local rainfall condition potentially influences the annual variability of water availability in reservoir.

Potential of hydrometeorology disasters depend on rainfall volume, intensity and duration of rainfall. Rainwater increases river flow and causes flooding [5,6]. Extreme changes of low to high rainfall intensity in a short period of time are often indications of flash floods or landslides. The frequency and intensity of hydrometeorology disasters and their impacts in Indonesia are increasing and resulting in greater losses, even claiming human lives. The Situ Gintung tragedy on March 27, 2009 is as an example

Extreme changes in rainfall intensity from high to low over a long duration of time lead to drought. Water scarcity impacts all aspects of life and the economy [7, 8]. Floods, landslides and drought are related with the characteristics of local rainfall as a consequence of climate change. The Climate change impacts the changes in hydrological cycle and water availability in reservoirs [9], the frequency and extent of flooding [6], and also makes it difficult to form reservoirs operating patterns [10].

The agricultural sector is directly affected [11], so food security is very vulnerable to this disaster, both droughts in the dry season and flooding in the rainy season. Failure to harvest can occur due to lack of water, as well as when the plants approaching the harvest hit by flooding. The agricultural sector is very dependent on climatic conditions and its variability that can change agricultural ecosystems, which have an impact on agricultural production and food security [12]. Deforestation in the Muria Mountain by human activities in land use can trigger a potential of hydrometeorology disasters in reservoir areas at any time [13]

Reservoir operation patterns in Indonesia are based on the results of rainfall estimates [14], so that an understanding of the characteristics (volatility, trends, patterns) of local rainfall [15] is needed in its preparation. This article contains the results of research aimed at identifying and analyzing the volatility of local rainfall and its predictions in the future in rain-fed reservoirs. Local rainfall is a key component of climate, changing patterns and intensity having a direct impact on the availability of water in rain-fed reservoirs. The results of this research provide a very important and valuable input on sustainable management of water resources and planning for hydrometeorology disaster mitigation strategies [3, 5]. The Gembong Reservoir, built in 1933, is very vulnerable to climate change, especially the local rainfall volatility.

Therefore, the main discussion in this article includes (1) detecting and analyzing the characteristics of local rainfall in the Gembong reservoir area, (2) predicting local rainfall using a time series forecasting model, (3) identifying a potential hydrometeorology disaster, and (4) conclusions.

2. Methods

2.1. Study Area

Gembong Reservoir belongs to Gembong District, Pati Regency, Central Java Province, Indonesia and on the eastern of Mount Muria with altitude of 500-1000 meters above sea level. Gembong Reservoir is located on 06°41'47.99" South Latitude and 110°57'21.79" East Longitude. The main benefit of this reservoir is as irrigation water supply to agricultural land area of around 4,606 ha. The Gembong Reservoir was chosen as the location of www.astesj.com

the study because it was the same age as the Situ Gintung Reservoir, Banten, Indonesia, which had experienced a flash flood disaster. The tragedy occurred on March 27, 2009 not only due to rainfall factors but also by the age of the old embankment [16].

2.2. Research Design

Daily rainfall data from the rainfall gauge station in the Gembong reservoir area in 2007-2019 was obtained from the reservoir operator's daily records through observation. Raw data is then presented in an Excel worksheet. It was processed and analyzed by using the Statistics methods. Software R version 3.6.1 for descriptive statistics dan Minitab version 17 for forecasting.

Rainfall volatility analysis is based on calculating the coefficient of variation (CV), the standardized precipitation anomaly index (SPAI), and the precipitation concentration index (PCI) [17], while the trend is analyzed by using Mann-Kendall (MK) test [2,18-21]. SPAI was used to display quantitative analysis related to changes in rainfall and identify potential floods or droughts [8, 22], where rainfall is seasonal and periodic in a monsoon-dominated climate context [23], like the climate in Indonesia. The PCI is an important feature in water resource management planning, prediction or serves as a warning tool for potential hydrometeorology disasters [24]. A positive value of Z-statistic indicates that there is the upward trend in time series data and vice versa [3].

The ARIMA (Autoregressive Integrated Moving Average)-Box Jenkins method is one of the prediction methods that is often applied for the prediction of rainfall time series data [25, 26]. In this research, the ARIMA model for wet months is used for forecasting. The MASE (Mean Absolute Scaled Error) value measures the accuracy of the prediction, because there are zero data in the dry month [27].

3. Results and Discussion

3.1. Description

The results of in-depth interviews with the community around the reservoir, sub-district and village officials, reservoir operator and reservoir care communities, and observations onsite confirm that there are no spring water sources in the river basin. This statement confirms that Gembong Reservoir is a rain-fed reservoir. The daily local rainfall graph in the Gembong Reservoir area is presented in Figure. The pattern of each year illustrates the form of periodic sinusoidal waves as generally occurs in other countries [3], although extreme rainfall of 425 mm occurred at the beginning of the month in 2018. Extreme fluctuations of rainfall increase the inflow of the river into the reservoir. Such conditions have the potential for flooding [5, 6].

Descriptive statistics in Table 1 can explain the average rainfall on the Gembong Reservoir has a wide data distribution. It shows the interval of changes in the rise and fall of the local daily rainfall is wide. Rainfall in the Gembong Reservoir has fluctuations in the data distribution is quite large. Table 2 shows the maximum rainfall occurring in January. as in Figure 1.

Table 1. Descriptive Statistics of Rainfall (2011-2018)

Min	Max	Mean	Standard deviation
0	425	5.6	17.2

3.2. Analysis of rainfall trends and volatility in the Gembong Reservoir

Non-parametric tests are used to detect trends where the results are shown in Table 2. Results on Table 2 explain that the largest CV occurred in August and September, which shows that the variability between years is greater than in other months, but it has the smallest volatility. The level of variability in rainfall events is high according to the classification developed by Hare [2]. At the significance level $\alpha=0.05$, the rainfall data do not follow monotonic trend patterns at all months. Only in January, August and December, it was significant with $\alpha=0.1$, with approaching

the pattern of ascending monotonic trends. It illustrates that the rainfall data fluctuate with the uncertainty in the sense of the monotonic trend patterns.

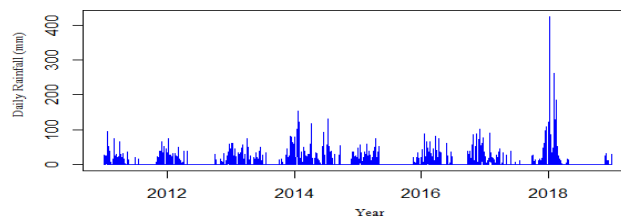


Figure 1: Daily Rainfall on Gembong reservoir 2011-2019

Table 2: Rainfall characteristics at Gembong station

Month	Min	Max	Mean	CV	Volatility	S	Z	Tau MK-test (Sig)
Jan	0	425	15.9	199.2	31.7	3341	1.5	0.06 (0.10)*
Feb	0	263	11.3	211.5	23.8	-693	-0.3	-0.01 (0.74)
Mar	0	98	6.4	218.8	14.0	-442	-0.3	-0.01 (0.82)
Apr	0	118	6.7	234.8	15.7	-356	-0.2	-0.01 (0.84)
May	0	61	2.8	324.3	9.1	-525	-0.4	-0.02 (0.71)
Jun	0	93	1.7	475.0	7.9	138	0.1	0.01 (0.90)
Jul	0	75	1.1	587.4	6.3	-285	-0.4	-0.02 (0.73)
Aug	0	68	0.6	855.6	4.7	-950	-1.6	0.07 (0.10)*
Sep	0	86	0.7	833.2	6.0	-447	-0.7	-0.03 (0.51)
Oct	0	86	2.4	347.3	8.3	1227	0.9	0.04 (0.38)
Nov	0	88	4.3	261.7	11.3	719	0.4	0.02 (0.67)
Des	0	108	10.4	180.7	18.7	3445	1.6	0.06 (0.10)*

Note: * statistically significant at the level of significance $\alpha = 0.1$

Table 3: Classification of meteorological drought severity

Month	Classification of Meteorological Drought Severity			
	Extreme ($Z < -1.65$)	Severe ($-1.65 < Z < -1.28$)	Moderate ($-1.28 < Z < -0.84$)	No drought ($Z > -0.84$)
Jan				v
Feb			v	
Mar			v	
Apr			v	
May			v	
Jun				v
Jul			v	
Aug		v		
Sep			v	
Oct				v
Nov				v
Dec				v

Volatility, the magnitude of fluctuations in rainfall data measured using standard deviations, values ranging from 4 to 32. January has the highest volatility value by 31.7 with an upward monotonic trend ($Z > 0$) with significance level. Likewise, in the month of December, however, the value of volatility is smaller than January by 18.7. In June, October and November, it has an upward trend ($Z > 0$) but does not follow a monotonic trend pattern. August has a downward monotonic trend with a level of significance, while in other months it has a downward trend but does not follow a monotonic trend pattern. Based on the calculation of rainfall anomalies standardized as Z in Table 2, the severity class of meteorological drought is presented in Table 3. It can be seen that severe meteorological drought is occurred only in August

A strong indicator for the distribution of temporal rainfall is shown by PCI [17, 20]. Table 4 present the annual PCI values based on monthly rainfall data each year. The classification criteria using the criteria introduced by Oliver in 1980 [21] are presented in Table.5. The average PCI in the 2007-2018 periods was 19.67, which was classified as a high concentration of rainfall. Table 5 shows those 9 out of 12 years have a heterogeneous distribution of rainfall, so the occurrence of rainfall is less predictable. In 2018, extreme rainfall was confirmed, as shown in Figure. Furthermore, the amount of rainfall for the future is predicted using the time series model [26]. Therefore, forecasting local rainfall in the wet months is needed to analyze the potential for hydrometeorology disasters.

Table 4: The Annual PCI 2007-2018

Year	PCI	Year	PCI
2007	16	2013	15
2008	25	2014	19
2009	20	2015	20
2010	11	2016	15
2011	17	2017	17
2012	20	2018	42

Table 5: The PCI Classification

Index	Precipitation Concentration Class*	Number of Years (2007-2018)
<10	Low (almost uniform)	0
11-15	Moderate	3
16-20	High	7
≥21	Very high	2

Note: *Processing results based on Oliver's criteria

3.3. ARIMA modeling for wet months

In this modeling, wet months are months where rainfall exceeds 100 mm/ month, otherwise called dry months. Modeling is based on rainfall data in the Gembong reservoir in 2007-2018, while testing uses rainfall data for 2019. The iterative steps that have been carried out in ARIMA modeling for the wet month include:

- Data stationary test shows that the rainfall data of the Gembong reservoir is stationary in the mean.
- The charts of the autocorrelation function(ACF) and partial autocorrelation function (PACF) describe that the Autoregressive (AR) model was interrupted at lag 3, 4 and the Moving Average (MA) model was also interrupted in lag 3, 4. Therefore, the possible ARIMA models are ARIMA([3, 4], 0, [3, 4]), ARIMA([3, 4], 0, 0) or ARIMA(0, 0,[3, 4]).
- Testing parameter ϕ in the AR model. The hypothesis $H_0: \phi = 0$ (the parameter ϕ is not significant in the model) is

rejected based on the t test statistic with level of significance $\alpha = 5\%$. Table 6 displays the test results for significance of the model parameters. The parameters of the ARIMA ([3, 4], 0, 0) or ARIMA (0, 0, [3, 4]) are significant.

- The white-noise process can be detected using the residual autocorrelation test in its error analysis, which is detecting whether there is a residual correlation between lags. The H_0 hypothesis is rejected with level of significance $\alpha = 5\%$. The ARIMA([3,4], 0,0) has the R^2 greater than the ARIMA(0,0, [3,4]).
- The ARIMA([3,4],0,0) or AR([3,4]) model was chosen as the best model because it fulfills all assumptions. Based on the coefficient values in Table 6, the ARIMA([3,4],0,0) model can be written

$$X_t = \phi_0 + \phi_3 X_{t-3} + \phi_4 X_{t-4} \tag{1}$$

$$X_t = 0.024623 - 0.24784 X_{t-3} - 0.35878 X_{t-4} \tag{2}$$

Table 6: Test results for the significance of model parameters

Model	Parameter	Coefficient	p-value	Explanation
ARIMA([3,4],0,0)	c	0.024623	0.0000	Significant
	AR 3	-0.24784	0.0280	
	AR 4	-0.35878	0.0018	
ARIMA(0.0,[3,4])	c	0.024404	0.0000	Significant
	MA 3	-0.31918	0.0039	
	MA 4	-0.31767	0.0042	

Table 7: Validations and Forecasting

Wet Month	Average Rainfall in 2019 (data)	Average Rainfall in 2019 (model)	Average Rainfall in 2020 (forecasted)
1	14.52	6.25	7.83
2	4.71	5.13	7.54
3	6.87	5.08	6.64
4	8.40	5.55	5.98
5	3.23	6.91	5.84
Accumulation	37.73	28.92	33.83

MASE (model)= 0.73 < 1 ; MASE (forecasted) = 0.63 < 1

where X_t is the observation data at time t .

- f. The model is validated using rainfall data for 2019. Because there is monthly rainfall data that is the amount of 0 (dry season), then forecasting accuracy is based on value of MASE studied in [27]. Table 7 presents the model validations and forecasting.

The calculation results show that the MASE value is smaller than one, which indicates it is a good forecasting [27]. Forecasting results show that the accumulation of average rainfall in a wet month is increased by 17%. This condition can trigger a hydrometeorology disaster in the reservoir area.

3.4. Potential of Hydrometeorology Disasters

Analysis of rainfall data is also intended to be able to assess the dynamics of extreme fluctuations in rainfall. Potential of hydrometeorology disasters generally occur in extreme conditions that cause very heavy rain to fall in a short time or long drought. From Figure, an extreme rainfall was occurred on the Gembong Reservoir area in 2018. Therefore, studies in extreme conditions need to be conducted in order to anticipate the potential hydrometeorology disasters.

From Table 2, the volatility value is ranging from 4 to 32, and the rainfall data fluctuate with an uncertainty of the monotonic trend patterns. This condition of rainfall that fluctuate in this way with irregular changes between water abundance and water scarcity make it difficult to manage water allocation [4]. Changes

in rainfall patterns and intensity have a direct impact on the availability of water in reservoirs, which has the potential of hydrometeorology disasters. Increased rainfall trends can lead to an increase in the frequency of floods, but a downward trend can increase meteorological drought [23, 28]. Hydrological behavior in catchments is influenced by changes in local rainfall patterns that trigger erosion [15]. Uncertainty in local rainfall trend patterns makes it more difficult to anticipate disaster risk. The capacity of the Gembong Reservoir as a rain-fed reservoir experiences environmental stress [13, 29]. The pressure is even greater with the uncertainty of this rainfall trend pattern, and the occurrence of rainfall is less predictable. Its condition will complicate the management of reservoir water availability.

According to the classification in Table 5, the frequency of both High and Very High Concentration of Rainfall is 75%, and there is no Low Concentration. This high concentration triggers erosion and landslides [15, 24], the material of which will increase sedimentation in reservoirs and upstream reservoirs. The smaller the capacity due to increased sedimentation and is triggered by high rainfall concentrations, the chance of flooding through the watershed or even the sinking Gembong reservoir will increase. When rain with high intensity flushes the slopes of Muria experienced deforestation and land conversion in the catchment area, the absorption capacity of rainwater in the area is getting smaller, resulting in erosion and the material conducted by water flow on the slopes of Mount Muria which is included in the Gembong District area of 342 ha [30]. The occurrence of a flash

floods caused by a broken embankment in Situ Gintung on March 27, 2009 is a warning to raise awareness [16].

Potential of hydrometeorology disasters can be caused by the inability of a reservoir to collect direct rainwater or surface water from a watershed. When the flow of inflow and precipitation into the reservoir exceeds the reservoir capacity, it will overflow through the spillway which has the potential to cause flash floods. The results of in-depth interviews with community leaders confirmed that there was a flash flood in 2006 that affected settlements adjacent to the river that tipped on the slopes of Muria. The flash floods occurred due to the upstream river on the slopes of Mount Muria being unable to support rainwater, what if the Gembong Reservoir which volume of water is 9.5 million meters³ is broken down? In fact, the Situ Gintung tragedy occurred with a volume of water of around 2.1 million m³ which caused huge losses and caused many casualties [16].

When the water inflow discharge and precipitation into the reservoir do not exceed the capacity, even smaller than the capacity, there will be a shortage of water that starts a drought. Lack of water will result in drought in the area of irrigation. Forecasting results with the ARIMA model for wet months show that an increase of 17% in 2020 (Table 7). The average PCI is high (Table 4) and the reservoir capacity that is shrinking due to sedimentation will increase the potential for flooding. The potential for flooding is greater than the potential for drought, especially since the reservoir is getting older.

Conservation of the catchment areas and protected forests to increase the absorption capacity of rainwater on the slopes of Mount Muria is prioritized as flood mitigation strategy in the highlands. Directives for the management of protected areas have been outlined in the Pati Regency Spatial Planning in 2010-2030 [30]. Protecting forest conservation should be implemented immediately when the mountainous area is not yet barren and loses the topsoil.

4. Conclusions

The characteristics of local rainfall in the Gembong reservoir area based on empirical studies with statistical methods are (1) volatility of rainfall with irregular changes in high variability between maximum and minimum rainfall, (2) there is no extreme meteorological drought, (3) rainfall trends follow fluctuating patterns and do not follow monotonic trends with a significance level $\alpha=0.05$, but high rainfall concentrations have the potential to cause erosion or flooding, (4) forecasting with the ARIMA model for wet months shows that there is an increase in rainfall accumulation. The ARIMA's prediction results warn that the vigilance against potential flood disasters is greater than the potential for drought disasters.

These findings indicate that the Gembong Reservoir is sensitive to climate change and vulnerable to potential hydrometeorology disasters. Preparation of mitigation strategies is more directed at reducing the risk of flooding because rainfall is not the only environmental factor that causes disasters. Disaster mitigation strategies need to pay attention to other triggering factors, such as reservoir infrastructure and surrounding land use. Reducing sedimentation in watersheds and reservoirs, accelerating conservation of water catchment areas and reforestation of

protected forests on the slopes of Mount Muria are mitigation strategies on flood that can be preferred at this time.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Authors thank Universitas Diponegoro for funding this research through RPP Research Grant 2019 contract: 329-55/UN7.P4.3/PP/2019.

References

- [1] IPCC, Climate Change 2007: Impacts, Adaptation and Vulnerability, Cambridge University Press, Cambridge, USA, 2007
- [2] A. Asfaw, B. Simane, A. Hassen, A. Bantider, "Variability and time series analysis of rainfall and temperature in northcentral Ethiopia: A case study in Woleka sub-basin", *Weather and Climate Extremes* 19, 29-41, 2018. <https://doi.org/10.1016/j.wace.2017.12.002>.
- [3] R. Mahmood, S. Jia, W. Zhu, "Analysis of climate variability, trends, and prediction in the most active parts of the Lake Chad basin, Africa", *Scientific Reports* 9:6317, 2019 <https://doi.org/10.1038/s41598-019-42811-9>
- [4] A. Kuhn, W. Britz, D.K. Willy, P. van Oel, "Simulating the viability of water institutions under volatile rainfall conditions- The case of the Lake Naivasha Basin", *Environmental Modelling & Software* 75, 373-387, 2016. <http://dx.doi.org/10.1016/j.envsoft.2014.08.021>.
- [5] L. Malinowski, I. Skoczko, "Impacts of Climate Change on Hydrological Regime and Water Resources Management of the Narew River in Poland", *Journal of Ecological Engineering* 19(4), 167-175, 2018. <https://doi.org/10.12911/22998993/91672>.
- [6] S. Szewranski, J. Kazak, M. Szkaradkiewicz, J. Sasik, "Flood risk factors in suburban area in the context of climate change adaptation policies-Case study of Wroclaw Poland", *Journal of Ecological Engineering* 16(2), 13-18, 2015. doi:10.12911/22998993/1854
- [7] Y. Uttarak, T. Laosuwan, "Drought Detection by Application of Remote Sensing Technology and Vegetation Phenology", *Journal of Ecological Engineering* 18(6): 115-121, 2017. DOI:10.12911/22998993/76326.
- [8] A. Lopez-Nicolas, M. Pulido-Velazquez, H. Macian-Sorribes, "Economic risk assessment of drought impacts on irrigated agriculture", *Journal of Hydrology* 550:580-589, 2017. <http://dx.doi.org/10.1016/j.jhydrol.2017.05.004>.
- [9] A.W. Salami, H. Ibrahim, A.O. Sojobi, "Evaluation of impact of climate variability on water resources and yield capacity of selected reservoirs in the north central Nigeria". *Environ. Eng. Res.* 20(3): 290-297, 2015. <http://dx.doi.org/10.4491/eer.2015.0041>.
- [10] N. Ehsani, C.J. Vorosmarty, B.M. Fekete, E.Z. Stakhiv, "Reservoir operations under climate change: Storage capacity options to mitigate risk", *Journal of Hydrology* 555, 435-446, 2017. <https://dx.doi.org/10.1016/j.jhydrol.2017.09.008>.
- [11] K. Ryzuma, E. Radzka, T. Lenartowicz, "The Impact of Precipitation Conditions on Medium-Early Cultivars of Potato Yielding", *Journal of Ecological Engineering* 16(3): 206-210, 2015. DOI: 10.12911/22998993/2958.
- [12] R.K. Rai, L.D. Bhatta, U. Acharya, A.P. Bhatta, "Assessing climate-resilient agricultural for smallholders", *Environmental Development* xxx, 2018. <http://doi.org/10.1016/j.envdev.2018.06.002>.
- [13] K. Kartono, P. Purwanto, S. Suripin, "Potential environmental pressures on water availability in Gembong reservoir in Pati District for the development of agropolitan area", *Journal of Physics: Conf. Series* 1217- 012061, 2019. doi:10.1088/1742-6596/1217/1/012061.
- [14] Government Regulation of the Republic of Indonesia number 37 of 2010 concerning Dams.
- [15] E. Aydin, J. Antal, "Introduction to Precipitation Runoff Process and Soil Erosion Risk Analysis in a Specific Area of Interest to design Control Measures", *Journal of Ecological Engineering* 20(2), 44-50, 2019. <https://doi.org/10.12911/22998993/94921>.
- [16] B. Harsoyo, "Jebolnya Tanggul Situ Gintung (27 Maret 2009) Bukan Karena Faktor Hujan Ekstrem", *Jurnal Sains dan Teknologi Modifikasi Cuaca*, 11(1), 9-17, 2010. <http://ejurnal.bppt.go.id/index.php/JSTMC/article/view/2176/1814>.
- [17] E. Radzka, "Classification of Precipitation Intensity During Vegetation Season in Central-Eastern Poland (1971-2005)", *Journal of Ecological Engineering* 15(3):51-55, 2014. DOI:10.12911/22998993.1109123

- [18] A.Mondal, S.Kundu, Mukhopadhyay, "Rainfall Trend Analysis by Mann-Kendall Test: A case study of North-Eastern Part of Cuttack District, Orissa", *International Journal of Geology, Earth and Environmental Sciences* ISSN:227-2081, 2(1):70-78, 2012. <http://www.cibtech.org/jgee.htm>.
- [19] E.Gasiorek, E.Musial, "Evaluation of the precision of standardized precipitation index (SPI) based on years 154-1995 in Lodz". *Journal of Ecological Engineering* 16(4): 49-53, 2015. <https://doi.org/10.12911/22998993/59347>
- [20] K.Zhang, Y.Yao, X. Qian, J.Wang, "Various characteristics of precipitation concentration index and its cause analysis in China between 1960 and 2016", *Int J Climatol*; 1-11. 2019 <https://doi.org/10.1002/joc.6092>..
- [21] B.Lamboni, L.A.Emmanuel, C.Manirakiza, Z.M.Djibib, "Variability of Future Rainfall over the Mono River Basin of West-Africa", *American Journal of Climate Change* 8:137-155, 2019 <http://www.scirp.org/journal/ajcc>.
- [22] A.H.Nury, K.Hasan, "Analysis of drought in northwestern Bangladesh using standardized precipitation index and its relation to Southern oscillation index". *Environ.Eng.Res.* 21(1):58-68, 2016. <http://dx.doi.org/10.4491/eer.2015.115>.
- [23] K.Chanda, R.Maity, "Meteorological drought quantification with standardized precipitation anomaly index for the regions with strongly seasonal and periodic precipitation", *J. Hydrol. Eng.*, 2015. doi:10.1061/(ASCE)HE.1943-5584.0001236.
- [24] E.E.Ezenwaji, C.P.Nzolwu, G.N. Chima, "Analysis of Precipitation Concentration Index (PCI) for Awka Urban Area, Nigeria", *Hydrology Current Research* 8(4), 2016. doi:10.4172/2157-7587.1000287
- [25] W.W.S.William, "Time Series Analysis: Univariate and Multivariate Method", New York : Pearson Education, Second Edition, 2006.
- [26] J.Farajzadeh, A.F. Fard, S. Lotfi, " Modeling of monthly rainfall and runoff of Urmia lake basin using "feed-forward neural network" and "time series analysis" model", *Water Resources and Industry* 7-8, 38-48, 2014. <http://dx.doi.org/10.1016/j.wri.2014.10.003>
- [27] R.J.Hyndman, A.B. Koehler, "Another look at measures of forecast accuracy", *International Journal of Forecasting* 22, 679-688, 2006. doi:10.1016/j.ijforecast.2006.03.001
- [28] C.M. Tfwala., L.D van Rensburg, R.Schall, P.Dlamini, " Drought dynamics and interannual rainfall variability on the Ghaap plateau, South Africa, 1918-2014", *Physics and Chemistry of the Earth* 107:1-7, 2018. <https://doi.org/10.1016/j.pce.2018.09.003>.
- [29] Pemali-Juana River Region Center, " Final Report of the infrastructure investigation in the Gembong Reservoir on 2015", 2015
- [30] Pati Regency Regulation Number 5/2011 concerning Pati Regency Spatial Planning in 2010-2030

Multimode Control and Simulation of 6-DOF Robotic Arm in ROS

Rajesh Kannan Megalingam*, Raviteja Geesala, Ruthvik Rangaiah Chanda, Nigam Katta

Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, 690525, India

ARTICLE INFO

Article history:

Received: 15 January, 2020

Accepted: 02 May, 2020

Online: 29 May, 2020

Keywords:

Inverse Kinematics

Robot Operating System (ROS)

Graphical User Interface

6 DOF (Degree of freedom)

Control Interfaces.

ABSTRACT

The paper proposes the design and simulation of a 6 Degree of Freedom (DOF) robotic arm, tailored for the coconut crop harvesting, assistive robots like wheelchairs and Home robots, Search and rescue robots for disastrous environments, Collaborative robots for research use. A kinematics-based solution has been developed for the robotic arm which makes it easier to operate and use. Keyboard, GUI, Joystick are the three control interfaces used in the paper. The robotic control interfaces proposed in the paper were developed using the Robot Operating System (ROS). The 6- DOF articulated robotic arm was designed and visualized in RVIZ. The kinematics helped for the easy manipulation of the robotic arm with the end effector. The methodologies proposed in the research work are easy to operate and inexpensive. The designed 6 DOF robotic arm, the first three DOFs are for positioning of the robot's arm, while the residual three are used for manipulation of the gripper.

1. Introduction

The robotic arm is a programmable mechanical arm that works similar to a human arm. Robotic arms played a significant role in the process of industrial automation. The human-like dexterity of these robotic arms makes them efficacious in diverse applications in a variety of industries - manufacturing, atomic power plants, space exploration, material handling, painting, drilling, agriculture deployments [1], assistive robotics applications [2] and numerous other applications. The robotic arm typically comprises an end effector that is designed to manipulate and govern with the surroundings. The 6 DOF robot arm is designed to manipulate and govern with the surroundings. The 6 DOF is to pivot in 6 different ways that mimic a human arm. The major issues concerned in an industrial robotic arm are its mechanical structure and the control mechanism. The control mechanisms can be effectuated by 3 options: keyboard, joystick and slider-based control. Design of a lightweight robotic arm which can be compatible with any kind of robotic system. In the research, all the proffered control mechanisms adopted inverse kinematics, which makes it easier to control. The proposed control mechanisms are compatible with any other complex robotic systems of the same degrees of freedom. The dexterity of the robotic manipulator depends on the degree of freedom.

Kinematic analysis [3] is one of the important steps in the design of the robotic arm. The loop equations of complex robotic systems

can be deduced from the kinematic equations. Joint angles regulate the motion of the robotic arm. Computer-Aided Design (CAD) software is the platform to create models with the given set of geometrical parameters. The proposed robotic arm was designed in Solid Works CAD software considering all the given sets of geometrical parameters. Robot Operating System (ROS) provides an integrated platform for the control of robotic systems. ROS is a special kind of framework initially developed with the purpose of working on robots in the research domains. In order to understand how the ROS framework works one should be clear about the concept of communication of messages through topic between nodes. Simulation is one of the ways to optimize the design and improve the control of robotic systems. ROS provides a 3-D visualizer (RVIZ) which helps in the visualization of the pose. With a properly-set URDF (Unified Robot Description Format) file, one can visualize the robot model in RVIZ. RVIZ is the simulation software used for the control of the 6-DOF robotic arm.

The design and development of 7-DOF arms [4] and even successful simulations on 11-DOF [5] arm are done by other authors with valid proofs of classic IK algorithms [6] in the discrete-time domain for robots. Taking those inputs into consideration the 6-DOF robotic arm in the paper was designed and tested with various goal positions using different modes of control. Algorithm developed for the keyboard based control [7] of the 6 DOF robotic arm using ROS which makes the control easier. The design and simulation of this robotic arm which is

*Rajesh Kannan Megalingam, India, +91-9496120900, rajeshkannan@ieec.org

visualized in RVIZ. Simulation of a robot is also possible on the Gazebo platform [8] with the help of ROS control packages and the generation of the robot configuration is in Gazebo, it uses the URDF file and this is extensive use of a simulation platform for a real-life scenario. Simulation of the actuators in the arm can be done using various simulation software's [9] like Matlab and Simulink [10], [11] but choosing RVIZ is preferable as there are many references to proceed. An open-source robot simulator called USARSim can also be used for both research and education of a robot's general architecture [12]. An inverse kinematics algorithm [13] needs to be developed to achieve the goal positions in simulation over these platforms using the traditional method the DH parameter [11], [14] the inverse kinematics solutions [1], [15] for the arm have been derived. Few other authors used ADAMS [16] to simulate and evaluate the arm design but this paper deals with RVIZ simulation with multimode controls. With the help of ROS platform and MoveIt, the control [17] of the simulated arm was done and the three control modes [18] were tested. The sliding mode controller mechanism can be helpful for regulation of robotic arms with unknown behaviors [19]. Laser range method [20] can also be used for positioning and tracking tasks for a 6-DOF robotic arm. This method helps in the determination of the distance from the camera to the target. The workspace [21] of the arm and the path for reaching the goal position varies on the approach and techniques used while modelling the simulation. There are various other control mechanisms like trajectory tracking [22], master slave control [23], vibration control of a robotic arm with input constraints. Testing in real time doesn't show accurate results due to backlash and electric issues. Considering this problem some authors proposed a self-learning algorithm [24] that uses positioning error after each trail. Simulation gives an idea about the performance of the algorithm, but in case if real time implementation performance changes due to several factors. One such thing is the connectivity and the architecture. Few authors explicitly gave intuition into the master-slave configuration based tele existence concept [25]. Considering all this leads, a 6 DOF robotic arm is simulated and tested by using different control mechanisms.

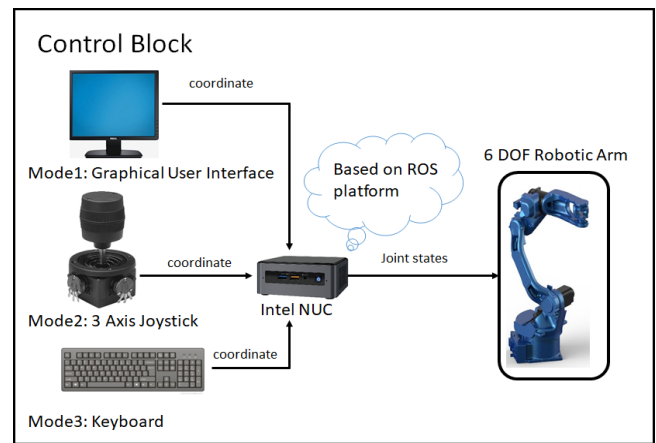


Figure 1: General Architectural diagram.

2.1. Control Block

The control block is the interface between the user and the robotic arm. It consists of three modes of control. The user can control a robotic arm using any of the below modes.

- GUI
- Joystick
- Keyboard

The robotic arms is equipped with a manipulator. A manipulator is a robotic gripper which is used to do dexterity. Figure 2 represents the software architecture of the 6 DOF robotic arm in the ROS platform.

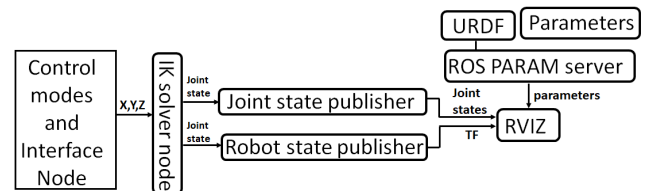


Figure 2: Software Architecture diagram.

The main aim is to control the robotic arm's manipulator position in the RVIZ. Figure. 11, describes an RVIZ axis at the base of the robotic arm. The user can control the manipulator by using any of the control modes. These users input x, y, z coordinates will be published to IK solver node. The IK solver node will compute the joint states which are required for the manipulator to reach to the desired coordinates.

2.1.1 Control modes and Interface nodes

As there are different optional modes to command the robotic arm, interface node acts as an intermediate module that prioritizes the modes. The interface node will act as a bridge between the modes of control and the computing block IK_solver_node. The user can define the order of priority for the modes in the definition file in the interface module. On uploading the entire program, the compiler uses the updated definition file. On simultaneous publishing of data from the three control modes, always the

Contributions of the research is as follows:

- Kinematics based solution for the 6 DOF robotic arm is proposed.
- Validation of three control methodologies of the robotic arm.
- Simulation and evaluation of the 6 DOF robotic arm in RVIZ.

2. Architecture

The paper is primarily concerned about the different modes of control. Figure 1 represents the general architectural diagram which consists of control modes, a processing unit, and a 6 DOF robotic arm. ROS is used as the platform to develop software programs.

Control in testing the 6-DOF robotic arm by simulating it with an IK-based solution in RIVZ. The programs and files for the control modes and processing unit entirely return in CPP.

highest priority mode values will be published to the IK_solver_node from the interface node.

2.1.2. IK Solver node

The Inverse Kinematics equations are generated by using Denavit-Hartenberg. An IK function is written in a computing file where it is passed with x, y, z parameters for computing. And there is a threshold evaluating block of code where the arms reach limit is checked whether it could reach the x, y, z coordinates. If the coordinate is present within the workspace of the arm it generates the individual joint values else, it generates the joint values which head to the maximum point it could reach in the direction heading to the input point. The generated joint values for the desired x, y, z coordinate point will be conveyed to the joint state publisher. Joint state publisher is the one which updates the joint values of the robotic arm simulation.

2.1.3. Graphical User Interface (GUI)

A user-friendly GUI was designed in QT Designer software [26] of version 4.8. Qt Designer is used to create user interface files containing windows and controls. Figure 3 depicts the representation of the graphical user interface which was developed in Qt Designer. The designed window consists of 3 fill in blocks, a SEND COORDINATES toggle button, a HOME toggle button, and a HALT toggle button.

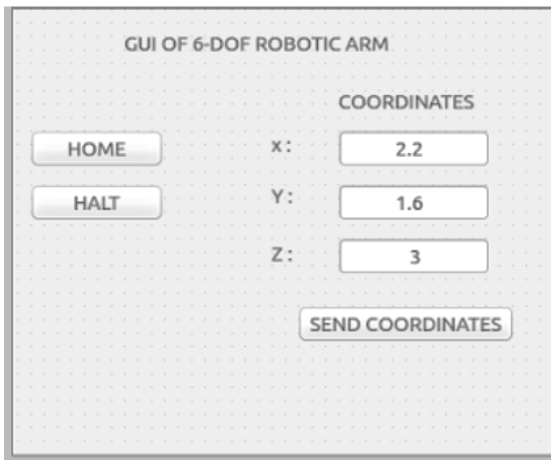


Figure 3: Representation of the graphical user interface.

The user can enter their x y z coordinates to which the user wanted the robotic arm to reach. On clicking the SEND COORDINATE button, the values will get published to the IK solver node. For resetting the robot, a HOME toggle button is designed. Whenever the user presses the toggle button, the robotic arm will reach the predefined pose. A HALT toggle button is also designed which will stop the motion of the robotic arm in case of emergency.

2.1.4. 3- Axis Joystick

The Joystick is another option for controlling the robotic arm. Figure 4 depicts a model of the 3-axes joystick. Arduino is the microcontroller used for taking the joystick values. The joystick will be connected to Arduino analog pins externally; through analog pins the joystick values are fetched.

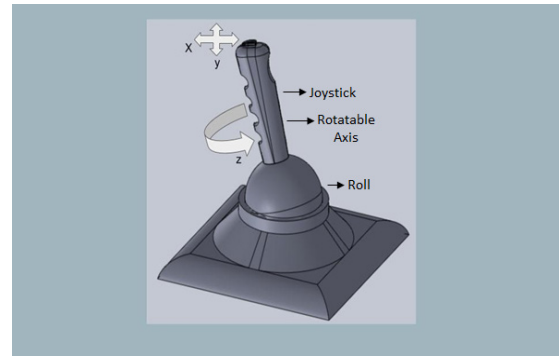


Figure 4: representation of 3 axes joystick

The three axes of the joystick are used to control the x y z coordinates position of the manipulator of 6 DOF robotic arm. If the user moves the joystick's axis in forward direction, x value in the coordinate increases, similarly y and z can also be changed with the other two axes of the joystick. In this way, the user can control the manipulator position.

2.1.5. Keyboard

The keyboard is another mode of control. The keys 'a', 's', 'd', 'z', 'x', 'c', 'h', 'j' are used to control the x, y, z values. When a key from 'a' or 's' or 'd' on the keyboard is pressed, the x, y, z coordinate value increases respectively. When keys from 'z' or 'x' or 'c' are pressed then the coordinate variables x, y, z decreases respectively. 'h' and 'j' are the home pose and halt buttons. These coordinates (x, y, z) are published to the interface node. And from the interface node to the IK solver node for further computation.

3. Design and Implementation

Robotic arms play a vital role in the industrial and social applications, where the design of the arm plays an important role in its function. According to the functionality, there are six main types of industrial robots: Cartesian, SCARA, cylindrical, delta, polar and vertically articulated. The design of the robotic arm that was developed for the research is not compromised in terms of the workspace. Any arm with revolute joints between all its base members falls in the category of the articulated robotic arm. An articulated arm has the ability to rotate all the joints (number of joints is from two to ten). The reason behind the choice for the articulated robotic arm for the research is that it has more work envelope compared to SCARA, Delta, and Cylindrical.

Arm base is called the waist which is vertical to the ground and the upper body of the robot base is connected to the waist through a revolute joint which rotates along the axis of the waist and to which the second DOF (shoulder) is attached as similar to the arm in Figure 5. Third DOF which resembles the elbow of a human arm. The fourth DOF gives yaw movement, fifth gives the pitch and the last DOF gives the roll movement. The 6-DOF articulated robotic arm was designed on Solid works and dimensions are mentioned in the Table 1. The first 3-DOFs give positional (i.e. x, y, and z) coordinates to the end-effector and the last 3-DOFs give the orientation of the end-effector. In the Table 1 base footprint is the base mounting stand for the robotic arm. The arm has 6 links namely base_link, Link1, Link2, Link3, Link4, Link5 and

followed by the gripper. Link1 is the link between joints J1 and J2, Link2 is between joints J2 and J3. Similarly, for Link3, Link4, Link5 are the links with their respective joints.

Assembly file in Solid works is converted into a URDF file using Solid works with a maximum reach length of 1.248 m. Figure 5 and Figure 6 represents the design of the model of a 6-DOF robotic arm in Solid works. URDF is the conventional format for describing robots in ROS. The format is used to simulate a virtual arm. The joint positions, link lengths, and global origin are defined in the Solid works before exporting the assembly file of Solid works into URDF for simulation. The designed arm is simulated and controlled in Rviz using ROS. The evaluation of the designed URDF is done using MoveIt setup assistant which is ROS platform. In this section, the author first loaded the URDF file and created the collision matrix. MoveIt divides the arm joints and manipulators into groups, which helps to have a safer control of the robotic arm by resolving errors due to singularity issues.

Table 1: Joint link lengths.

i	joint	parent	child	Link length “(m)”	Shift in Origin “(m)”
0	Fixed base	Base footprint	base_link	0	0
1	J1	base link	link1	0.15	0
2	J2	link1	link2	0.14	0
3	J3	link2	link3	0.45	0
4	J4	link3	link4	0.216	0.04
5	J5	link4	link5	0.091	0
6	J6	link5	link6	0.106	0.03
7	gripper	link6	Gripper link	0.095	0
	Total			1.248	

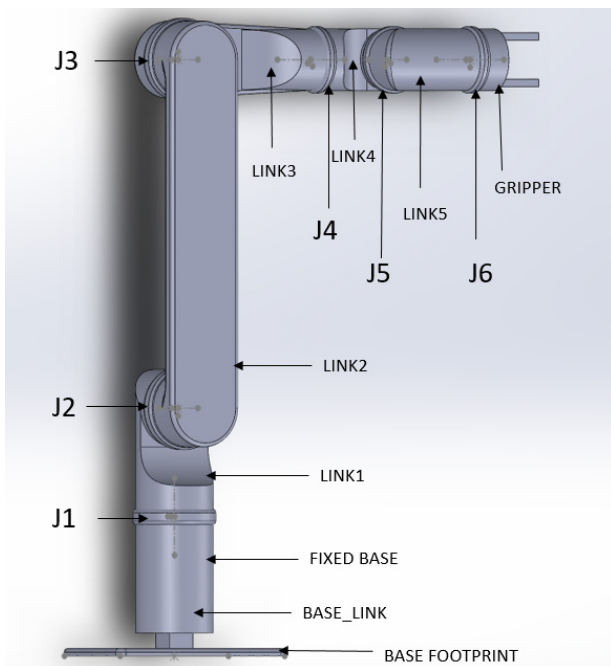


Figure 5. Front view of the Robotic arm

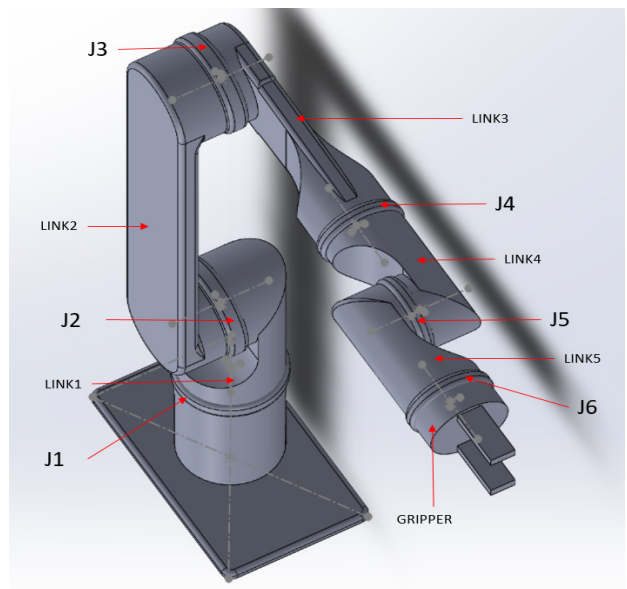


Figure 6. Top view of the Robotic arm

4. Robotic Arm Kinematics

The kinematics of the robotic arm are categorized into forward kinematics and Inverse kinematics. To calculate the kinematics solution the concepts of reference frames, Euler angles and Homogenous transforms are important.

4.1. Reference Frame representation

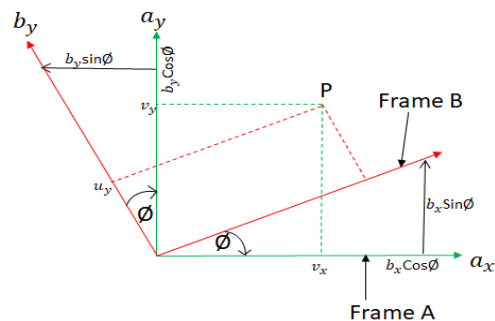
Reference frames play an important role for the kinematics [27] of any arm. The general concept of reference frames is explained with the illustration in Figure 7. Consider point P, in the 2-D Cartesian as shown in Figure 7, is expressed with vector U relative to coordinate frame B. The same point P needs to be represented in frame A with vector V as shown in equations (1), (2) and (3). The matrix representation of the point P in frame A is given in (3). Any point in frame B is multiplied with a_bR will project it onto frame A.

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \hat{a}_x \hat{b}_x & \hat{a}_x \hat{b}_y \\ \hat{a}_y \hat{b}_x & \hat{a}_y \hat{b}_y \end{bmatrix} \begin{bmatrix} U_x \\ U_y \end{bmatrix} \quad (1)$$

$$\begin{aligned} \hat{a}_y &= \hat{b}_x \sin \phi = \hat{b}_y \cos \phi \\ \hat{a}_x &= \hat{b}_x \cos \phi = -\hat{b}_y \sin \phi \end{aligned} \quad (2)$$

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} U_x \\ U_y \end{bmatrix} \quad (3)$$

$${}^a_bR = \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix}$$



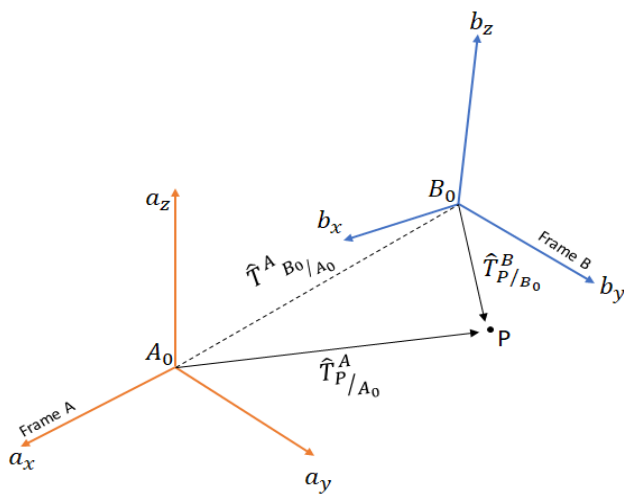
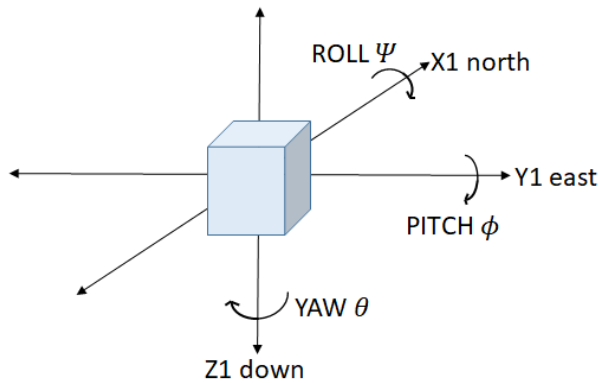
4.2. Euler angles

For a rigid body in space as shown in Figure 8 a box is considered as a rigid body, it has 3DOF for rotation those are roll, pitch, and yaw. These are called the Euler angles [28] and the representation of this rotational DOF. Euler angles are a system to describe a sequence or a composition of rotations. Conventionally, the movements about the three axes of rotations and their associated angles are described by the 3D rotation matrices. Equations (4), (5) and (6) show the matrix representation of roll, pitch, and yaw. These are 3D- rotation matrices and anticlockwise in direction.

$$R_x(\Psi) = \begin{bmatrix} \cos \Psi & -\sin \Psi & 0 \\ \sin \Psi & \cos \Psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

$$R_y(\phi) = \begin{bmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$R_z(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$



4.3. Homogeneous transforms

Homogeneous transform is the representation of the transform matrix of one frame which is both rotated and translated with

respect to some other unknown frame. In Figure 9 point in frame, B is transformed onto frame A. And the homogeneous transform is generated using rotational and translational matrices.

Equation (7) the term \hat{T}_{P/A_0}^A is the representation of point P vector with respect to frame A and this is 3*1 matrix. On the other end in the term R_B^A is the rotational matrix of frame B on to frame A and this is a 3*3 matrix. \hat{T}_{B_0/A_0}^A Is the translation vector and this is 3*1 matrix. The \hat{T}_{P/B_0}^B is a representation of point P vector with respect to B frame.

$$\begin{bmatrix} \hat{T}_{P/A_0}^A \\ 1 \end{bmatrix} = \begin{bmatrix} R_B^A & \hat{T}_{B_0/A_0}^A \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{T}_{P/B_0}^B \\ 1 \end{bmatrix} \quad (7)$$

4.4. Forward kinematics

The forward kinematics [27] deals with location and pose of the end effector with the joint variables and parameters. Here the end-effector frame is mapped to the base joint of the arm using homogenous reference frames as discussed above. Denavit–Hartenberg parameters [14], [11] (also called DH parameters) are the four parameters associated with a particular convention for attaching reference frames to the links of a spatial kinematic chain, or robot manipulator. These are calculated for all the joints using the Figure 10 as a reference and data given in Table 2. Origin O(i) is the intersection between x(i) and z(i) axes. The joint angle ($\theta(i)$) is t, the angle from axis x(i-1) to x(i) about z(i) using the right-hand rule. In joint 3 it has an offset of 180-degree constant between x (2) and x(3) and in joint 4 it has an offset of 90-degree constant between x(3) and x(4). Link twist ($\alpha(i-1)$) is the angle from z(i-1) to z(i) measured about x(i-1) using right hand rule. In Table2 link length (a(i-1)) is equal to the difference between z(i-1) and z(i) along the x(i) direction. Link offset (d(i)) in Table 2 is the difference between x(i-1) and x(i) along the z(i) axis. Using equation 8, which is the general representation of the transform matrix from one frame to another. So by using the data available in Table 2 transform matrix for link 0 and link 1 is given in T_0^1 in (8) Similarly, (9) is the transform for link 1 and link 2. $T_1^2, T_2^3, T_3^4, T_4^5, T_5^6, T_6^7$ These are also similar kinds of transform matrices with their respective links. Finally, the transform matrix from base joint to the end effector is derived by T_0^7 shown in (16). The orientation difference between the definitions of the gripper in URDF to the DH conventions so the rotation need to be done around y axis (17) and then need to be done around z axis (18) that is R_c shown in (19).

Table 2: DH-Parameters for the arm.

Links	O(i)	$\theta(i)$	$\alpha(i-1)$	a(i-1)	d(i)
0-1	1	θ_1	0	0	0.15
1-2	2	θ_2	-90	0	0.14
2-3	3	θ_3+180	180	0.45	0.45
3-4	4	θ_4+90	-90	0.04	0
4-5	5	θ_5	-90	0	0.091
5-6	6	θ_6	90	0.03	0.106
6-7	7	θ_7	0	0	0.095

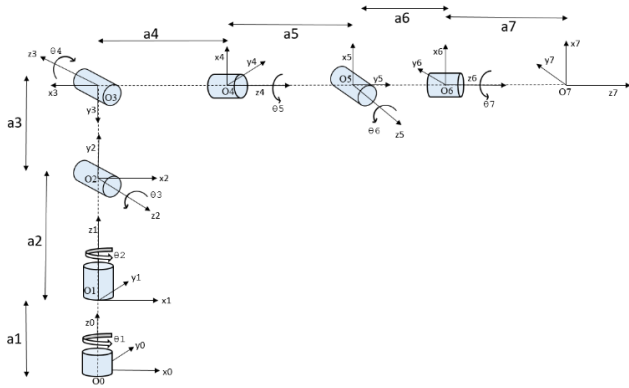


Figure 10: Cartesian representation of joints with rotational angles for DH-parameters.

$$T_i^{i-1} = \begin{bmatrix} c\theta_i & -s\theta_i & 0 & a_{i-1} \\ s\theta_i c\alpha_{i-1} & c\theta_i c\alpha_{i-1} & -s\alpha_{i-1} & -s\alpha_{i-1} d_i \\ s\theta_i s\alpha_{i-1} & c\theta_i s\alpha_{i-1} & c\alpha_{i-1} & c\alpha_{i-1} d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (8)$$

$$c\theta_i = \cos \theta_i$$

$$s\theta_i = \sin \theta_i$$

$$c\alpha_{i-1} = \cos \alpha_{i-1}$$

$$T_0^1 = \begin{bmatrix} c\theta_1 & -s\theta_1 & 0 & 0 \\ s\theta_1 & c\theta_1 & 0 & 0 \\ 0 & 0 & 1 & 0.15 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

$$T_1^2 = \begin{bmatrix} c\theta_2 & -s\theta_2 & 0 & 0 \\ 0 & 0 & 1 & 0.14 \\ -s\theta_2 & -c\theta_2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (10)$$

$$T_2^3 = \begin{bmatrix} -c\theta_3 & s\theta_3 & 0 & 0.45 \\ s\theta_3 & c\theta_3 & 0 & 0 \\ 0 & 0 & -1 & -0.45 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11)$$

$$T_3^4 = \begin{bmatrix} s\theta_4 c\theta_4 & 0 & 0 & 0.04 \\ 0 & 0 & 1 & 0 \\ -c\theta_4 & -s\theta_4 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (12)$$

$$T_4^5 = \begin{bmatrix} c\theta_5 & -s\theta_5 & 0 & 0 \\ 0 & 0 & 1 & 0.0911 \\ -s\theta_5 & -c\theta_5 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (13)$$

$$T_5^6 = \begin{bmatrix} c\theta_6 & -s\theta_6 & 0 & 0.03 \\ 0 & c\theta_6 & -1 & -0.106 \\ c\theta_6 & c\theta_6 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (14)$$

$$T_6^7 = \begin{bmatrix} c\theta_7 & -s\theta_7 & 0 & 0 \\ c\theta_7 & c\theta_7 & 0 & 0 \\ 0 & c\theta_7 & 1 & 0.095 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (15)$$

$$T_7^0 = T_1^0 * T_2^1 * T_3^2 * T_4^3 * T_5^4 * T_6^5 * T_7^6 \quad (16)$$

$$R_y = \begin{bmatrix} \cos(-\pi/2) & 0 & \sin(\pi/2) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(-\pi/2) & 0 & \cos(\pi/2) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (17)$$

$$R_z = \begin{bmatrix} \cos(\pi) & -\sin(\pi) & 0 & 0 \\ \sin(\pi) & \cos(\pi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (18)$$

$$R_c = R_y * R_z \quad (19)$$

4.5. Inverse kinematics (IK)

Inverse kinematics analysis is for obtaining the Joint angles by using end effector Cartesian space or position coordinates. Since the last three joints J4, J5, J6 in Figure5 and Figure 6 are revolute joints with joint axis intersection at J5 which would be wrist center (WC). Thus the kinematic of the IK is now evaluated by calculating Inverse position and Inverse Orientation.

4.5.1. Inverse Position

The inverse position problem is for obtaining the first three joint angles ($\theta_1, \theta_2, \theta_3$). To evaluate Inverse Position the end effector position (Px, Py, Pz) and Orientation (Roll, Pitch, Yaw) need to be taken from the test case input data. Thus the rotation matrix for the end effector R_rpy shown (20). After the error correction the actual end effector rotation matrix R_end is given in (21). In (22) the obtained matrix R_6^0 is the rotation part of the full homogeneous transform matrix T_0^{EE} which is a transform matrix of end effector to the base joint. Using these concepts the joint angles ($\theta_1, \theta_2, \theta_3$) are obtained.

$$R_{rpy} = \text{Rot}(Z, \text{yaw}) * \text{Rot}(Y, \text{pitch}) * \text{Rot}(X, \text{Roll}) \quad (20)$$

$$R_{\text{end}} = R_{rpy} * R_c \quad (21)$$

$$T_0^{EE} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & p_x \\ r_{21} & r_{22} & r_{23} & p_y \\ r_{31} & r_{32} & r_{33} & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} R_6^0 & r_{0/EE}^0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (22)$$

4.5.2. Inverse Orientation

Inverse orientation problem is for obtaining the final three joint angles ($\theta_4, \theta_5, \theta_6$). The resultant transform is obtained using the individual DH transforms. Hence the resultant rotation is calculated as shown in (23). Since the overall the RPY (roll, pitch and yaw) rotation between base_link and gripper_link is equal to the product of individual rotations between the links as shown in (24). Multiply inverse matrix of R_3^0 ($\text{inv}(R_3^0)$) on either sides in (24) the resultant is shown in (25). The resultant matrix (25) on RHS (right hand side of the equation) does not have any variables after substituting the joint angles, and hence comparing LHS (left hand side of the equation) with RHS will result the equations for the last three joint angles ($\theta_4, \theta_5, \theta_6$). Solving the obtained equations which results the joint angles.

$$R_6^0 = R_1^0 * R_2^1 * R_3^2 * R_4^3 * R_5^4 * R_6^5 \quad (23)$$

$$R_{\text{end}} = R_6^0 \quad (24)$$

$$R_6^3 = \text{inv}(R_3^0) * R_{\text{end}} \quad (25)$$

5. Experiment and results

The robotic arm in the research is designed for the manipulation tasks for various applications such as Coconut crop harvesting, Assistive robots like self-driving wheel chairs and Home robots, Search and Rescue robots for disastrous environments, Collaborative robots for research use and etc. The evaluation of the design and software was first done on RVIZ simulation with few goal positions, these goal positions are the most desired positions of the end effector during the general operations. Once the software simulation has given satisfactory results, the testing was done on hardware which was already built. The designed 6 DOF robotic arm is simulated in RVIZ. An axis plugin has been added to visualize a coordinate axis at the base of the model robotic arm. The blue line indicates the z-axis, the red line indicates the x-axis and green line indicates the y-axis. The silver-coloured arm represents the goal pose of the arm for the user input values.

Figure 11 represents the different poses of the robotic arm when different coordinates are input to IK solver. Different x, y, and z coordinates are published to the interface module to manipulate the pose of the model arm. Table 3, 5 and 6 represents the outcomes of the IK solver when different goal points are input into the solver using different control methods on the RVIZ simulation platform. The output joint state values of all the goals positions in Figure 11 are given in Table 3, 5 and 6. The input point is considered with respect to the RVIZ [17] world axis frame. With the IK based approach, the robotic arm was manipulated at the desired points in the RVIZ world frame simulation. The three control modes have been tested on 16 different goal positions. The input data from the interface block enters the IK solver node where the required joint values are calculated to achieve the respective input pose. These generated joint values are passed to Joint state publisher and robot state publisher where the required Tf and combined joint state array will be passed to the RVIZ for the simulation.

On comparing the accuracy between the three methods keyboard and GUI were able to position the arm to the desired goal pose. Joystick failed in some cases as the control mode was difficult to manipulate the robotic arm. Since the Joystick PWM values need to be transferred to ROS via a microcontroller, due to the disturbances in the hardware, in a few cases joystick input readings varied. There was 1%-3% error in the input and hence the output was also affected due to the change in the input values. Comparing the results from Table 3, 4, 5 it is clear that GUI and Keyboard are two ideal methods to accurately position the robotic arm at a desired position.

The Figure 12 and Figure 13 illustrates the error percentage values while conducting a simulation experiment through a joystick. A survey has been conducted by testing the system with a 20 number of individuals. Based on the outputs that were obtained, an analysis has been made on the error percentage in reaching the goal position. The error is calculated by comparing the correct outcomes obtained by an individual to the total no of test cases in Figure 12. And an analysis has been made which showed the error percentage for each goal pose. This is calculated based on the pass

and fail cases obtained while conducting the survey by 20 individuals. The test results are plotted in Figure 13. Considering these errors, joystick control was avoided for the hardware test. The hardware test has been implemented using GUI as a control interface since the simulation results of keyboard based control and GUI based control are the same. Table 6 represents the outcomes of IK solver when implemented on hardware using GUI based control. It is observed that the output in the hardware testing using GUI based control is similar to the simulation testing using GUI.

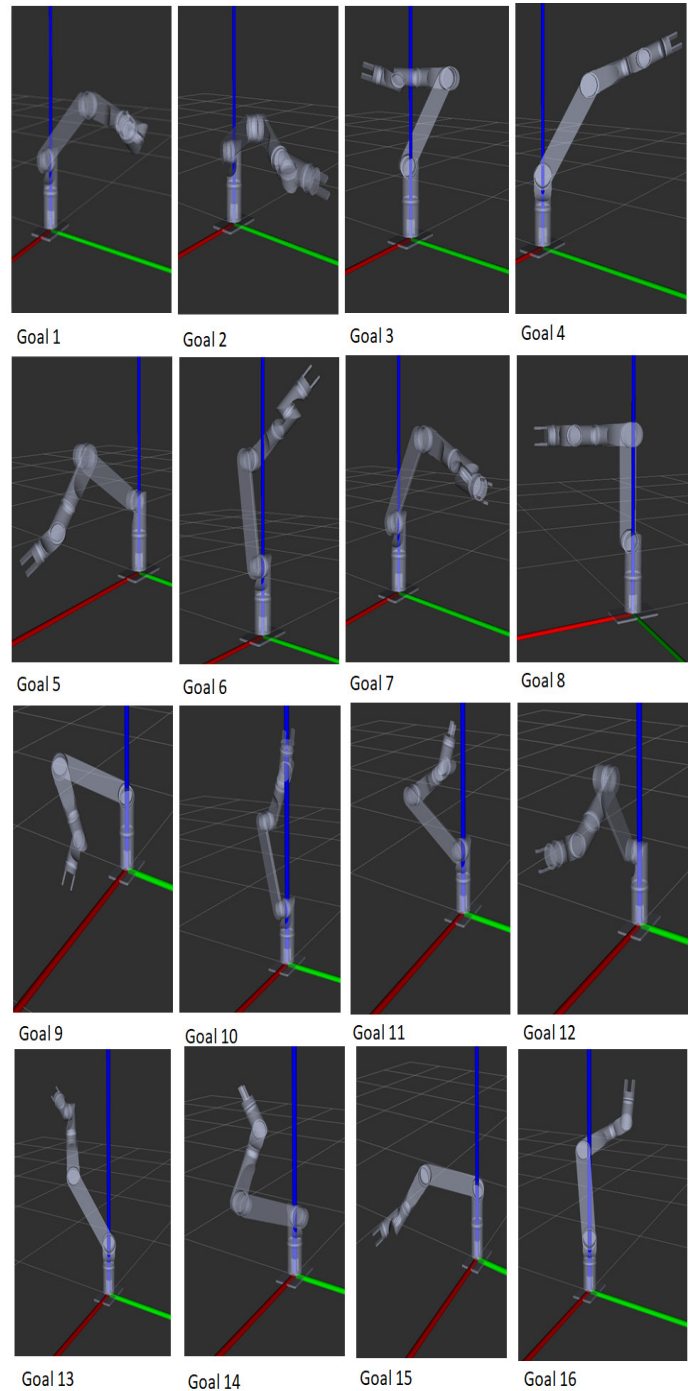


Figure 11: The goal pose of the robotic arm on different user inputs.

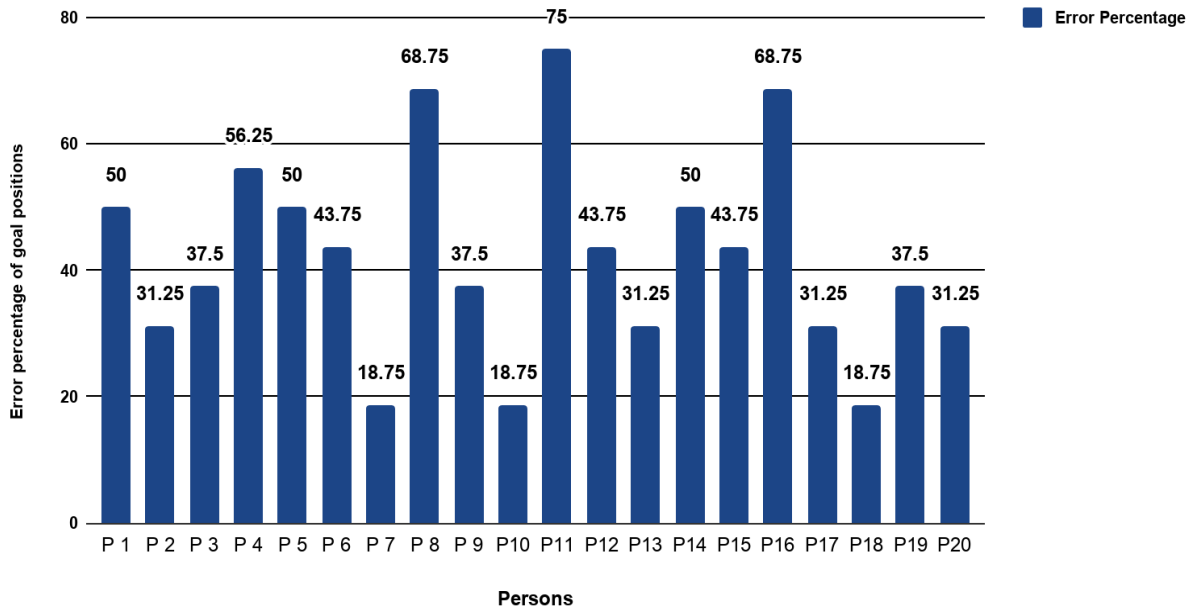


Figure 12: Error percentage for individual person

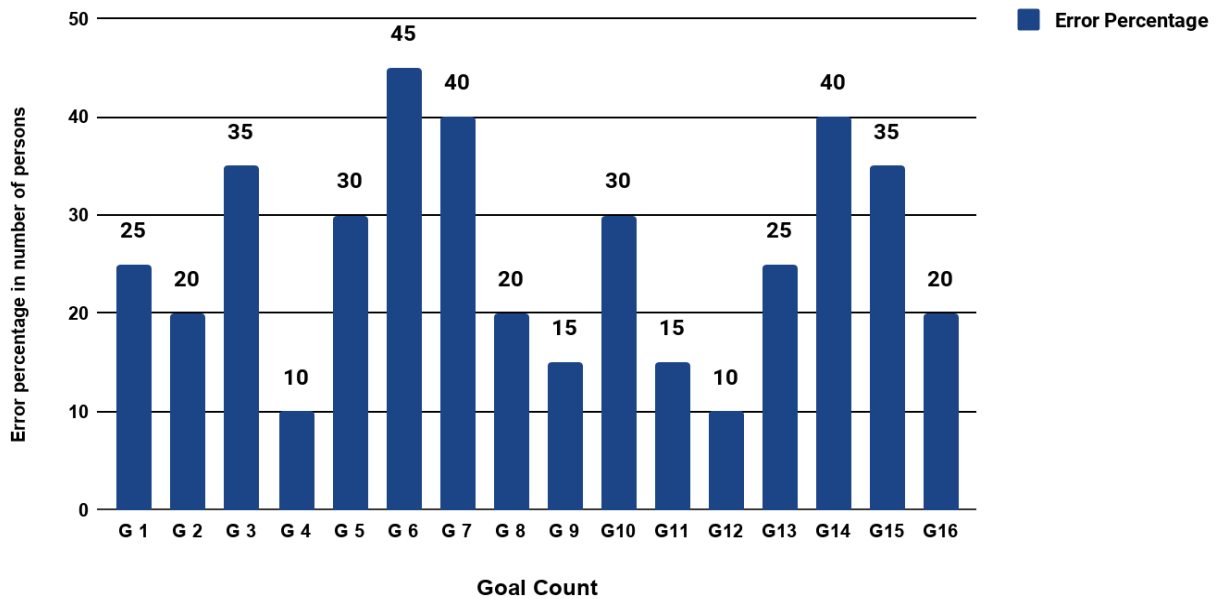


Figure 13 Error Percentage for individual goal position

Table 3: Outcomes of the joint state values from the IK solver node in simulation for the user input coordinates using GUI based control.

Goal Count	X position	Y position	Z position	DOF1 (radians)	DOF2 (radians)	DOF3 (radians)	DOF4 (radians)	DOF5 (radians)	DOF6 (radians)
Goal 1	0.301	0.651	0.525	1.2704	-0.9539	0.7554	-1.0660	1.3424	-1.185
Goal 2	0.402	0.730	0.376	1.0098	-1.222	0.9088	1.2786	0.476	1.4327
Goal 3	0.133	- 0.116	0.699	-0.5268	0.5868	-0.567	-0.6876	0.3067	0.0028
Goal 4	-0.392	0.532	0.825	2.2363	-0.6110	0.9545	0.043966	0.0151	0.6119
Goal 5	0.675	- 0.095	0.220	-0.0797	-1.0676	0.2312	-0.8217	0.4022	0.9090
Goal 6	- 0.019	0.242	1.020	1.4644	0.1755	0.3900	1.0424	0.5786	-1.297
Goal 7	0.204	0.644	0.529	1.3112	0.6556	0.3388	-0.9354	0.2892	1.14233
Goal 8	0.428	- 0.020	0.760	0	0	0	0	0	0

Goal 9	0.255	- 0.267	0.067	-0.89364	-1.23433	-0.70224	0.579329	0.924559	-0.41793
Goal 10	0.043	0.027	1.079	1.12537	0.424841	0.580440	-0.63930	0.2128	-0.99506
Goal11	- 0.017	- 0.110	0.759	1.396672	1.164312	-0.62972	0.222626	0.77854	0.505854
Goal 12	0.671	- 0.020	0.510	0.028259	-0.67949	0.223439	-0.42861	0.52486	0.095769
Goal 13	0.141	- 0.338	1.038	2.177904	0.639303	0.979159	1.05975	1.08808	-0.56990
Goal 14	0.281	- 0.123	0.844	2.831651	1.213924	-0.16744	0.57619	0.93674	-1.19414
Goal 15	0.303	- 0.716	0.108	-1.25223	-1.48647	0.71624	1.17593	0.725	-0.75893
Goal 16	-0.116	0.235	1.001	2.050420	0.0916879	0.3374000	0.156057999	1.03124	0.649038000

Table 4: Outcomes of the joint state values from the IK solver node in simulation for the user input coordinates using Keyboard.

Goal Count	X position	Y position	Z position	DOF1 (radians)	DOF2 (radians)	DOF3 (radians)	DOF4 (radians)	DOF5 (radians)	DOF6 (radians)
Goal 1	0.301	0.651	0.525	1.2704	-0.9539	0.7554	-1.0660	1.3424	-1.185
Goal 2	0.402	0.730	0.376	1.0098	-1.222	0.9088	1.2786	0.476	1.4327
Goal 3	0.133	- 0.116	0.699	-0.5268	0.5868	-0.567	-0.6876	0.3067	0.0028
Goal 4	-0.392	0.532	0.825	2.2363	-0.6110	0.9545	0.043966	0.0151	0.6119
Goal 5	0.675	- 0.095	0.220	-0.0797	-1.0676	0.2312	-0.8217	0.4022	0.9090
Goal 6	- 0.019	0.242	1.020	1.4644	0.1755	0.3900	1.0424	0.5786	-1.297
Goal 7	0.204	0.644	0.529	1.3112	0.6556	0.3388	-0.9354	0.2892	1.14233
Goal 8	0.428	- 0.020	0.760	0	0	0	0	0	0
Goal 9	0.255	- 0.267	0.067	-0.89364	-1.23433	-0.70224	0.579329	0.924559	-0.41793
Goal 10	0.043	0.027	1.079	1.12537	0.424841	0.580440	-0.63930	0.2128	-0.99506
Goal 11	- 0.017	- 0.110	0.759	1.396672	1.164312	-0.62972	0.222626	0.77854	0.505854
Goal 12	0.671	- 0.020	0.510	0.028259	-0.67949	0.223439	-0.42861	0.52486	0.095769
Goal 13	0.141	- 0.338	1.038	2.177904	0.639303	0.979159	1.05975	1.08808	-0.56990
Goal14	0.281	- 0.123	0.844	2.831651	1.213924	-0.16744	0.57619	0.93674	-1.19414
Goal 15	0.303	- 0.716	0.108	-1.25223	-1.48647	0.71624	1.17593	0.725	-0.75893
Goal16	-0.116	0.235	1.001	2.050420	0.0916879	0.3374000	0.156057999	1.03124	0.649038000

Table 5: Outcomes of the joint state values from the IK solver node in simulation for the user input coordinates using Joystick.

Goal Count	X position	Y position	Z position	DOF1 (radians)	DOF2 (radians)	DOF3 (radians)	DOF4 (radians)	DOF5 (radians)	DOF6 (radians)
Goal 1	0.305	0.652	0.525	1.2704	-0.9540	0.7556	-1.0660	1.3424	-1.185
Goal 2	0.406	0.7304	0.376	1.0098	-1.230	0.9089	1.2786	0.476	1.4327
Goal 3	0.135	- 0.116	0.699	-0.5268	0.5871	-0.567	-0.6876	0.3067	0.0028
Goal 4	-0.394	0.530	0.825	2.2363	-0.6116	0.9548	0.043966	0.0151	0.6119
Goal 5	0.67	- 0.095	0.220	-0.0797	-1.0681	0.2312	-0.8217	0.4022	0.9090
Goal 6	- 0.015	0.242	1.020	1.4644	0.1755	0.3900	1.0424	0.5786	-1.297
Goal 7	0.209	0.645	0.529	1.3112	0.6558	0.3389	-0.9354	0.2892	1.14233
Goal 8	0.421	- 0.020	0.760	0	0	0	0	0	0
Goal 9	0.255	- 0.267	0.067	-0.89364	-1.23433	-0.70224	0.579329	0.924559	-0.41793
Goal 10	0.043	0.028	1.079	1.12537	0.424841	0.58048	-0.63930	0.2128	-0.99506
Goal11	- 0.017	- 0.110	0.759	1.396672	1.164312	-0.62972	0.222626	0.77854	0.505854
Goal 12	0.671	- 0.020	0.510	0.028259	-0.67949	0.223439	-0.42861	0.52486	0.095769
Goal 13	0.141	- 0.338	1.038	2.177904	0.639303	0.979159	1.05975	1.08808	-0.56990
Goal14	0.281	- 0.123	0.844	2.831651	1.213924	-0.16744	0.57619	0.93674	-1.19414
Goal 15	0.303	- 0.716	0.108	-1.25223	-1.48647	0.71624	1.17593	0.725	-0.75893
Goal16	-0.116	0.234	1.001	2.050420	0.0916879	0.3375	0.156057999	1.03124	0.649038000

Table 6: Outcomes of the joint state values from the IK solver node on hardware testing for the user input coordinates using GUI

Goal Count	X position	Y position	Z position	DOF1 (radians)	DOF2 (radians)	DOF3 (radians)	DOF4 (radians)	DOF5 (radians)	DOF6 (radians)
Goal 1	0.301	0.651	0.525	1.2799	-0.9538	0.7600	-1.0659	1.3426	-1.184
Goal 2	0.402	0.730	0.376	1.0095	-1.222	0.9086	1.2786	0.480	1.4327
Goal 3	0.133	- 0.116	0.699	-0.5268	0.5868	-0.567	-0.6876	0.3067	0.0028
Goal 4	-0.392	0.532	0.825	2.2363	-0.611015	0.95467	0.0439666	0.01561	0.6119
Goal 5	0.675	- 0.095	0.220	-0.07978	-1.0676	0.2312	-0.82172	0.4022	0.9090
Goal 6	- 0.019	0.242	1.020	1.464454	0.1755	0.3900	1.0424	0.57862	-1.29711
Goal 7	0.204	0.644	0.529	1.3112	0.6556	0.3388	-0.9354	0.2892	1.14233
Goal 8	0.428	- 0.020	0.760	0	0.	0	0	0	0
Goal 9	0.255	- 0.267	0.067	-0.89364	-1.234335	-0.702246	0.579329	0.924559	-0.41793
Goal 10	0.043	0.027	1.079	1.12537	0.424841	0.5804405	-0.63930	0.21285	-0.99506
Goal11	- 0.017	- 0.110	0.759	1.3966725	1.1643121	-0.629728	0.2226262	0.77854	0.505854
Goal 12	0.671	- 0.020	0.510	0.028259	-0.67949	0.223439	-0.42861	0.52486	0.095769
Goal 13	0.141	- 0.338	1.038	2.177904	0.639303	0.97915954	1.0597566	1.088084	-0.56990
Goal14	0.281	- 0.123	0.844	2.831651	1.2139244	-0.16744	0.576198	0.93674	-1.19414584
Goal 15	0.303	- 0.716	0.108	-1.252234	-1.48647	0.7162445	1.175932	0.72512	-0.75893
Goal 16	-0.116	0.235	1.001	2.050420	0.0916879	0.3374000	0.156057999	1.03124	0.649038000

6. Conclusion

In this research work, the authors proposed and evaluated reliable methods for controlling a robotic arm by testing it in both hardware and simulation using RVIZ. The authors in [17] used the MoveIt to build a kinematics library for the IK of the robotic arm, but as an extension of that, this research designed and developed a kinematics-based solution using DH parameters method. Using the derived equations, successfully tested the designed robotic arm in Rviz using the proposed control mechanisms. A survey is conducted in evaluating the best control methodology where Figure [12] and Figure [13] depict the results. Based on the survey, the research validates that the joystick failed in achieving the desired input coordinate positions due to signaling and hardware issues. But GUI and Keyboard showed better results in controlling the arm.

As a rule, simulations do not reproduce the exact real-time behavior of an entity or a system. PID based control can reduce the error. The research can be enhanced by testing the proposed design in the Gazebo simulation software. Gazebo provides the flexibility to use a PID-based controller, which helps in smooth and exact mimicking of the real-time robotic arm as per simulation in RVIZ. The proposed testing gives the developer good results. Design of the end effector can be improved for performance of multiple, divergent tasks in a real time environment. The singularity issues can be reduced for better performance and enhancement of the task.

Acknowledgment

We are grateful for Amrita Vishwa Vidyapeetham and Humanitarian and Technology Labs for providing us with all the sophisticated requirements to develop and complete this paper.

References

- [1] Megalingam R.K, Sivanantham V, Kumar K.S, Ghanta S, Teja P.S, Gangireddy R, Sakti Prasad K.M, Gedela V.V, “Design and development of inverse kinematic based 6 DOF robotic arm using ROS”, International Journal of Pure and Applied Mathematics, 118, pp. 2597-2603, ISSN: 1314-3395
- [2] Xinquan Liang, Haris Cheong, Yi Sun, Jin Guo, Chee Kong Chui, Chen-Hua Yeo, “Design, Characterization, and Implementation of a Two-DOF Fabric-Based Soft Robotic Arm”, IEEE Robotics and Automation letters(Volume:3,Issue:3,July2018),<https://doi.org/10.1109/LRA.2018.2831723>
- [3] Asghar Khan, Wang Li Quan. “Structure design and workspace calculation of 6-DOF underwater manipulator”. 2017 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST), 10-14 Jan 2014, <https://doi.org/10.1109/IBCAST.2017.7868119>
- [4] Elkin Yesid Veslin; Max Suell Dutra; Omar Lengerke; Edith Alejandra Carreño; Magda Judith Morales, “A Hybrid Solution for the Inverse Kinematic on a Seven DOF Robotic Manipulator”, IEEE Latin America Transactions (Volume: 12, Issue: 2, March 2014), <https://doi.org/10.1109/TLA.2014.6749540>
- [5] Pietro Falco; Ciro Natal, “On the Stability of Closed-Loop Inverse Kinematics Algorithms for Redundant Robots”, 05 May 2011, IEEE Transactions on Robotics (Volume: 27, Issue: 4, Aug. 2011),<https://doi.org/10.1109/TRO.2011.2135210>
- [6] Seo-Wook Park, Jun-Ho Oh, “Hardware Realization of Inverse Kinematics for Robot Manipulators”, IEEE Transactions on Industrial Electronics Vol.41, No 1, February 1994, <https://doi.org/10.1109/41.281607>
- [7] Rajesh Kannan Megalingam, Nigam Katta, Raviteja Geesala, Prasant Kumar Yadav, “Keyboard Based Control and Simulation of 6 DOF Robotic Arm using ROS”, 2018 4th International Conference on Computing Communication and Automation(ICCA), <https://doi.org/10.1109/CCAA.2018.8777568>
- [8] Wei Qian, Ziyang Xia, Jing Xiong, Yangzhou Gan, Yangchow Guo, Shaokui Weng, Hao Deng, Ying Hu, Jiawei Zhang, “Manipulation Task simulation using ROS and Gazebo”, Robotics and Biomimetics (ROBIO), 2014 IEEE International Conference”, <https://doi.org/10.1109/ROBIO.2014.7090732>
- [9] Žlajpah Leon, Simulation in robotics (2008) Mathematics and Computers in Simulation, 79 (4), 15 December 2008, pp. 879-897, <https://doi.org/10.1016/j.matcom.2008.02.017>
- [10] Kichang Lee ; Jiyoung Lee ; Bungchul Woo ; Jeongwook Lee ; Young-Jin Lee, Kimhae-si ; Syungkwon Ra “Modeling and Control of an Articulated Robot Arm with Embedded Joint Actuators ”,2018 International Conference on Information and Communication Technology Robotics (ICT-ROBOT), <https://doi.org/10.1109/ICT-ROBOT.2018.8549903>
- [11] Alla N Barakat, Khaled A. Gouda, Kenza.Bozed, “ Kinematics analysis and simulation of a robotic arm using MATLAB”,2016 4th International Conference on Control Engineering & Information Technology (CEIT),16-18 Dec. 2016, Hammamet, Tunisia, <https://doi.org/10.1109/CEIT.2016.7929032>

- [12] Stefano Carpin, Mike Lewis, Jijun Wang, "USARSim: a robot simulator for research and education" Proceedings 2007 IEEE International Conference on Robotics and Automation, 10-14 April 2007, Roma Italy, <https://doi.org/10.1109/ROBOT.2007.363180>
- [13] J Jun-Di Sun; Guang-Zhong Cao; Wen-Bo Li; Yu-Xin Liang; Su-Dan Huang, "Analytical inverse kinematic solution using the D-H method for a 6-DOF robot", 2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), <https://doi.org/10.1109/URAI.2017.7992807>
- [14] Akhilesh Kumar Mishra; Oscar Meruvia-Pastor; " Robot arm manipulation using depth-sensing cameras and inverse kinematics "2014 Oceans-St John's,14-19 September 2014 IEEE International Conference, St. John's, NL, Canada, <https://doi.org/10.1109/OCEANS.2014.7003029>
- [15] C. A. G. Gutiérrez; J. R. Reséndiz; J. D. M. Santibáñez; G. M. Bobadilla, "A Model and Simulation of a Five-Degree-of-Freedom Robotic Arm for Mechatronic Courses", IEEE Latin America Transactions (Volume: 12, Issue: 2, March 2014),<https://doi.org/10.1109/TLA.2014.6749521>
- [16] Zenghua Bian ; Zhengmao Ye ; Weilei Mu," Kinematic analysis and simulation of 6-DOF industrial robot capable of picking up die-casting products", 2016 IEEE International Conference on Aircraft Utility Systems (AUS), <https://doi.org/10.1109/AUS.2016.7748017>
- [17] S. Hernandez-Mendez, C. Maldonado-Mendez, A. Marin-Hernandez, H. V. Rios-Figureueroa, H. Vazquez-Leal and E. R. Palacios-Hernandez, "Design and implementation of a robotic arm using ROS and MoveIt!", 2017 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), Ixtapa, 2017, pp. 1-6, <http://doi.org/10.1109/ROPEC.2017.8261666>
- [18] Weimin Shen, Jason Gu, Yide Ma, "3D Kinematic Simulation for PA10-7C Robot Arm Based on VRML", 18-21 Aug. 2007, Jinan, China. <https://doi.org/10.1109/ICAL.2007.4338637>
- [19] José de Jesús Rubio, "Sliding mode control of robotic arms with dead zone", IET Control Theory & Applications (Volume: 11, Issue: 8, 5 12 2017), <https://doi.org/10.1049/iet-cta.2016.0306>
- [20] Megalingam R.K, Rajesh Gangireddy, Gone Sriteja, Ashwin Kashyap, Apuroop Sai Ganesh "Adding intelligence to the robotic coconut tree climber" 2017 International Conference on Inventive Computing and Informatics (ICICI), 23-24 Nov. 2017, Coimbatore India, <https://doi.org/10.1109/ICICI.2017.8365206>
- [21] Teerawat Thepmanee, Jettiya Sripituk, Prapart Ukakimapun, "A simple technique to modeling and simulation four-axe robot-arm control" 17-20 Oct. 2007, Coex, Seoul, Korea, <https://doi.org/10.1109/ICCAS.2007.4406694>
- [22] Tingting Meng; Wei He, "Iterative Learning Control of a Robotic Arm Experiment Platform with Input Constraint", IEEE Transactions on Industrial Electronics (Volume: 65, Issue: 1, Jan. 2018), <https://doi.org/10.1109/TIE.2017.2719598>
- [23] Gourab Sen Gupta, Subhas Chandra Mukhopadhyay, Christopher H. Messom, Serge N. Demidenko, "Master-Slave Control of a Teleoperated Anthropomorphic Robotic Arm With Gripping Force Sensing ", IEEE Transactions on Instrumentation and Measurement (Volume: 55, Issue: 6, Dec. 2006),<https://doi.org/10.1109/TIM.2006.884393>
- [24] P. Lucibello," Repositioning control of robotic arms by learning ", IEEE Transactions on Automatic Control (Volume: 39, Issue: 8, Aug 1994), <https://doi.org/10.1109/9.310053>
- [25] Qingyuan Sun; Lingcheng Kong; Zhihua Zhang; Tao Mei, "Design of wireless sensor network node monitoring interface based on Qt", 2010 International Conference on Future Information Technology and Management Engineering, Changzhou, 2010, pp. 127-130
- [26] R. Takuma; Y. Asahara ; H. Kajimoto ; N. Kawakami ; S. Tachi, "Development of anthropomorphic multi-D.O.F master-slave arm for mutual teleexistence ", IEEE Transactions on Visualization and Computer Graphics (Volume: 11 , Issue: 6 , Nov.-Dec. 2005), <https://doi.org/10.1109/TVCG.2005.99>
- [27] Ramesh, S. B. Hussain and F. Kangal, "Design of a 3 DOF robotic arm," 2016 Sixth International Conference on Innovative Computing Technology (INTECH), Dublin, 2016, pp. 145-149, <https://doi.org/10.1109/INTECH.2016.7845007>
- [28] Jose de Jesus Rubio; Adrian Gustavo Bravo; Jaime Pacheco; Carlos Aguilar;"Passivity analysis and modeling of robotic arms", IEEE Latin America Transactions (Volume: 12 , Issue: 8 , Dec. 2014),<https://doi.org/10.1109/TLA.2014.7014505>

A Hybrid Approach for Intrusion Detection using Integrated K-Means based ANN with PSO Optimization

Jesuretnam Josemila Baby*, James Rose Jeba

Noorul Islam Centre for Higher Education, Department of Computer Applications, Noorul Islam University, Kumaracoil, 629180, India

ARTICLE INFO

Article history:

Received: 04 February, 2020

Accepted: 24 May, 2020

Online: 29 May, 2020

Keywords:

Particle Swarm Optimization

Intrusion Detection System

Artificial Neural Networks

ABSTRACT

Many advances in computer systems and IT infrastructures increases the risks associated with the use of these technologies. Specifically, intrusion into computer systems by unauthorized users is a growing problem and it is very challenging to detect. Intrusion detection technologies are therefore becoming extremely important to improve the overall security of computer systems. In the past decades, most of the intrusion detection systems designed suffer from the problem of high false negative and low efficiency rate. A powerful intrusion detection system (IDS) should be implemented to solve these issues and it is necessary to collect, reduce and analysis the data automatically. The integration of machine learning and artificial intelligence techniques serves this purpose in this paper. A use of particle swarm optimization (PSO) selects the optimal number of clusters and the integration of k-means based artificial neural network (ANN) achieves maximum efficiency when the number of clusters selected optimally. The proposed IDS are t bested with NSL-KD dataset and the experiment result shows the significance of the proposed IDS.

1. Introduction

Due to the advancement of computer and communication technology, the reliance on Internet and worldwide connectivity, damages caused by unexpected intrusions. These crimes related to computer systems have been increased rapidly; a computer system should provide confidentiality, integrity and availability against denial of service; therefore, it is very important that the security mechanisms of systems be designed to prevent unauthorized access to system resources and data [1]. Firewalls are hardware or software systems placed in between two or more computer networks to stop the committed attacks by isolating these networks using the rules and policies determined for them [2]. However, it is very clear that firewalls are not enough to secure a network completely because the attacks committed from outside of the network are stopped whereas inside attacks are not [3]. This is the situation where intrusions detection systems (IDSs) are in charge.

Intrusion detection Systems (IDSs) is a software or device that helps to resist network attacks [4]. The goal of IDS is to have defense wall, which does not allow such types of attacks [5]. It detects unauthorized activities of a computer system or a

network. IDS are an active and secure technology there are two categories of intrusion detection system [6]. Anomaly detection system creates a database of normal behavior and any deviations from the normal behavior are occurred an alert is triggered regarding the occurrence of intrusions [7]. Misuse Detection system stores the Predefined attack patterns in the database if a similar data and if similar situations occur it is classified as attack. Based on the source of data the intrusion detection system is classified to Host based IDS and Network based IDS [8]. In network-based IDS the individual packet flowing through the network are analyzed [9]. The host-based IDS analyzes the activities on the single computer or host [10]. The main disadvantage of the misuse detection (signature detection) method is that it cannot detect novel attacks and variation of known attacks [11]. To avoid these drawbacks, we proposed anomaly-based detection methods.

Most unsupervised anomaly detection methods are established on two basic assumptions about data [12]. First, the number of normal instances vastly outnumbers that of anomalies. Second, data instances of the same classification (type of attack or normal) should be close to each other in the feature space under some reasonable metrics, and instances of different classifications are far apart [13]. Data mining technique is used to find the interesting rules from a large database depending

* Jesuretnam Josemila Baby, Email: josemilakissinger@gmail.com

upon the user defined support and confidence [14]. It will be useful for the decision maker to differentiate between data as useful or irrelevant [15]. Clustering is the unsupervised classification of input items into groups (clusters) without any prior knowledge [16].

Although many kinds of clustering methods, such as Fuzzy C-Means (FCM), K-means, are widely used in intrusion detection, few clustering algorithms guarantee a global optimal solution [17]. Based on this intention we developed a new anomaly intrusion detection mechanism employed with multi-dimensional hierarchical k-means algorithm in this research work in order to overcome all the above issues [18]. To integrate K-means and Artificial Neural Network (ANN) technique with optimization to propose an efficient IDS with High True Positive Rate (TPR) and Low False Negative Rate (FNR) [19].

The organization of the paper is as follows: In Section 2, discuss about the related work with emphasis on various methods and frameworks used for intrusion detection. Section 3 covers the Integrated K-Means based ANN with PSO Optimization algorithm. Section 4 presents the experimental results and comparison of the proposed method with other approaches. It is observed that the proposed system. Section 5 gives some conclusions.

2. Related Work

Sharma, Ruby, and Sandeep Chaurasia (2018) [20] proposed an Intrusion Detection System based on the density maximization-based fuzzy c-means clustering (DM-FCC). In that approach, cluster efficiency was improved through a membership matrix generation (MMG) algorithm. Dissimilarity Distance Function (DDF) has been used to compute the distance metric while creating a cluster in proposing IDS. The proposed enhanced fuzzy c-means algorithm has been tested up on ADFA Dataset and the model performs highly appreciable in terms of accuracy, precision, detection rates, and false alarms.

Chung, Yuk Ying, and Noorhaniza Wahid (2012) [21] proposed a new hybrid intrusion detection system by using intelligent dynamics warm based roughest (IDS-RS) for feature selection and simplified swarm optimization for intrusion data classification. IDS-RS was proposed to select the most relevant features that can represent the pattern of the network traffic. In order to improve the performance of SSO classifier, a new weighted local search (WLS) strategy incorporated in SSO was proposed. The purpose of this new local search strategy was to discover the better solution from the neighborhood of the current solution produced by SSO.

Thaseen, Ikram Sumaiya, and Cherukuri Aswani Kumar (2017) [22] proposed an intrusion detection model using chi-square feature selection and multi class support vector machine (SVM). A parameter tuning technique was adopted for optimization of Radial Basis Function kernel parameter namely gamma represented by ' γ ' and over fitting constant ' C '. These are the two important parameters required for the SVM model. The main idea behind this model was to construct a multi class SVM which has not been adopted for IDS so far to decrease the training and testing time and increase the individual classification

accuracy of the network attacks. The investigational results on NSL-KDD dataset which was an enhanced version of KDD Cup 1999 dataset shows that the proposed approach results in a better detection rate and reduced false alarm rate.

Çavuşoğlu, Ünal (2019) [23] developed a hybrid and layered Intrusion Detection System (IDS) that uses a combination of different machine learning and feature selection techniques to provide high performance intrusion detection in different attack types. In the developed system, initially, data preprocessing was performed on the NSL-KDD dataset, then by using different feature selection algorithms, the size of the dataset was reduced. Two new approaches have been proposed for feature selection operation. The layered architecture was created by determining appropriate machine learning algorithms according to attack type. Performance tests such as accuracy, DR, TP Rate, FP Rate, F-Measure, MCC and time of the proposed system are performed on the NSL-KDD dataset.

Aswani, Reema et al., (20-17) [24] introduced a hybrid artificial bee colony approach integrated with k-nearest neighbors to identify and segregate buzz in Twitter. A set of metrics comprising of created discussions, increase in authors, attention level, burstiness level, contribution sparseness, author interaction, author count and average length of discussions are used to model the buzz. The proposed approach considers the buzz discussions as outliers deviating from the normal discussions and identifies the same using the proposed hybrid bio inspired approach. Findings may be useful in domains like e-commerce, digital and influencer marketing to explore the factors that might create buzz along with the difference between the impact of buzz and normal discussions on the consumers.

3. Intrusion Detection Model

Over the last two decades, computer threats and cybercrimes have proliferated at the disadvantage of the general public, and newer threats are introduced each day that compromise the integrity, validity and confidentiality of data. Malicious activities in the internet are also known as intrusion.

Intrusion detection system (IDS) is software and hardware deployed to carry out the process of detecting unauthorized use of, or attack upon, a computer or a telecommunications network which is supposed to bridge the gaps in firewall and anti-viruses. An IDS provides monitoring [25] and analysis of user and system activity, can audit system configuration and vulnerabilities, assess the integrity of critical system and data files, provide statistical analysis of activity patterns based on the matching with known attacks, analyze abnormal activity, and operate system audit.

One advantage of the IDS is its ability to document the intrusion or threat to an organization, thereby providing bases for informing the public regarding the latest attack patterns through system logs. The proposed IDS model is the integration of K means-based ANN and PSO. Initially the features of the intruded networks are extracted from a benchmark dataset and it was trained to the proposed classifier. But it is difficult to find the number of clusters and the detection accuracy will be maximum if the number of clusters equal to the number of data types in the

dataset. This problem is formulated as a multi-objective function in PSO which optimally finds the cluster numbers that maximizes the detection accuracy and lower the false negatives.

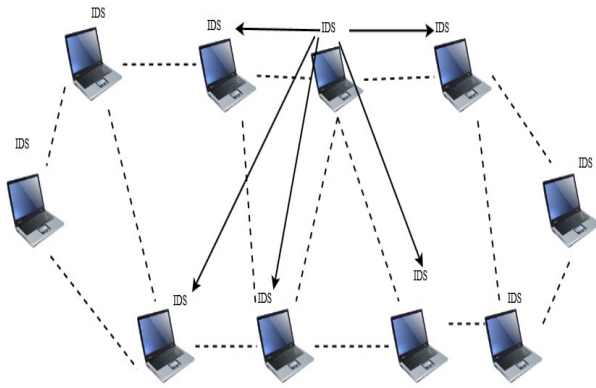


Figure 1: The IDS Architecture for Wireless Ad-Hoc Network

4. Proposed IDS Model

The study integrates artificial intelligence and machine-learning techniques with k-means data mining algorithm to develop an IDS [26] model with higher efficiency and lower false negatives. A common problem shared by current IDS is the high false positives and low detection rate. In the proposed work, we integrate the advantage of Artificial intelligence and machine learning techniques to overcome this issue. The proposed IDS model uses K-means algorithm based Artificial Neural Network (ANN) [27] integrated with Particle swarm optimization (PSO) algorithm [28] to increase the efficiency rate.

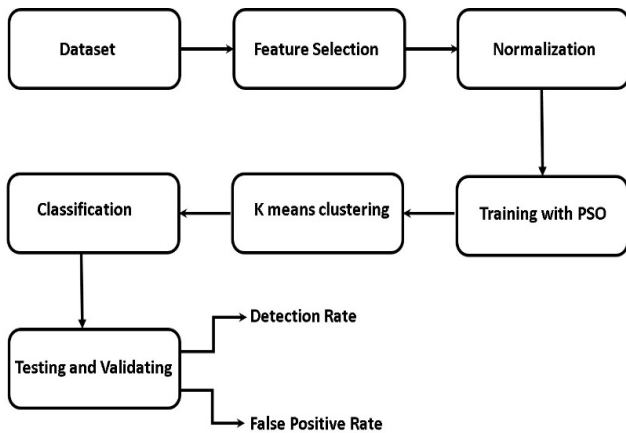


Figure 2: Proposed Block Diagram

4.1. Pre-processing

Pre-processing involves cleaning the data of inconsistencies and/or noise and combining or removing redundant entries. Pre-processing also involves converting the attributes of the dataset into numeric data and saving in a format readable because k-means works only on numerical data.

4.2. Feature Extraction [29]

After pre-processing the initial data, the useful information's or features from each data are extracted. We employed the simple features (attributes) that are extracted from the header's area of the selected network packets. These intrinsic features are

available in many networks, for example, the duration (length of the connection), source host, destination host, source interface, and destination interface. We also used three features in each 2 seconds time interval:

- Total number of packets sent from and to the given interface in the considered time interval,
- Total number of bytes sent from and to the given interface in the considered time interval,
- Number of different source- destination pairs matching the given hostname-interface that are observed in the considered time interval.
- The number of packets and bytes allows to detect anomalies in traffic volume, and the third features allows detecting the network and the interface scans as well as the distributed attacks, which both result in an increased number of source-destination pairs.
- An efficient classifier further utilizes the extracted features in order to detect Intrusion present in the network.

4.3. K-Means Clustering Algorithm

Clustering is the method of grouping objects into meaningful subclasses so that the members from the same cluster are quite similar, and the members from different clusters are quite different from each other. Until now, the clustering algorithms can be categorized into four main groups partitioning algorithm, hierarchical algorithm, density-based algorithm and grid-based algorithm. Partitioning algorithms construct a partition of a database of N objects into a set of K clusters. Usually they start with an initial partition and then use an iterative control strategy to optimize an objective function. K-means represents a type of useful clustering techniques by competitive learning, which is also proved promising techniques in intrusion detection. K-Means is one of the simplest unsupervised learning algorithms that solve the clustering problem. The objective is to classify a given data set into a certain number of clusters (assume initial clusters) fixed a priori.

Algorithm 1 The pseudo code for the adapted K-Means clustering

1. Choose random k data points as the initial Cluster Centroids.
 2. Repeat
 3. For each data point x from D
 4. Then compute the distance of x from each cluster mean (centroid)
 5. Assign x to the nearest cluster.
 6. End for loop.
 7. Again compute the mean for current cluster collections.
 8. Until reaching stable cluster
 9. Use these centroids for normal and anomaly traffic.
 10. Calculate the distance of centroid from normal and anomaly centroid points.
 11. If $\text{distance}(d, D_m) \geq 5$
 12. Then anomaly found, exit 2.
 13. Else 3.
 14. d is a normal and it is not an Intrusion;
-

K-means clustering module can be summarized as follows:

$$K = \sum_{m=1}^x \sum_{n=1}^y \|d_n^{(m)} - c_m\| \quad (1)$$

$\|d_n^{(m)} - c_m\|$ is a chosen distance measure between a data point and the cluster centroid, is an indicator of the distance of the n data points from their respective cluster Centroids. In order to apply the K-means algorithm to intrusion detection system, we design and realize the K-means algorithm analyze module, the graph shows the process flow:

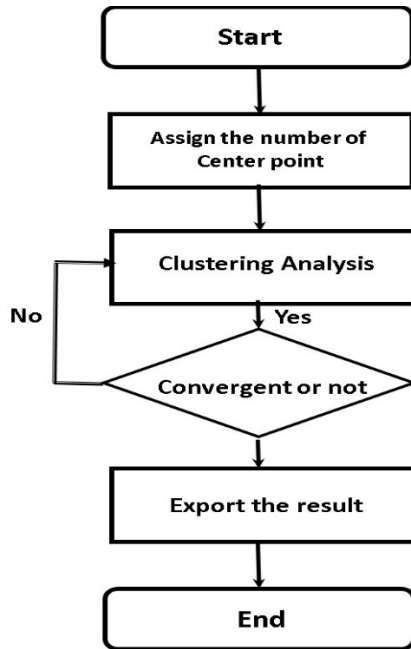


Figure 3: Working flow for K- means Algorithm

4.4. Review of Artificial Neural Network [30]

Feed forward neural network training is usually carried out using the called back propagation algorithm. Training the network with back propagation algorithm results in a non-linear mapping between the input and output variables. Thus, given the input/output pairs, the network can have its weights adjusted by the back propagation algorithm to capture the non-linear relationship. After training, the networks with fixed weights can provide the output for the training the network is based on the minimization of an energy function representing the instantaneous error.

ANN analysis was carried out using software Easy NN version 8.01. The network software uses back propagation algorithm and logistic function as activation function. The ANN used has three layers: an input layer that consists of five nodes (variables), one hidden layer consisting of five hidden nodes and an output layer that has one output node. To train an ANN model, a set of data containing input nodes and output nodes are fed. Once the training is over, ANN is capable of predicting the output when any input similar to the pattern that it has learned is fed. The ANN is tested for the remaining set of experimental data. The learning rate and momentum value of the network is set to optimize with a targeted error value of 0.05.

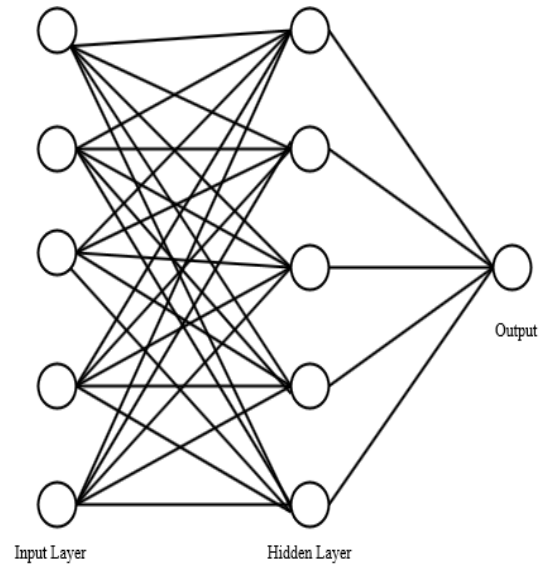


Figure 4: Schematic diagram of artificial neural network

4.5. K-Means based ANN-PSO

For implementation of ANN algorithm on the dataset, training data is divided into several subsets using k-means clustering technique. Subsequently, it trains different ANN using different subsets. Then, it determines membership grades of all these subsets and combines them using a new ANN to get final results. The whole framework of K-Means ANN is illustrated in Fig. 1. As typical machine learning framework, K-Means ANN incorporates both the training phase and testing phase.

Algorithm 2 The pseudo code for the adapted K-Means ANN

1. Start
2. Get the input features and Initialize the neural network parameters.
3. Define the relationships between input and output.
4. Divide the input for testing and training.
5. Train ANN using the training set.
6. Execute the Neural Network for the testing set.
7. Set $j=1$;
8. While($j < x$)
9. Cluster the ANN outputs into K-clusters.
10. Calculate the weight of each cluster.
11. Calculate the error between the target output and ANN estimated output.
12. Increase j by 1.
13. Calculate false negatives.
14. End

The results of K-means ANN is optimized using PSO. In PSO, each particle is a point of N -dimensional solution space and has a speed($N - dimensional vector$). Different particle has individual fitness associated with objective function. Each particle adjusts their flight path according to its flying experience and flying experience of group and move closer to optimal point. The position of i -particle is denoted as, $X_i = (x_{i1}, x_{i2} \dots x_{iN})$ Flight speed is denoted as: $V_i = (v_{i1}, v_{i2} \dots v_{iN})$ the best position which i particle passed is denoted as $P_{ibest} = (p_{i1}, p_{i2} \dots p_{iN})$ the

groups' best position which it can get is denoted as $G_{best} = (g_1, g_2 \dots g_N)$. Particle Swarm has two primary operators: Velocity update and Position update. During each generation each particle is accelerated toward the particles previous best position and the global best position. At each iteration a new velocity value for each particle is calculated based on its current velocity, the distance from its previous best position, and the distance from the global best position. The new velocity value is then used to calculate the next position of the particle in the search space. This process is then iterated a set number of times or until a minimum error is achieved. In each step, according to PSO algorithm formula, which is proposed by Kennedy, particles update their velocity and position according to the following formula:

$$V_i(t) = wV_i(t - 1) + C_1r_1(P_i - X_i(t - 1)) + C_2r_2(G - X_i(t - 1)) \quad (2)$$

$$X_i(t) = X_i(t - 1) + V_i(t - 1) \quad (3)$$

C_1 And C_2 denote accelerating factor. According to the experience of PSO algorithm, they are usually set $C_1 = C_2 = 2$. r_1 and r_2 are two random number between zero and one, w is called inertia weight. Researchers often use a constant V max to limit the speed of particles and improve search results. w plays a role which balance global search ability and local search ability. It is essential for the success of the algorithm. Shi and Eberhart, study on the effect of the w for optimize performance. They found that the larger the w is, the more easily escape from local minima, and the smaller the w is, the more favorably algorithm converges. Then they present a method, which makes inertia weight decrease linearly according to number of iterations.

In the beginning algorithm uses large inertia weight, it has a strong overall search capability. The later smaller inertia weight is used, and local search ability is improved. w is calculated as follows:

$$w = (w_1 - w_2) * \frac{Max_i - 1}{Max_i} + w_2 \quad (4)$$

w_1 and w_2 are the initial value and final value of inertia weight. Max_i and i are the maximum number of iterations and the current number of iterations for the algorithm, w reduces from 0.9 to 0.4 with the conduct of iteration. The objective function for PSO is set as

$$f(x) = \max\left(\frac{Detection\ Accuracy}{False\ Negatives}\right) + \max(1/Time) \quad (5)$$

The PSO algorithm proceeds as follows:

Algorithm 3 The pseudo code for the adapted PSO

1. Begin
2. Initialize the total number, velocity and position of the particles.
3. Calculate pbest and gbest.
4. Calculate the objective function.
5. Update the velocity and position of the particle.
6. Update pbest and gbest .
7. Repeat the steps until the termination condition reached.
8. End

5. Performance Analysis

NSL-KDD dataset is an improved version of the popular KDD Cup'99 dataset. It solves some inherent problems of the KDD'99 dataset (Tavallaee et al. 2009; McHugh 2000). Due to lack of public datasets for network-based IDS, the current version of NSL-KDD may be applied as an effective benchmark dataset for this work. Furthermore, major improvements are carried out on KDD'99 to obtain NSL-KDD and this is more advantageous over KDD'99. The number of instances in the NSL-KDD train and test sets is reasonable, which makes experimentation bias less by running experiments on whole dataset instead of running on randomly selected short portion of KDD'99 dataset. It is free of redundant records in train and test dataset, so the classifiers will not be biased towards repeated instances in the dataset.

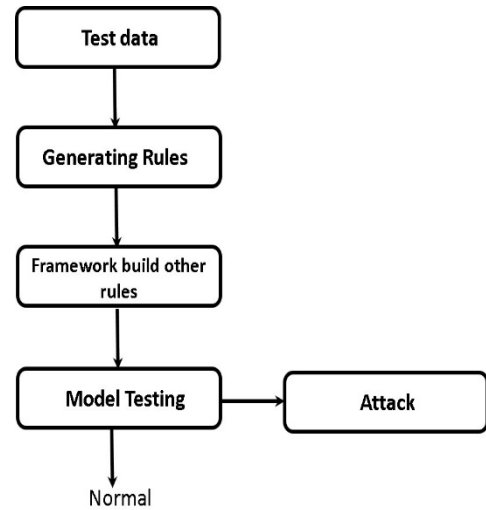


Figure 5: Implementation Setup Model

The NSL-KDD dataset contains two different files for training and testing, and hence, there is no overhead of dividing the dataset into training and testing which also makes a slight contribution towards performance evaluation of the learning techniques.

Table 1: Datasets with Attributes

Dataset	No. of records	No of attributes
KDDTrain+.txt	125,973	42
KDDTest+.txt	22,544	42

Table 1 shows the properties of the NSL-KDD train and test datasets (both are obtained as .txt files). It should be noted that there are no missing values in any attribute, and number of attributes in the table includes a class attribute. One approach during the training phase refers to processing of all the patterns or instances present in the KDDTrain+.txt dataset. However, size of testing set is 15% of size of whole dataset. Since these datasets are large and computation in ordinary machines might take too much time or sometimes might not support the memory requirements, randomly chosen 10% of the training dataset was used during training. The random selection of 10% of the training dataset was repeated for different executions so that there would be comparatively less chance of repeating the training data in different simulations. There are different types of attacks in the

dataset. However, in this work, it is considered as a two-class problem where patterns may belong to either ‘normal’ or ‘anomaly’ class.

5.1 Normalization of dataset

In the NSL-KDD datasets, the values for each attribute are often not distributed uniformly. It is wise to maintain a uniform distribution of each input attributes in the dataset before processing in the neural network. Hence, to ensure that the input values were compatible despite significant differences in their values, the dataset is normalized with respect to each input value where d_i the original value is; d_{max} and d_{min} are the maximum and minimum value, respectively, in the input attribute from which d_i is obtained. Then, normalized value of d_i is denoted as d'_i . However, it was seen that the normalized dataset some-times contained majority of zeros and in such cases, it was preferred to use the in-built data normalization function of MATLAB called *mapminmax*, which normalizes data in the range $[-1,1]$.

$$d'_i = \frac{d_i - d_{min}}{d_{max} - d_{min}} \tag{6}$$

5.1.1 Testing and Validation

For our experiments, we are using NSL-KDD dataset. NSL-KDD contains 42 fields as an attribute. In our algorithm, we have taken selected features. The performances of each method are measured according to,

- Accuracy
- False Positive Rate

A false positive occurs when the system classifies an action as anomalous (a possible intrusion) when it is a legitimate action. Although this type of error may not be completely eliminated, a good system should minimize its occurrence to provide useful information to the users. A false-negative occurs when an actual intrusive action has occurred but the system allows it to pass as non-intrusive behavior. While the true-positives (TP) and true-negatives (TN) are correct classifications. Recall Rate measures the proportion of actual positives that are correctly identified.

a. Accuracy

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{7}$$

Where FN is False Negative, TN is True Negative, TP is True Positive, and FP is False Positive

b. False Alarm Rate

The false positive rate is the number of normal connections that are misclassified as attacks divided by the number of normal connections in the data set.

$$False\ alarm = \frac{FP}{FP+TN} \tag{8}$$

5.1.2 Performance Comparison

The performance of the proposed IDS model is compared with the performance of the K-means algorithm and K-means

based ANN algorithm in this section in terms of accuracy and false alarm rate.

Figure 6 shows the accuracy comparison between proposed IDS model with K-means algorithm and K-means based ANN algorithm.

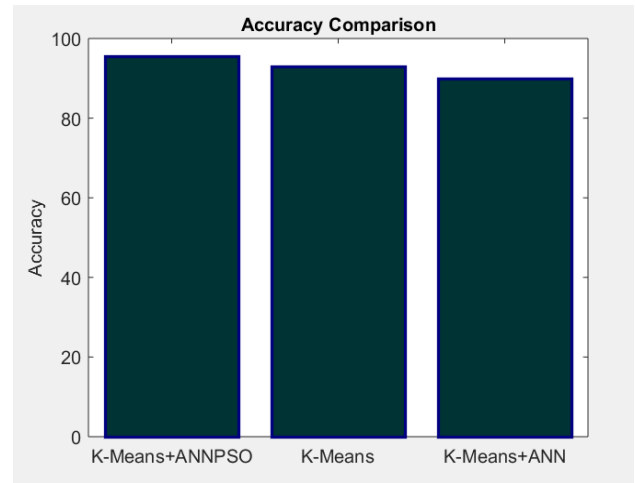


Figure 6: Accuracy comparison between existing techniques

Figure 6 shows the accuracy of proposed IDS model is 88.2321% and the K-means algorithm attains 85% accuracy and K-means based ANN algorithm attains 83% accuracy with NSL-KDD dataset. This shows the significance of the proposed IDS model than the existing IDS models.

Figure 7 shows the false alarm rate comparison between proposed IDS model with K-means algorithm and K-means based ANN algorithm.

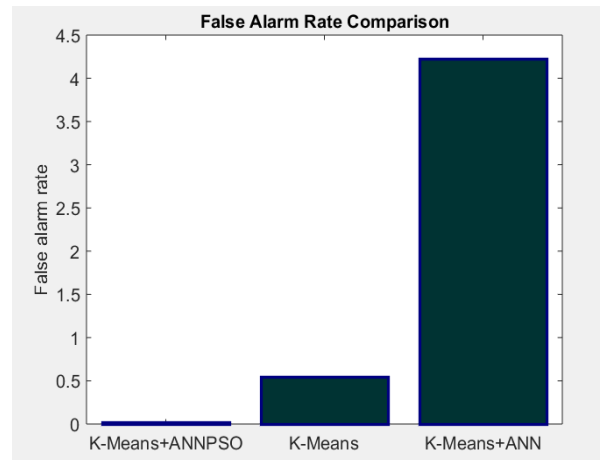


Figure 7: False Alarm Rate Comparison between existing techniques

Figure 7 shows the false alarm rate of proposed IDS model is ranging between 0-0.5 and the K-means algorithm attains false alarm rate of 0.5 and K-means based ANN algorithm attains maximum false alarm rate above 4 with NSL-KDD dataset. The significant reduce in false alarm rate shows the proposed IDS model is efficient.

6. Conclusion

Intrusion Detection is a process of detecting Intrusion in a computer system in order to increase the security. Intrusion detection is an area in which more and more sensitive data are stored and processed in networked system. After reading several research works, we come with several advantage and disadvantage. In this paper, a novel method for Intrusion detection was proposed. The proposed Intrusion detection system combines the advantages of machine learning and artificial intelligence to overcome the general issues present in the intrusion detection systems. Using the PSO algorithm, a multi-objective problem is formulated to find the optimal clusters and the issue is solved. The proposed system is tested in NSL-KD dataset and it achieved maximum detection accuracy of 88%. In addition to this, it is observed that the false negative rate of the proposed IDS is significantly reduced when compared with the results obtained with K-means algorithm and K-means based ANN algorithm.

References

- [1] Sobh, Tarek S. "Wired and wireless intrusion detection system: Classifications, good characteristics and state-of-the-art." *Computer Standards & Interfaces* 28, no. 6 (2006): 670-694.
- [2] Aydın, M. Ali, A. Halim Zaim, and K. GökhanCeylan. "A hybrid intrusion detection system design for computer network security." *Computers & Electrical Engineering* 35, no. 3 (2009): 517-526.
- [3] Raymond, Jean-François. "Traffic analysis: Protocols, attacks, design issues, and open problems." In *Designing Privacy Enhancing Technologies*, pp. 10-29. Springer, Berlin, Heidelberg, 2001.
- [4] Tsai, Chih-Fong, Yu-Feng Hsu, Chia-Ying Lin, and Wei-Yang Lin. "Intrusion detection by machine learning: A review." *expert systems with applications* 36, no. 10 (2009): 11994-12000.
- [5] Debar, Hervé, Marc Dacier, and Andreas Wespi. "A revised taxonomy for intrusion-detection systems." In *Annales des télécommunications*, vol. 55, no. 7-8, pp. 361-378. Springer-Verlag, 2000.
- [6] Liao, Hung-Jen, Chun-Hung Richard Lin, Ying-Chih Lin, and Kuang-Yuan Tung. "Intrusion detection system: A comprehensive review." *Journal of Network and Computer Applications* 36, no. 1 (2013): 16-24.
- [7] Patcha, Animesh, and Jung-Min Park. "An overview of anomaly detection techniques: Existing solutions and latest technological trends." *Computer networks* 51, no. 12 (2007): 3448-3470.
- [8] Depren, Ozgur, Murat Topallar, EminAnarim, and M. Kemal Ciliz. "An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks." *Expert systems with Applications* 29, no. 4 (2005): 713-722.
- [9] Labib, Khaled, and V. Rao Vemuri. "An application of principal component analysis to the detection and visualization of computer network attacks." In *Annales des télécommunications*, vol. 61, no. 1-2, pp. 218-234. Springer-Verlag, 2006.
- [10] Verwoerd, Theuns, and Ray Hunt. "Intrusion detection techniques and approaches." *Computer communications* 25, no. 15 (2002): 1356-1365.
- [11] Syarif, Iwan, Adam Prugel-Bennett, and Gary Wills. "Unsupervised clustering approach for network anomaly detection." In *International conference on networked digital technologies*, pp. 135-145. Springer, Berlin, Heidelberg, 2012.
- [12] Jiang, ShengYi, Xiaoyu Song, Hui Wang, Jian-Jun Han, and Qing-Hua Li. "A clustering-based method for unsupervised intrusion detections." *Pattern Recognition Letters* 27, no. 7 (2006): 802-810.
- [13] Fan, Cheng, Fu Xiao, Yang Zhao, and Jiayuan Wang. "Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data." *Applied energy* 211 (2018): 1123-1135.
- [14] Aggarwal, Charu C., and S. Yu Philip. "Data mining techniques for associations, clustering and classification." In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 13-23. Springer, Berlin, Heidelberg, 1999.
- [15] Stewart, Theo J. "A critical survey on the status of multiple criteria decision making theory and practice." *Omega* 20, no. 5-6 (1992): 569-586.
- [16] Abraham, Ajith, Swagatam Das, and Sandip Roy. "Swarm intelligence algorithms for data clustering." In *Soft computing for knowledge discovery and data mining*, pp. 279-313. Springer, Boston, MA, 2008.
- [17] Zhang, Zhongxing, and Baoping Gu. "Intrusion detection network based on fuzzy c-means and particle swarm optimization." In *Proceedings of the 6th International Asia Conference on Industrial Engineering and Management Innovation*, pp. 111-119. Atlantis Press, Paris, 2016.
- [18] Patel, Ahmed, Mona Taghavi, KavehBakhtiyari, and Joaquim CelestinoJúNior. "An intrusion detection and prevention system in cloud computing: A systematic review." *Journal of network and computer applications* 36, no. 1 (2013): 25-41.
- [19] Folino, Gianluigi, and Pietro Sabatino. "Ensemble based collaborative and distributed intrusion detection systems: A survey." *Journal of Network and Computer Applications* 66 (2016): 1-16.
- [20] Sharma, Ruby, and Sandeep Chaurasia. "An enhanced approach to fuzzy C-means clustering for anomaly detection." In *Proceedings of First International Conference on Smart System, Innovations and Computing*, pp. 623-636. Springer, Singapore, 2018.
- [21] Chung, Yuk Ying, and Noorhaniza Wahid. "A hybrid network intrusion detection system using simplified swarm optimization (SSO)." *Applied Soft Computing* 12, no. 9 (2012): 3014-3022.
- [22] Thaseen, IkramSumaiya, and CherukuriAswani Kumar. "Intrusion detection model using fusion of chi-square feature selection and multi class SVM." *Journal of King Saud University-Computer and Information Sciences* 29, no. 4 (2017): 462-472.
- [23] Çavuşoğlu, Ünal, Shirin Panahi, AkifAkgül, SajadJafari, and SezginKaçar. "A new chaotic system with hidden attractor and its engineering applications: analog circuit realization and image encryption." *Analog Integrated Circuits and Signal Processing* 98, no. 1 (2019): 85-99.
- [24] Aswani, Reema, S. P. Ghrera, Arpan Kumar Kar, and Satish Chandra. "Identifying buzz in social media: a hybrid approach using artificial bee colony and k-nearest neighbors for outlier detection." *Social Network Analysis and Mining* 7, no. 1 (2017): 38.
- [25] Duque, Solane, and Mohd Nizam bin Omar. "Using data mining algorithms for developing a model for intrusion detection system (IDS)." *Procedia Computer Science* 61 (2015): 46-51.
- [26] Benmouiza, Khalil, and Ali Cheknane. "Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models." *Energy Conversion and Management* 75 (2013): 561-569.
- [27] Abd-El-Wahed, W. F., A. A. Mousa, and M. A. El-Shorbagy. "Integrating particle swarm optimization with genetic algorithms for solving nonlinear optimization problems." *Journal of Computational and Applied Mathematics* 235, no. 5 (2011): 1446-1453.
- [28] Hart, Christopher G., and Nickolas Vlahopoulos. "An integrated multidisciplinary particle swarm optimization approach to conceptual ship design." *Structural and Multidisciplinary Optimization* 41, no. 3 (2010): 481-494.
- [29] Lunt, Teresa F. "A survey of intrusion detection techniques." *Computers & Security* 12, no. 4 (1993): 405-418.
- [30] Manohar, Balaraman, and SoundarDivakar. "An artificial neural network analysis of porcine pancreas lipase catalysed esterification of anthranilic acid with methanol." *Process Biochemistry* 40, no. 10 (2005): 3372-3376.

ANN Based MRAC-PID Controller Implementation for a Furuta Pendulum System Stabilization

Efrain Mendez*, German Baltazar-Reyes, Israel Macias, Adriana Vargas-Martinez, Jorge de Jesus Lozoya-Santos, Ricardo Ramirez-Mendoza, Ruben Morales-Menendez and Arturo Molina

Tecnologico de Monterrey, School of Engineering and Sciences, Mexico

ARTICLE INFO

Article history:

Received: 15 January, 2020

Accepted: 09 May, 2020

Online: 30 May, 2020

Keywords:

Furuta Pendulum

Robotics

Robotic systems

UMS

Artificial Intelligence

ANN

MRAC-PID

Non-linear Systems

ABSTRACT

Nowadays, process automation and smart systems have gained increasing importance in a wide variety of sectors, and robotics have a fundamental role in it. Therefore, it has attracted greater research interests; among them, Underactuated Mechanical Systems (UMS) have been the subject of many studies, due to their application capabilities in different disciplines. Nevertheless, control of UMS is remarkably more difficult compared to other mechanical systems, owing to their non-linearities caused by the presence of fewer independent control actuators with respect to the degrees of freedom of the mechanism (which characterizes the UMS). Among them, the Furuta Pendulum has been frequently listed as an ideal showcase for different controller models, controlled often through non-linear controllers like Sliding-Mode and Model Reference Adaptive controllers (SMC and MRAC respectively). In the case of SMC the chattering is the price to be paid, meanwhile issues regarding the coupling between control and the adaptation loops are the main drawbacks for MRAC approaches; coupled with the obvious complexity of implementation of both controllers. Hence, recovering the best features of the MRAC, an Artificial Neural Network (ANN) is implemented in this work, in order to take advantage of their classification capabilities for non-linear systems, their low computational cost and therefore, their suitability for simple implementations. The proposal in this work, shows an improved behavior for the stabilization of the system in the upright position, compared to a typical MRAC-PID structure, managing to keep the pendulum in the desired position with reduced oscillations. This work, is oriented to the real implementation of the embedded controller system for the Furuta pendulum, through a Microcontroller Unit (MCU). Results in this work, shows an average 58.39% improvement regarding the error through time and the effort from the controller.

1 Introduction

In recent decades, robotics have gained increasing importance in endless applications for different disciplines (as explained in [1]); such as, robotic manipulators for industrial automation, precision robots to perform surgeries, automation of assembly lines, among others. Thus, robotics have attracted greater interest multidisciplinary researchers.

Moreover, according to [2], it is defined in robotics that an Underactuated Mechanical System (UMS), is a scheme with more Degrees of Freedom (DOF) compared to its control actuators. Additionally, [3] explains that the control of UMS's is a very active research topic, due to their broad applications in Robotics, Aerospace, and Marine Vehicles.

Therefore, in spite of the complexity caused by the lack of actuators to control the movement of the system, it is precisely the low number of required actuators that makes UMS's ideal for applications where energy efficiency is sought. Additionally, UMS's are systems that allow to decrease the size of the manipulators, and even simplify the amount of elements of a more complex system. The above, resulting also in cost reduction with an increased process efficiency, which is the key element that attracts the most interest for its development in the industry.

The non-linearities caused by the relation between the actuators and the DOF to be controlled, makes the complexity of these systems attractive as testbeds, for the research of different control structures. Among them, this work is focused in the rotatory pendulum, also known as the *Furuta Pendulum* (in honor of its inventor K.

*Corresponding Author: Efrain Mendez Flores; Tecnologico de Monterrey, School of Engineering and Sciences, Mexico City 14380, Mexico; Correspondence: efrain@tec.mx; Tel.: +52-5521-333-361.

Furuta [4]), which is a highly non-linear system that is unstable at the desired upright position, as described in [3].

Hence, different solutions have been proposed addressing the swinging-up and the upright issues of the *Furuta Pendulum*, which are the main control objectives to fulfill in order to achieve the desired position. In the case of the stabilization issue, [5] proposes a Linear Quadratic Regulator (LQR) approach, showing a better settling time with low overshoots, regarding a Sliding-Mode Controller (SMC). Nevertheless, solutions that require linearized equations are more effective when the non-linear components of the system, are small regarding the linear predominant behavior; therefore, in high non-linear systems like the *Furuta Pendulum*, the control structure may be highly sensitive to parametric variations, compromising the behavior of the controller when additional friction coefficients, unexpected center of mass changes in the real, or even the uncertainty of the linearized model, are presented.

Therefore, many other approaches based on nonlinear controllers are still being studied, as shown in [6] and [7] applications, where the implementation of the SMC and an Integral Sliding Mode Controller (ISMC) for pendulum stabilization are presented respectively, showing that the (ISMC) has the best response regarding perturbations rejection. Nevertheless, the chattering continues to be a limitation regarding to practical limitations of actuators, as explained in [8].

On the other hand, [9] presents a Model Reference Adaptive Controller (MRAC) implementation, discussing the implementation of the technique compared again to an LQR approach. In this particular case, it is shown that the LQR controller had an improved behavior regarding the settling time of the MRAC controller through the MIT rule, nevertheless perturbation rejection tests were not clearly analyzed.

Additionally, other studies such as [10], validates that the adaptive features of the MRAC can create a more robust system, features that are also translated to all the MRAC derivations. The previous premise can be clearly validated by the work presented in [11], where an MRAC combined with a Proportional Integrative Derivative (MRAC-PID) is implemented in a real *Furuta pendulum* testbed, achieving through the simulated and the experimental responses the validation of the method, for the upright stabilization issue of the pendulum.

Henceforth, it is important to highlight that this paper is an extension of the [11] work, originally presented in the *ICMEAE 2019* (6th International Conference on Mechatronics, Electronics and Automation Engineering) and *ICCRE 2019* (4th International Conference on Control and Robotics Engineering); where the design, modelling and a first control approach for the *Furuta Pendulum* was presented.

Nevertheless, the [11] implementation had its main drawback, in the great amount of processing resources consumed from the Microcontroller Unit (MCU), which in some instances compromised the correct and precise sampling rate of the data, which led to unexpected noise in the error signal and therefore to practical issues in the implementation. Therefore, this proposal takes the obtained results from the MRAC-PID approach, improving the behavior of the system through the implementation of an ANN based controller.

Controllers based on ANN are able to provide robust, non-fluctuating performances in the presence of parametric uncertainty

and extraneous disturbances, as explained in [12] where an ANN was simulated in order to validate its capabilities for UMS. However, the methodology was only simulated, and therefore there was none experimental validation of the performance of the control theory.

On the other hand, the main objective of designed ANN in this work, is to emulate the behavior of the MRAC-PID controller implemented in [11], extending the operating margins and maintaining the stability gained by the original MRAC-PID, but with lower computational effort; which is indeed, a common characteristic of the implementation of ANN's (as discussed in [13] and [10]).

In this particular proposal, the ANN is not trained through a model oriented approach as in [12]; in this work, the model of the *Furuta pendulum* was taken for the MRAC-PID controller calibration and implementation, and the ANN was trained through the acquired data from the experimental testbed, expecting to achieve a correct emulation of the controller's behavior, but improving its performance by increasing response speed of the controller through the reduction of the required embedded processing effort.

Then, the results in this work were acquired through the embedded implementation in the real experimental testbed. It is critical for this work to highlight that, the ANN is trained through the input/output data acquired from the stabilization MRAC-PID controller implemented in [11], in order to achieve the improved behavior of the controller with less computational effort.

Therefore, the control effort is validated through the evaluation of the control signal through time, where the duty cycle sent to the driver of the actuator can be evaluated for the task. Meanwhile, the performance is evaluated by the evolution of the error through time. Together, both signals allow evaluating the performance of the system and, the actuator stress required to achieve the control objective.

Hence, this work presents in Section 2 the control objectives of the *Furuta Pendulum* with greater details, meanwhile the experimental testbed is presented in Section 3; Section 4 discusses the structure of the MRAC-PID control topology, leading to the ANN concepts discussed in Section 5. Finally, Section 6 presents the results of this work and Section 7 presents the final conclusion.

2 Furuta Pendulum

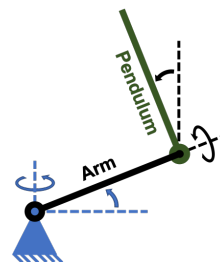


Figure 1: Basic *Furuta Pendulum* diagram.

As explained in [14], the rotatory inverted pendulum (also known as the *Furuta Pendulum*), is a rotatory system composed by an homogeneous pendulum attached to a perpendicular arm, which rotates in a horizontal plane. Fig. 1 shows the basic configuration of the pendulum, where the blue triangle represents the fixed point in which

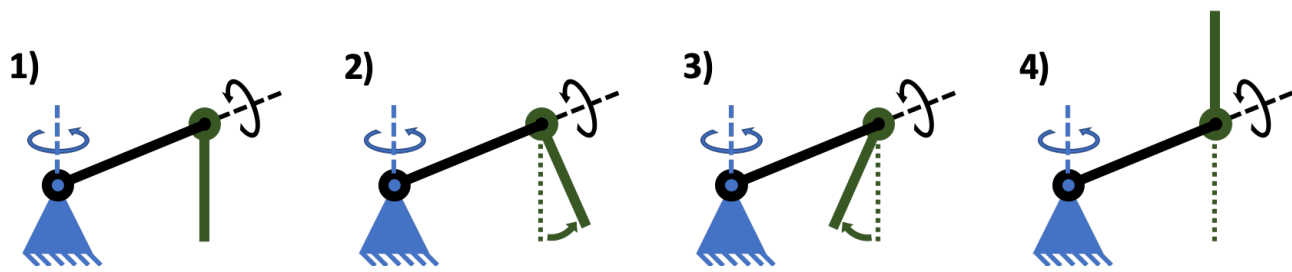


Figure 2: Swinging-up sequence for the Pendulum.

the arm is attached, therefore the first DOF of the system is in the arm, that can only rotate around the vertical axis, in the horizontal plane.

Moreover, also from Fig. 1 can be seen that the pendulum is coupled to the arm by means of a bearing, which allows it to rotate around the arm axis and therefore, it is there where the next degree of freedom of the system is found. Henceforth, attending to the UMS classification in which the Furuta Pendulum is located in, the only actuator of the system is attached in the fixed point at the beginning of the arm.

2.1 Control Objectives

As implied from the structure of the Furuta Pendulum, it is clear that it is a system that is unstable without a controller (as also explored in [15]). Thus, the control objectives of a Furuta Pendulum can be divided into two main issues, as addressed in [11], where they have been extensively studied both individually and together.

On the one hand, the initial state of the Furuta pendulum is just as shown in the first diagram of Fig. 2, where the pendulum is in its resting position. The control objective there, is located in the natural equilibrium point; in this scenario, a controller needs to be applied in order to destabilize the system by applying torque to the arm through the actuator, traduced into a swing-up routine, by then, arising the pendulum into the working range of the main stabilization controller.

The swing-up routine can be summarized by the second and third diagrams of Fig. 2, finally leading to the stage of the fourth diagram, where the pendulum switches to the controller in charge of maintaining the upright position. The swinging-up routine can be made through energy control, as explored by Furuta itself in [4]. Other proposed solutions for the task is the predictive controller ([16]), and even solved through feedforward approaches such as presented in [17].

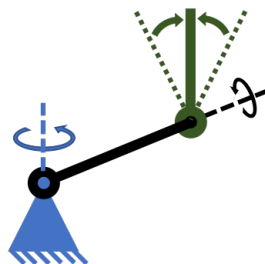


Figure 3: Representation of the stabilization objective.

On the other hand, the second control objective lies in the stabilization of the pendulum, in order to maintain it in an upright position by correcting any disturbance, only through the applied torque of the arm. Fig. 3, shows the basic outline of the control objective, where the green dotted lines represent the operating range that delimits the control actions, knowing that most of the state of the art recommends it to be maximum $\pm 15^\circ$ from the upright position.

Therefore, summarizing the control objectives, out of the stabilization range the swing-up controller is activated, in order to take the pendulum into the area in which the stabilization controller takes the wheel, achieving then the expected upright position.

In addition, the study of controllers in charge of maintaining the pendulum in the upright position, have served for inspiration of many applications, such as the stability of walking humanoid robots among others. Hence, this work proposal is focused in a solution regarding the stabilization control objective, where ANN based controller seeks to achieve an effective solution for its microcontroller implementation.

2.2 System Modelling

As explained in [11], the mathematical model of the system can be delivered by the segmentation of the mechanical system, through its kinetic and potential energy (as discussed in [4]). Therefore, the energies relation from which the mathematical model of the pendulum emerges, is given by (1).

$$L = \{K_{arm} + K_p\} + \{V_{arm} + V_p\} \quad (1)$$

where L is the Lagrange relation obtained by combining both energy equations, K_{arm} and k_p are the kinetic energy equations of the arm and the pendulum respectively, meanwhile V_{arm} and V_p are their potential energy equations. After the kinetic and potential energies analysis, as explained in [18] and also validated by [11], the expansion of (1) leads to second order differential equations; the first one, is the differential equation of the motion of the pendulum in terms of the angle generated between the horizontal plane and the pendulum's arm (θ_{arm}). Meanwhile, the second one is the differential equation obtained from the motion of the system, in terms of the pendulum and the angle between the vertical axis and the pendulum itself (θ_p).

The simplified expression that describes the behavior of (θ_{arm}), is given by (2). Where α and β are products of a change of variable, since α is the expansion given by (3), and β is developed through (4).

$$\ddot{\theta}_{arm} = \frac{\alpha}{\beta} \tag{2}$$

$$\begin{aligned} \alpha = & \tau (m_p l_p^2 + I_p) - \sin(\theta_p) (l_p^3 l_r m_p^2 \dot{\theta}_p^2 \cos(\theta_p)^2 \\ & - \dot{\theta}_p^2 (l_r l_p^3 m_p^2 + I_p l_r l_p m_p) + g l_p^2 l_r m_p^2 \cos(\theta_p)) + \\ & \dot{\theta}_p \sin(2\theta_p) (\dot{\theta}_{arm} l_p^4 m_p^2 + l_p \dot{\theta}_{arm} l_p^2 m_p) \end{aligned} \tag{3}$$

$$\begin{aligned} \beta = & m_p (I_r l_p^2 + l_p l_r^2) + \sin(\theta_p)^2 (l_p^4 m_p^2 + l_p l_p^2 m_p) \\ & + l_p I_r + l_p^2 l_r^2 m_p^2 - l_p^2 l_r^2 m_p^2 \cos(\theta_p)^2 \end{aligned} \tag{4}$$

On the other hand, the behavior of (θ_p) is given by (5). Where another change of variable was made in order to simplify the expression, as explored in (2). Therefore in this case, γ is given by the expression given by (6), meanwhile λ is traduced as (7).

$$\ddot{\theta}_p = \frac{\gamma}{\lambda} \tag{5}$$

$$\begin{aligned} \gamma = & \sin(\theta_p)^3 (\cos(\theta_p) l_p^4 m_p^2 \dot{\theta}_{arm}^2 + g l_p^3 m_p^2) + \\ & \sin(\theta_p) (\cos(\theta_p) (l_p^2 l_r^2 m_p^2 \dot{\theta}_{arm}^2 \\ & - l_p^2 l_r^2 m_p^2 \dot{\theta}_p^2 + I_r l_p^2 m_p \dot{\theta}_{arm}^2) \\ & + g l_p l_r^2 m_p^2 + I_r g l_p m_p) - l_p l_r m_p \tau \cos(\theta_p) \\ & + l_p^3 l_r m_p^2 \dot{\theta}_{arm} \dot{\theta}_p \sin(2\theta_p) \cos(\theta_p) \end{aligned} \tag{6}$$

$$\begin{aligned} \lambda = & m_p (I_r l_p^2 + I_p l_r^2) + \sin(\theta_p)^2 (l_p^4 m_p^2 + I_p l_p^2 m_p) \\ & + I_p I_r + l_p^2 l_r^2 m_p^2 - l_p^2 l_r^2 m_p^2 \cos(\theta_p)^2 \end{aligned} \tag{7}$$

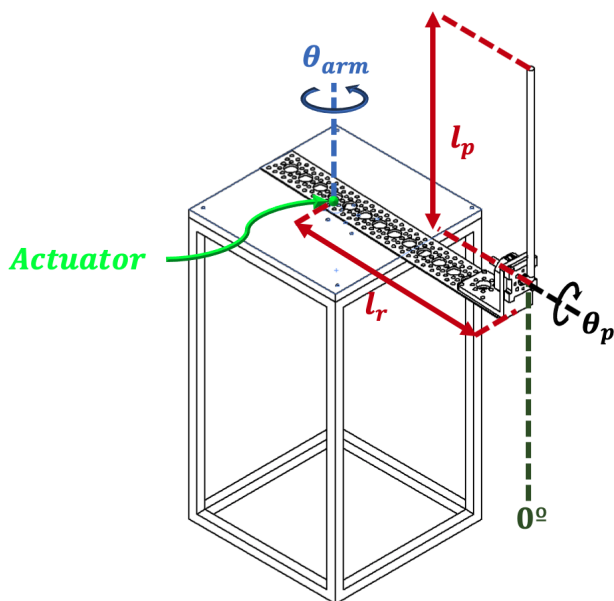


Figure 4: Representation of the stabilization objective.

In terms of the testbed analyzed in this paper, the variables from (2) and (5) are sketched in the diagram shown in Fig. 4, where an

isometric view of the designed 3D model of the pendulum is used, design that will be explained in more detail later in this document.

Nevertheless, all the variables of the equations presented in this section and analyzed in Fig. 4, are summarized in Table 1. Additionally, Table 1 shows the parameters of the each variable in the implemented experimental testbed, which is explained in Section 3.

Table 1: Furuta Pendulum Variables

Variable	Description	Value
m_r	Arm mass	0.35 [kg]
m_p	Pendulum mass	0.12 [kg]
l_r	Arm length	0.15 [m]
l_p	Pendulum length	0.25 [m]
I_r	Arm moment of inertia	0.02 [kgm ²]
I_p	Pendulum moment of inertia	0.001 [kgm ²]
τ	Nominal Torque	0.09 [Nm]
θ_{arm}	Arm angle	- [rad]
θ_p	Pendulum angle	- [rad]

It is important to highlight that in this application, the arm length l_r is measured as the distance from the center of the actuator shaft to the location of the pendulum. On the other hand, the length of the pendulum is taken as the length of the pendulum bar, from the point where it is attached to the system. Nevertheless, the center of mass of the pendulum is found as the half of the pendulum's length.

3 Experimental Testbed

The designed testbed, was inspired by the one implemented in the short version of this work ([11]), where basically the mechanical design is described in Fig. 5.

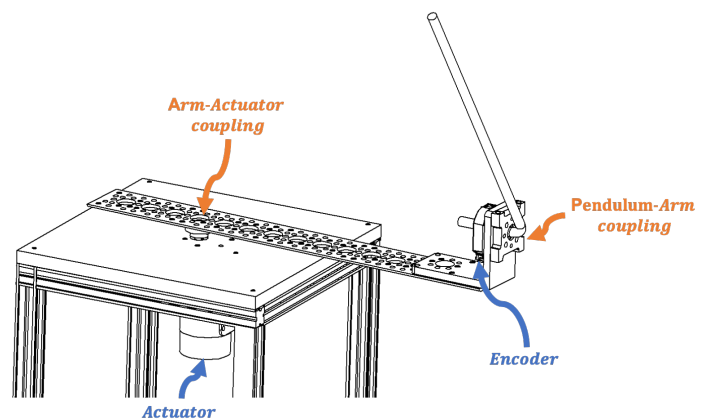


Figure 5: Schematic of the component assembly for the testbed.

The selected actuator, was the *Pololu 12V, 19:1 Gear Motor w/ 64 CPR Encoder*, which is a DC motor with a gearbox, that already has a two phase encoder for speed measurements that also has a 5A stall Current, same that is activated and controlled through a *MD30C 30A DC Motor Driver*.

Additionally, the control signal of the system, the signals processing and data acquisition, are made by the *FRDM-K64F* development board, which has an *MK64FN1M0VLL12* microcontroller that

works at 120MHz, with 1024KB of Flash memory, 256KB SRAM memory, and also allows an easier implementation of the equations through its integrated Floating Point Unit (as explored in [11]).

It is important to highlight that the original version of this work ([11]), was programmed using *FreeRTOS* (Free Real-Time Operating System), in order to work with a better sample time response from the microcontroller and a faster control output signal. Nevertheless, several changes were made in order to improve the sampling rate and therefore the error post-processing. The main changes, lie in the implementation of a Periodic Interrupt Timer (PIT), which generates highly precise interrupts at regular intervals with minimal processor intervention, therefore allowing to exploit more resources for the controller implementation.

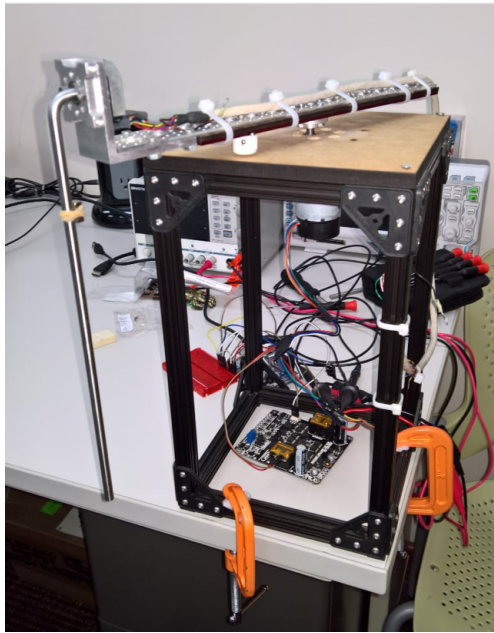


Figure 6: Experimental testbed designed and implemented.

On the other hand, for the pendulum’s data acquisition, a modular-incremental encoder was used to acquire the data using the PIT, in order to determine the speed and the position signals of the pendulum. The encoder *Cui-AMT113Q* (also highlighted in Fig. 5) was selected for the task, since it is a quadrature encoder with two phases configured with 1024 counts per revolution.

Therefore with the selected hardware, the measurement of θ_p has a resolution of 0.35° , which allows monitoring the pendulum’s position with a correct error estimation for the controller implementation.

Finally, Fig. 6 shows the experimental testbed implemented for the controller evaluation, where the implemented hardware was taken as shown in Fig 5. It can be noted in Fig. 6 in black, the frame that structurally supports the system is observed, with a wooden surface where the motor is attached. Consequently, as mentioned above the arm is directly coupled to the motor shaft, and finally in the arm the bearing that holds the pendulum is attached together with the encoder.

4 MRAC-PID controller

As highlighted all along this work, the main objective addressed in this work is to maintain a steady-state around the upright position of the pendulum. Therefore, as analyzed and validated in [11], the position of the pendulum can be controlled through a Model Reference Adaptive Controller (MRAC) with a Proportional Integrative Derivative (PID) Controller. Thus, the structure of the controller implemented in [11], was inspired by the topology shown in [19].

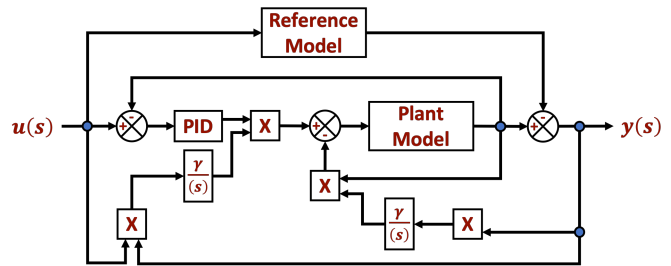


Figure 7: Blocks diagram of the MRAC-PID structure.

Then, Fig. 7 shows the general structure of the implemented controller, where the PID block from the diagram is taken as the classic PID parallel topology, which is represented by Fig. 8.

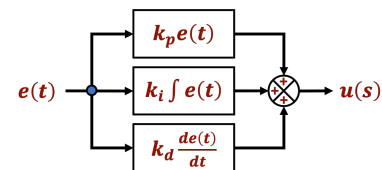


Figure 8: Blocks diagram of the Parallel PID structure.

Henceforth, the PID structure of the controller described by Fig. 8, can be expressed as shown by (8).

$$U(t) = k_p e(t) + k_i \int e(t)dt + k_d \frac{de(t)}{dt} \quad (8)$$

where k_p, k_i and k_d are the proportional, integral and derivative gains respectively, and $e(t)$ is the error signal. On the other hand, the reference model described in the blocks diagram from Fig. 7, is taken as a general second order transfer function, which is implemented through the expression defined in (9).

$$ReferenceModel = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (9)$$

where, ζ is the dampening ratio of the reference system, which can be delivered through (10).

$$\zeta = \frac{-\ln(M_p)}{\sqrt{\pi^2 + \ln(M_p)^2}} \quad (10)$$

where in this case, M_p is taken as the maximum allowed overshoot of the system. Additionally from (9), ω_n is the natural frequency of the reference model, which can be defined by (11).

$$\omega_n = \frac{4.6}{\zeta * t_s} \quad (11)$$

where t_s represents the settling time. Moreover, for this application, the settling time is taken as $t_s = 0.001$, with a maximum overshoot $M_p = 0.02$, where the learning rate for the MRAC is $\gamma = 0.25$.

Nevertheless, according to the validation shown in [11], the proposed MRAC-PID controller can be implemented in embedded systems, by using the equations developed in [19], where the constants allow real-time tuning of the controller that learns from the actual behavior of the error. Therefore, the initially implemented equations in the embedded system are:

$$k_p = \frac{-\gamma * err * (u_c - y) * w_n}{p^2 + 2\zeta w_n p + w_n^2} \quad (12)$$

$$k_i = \frac{-\gamma * err * (u_c - y) * w_n^2}{(p^2 + 2\zeta w_n p + w_n^2) * p} \quad (13)$$

$$k_d = \frac{-\gamma * err * y * p}{p^2 + 2\zeta w_n p + w_n^2} \quad (14)$$

where, err is the normalized error of the system, u_c the output of the system, y the reference model output, and p is the slope defined by $p = \frac{d(err)}{dt}$.

Consequently, in order to validate the behavior of the controller from which the ANN is going to learn from, Fig. 9 shows the the pendulum using the proposed control topology, which as validated before in [11], it is clear the capabilities of the MRAC-PID to maintain the pendulum in a steady-state in the upright position.

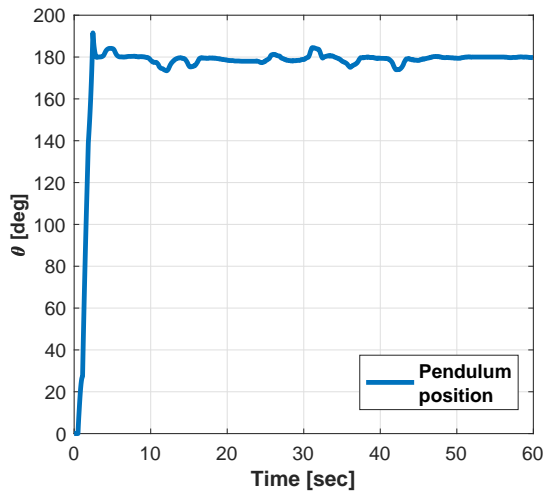


Figure 9: Validation through pendulum’s position against time.

Basically, the Fig. 9 was delivered after sampling a minute of the pendulum’s behavior, using a $50[mSec]$ sampling rate. In every performed test for this work, the pendulum starts at a 0° initial point (as shown in Fig. 4), and the reference is in the upright position at 180° . Subsequently, with the control objective already fulfilled, the next stage of this work is to take a step into the design based on ANN.

5 Artificial Neural Networks

Artificial Neural Networks (ANN) were first introduced by McCulloch and Pitts in [20] as a propositional logic model that intended to

imitate biological neurons. The main objective of their proposal was to use models of artificial neurons, in order to evaluate any logical computation based on AND, OR, and NOT assignments.

Nevertheless, the interconnection of multiple neuron models as seen in Fig. 10, enables to assign different weights at each connection, which allows a better prediction of the training model through the implementation of backpropagation models between them.

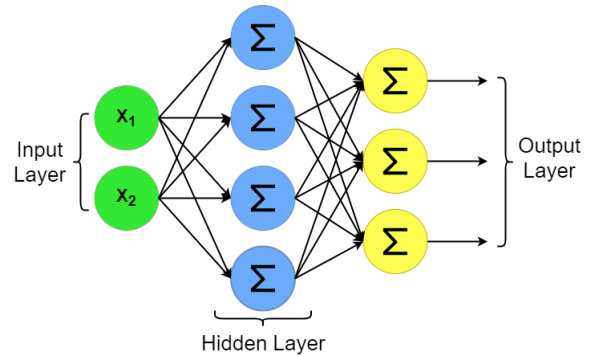


Figure 10: Multiple-layer neuron model

Therefore, knowing that it is possible to use multi-layer neural networks for regression purposes, it can be designed an ANN according to the input-output relationship of the control system (as explained in [21]). Moreover, in the case of a Multiple Input-Single Output (MISO) model, the input layer of the network will assign one neuron to each input, while the final layer will include only one single neuron for the output.

Therefore, regarding the layers of the network, it is commonly used a Rectified Linear Unit (ReLU) function for the hidden layers, which can be described by the expression shown in (15).

$$y = \max(0, x) \quad (15)$$

Meanwhile, any linear equation can be used to guarantee positive values; although the selection of the function could vary depending on the nature of the system, for this implementation a Sigmoid given by (16)), was selection for the activation task.

$$y = \frac{1}{1 + e^{-x}} \quad (16)$$

For the regression, the network loss function seeks to minimize the error from the ANN, and can be performed (as explored in [22, 23, 24]) through the mean-squared error (MSE) as given by (17).

$$L = E [||\epsilon||^2] = E [||y - f(x)||^2] \quad (17)$$

where the estimation of $f(x)$ can be expressed as the product of the input values with the networks weights, as seen in (18).

$$\hat{y} = f(x) = x^T w \quad (18)$$

Where the optimization objective, is to find the best values for the network weights w , that together with the neurons, brings the behavior of the ANN closer to what is expected; which as explained in [25], can be expressed as shown by (19).

$$\arg \min_{\hat{w} \in \mathbb{R}^M} E [||y - \hat{w}^T x||^2] \quad (19)$$

Subsequently, after presenting the basic elements for the ANN to be designed for the controller of the Furuta Pendulum, the network training will be presented below, based on the data acquired from the experimental testbed.

5.1 MRAC-PID through ANN

The original MRAC-PID controller from [11], implemented in the experimental testbed shown in Fig. 6, was used to generate the dataset of the error measurements obtained during the settling process of the pendulum.

Hence, a total of 11,330 values of the pendulum’s angle error were acquired, which are used as input values for the ANN. Simultaneously, the data from the motor’s duty cycle through time was also acquired, since the duty cycle for the actuator represents the response of the controller (control signal) against the error signal. Thus, output of the network is taken as the control signal for the ANN.

Moreover, the relation between these two variables and their behavior is shown in Fig. 11, where it can be observed that the ideal behavior of the model would be to reduce the duty cycle of the motor when the pendulum is located at 180°.

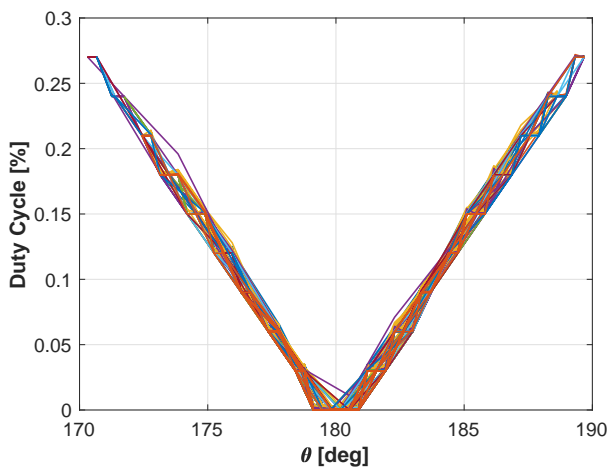


Figure 11: Acquired training data.

Both values (error and duty cycle) were normalized and separated into training, validation, and testing sets, with the 70%, 15%, and 15% relation of the data respectively. The ANN was initialized as described in Table 2, with a MSE loss function and an Adam optimizer with $\alpha = 0.0001$ inspired by [26].

Table 2: ANN architecture

Layer	Activation func.	Neurons
Input	ReLU	6
Output	Sigmoid	1

Consequently, the network was trained using 100 epochs with a 128 batch size. After the network was trained, the weight matrix was used to evaluate arbitrary error readings in order to generate the regression function, which later was implemented into the embedded system.

Nevertheless, before the experimental test, the resulting matrix was used in order to evaluate the error in a range between 170° and 190°, which is the operating range for the controller in the testbed, seeking to simulate the the resulting duty cycle that would be programmed into the microcontroller. Therefore, the simulated predictions of the control signal against the error signal are shown in Fig. 12.

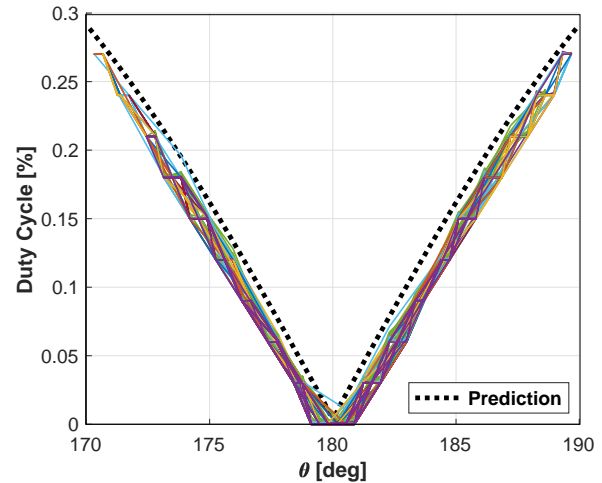


Figure 12: Acquired training data compared to the ANN prediction.

The simulated behavior of the trained network architecture, achieved a similar behavior than the real MRAC-PID controller, with a slightly more aggressive response against the error signal. Which allows predicting a faster response by the ANN-based controller, with a greater performance specially when the pendulum reaches the limit inclination at 170° and 190°.

In addition, the behavior of the predicted results from Fig. 12, demonstrated how the use of an ANN allows more freedom in the system, since the behavior of the net achieves a more regular response of the system, instead of constant changes produced by the modification of the controller coefficients.

Besides, it is clear that in simulation the main advantages of the ANN are not clearly demonstrated, due to the fact that the processing reduction is not evaluated. Thus, it is important to consider and highlight the fact that the obtained prediction, was with simpler arithmetical operations compared to the MRAC-PID controller model; which in a Microcontroller, is traduced into a faster response and into a greater reliability regarding the correct sampling of the error signal, due to the fact that there would be no loss of information or lag in the measurement caused by an improper use of the ticks in the embedded system.

6 Results

The results were achieved through the implementation of the ANN in the embedded system, where one of the main advantages of the proposal is the decrease of the microcontroller’s computational effort, when evaluating real-time samples, due the reduction of computational operators through the the simpler mathematical expressions of the ANN.

In other words, it is faster for the hardware to evaluate just a single matrix operation than the three equations that rule the behavior of the coefficients of the PID controller. Such assumption is observed in Fig. 13 and Fig. 14, where the pendulum's angle and system's duty cycle are compared between the original MRAC-PID controller and the ANN model.

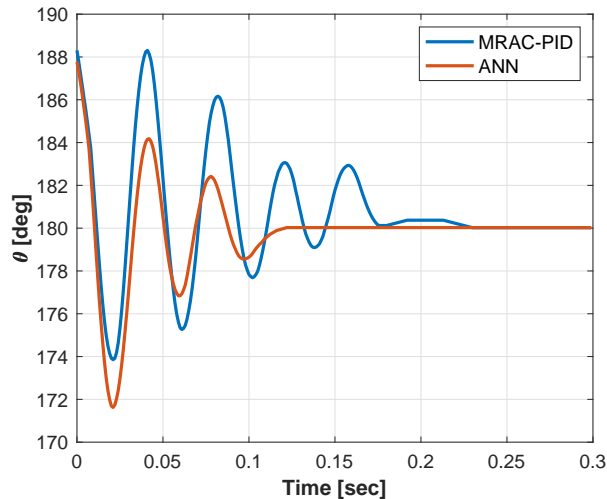


Figure 13: Pendulum's position through time.

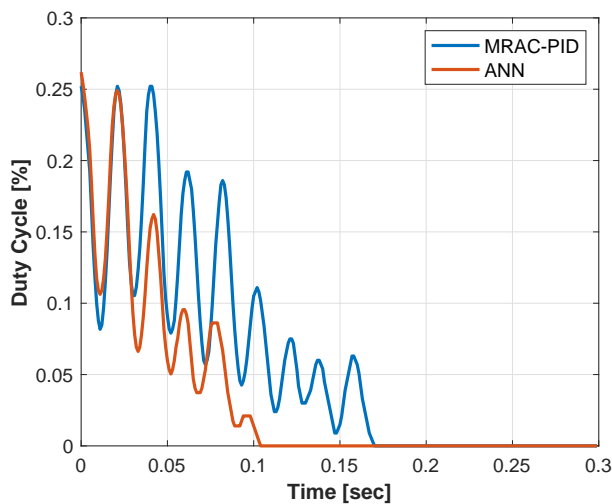


Figure 14: Control signal through time.

Additionally, Fig. 13 shows that the Furuta Pendulum achieves a reduction of the transitional stage, improving the establishment time. As can be clearly seen also in Fig. 13, the controllers were implemented in order to maintain the pendulum in the upright position, therefore the controller begins its operation in the established range ($180 \pm 10^\circ$). Thus, both controllers begin to operate below the 190° for the analysis, stipulated in order to perform fair trials for both controllers.

Therefore, in order to quantify the improvement of the pendulum's behavior, the error between the reference angle (180°) and the actual angle through time, can be measured by performance indexes

as shown in [27] and explored by [28]; from where, Integral of Time multiplied by Squared Error (ITSE) and Integral of time multiplied by Absolute Error (ITAE) were the selected indexes.

The first one, is associated to the error energy it is weighted by time, giving more importance to the most recent errors, and the second one related to the absolute error value in time, giving more importance to the errors in the steady-state. Table 4, summarizes the error indexes obtained after evaluating results from Fig. 13.

Table 3: Error indexes of the controllers.

	MRAC-PID	ANN
ITSE	2.3942	1.4032
ITAE	0.5381	0.3114

Hence, by taking the ITSE performance index, the system achieved a 58.61% of error reduction, meanwhile a 57.87% of error reduction was achieved respect to the ITAE index. Then, averaging both results it can be concluded consistently that a 58.24% error reduction was achieved.

On the other hand, last but not the least, Fig. 14 allows to validate the control signal of the controllers, where the [%] of duty cycle through time is shown. Additionally, it can be seen the signal consistency, that is traduced into the reduction of the control effort in time.

The effort reduction can be observed by the way in which the oscillations due to the change in time are compacted, that the ANN controller responds lightly faster than the MRAC-PID, which is enough to allow the faster correction of the pendulum's angle shown in 13. By integrating the response of the controller through time, in can be quantified the reduction in the control effort, achieving then a 58.69% of control effort reduction from the values summarized in Table 4.

Table 4: Control effort.

MRAC-PID	ANN
17.8285	10.4642

Meanwhile, in Fig. 14 it can be analyzed with more detail the settling time improvement, where the system passes from the $T_s \approx 0.17[sec]$ of the MRAC-PID controller, into the $T_s \approx 0.1[sec]$ settling time obtained by the proposed solution.

In spite of the consistency of the results, the authors suggest that another possible implementation to evaluate, could be the use of a Recursive Neural Network (RNN) to evaluate the data as a complete sequential variable, instead of evaluating one measurement at a time; however, the RNN case could result in an equal or more complex model than the original PID controller, which could significantly affect the speed in which the controller responds to the error signal.

Nevertheless, the quantified results from the behavior of the controllers in terms of the angle of the pendulum and the controller effort, validated that the proposed controller achieved a better response as expected, due to the faster response of the system and the computational cost reduction.

Finally, it is important to highlight that the duty cycle signal from Fig. 14, it shown in an absolute value due physical limitations. In real applications such as this one, drivers like the MD30C

DC motors driver from the designed testbed, receives two control signals: the first one is the PWM whose modulation leads to the amplitude of the control signal, and the second one is the direction for the motor. Therefore, the direction of the pendulum does not depend on the duty cycle, it is defined through the error's sign.

7 Conclusion

It was observed from the quantification of the results that, not also the behavior of the system followed the original controller's behavior, but also added a more aggressive response when the pendulum was reaching a limit angle that could result in losing equilibrium.

Additionally, results show how a simple neural network model can be narrowed into a simpler regression model, capable of executing similarly as an original and complex controller; such as the MRAC-PID. Nevertheless, the nature of ANN allowed to achieve the stabilization of the pendulum in the desired state, but through less computational effort from the Microcontroller.

The obtained results showed a 58.61% of error reduction respect to the ITSE, a 57.87% reduction respect the ITAE, and a control effort reduction of 58.69%; those three results remarkably allow to conclude that the implementation of the ANN as substitute for the MRAC-PID controller, achieved a better response for the proposed testbed.

Therefore, this works shows an approach methodology for UMS that could allow to improve other designs, due the faster response of the system and its clear lower computational cost from its mathematical expression. Thus, this work enables the method for simpler implementation or even improvements for existing models through artificial intelligence.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment This research is a product of the Project 266632 "Laboratorio Binacional para la Gestión Inteligente de la Sustentabilidad Energética y la Formación Tecnológica" ("Bi-National Laboratory on Smart Sustainable Energy Management and Technology Training"), funded by the CONACYT (Consejo Nacional de Ciencia y Tecnología) SENER (Secretaría de Energía) Fund for Energy Sustainability (Agreement S0019201401).

References

- [1] Emese Gincsaïne Szadeczky-Kardoss Daniel Szabo. "Novel Cost Function based Motion-planning Method for Robotic Manipulators". In: *Advances in Science, Technology and Engineering Systems Journal* 4.6 (2019), pp. 386–396. doi: 10.25046/aj040649.
- [2] Yang Liu and Hongnian Yu. "A survey of underactuated mechanical systems". In: *IET Control Theory & Applications* 7.7 (2013), pp. 921–935.
- [3] JA Acosta. "Furuta's pendulum: a conservative nonlinear model for theory and practise". In: *Mathematical Problems in Engineering* 2010 (2010).
- [4] Karl Johan Åström and Katsuhisa Furuta. "Swinging up a pendulum by energy control". In: *Automatica* 36.2 (2000), pp. 287–295.
- [5] Navin John Mathew, K Koteswara Rao, and N Sivakumaran. "Swing up and stabilization control of a rotary inverted pendulum". In: *IFAC Proceedings Volumes* 46.32 (2013), pp. 654–659.
- [6] Ali Wadi, Jin-Hyuk Lee, and Lotfi Romdhane. "Nonlinear sliding mode control of the Furuta pendulum". In: *2018 11th International Symposium on Mechatronics and its Applications (ISMA)*. IEEE. 2018, pp. 1–5.
- [7] Saqib Irfan et al. "Advanced sliding mode control techniques for inverted pendulum: Modelling and simulation". In: *Engineering science and technology, an international journal* 21.4 (2018), pp. 753–759.
- [8] Asif Chalanga, Shyam Kamal, and B Bandyopadhyay. "Continuous integral sliding mode control: A chattering free approach". In: *2013 IEEE International Symposium on Industrial Electronics*. IEEE. 2013, pp. 1–6.
- [9] Tamen Thapa Sarkar and Lillie Dewan. "Application of LQR and MRAC for swing up control of Inverted Pendulum". In: *2017 4th International Conference on Power, Control & Embedded Systems (ICPCES)*. IEEE. 2017, pp. 1–6.
- [10] Adriana Vargas-Martínez and Luis E Garza-Castañón. "Combining adaptive with artificial intelligence and nonlinear methods for fault tolerant control". In: *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer. 2010, pp. 31–41.
- [11] Efrain Mendez-Flores et al. "Design, Implementation and Nonlinear Control Analysis of a Furuta Pendulum System". In: *2019 4th International Conference on Control and Robotics Engineering (ICCRE)*. IEEE. 2019, pp. 65–69.
- [12] Seyed Hassan Zabihifar, Arkady Semenovich Yushchenko, and Hamed Navvabi. "Robust control based on adaptive neural network for Rotary inverted pendulum with oscillation compensation". In: *Neural Computing and Applications* (2020), pp. 1–13.
- [13] Efrain Mendez et al. "Mobile Phone Usage Detection by ANN Trained with a Metaheuristic Algorithm". In: *Sensors* 19.14 (2019), p. 3110.
- [14] Slávka Jadlovský and Ján Sarnovský. "Modelling of classical and rotary inverted pendulum systems—a generalized approach". In: *Journal of Electrical Engineering* 64.1 (2013), pp. 12–19.
- [15] Daniel Galan et al. "Customized Online Laboratory Experiments: A General Tool and Its Application to the Furuta Inverted Pendulum [Focus on Education]". In: *IEEE Control Systems Magazine* 39.5 (2019), pp. 75–87.
- [16] Pavol Seman, Martin Juh, Michal Salaj, et al. "Swinging up the Furuta pendulum and its stabilization via model predictive control". In: *Journal of Electrical Engineering* 64.3 (2013), pp. 152–158.

- [17] Yun Feng Wu, Zhu Ming, and Ke Chang Fu. "Feedforward and Feedback Control of an Inverted Pendulum". In: *Advanced Materials Research*. Vol. 328. Trans Tech Publ. 2011, pp. 2194–2197.
- [18] Mayra Antonio-Cruz et al. "Modeling, simulation, and construction of a furuta pendulum test-bed". In: *2015 International Conference on Electronics, Communications and Computers (CONIELECOMP)*. IEEE. 2015, pp. 72–79.
- [19] Ai Xiong and Yongkun Fan. "Application of a PID Controller using MRAC Techniques for Control of the DC Electromotor Drive". In: *2007 International Conference on Mechatronics and Automation*. IEEE. 2007, pp. 2616–2621.
- [20] Warren S McCulloch and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity". In: *The bulletin of mathematical biophysics* 5.4 (1943), pp. 115–133.
- [21] Aurélien Géron. *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems.* O'Reilly Media, Inc., 2017.
- [22] Thomas Kailath, Ali H Sayed, and Babak Hassibi. *Linear estimation*. BOOK. Prentice Hall, 2000.
- [23] Bernard C Levy. *Principles of signal detection and parameter estimation*. Springer Science & Business Media, 2008.
- [24] Ali H Sayed. *Fundamentals of adaptive filtering*. John Wiley & Sons, 2003.
- [25] Sun Yuan Kung. *Kernel methods and machine learning*. Cambridge University Press, 2014.
- [26] Diederik P Kingma and Jimmy Ba. "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980* (2014).
- [27] Manuel A Duarte-Mermoud and Rodrigo A Prieto. "Performance index for quality response of dynamical systems". In: *ISA transactions* 43.1 (2004), pp. 133–151.
- [28] E Mendez et al. "Electric machines control optimization by a novel geo-inspired earthquake metaheuristic algorithm". In: *2018 Nanotechnology for Instrumentation and Measurement (NANOIM)*. IEEE. 2018, pp. 1–6.

Balance as One of the Attributes in the Customer Segmentation Analysis Method: Systematic Literature Review

Uus Firdaus*, Ditdit Nugeraha Utama

Computer Science Department, BINUS Graduate Program - Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480

ARTICLE INFO

Article history:

Received: 30 April, 2020

Accepted: 23 May, 2020

Online: 29 May, 2020

Keywords:

Customer Balance

Customer Segmentation

RFM

K-Means

ABSTRACT

The banking industry is very competitive. To utilize the information, they have in order to be a competitive advantage winner is reasonably very crucial for the company. At present, the company does not only focus on the company's strategy that prioritizes products (e.g. product or service oriented), however also necessitates to focus on the company's strategy in prioritizing customers. Customer segmentation, its attributes, and the appropriate analysis method are going to get accurate data segmentation results, so that it is able to be used as a reference by the company and as a basis for determining its products' marketing strategies. This systematic literature review discusses the types of attributes operated, including customer balance attributes, whether or not they can be included in segmentation. In addition, it also discusses what popular analytical methods are widely used in the customer segmentation process. Literature searching in the digital library resulted in a total are 592,363, 1,361, and 21 papers respectively in the first, second, and third stage. 10 papers found finally in the final stage that were considered capable of answering research questions. Based on 10 papers selected, it can be concluded that customer balances can be functioned scientifically as one of attributes for segmentation use. The popular analytic methods operated for customer segmentation are recency, frequency, monetary (RFM) model (4 times appeared), K-Means algorithm (6 times occurred), and C-Means (2 times emerged).

1. Introduction

In business competition, to use of existing capabilities as much as possible in order to compete with other companies is realistically required by the companies. They have to create and obtain the characteristics of their customers, the information is used as a strategy to develop and market their products. Thus, the companies are able to set target rightly and customers are going to be more interested and satisfied. Presently, the companies should not only focus on the companies' strategy prioritizing products, but also need to focus on the companies' strategy prioritizing customers (customer oriented) [1].

The banking industry is very aggressive. To optimizing its information for convincingly winning the competition is very imperative to be conducted by banking institution [2]. The marketing department has an important role to make strategy in the business competition of other companies. Service-oriented companies selling their products to customers will generally face marketing problems. The huge amount of data and parameters collected in the banking area can be one of the marketing problems.

In any company including financial institutions, data is a valuable asset that can be used for corporate strategy. Proper data processing from a collection of raw data is going to produce critical and beneficial information. The Company has a very large data set that can be used for commercial use [3]. Prominent banking data (i.e. transaction data, bank accounts, loan customer data, all payment data, etc.) stored in a database. In essence, this data is very useful and can be used to predict revenue and analyze sales that have been achieved [3].

Customer segmentation is demanded to classify customers who have similar characteristics to find out consumer behavior. This helps in controlling the right marketing strategy to expediently improve company revenue [1]. Numerous methods for segmenting customers are many. The most popular is RFM analysis. It is for behavioral-based data processing that extracts customer profiles using the recency (current), frequency, and monetary values [4].

This literature review aims to find out other customer segmentation attribute in the banking industry, the analytical methods frequently operated, and the parameters involved. The elements used as analysis in the segmentation process are customer transaction habits and customer nominal balances.

*Uus Firdaus, Computer Science Department, BINUS Graduate Program - Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480, +6281280118811, uus.firdaus@binus.ac.id

Customer segmentation and suitable analysis methods will get accurate data segmentation results so that it can be used as a reference by the company. The results of this customer segmentation can be used as a basis for determining or strategy in marketing its products. Systematic literature review (SLR) method is a mechanism adopted scientifically in conducting studies.

2. Research Methodology

There are three reasons in this study to conduct an SLR. First, gather knowledge from previous research on related topics. Second, identify the elements used from previous studies. And final one, support research in finding new methods or new topics with basic information needed. For this reason, SLRs are able to meet the needs of achieving various objectives while maintaining a strong evidence base [5]. Also apply five stages conducted in reviewing articles; topic selection, determination of the scope of the review, selection of online libraries, selection of literature, and complete review of much of the literature [6].

2.1. Research Question

The focus of this study is to show that the balance attribute as one of the attributes that can be included and an important part of the customer segmentation analysis process. This customer is part of a company / financial institution that has a balance component. This SLR is done also to answer two main research questions as a reference in conducting further research. The research question defined as a guidance for doing the SLR are:

1. RQ1: What can balance component be operated in customer segmentation?
2. RQ2: What are analytical methods functioned in customer segmentation?

RQ1 focuses on discovering previous research to get answer of how important and how many elements of customer balances are included in the customer segmentation process at financial institutions. This will require analysis and find in any research related to customer segmentation in companies or financial institutions that include nominal balances in the process, directly the nominal balance will affect the information and data on the final results of the study so that it can be useful for researchers. The right elements and methods of analysis are needed, especially in relation to the quality of this segmentation data. While RQ2 is for finding and showing most of the analytical methods used in customer segmentation, this research question strongly states that the appropriate analysis methods and parameters used in segmentation will produce accurate and useful information for researchers.

2.2. The Search Process

Several standard indexes are used in the process of finding literature related in online databases. The papers included is in January 2015 until currently. Three online digital libraries used in this process; IEEE Xplore, Scimedirect, and Springer link. The combination of words in the search criteria are:

- (customer segmentation OR bank OR balance)
- (models AND customer segmentation)

In search strings use the "find at least one in context or title" feature of an advanced digital library search. This is to find all about nominal balance paper. Whereas in the second search

string use "find exactly the same" in the advanced search. It's used to filter research papers that contain analytical methods for customer segmentation.

2.3. Study Selection

Inclusion and exclusion criteria were used to minimize the focus of study selection [5]. By including the article; research, journals, conferences, reviews, pattern grouping, bank channel filters, engineering, business management, IT in Business or banking & finance articles as much as possible to compare each method and element used and by avoiding research papers that are not related to customer segmentation and method of analysis. Inclusion criteria are: Papers related to customer segmentation, bank, balance, and analysis methods from 2015 to the present, case studies based on and / or research reports and experience. A year back is an overview that is operated to see trends that are usually operated through several studies.

Exclusion criteria are:

1. The paper focuses on segmentation but the object is not the customer.
2. Paper that discusses customers but does not have a balance component
3. Studies that do not discuss customer segmentation related to customer behavior.
4. Non-English papers or complete book papers.
5. Papers that contain methods without explaining certain parameters.

Table 1. Questions for Quality Assessment

Number	Questions
1	Does this paper fit the topic?
2	Is this paper about research?
3	Are examples discussed in this paper / implementations of methods used?
4	Are there research methods in the paper?
5	Are the findings clearly defined?

Start a search on a digital library by only using the search string without making changes to the search criteria. Next is to conduct the application of the inclusion criteria and exclusion criteria to get a smaller number of papers discovered. The next step is to filter papers from titles and abstracts to get papers that can answer research questions. The final step, the remaining papers are papers that are in accordance with quality assessment and with full text analysis. Questions for quality assessment as seen in Table 1, and complete text analysis, approach descriptively according to topic and read full text of papers.

3. Result of study

Overall selection results as shown in Fig. 1. Start by searching in a digital library by using search strings. Obtained a total of 42,122 papers from IEEEExplore, 549,074 papers from Scimedirect, and 1,167 papers from SpringerLink. The total number at the first search stage was 592,363 papers. Implementation of the inclusion and exclusion criteria resulted in 821 papers from IEEEExplore, 516 papers from Scimedirect, and 24 papers from SpringerLink. This stage provides 1,361 papers. The third stage, search by title and abstract, produced 21 papers. The final stage, filtering our search with quality assessment and full text analysis, found 10 papers that can answer research questions.

Each paper is reviewed to determine the fulfillment of the questions in the research mentioned earlier. Table 2 shows 10 papers obtained from the previous selection process. In the following paper the method used to segment the customer and the components included. Some of the topics found in this paper are about customer segmentation, models, and some of them about the methods used in the segmentation.

Table 2. Selected Papers

Paper	Year	Paper Type	Topics	Method(s)
[7]	2019	Conference (ICITBS)	Bank customers analysis for strategies and contributions to increase profits	PCC Model, Account Profit Analysis
[8]	2018	Conference (ICIMTech)	Customer Segmentation in Banking	RFM Score, K-Medoids and K-Means Clustering
[9]	2018	Conference (SIET)	Two-Step Mining Method for Customer Segmentation	RFM model, K-Means, Silhouette, Connectivity
[10]	2018	Journal	Effective approach to customer segmentation	RFM, Fuzzy C-Means and K-Means
[11]	2016	Conference (ICIC)	Customer potential segmentation	Customer C4.5 & K-Means Algorithm
[12]	2020	Conference (ICIC)	Marketing strategies and improving customer relationships with customer segmentation	RFM values and K-Means Algorithm, SSE, Elbow
[13]	2015	Journal	Profiling banking user Clustering	KDDM processes
[14]	2018	Conference (INNS)	segmentation helps marketing to achieve goals	Dataset, K-Means, SOM
[15]	2018	Conference (UV)	Bank Client Clustering	Fuzzy C-Means Clustering Bayesian
[16]	2018	Conference (ICAITI)	Financial performance of BPR Syariah	Information Criterion (BIC), CF Tree

In [7] the author used bank data as data objects, and had analyzed intensively on the bank customer profit contribution model. Focusing on income from the customer's loan business. The calculation of the main part of bank loans based on customer balances, loan rates, in the loan period. Furthermore, based on the profit contribution from assets, income and liability contributions from the middle business class, the basic model evaluates the contribution from the customer benefits built. The things proposed in this paper are verified to be in line with company needs. That's to increase customer loyalty to the bank; otherwise it can expand the company segment with effective customer support [7].

[8] discussed customer segmentation that has been applied to bank customer data in internet banking users by clustering. It is an unsupervised data mining technique that is used by segmentation of customers. Attribute names: Balance, Transaction Date, Posting Date, Account Number, CustomerID, Debit Code, AmountOf Transaction Code. This research was conducted to build a model of customer profile data based on the use of their IBs in banks and grouping methods using the K-Means and K-Medoids methods based on RFM scores from customer transactions. The results of his research show that the K-Means method outperforms the K-Medoids method, on the Davies-Bouldin index, K-Means performed slightly better than K-Medoids [8].

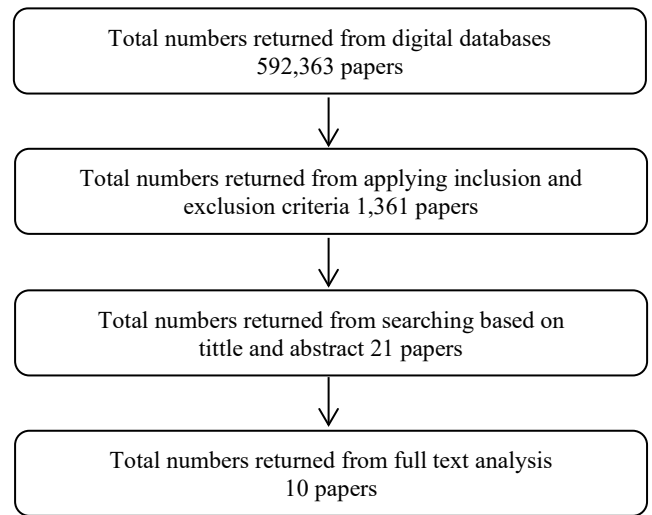


Fig. 1. Paper selection process

[9] explained customer relationship management (CRM) that has been researched to gain insight into customer requests and needs. The combination of user data such as the average customer balance was used for mining and computing methods, although this approach methodically still had several limitations. Also, several studies related to the RFM model in understanding customer habits. The model was functioned to segment customers in terms of customer transactions, frequencies, and monetary conditions. The first step, the results of the RFM model data and grouped with the k-Means algorithm. Then, data from each cluster was analyzed by association to make customer characteristics represented by IF-THEN rules. Clustering results were analyzed by silhouette size and connectivity [9].

In [10] the author examined the customer segmentation that has been run and efficiently in a company. Customers were categorized into the same behavior groups based on the customer's RFM values. Segmentation has provided a good understanding of customer needs and helped identify potential company customers. Dividing customers into several segments of the min-max balance component in an effort to increase company revenue. Retaining customers is considered more important than recruiting new customers. This can be in the form of implementing specific marketing strategies for individual segments to retain customers. This research begins with RFM analysis on transactional data and is developed into the same cluster using traditional K-means and Fuzzy C-Means algorithms. This paper also provided ideas for selecting initial centroids in the proposed K-Means algorithm [10].

In [11] the author reviewed companies that know that regular customers have costs even though they are lower than

the costs of getting new customers. Customer loyalty is important to know so that companies can project revenue as a reference in corporate planning and strategy. The process begins by using certain attributes to perform the segmentation process with the K-means algorithm. Attributes such as exchange rates, balances, the number of days past payment or the number of months of debt, and the age of the customer. Dataset without segmentation, is intended for termination, product, service disconnection, balance, age of the customer, and always decides that the class of loyalty is classified. This model combines the K-means algorithm and the C4.5 classification algorithm [11].

In [12] the author deliberated the segmentation of customers for marketing strategies and to improve the relationship between customers and companies. The behavior of loyal customers from certain services or products provides benefits to the company because customers will continue to use the services or products. This research is to find the value and types of customers that can be used to determine which customers provide the most profit for the company. RFM and balance value as monetary are used as basic criteria to identify customers in forming clusters and are called clustering. In this study the clustering method or algorithm used is K-Means and also uses the results of the Elbow Method from sum square error (SSE) from several existing clusters [12].

In [13] the author portrayed the analysis that has been carried out by analyzing data sets to obtain and ascertain the full potential of customers from realized techniques, it is called knowledge discovery and data mining (KDDM). This technique is used to process survey data from internet banking users in Jamaica that include demographics, attitude and behavior variables. The results of summarizing internet banking user data can find out the services and patterns of internet banking usage that are most frequently used. Balance inquiry and bill payment are two of the most frequently used features of a customer's account [13].

In [14] the author stated that customer segmentation whose concept has been designed in marketing to increase business and increase revenue. Also discussed are various data analytic algorithms, specifically K-Means and self-organized maps (SOM). The K-Means algorithm has shown promising clustering results, and SOM is beneficial in speed, quality in grouping, and visualization. Two levels of grouping have been applied to large customer transaction data sets, transaction history, transaction times, balances, transaction amounts, etc. K-means algorithm and grouping step are applied in this customer segmentation. Recent research proposes a segmentation framework of a customer's lifetime value (LTV). Researchers have prepared a framework for segmenting customers, calculating the value of each segment for life, and estimating the future value of each segment [14].

In [15] the author explained the groupings that have been carried out in detail. The techniques discussed mostly depend on the characteristic features of the database. In analyzing the actual data, the algorithm should be efficient in large multidimensional data sets, noise, and outliers. The researchers clearly explained this phenomenon in this paper discussing PFCVI, which took a lot of results from a shared partition and gave the advantage of analyzing information on the nominal amount of savings / balance in a bank account. Finally, we created profiles of each user and their categories and postulated references for banks [15].

In [16] the author explained Islamic BPRs in Indonesia which carry out their business activities are based on sharia principles and there is no payment service. The researcher explains that the company must evaluate its capital (balance) to survive and compete with customers. The level of bank performance can measure the health of the bank. The variables used in measuring bank performance are capital (C) and others that are part of the CAMELS method [17].

The results of the analysis from the financial side of the BPRS such as capital, productive assets, third party funds, and increasing credit have a great opportunity to change for the better. The results of the cluster feature (CF) tree were analyzed by hierarchical groups using the agglomeration method, and the calculation of the number of groups in the Bayesian information criteria (BIC) for each group [16].

4. Discussion of Research Questions

RQ1: What can balance component be operated in customer segmentation?

Based on the paper reviewed in this study, various types of attribute usage that can be interpreted as balance attributes are obtained. They were obviously described in Table 3. The use of these balance attributes described by this paper has the same goal of supporting the process of segmenting or grouping customers. [7] discussed that customer balance was calculated and was an object of research analysis which considered as revenue and contributes to company profits. While [8] discussed internet banking user balances functioned in customer segmentation and clustering with data mining techniques, and [9] technically expressed the average customer balance operated for mining and computing methods for specified CRM.

In addition, [10] described the Min-Max balance component in effort to increase company revenue. Customer balances in the dataset officially exploited to classify customer loyalty were discussed in [11], while [12] discussed the balance as monetary benefited in the clustering method. In [13], balances and transactions were two features that widely utilized to discover potential customers in the KDDM process.

Also, the balance before transaction discussed in [14]. It was used in the customer segmentation framework for customer lifetime value (LTV). [15] explained the analysis of information regarding the nominal amount of savings / account balance and becomes a reference for banks. Finally, [16] discoursed a capital (balance) as well as a variable or attribute exploited in practically measuring a company's financial performance.

Table 3. Approaches of customer segmentation methods

Number	Attribute	Description
1	Balance of Customer	Calculated, income from the customer's loan, for the bank profit contribution
2	Balance of IB users	Balance for customer segmentation and clustering in data mining technique
3	Average customer balance	Average customer balance used for mining and computing methods for CRM
4	Min-Max balance	Min-Max balance component in an effort to increase company revenue
5	Customer balance	Customer balance in the dataset to classify class of customer loyalty.

6	Balance as Monetary	Balance value as monetary for clustering method
7	Balances and Transactions	Two features are widely used to find potential customers in the KDDM
8	Balance before transaction	Used for the Bank Customer segmentation framework, on customer LTV
9	Value of savings	Create user profiles for each category and provide some advice for banks
10	Bank capital balance	To measure the health of the bank, and for the bank's financial performance

Discussing about the balance attributes used in the segmentation or clustering process will require analytic models that can process them and suitable methods so as to produce optimal information and results.

RQ2: What are analytical methods functioned in customer segmentation?

From some papers that have been reviewed, there are several models and methods used in segmentation. From some papers that have been reviewed, there are several models and methods used in segmentation as show in Table 4. Based on the paper that has been reviewed, there are several models and methods used in segmentation. The popular method used and appears in many discussions of this literature shows that the method is indeed suitable for use in customer segmentation. The RFM model appears 4 times from the papers selected for discussion, namely in the papers [8], [9], [10], [12], while the K-Means Algorithm appears most and is used in this literature as many as 6 times in papers [8], [9], [10], [11], [12], [14], and the most popular of which were used and appeared 2 times, namely C-Means in papers [10], [15].

Table 4. List of Popular Methods / Models

Number	Popular Methods	Paper
1	RFM models	[8], [9], [10], [12]
2	K-Means Algorithm	[8], [9], [10], [11], [12], [14]
3	C-Means	[10], [15]

Search results in the digital library showed 592,363 the first stage, filtered to 1,361 in the second stage, 21 papers in the third stage, and the final stage selected 10 papers that answered the research questions. In conclusion from the 10 papers, customer balances can be used as attributes for customer segmentation, and 3 analytical methods most often used; RFM models, K-Means algorithm, and algorithm and C-Means.

5. Limitations

The number of papers selected and reviewed may have a weak validity of the findings. One reason is in the number of digital libraries that discuss this topic is very much which will be able to show most of the other sophisticated methods or models in customer segmentation and / or the methods they use. However, this will also lead to a minimum number of relevant papers going through the screening process. Obviously, the search, selection, and quality assessment are very strict and reject some papers that do not meet some assessment questions.

6. Conclusion and Future Works

From the results of a systematic literature review that has been done, customer balances are included or used as attributes in the customer segmentation process. Responding to the RQ1

research question, all papers that have been reviewed show that balances are used as attributes in the customer segmentation process. This attribute is the total balance, the value of savings, the amount of capital, or the average value of the balance. There are 3 popular methods that are widely used in the customer segmentation process of the papers that have been reviewed. Future research can be carried out to improve the validity of the findings by including more digital libraries or expanding on inclusion criteria.

References

[1] B. E. Adiana, I. Soesanti and A. E. Permanasari, "Analisis Segmentasi Pelanggan Menggunakan Kombinasi RFM Model dan Teknik Clustering" JUTEI Edisi Volume 2, no. 2, 23-32, 2018. <https://doi.org/10.21460/jutei.2018.21.76>.

[2] S. M. Kostic, M. Duricic, M. I. Simic and M. V. Kostic, "Data Mining and Modeling Use Case in Banking Industry" in 2018 26th Telecommunications Forum (TELFOR), Belgrade, 1-4, 2018. <https://doi.org/10.1109/TELFOR.2018.8611897>.

[3] A. Subrahmanyam, "Big data in finance: Evidence and challenges" Borsa Istanbul Review, vol. 19, no. 4, pp. 283-287, 2019. <https://doi.org/10.1016/j.bir.2019.07.007>.

[4] M. Tavakoli, M. Molavi, V. Masoumi, M. Mobini and S. Etemad, "Customer Segmentation and Strategy Development based on User Behavior Analysis, RFM model and Data Mining Techniques: A Case Study" in 2018 IEEE 15th International Conference on e-Business Engineering (ICEBE), Xi'an, 119-126, 2018. <https://doi.org/10.1109/ICEBE.2018.00027>.

[5] R. P. Ghozali, H. Saputra, Suharjito, D. N. Utama and A. Nugroho, "Systematic Literature Review on Decision-Making of Requirement Engineering from Agile Software Development" in 4th International Conference on Computer Science and Computational Intelligence 2019 (IC2SCI), Jakarta, 274-281, 2019. <https://doi.org/10.1016/j.procs.2019.08.167>.

[6] R. Bria, A. Retnowardhani and D. N. Utama, "Five Stages of Database Forensic Analysis: A Systematic Literature Review" in 2018 International Conference on Information Management and Technology (ICIMTech), Jakarta, 246-250, 2018. <https://doi.org/10.1109/ICIMTech.2018.8528177>.

[7] Y. Jinping, "Analysis and Improvement Strategy for Profit Contribution of Bank Customer Under Big Data Background" in International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), China, 338-341, 2019. <https://doi.org/10.1109/ICITBS.2019.00089>.

[8] M. Aryuni, E. D. Madyatmadja and E. Miranda, "Customer Segmentation in XYZ Bank using K-Means and K-Medoids Clustering" in International Conference on Information Management and Technology (ICIMTech), Jakarta, 412-416, 2018. <https://doi.org/10.1109/ICIMTech.2018.8528086>.

[9] F. A. Bachtiar, "Customer Segmentation Using Two-Step Mining Method Based on RFM Model" in International Conference on Sustainable Information Engineering and Technology (SIET), Malang, 10-15, 2018. <https://doi.org/10.1109/SIET.2018.8693173>.

[10] A. J. Christy, A. Umamakeswari, L. Priyatharsini and A. Neyaa, "RFM Ranking – An Effective Approach to Customer Segmentation" Journal of King Saud University - Computer and Information, 2018. <https://doi.org/10.1016/j.jksuci.2018.09.004>.

[11] S. Moedjiono, Y. R. Isak and A. Kusdaryono, "Customer Loyalty Prediction In Multimedia Service Provider Company With K-Means Segmentation And C4.5 Algorithm" in International Conference on Informatics and Computing (ICIC), Mataram, 201-215, 2016. <https://doi.org/10.1109/IAC.2016.7905717>.

[12] Dedi, M. I. Dzulhaq, K. W. Sari, S. Ramdhan and R. Tullah, "Customer Segmentation Based on RFM Value Using K-Means Algorithm" in International Conference on Informatics and Computing (ICIC),

Semarang, 1-14,
2020.<https://doi.org/10.1109/ICIC47613.2019.8985726>.

- [13] G. Mansingh, L. Rao and K.-M. Osei-Bryson, "Profiling internet banking users: A knowledge discovery in data mining process model based approach" *Information Systems Frontiers* volume, vol. 17, no. 1, pp. 193-215, February 2015.<https://doi.org/10.1007/s10796-012-9397-2>.
- [14] W. Qadadeh and S. Abdallah, "Customers Segmentation in the Insurance Company (TIC) Dataset" in *INNS Conference on Big Data and Deep Learning* 2018, Edinburgh, 277-290, 2018.<https://doi.org/10.1016/j.procs.2018.10.529>.
- [15] J. Zheng, H. Cui, X. Li, L. Meng and T. Wang, "The Clustering for Clients in a Bank Based on Big Data" in *2018 4th International Conference on Universal Village (UV)*, Boston, 1-5, 2018.<https://doi.org/10.1109/UV.2018.8642136>.
- [16] M. F. Nazar, Maiyastri, D. Devianto and H. Yozza, "On the Clustering of Islamic Rural Banks Based on Financial Performance" in *International Conference on Applied Information Technology and Innovation (ICAITI)*, Padang, 108-113, 2018.<https://doi.org/10.1109/ICAITI.2018.8686755>.
- [17] A. Shaddady and T. Moore, "Investigation of the effects of financial regulation and supervision on bank stability: The application of CAMELS-DEA to quantile regressions" *Journal of International Financial Markets, Institutions and Money*, vol. 58, pp. 96-116, 2019.<https://doi.org/10.1016/j.intfin.2018.09.006>.

Estimation of Influential Parameter Using Gravitational Search Optimization Algorithm for Soccer

J. Vijay Fidelis^{1,*}, E. Karthikeyan²

¹Computer Application, Presidency College, 566024, India

²Department of Computer Application, Government Arts College, Udumalpet

ARTICLE INFO

Article history:

Received: 10 February, 2020

Accepted: 09 May, 2020

Online: 11 June, 2020

Keywords:

Evolutionary Algorithm

Gravitational Search Algorithm

Soccer

Heuristics

Optimization

ABSTRACT

Competitive sport has one phenomenal or fundamental aspect of selecting players into playing squad for a game that can influence a Club or a team in almost all major aspects. Various Characteristics or behavioral aspects of players will be instrumental towards the selection of a specific player into a team depending on the nature, level, or type of completion the club or team participates in. Many parameters such as medical, physical, technical and, Psychological aspects of players make the task of managers or coach a herculean to select 15 players out of 30 or 40 players available in his squad for a particular season. The role of managers or coaches is significantly challenging looking into the aspects most desirable towards the optimal contribution of players. Hence the parameters which are considered highly influential towards a Club or team cannot be analyzed manually due to various constraints such as time, the volume of players, or the limitation of human errors in decision making. The primary objective of this paper is towards assisting managers or coaches to see through this by applying Sports Parameter Estimation Gravitational Search Algorithm (SPEGSA) towards analytical ability in player selection considering minimal errors and time constraints using a stochastic approach. This paper gives an overview of how soft computing techniques help in optimization of selection procedures of team players for the matches to be played and competed in a soccer league for a given team at different levels of competition by measuring various influential parameters recorded at different point of juncture for every player in a team and estimating the parameter using the subset of evolutionary computation techniques and metaheuristic optimization algorithm.

1. Introduction

Soccer or the term synonym to Football has its evolution as a team sport way back more than 200 decades ago. It originated in Asia. Greeks, Romans, and parts of Central Americans also claim having initiated the sport Moosavian et al. in [1]. Eventually, it was the British who transformed and revolutionized the game of soccer into the game perceived and played today.

Any game requires a set of rules or guidelines which need to be institutionalized and British are the ones who will take this credit who regulated and imposed basic rules needed for the same, which included forbidding tumbling the opponents, to hold or touch the ball with hands when the game is at play Moosavian et al. in [1]. As the timeline increased there required more amendments in soccer and needed more rules which were to be

implemented. One of them being penalty kick which was introduced way back in 1891 which penalized players any offense inside the D area. When a spot-kick or goal kick is taken in the penalty area, no other player other than the goalkeeper is allowed inside (shown in "Figure 1").

The Federation International Football Association (FIFA) was registered as an official member of the International Football Association Board during the year 1913. To caution players for bad conduct on the field, referees were given the privilege of penalizing players at the game of play with caution cards. These cards were introduced during the 1970 World Cup finals which included a yellow card for caution and a red card for dangerous play. Recent changes included goalkeepers being banned from handling deliberate back passes towards the D Area, tackles from behind, and obstruction of the goalkeeper when a goal is being scored. These amendments did not bring down the viewership of

*Vijay Fidelis, Presidency College, Bengaluru, veefeed@gmail.com

the games but have helped gain more popularity and momentum over any other game. For instance, Cricket which started with Test format in its early stages came down to 50 over format then to 20 over and recently to 10 over format making it far more reachable to audiences throughout.

Over 250 million players spanning over 200 countries and dependencies play soccer, which makes it the world's most popular sport. As per the recent statistics on the top ten most popular sports in the world, with a fan following the base of almost 3.5 billion from every continent, the sport remains at the top with no other game coming close to it. And for the reasons above, soccer has its mainstay in most countries because of its cost-efficiency. Soccer just requires a ball, which will be tactically kept in procession by a team on the field which will be comprised of 11 players and their opponents in equal numbers. Another reason for soccer being such a popular game throughout the globe is that the rules have been very adaptive and simple to adhere to. Also, the rules have not been tampered or amended in terms of usage on a major basis, making the game more of a learning process than changing one.

The field area played in each game is 100 meters in length and 60 meters in width. As per the recent survey on the ground dimension of length and width, the conducive area of play depends on the minimum possibility of length and width for a full-fledged game. And hence every ground varies with the maximum amount of space available for a ground to be built keeping in mind the spectators to be seated while watching a game.

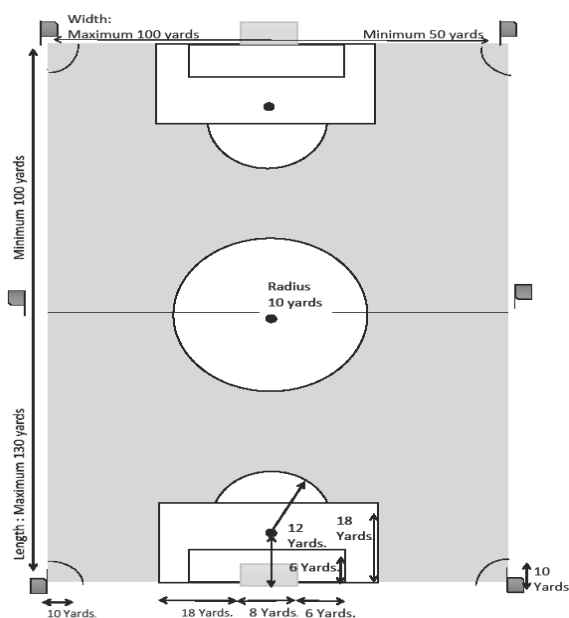


Figure 1: Soccer Field.

Soccer at school, College or, Club level has always generated interest only in a league format because 10 teams participating in the tournament. Each team tends to play each other minimum once or maximum twice as the 1st leg usually considered a home game and 2nd leg considered as away game. Teams usually have their home ground in their region of practice zone and only in a league format can we know the real quality of team performance. Such kinds of league games are played not only in Soccer but in other games such as Basketball, Hockey, Cricket, Kabaddi, and other www.astesj.com

competitive sport. These games have always drawn huge crowds from their respective home pitch audience and harness huge interest. These also involve a lot of sport management strategies.

1.1. Role of Coach/ Manager

Coach or manager for any team plays a very important role and spearheads the main responsibility of how to handle a potential squad of players in all possible means for an entire season. Coach also has the mentorship and administrative role to guide and mold players and bring in their best in all the matches to be inducted into. Soccer is taken in Club level matches; an elite team needs to play at different levels to attain mass recognition and importance in the said field. Fanfare has always surrounded the elite teams as they promise in building not only a team in terms of acquiring better and best players around the globe but keeps the money bank ticking for the club. Most of the time it is the fans that are responsible for any management based decision to rope in better coach or manager for the club in replacement of the other if the desired results are not met.

The yearly season is usually a time spanning from August of every year to June of consecutive year. Within which a winter break of 20 days will be applied in December. So when inking a deal with any club, a player needs to be available for the entire season and provide his optimal performance in every game of which the club represents at different levels. Elite clubs always give the complete role of responsibility to coach/manager in acquiring the best possible players and strengthening the squad and see through the possibility of player availability for crucial matches. The squad acquired by a coach/manager cannot be changed at every point of time once the transfer season is complete. The transfer season usually happens before the season commences and each club's coach/manager is liable to enable talks with other clubs in acquiring their desired players, if the other club obliges with the offer, well and good or else have to enhance their offer in terms of money to ink the deal. If in the worst case the squad selected by the coach/manager doesn't fare well in the league matches can avail his winter window session transfers which usually happen in December.

Playing a balanced squad for any match in whatsoever circumstances has always been an important aspect and a prerequisite of any coach/manager yielding to a win (3 Points) preferably and on worst-case scenario a draw (1 Point). A win won't be compromised at any cost and will be usually expected by the team owner and the main stakeholders that are committed and fans of the club who won't hold on to a loss on a more frequent note. Due to various reasons, there might be situations wherein the coach takes dubious decisions while picking the squad and in this regard, might elevate players of his choice while discriminating against others even when the credentials of the players not selected are well above the ones selected.

2. Review of literature

Darren J. Paul and Geroge P. Nassis in their paper titled, Testing strength and power in soccer players: The application of conventional and traditional methods of assessment Darren et al. in [2], have measured various factors such as physical, tactical and psychological aspects measuring the influences of them in Soccer and emphasize the importance of strength and power as key factors assessed in a soccer club/team. The work revolves around

selection factors such which are valid and reliable which makes a trustworthy analysis of implementing the system. Since the tests carried out in this process are having moderate to high levels in soccer players at several of either training or matches without compromising on the accuracy and minimized time delay. Hence the implementation phase of the paper provides monitoring of players at more specified intervals of time only in a specified season rather than specific time intervals. The major aspects taken from this paper finds a limited way as on how traditional or manual method might be a very moderate way of analyzing players evaluation factors as compared to various other parameters what SEPGSA can achieve and another main constraint is that the period of applicability is only confined to a particular season, rather than the complete history of a player once he has got into the professional level as discussed towards implementing SEPGSA method. Miguel Angel Perez Toledano, Francisco J. Rodriguez, Javier Garcia Rublo and Sergio Jose Ibanez in their work on Player's Selection for Basketball teams through performance Index rating, using multiobjective evolutionary algorithms Miguel et al. in [3], discuss on the stochastic methods deployed based on evolutionary techniques towards the selection of players for a basketball team. Financial limitations, sports characteristics, and participation of teams in various levels of competitions hence making the process of player selection very complex since multiple variables are involved in subjectivity.

Bukhari et al. in [4], discuss the various characteristics which enhance the players who are recruited in a team. Parameters include the composition of body parameters, vision-based tests, psychomotor tests and various and anaerobic activities to make a good sports team. The process is discussed in the paper as very difficult and time-consuming for team coaches towards individual players to be recruited in a team. To measure such parameters there is a huge limitation since human judgment is more biased and error-prone hence making a huge limitation on the same. To overcome such kind of limitations the authors have gone with fuzzy logic in application towards the same.

Bastien Talgorn et al. in [5] their work on statistical project formulation for simulation-based design and Optimization. In which the direct search algorithm is compared to that of surrogate-based algorithms that are completely oriented towards models based on surrogate methods. The main problem discussed here is on the evaluation of black boxes which will be used for Optimization and is quintessential simulation-based function since black boxes have various challenges deployed such as noise, high computational cost, and other constraints. The article has its major contributions towards formalizing surrogate problems which specifically Dyna-tree models. A total of 8 formulations were proposed in this paper out of which 3 are constrained while rest are considered using statistical criterion. The major work involved here is to compare the performance using 20 benchmark problems in which the functions are constrained towards simulation aspects of aircraft design. This novel problem using statistical surrogates combining it with a derivative-free Optimization algorithm towards simulation Optimization where the statistical features and exploring better design space. The experiments conducted towards appropriate use of surrogate models could be better improvised with considering Optimization concepts involved for the same.

Sethuraman sankaran et al. in [6] their article on a method for stochastic constrained optimization using derivative-free surrogate pattern search and collocation. The major aim of this work is towards developing computational Optimization technique which can be practically experimented on stochastic problems to evaluate cost functions. Since the complexity of the problem is for large scale equations. Various types of such problems such as biological applications of Solid Mechanics where surrogate models have been used and in the current article derivative-free surrogate Framework which embeds the methods for robust designing. The major drawbacks for gradient-based are not feasible for problems that are more complex on physics which have constraints such as large scale simulation, noise, and gradient information.

Pooriya Beyaghi et al. in [7] their article on "Delaunay-based derivate-free optimization via global surrogates, Part I: linear constraints, where a new optimization algorithm based on a derivative-free method for more non-convex functions are bounded by linear constraints. The algorithm which is developed in this research paper is extended to Part II since the problems which can be more convex constraints derivative-free methods are well suited where derivative or approximation is available on an instant nature. Direct search methods are a class of algorithms in which derivative-free methods are well suited for problems where derivative or approximation is available in an instant nature. Direct search methods are a class of algorithms in which, derivative-free algorithms are classified. The Nelder-Mesh Simplex algorithm was used for numerical optimization. The adaptive direction search algorithm was another category of direct search methods and more modern methods such as Rosenbrock and Powell methods, Pattern search methods that are characterized by various parameter spaces. The direct search methods are a model of the actual function of summarizing the data points. The Kriging method was one of the popular surrogate functions used in global optimization schemes which will automate in building estimate and uncertainty involved in this estimation.

Charles Audet et al. in [8] their paper on Mesh-based Nelder-Mead algorithm for inequality constrained optimization discusses the lack of support that is involved in theoretical and practical convergence and the Nelder-Mead algorithm considered towards solving unconstrained optimization problems. The current work proposes a method to enhance the Nelder Mead algorithm towards replacing the worst point of a simplex method. The methodology is to use the search step of mesh adaptive direct search which is inspired by the Nelder-Mead algorithm. It doesn't exhibit any limitations of convergence but overcomes this by using convergence analysis present in mesh adaptive direct search algorithm.

3. Need for Gravitational Search Algorithm for Soccer

When we consider a deterministic function $f: R^n \rightarrow R$. This could be used over any domain of interest on the lower and upper bound of the variables. Parameters considered for the work involves more of dynamic and constantly evolving variables and hence various parameters towards selecting the type of Optimization

technique which could be deployed for the current work. Comparison of bench-marking of the types of Optimization algorithms which is daunting and involves a high level of complexity first level of unbiased that could arise should be fair while evaluated. The number of Optimization algorithms and its implementation in various domains has evolved exponentially in relevance to the time taken. Comparative studies carried forward by researchers in various domains and Fields have been tested and experiment at different levels by carrying out experiments which have yielded result in an improved for the enhanced manner and considered as great help towards scrutinizing of algorithms based on Optimization for problems related to a specific domain of field and enhance the bench-marking of such technique in optimization-based algorithms.

There are various types or classes of Optimization algorithms that are distributed in network planning methods of which they are classified into two broader perspectives.

- Classical Optimization algorithms,
- Meta heuristic-based Optimization algorithms

In classical Optimization algorithm, there are a few which are listed below

- Transportation algorithm
- Branch and bound algorithm
- Mixed-integer programming
- Derivative free methods.

And in Meta heuristic-based, there are all the evolutionary algorithms which are sub-classified classified into,

- Trajectory based algorithm
- Population-based algorithm

The classical Optimization such as mixed-integer, branch and bound, derivative-free methods are summarized upon the lower and upper bounds on the problem variables which are of high importance. In a derivative-free Optimization technique which is applicable when the information of the derivative function is not available, not complete, not reliable or practically not feasible to obtain. When we consider a function derivative 'd' which contains noise or expensive to be evaluated practically confined to very few usages and differences, such type of problems is referred to as derivative-free Optimization. Another vision towards the derivative-free method is to that in the algorithm which involves derivatives of those particular functions other than the evaluating function 'd'. The problem that we consider has no such implications of uses of the derivative tree method since the function revolves more on data that is more certain unambiguous in many situations.

In the local search methods which are also one of the methods in classical Optimization, we see various methods of direct local search methods such as Hooke and Jeeves, Nelder-Mead Simplex algorithm, Mesh Adaptive direct search methods, Trust-region methods, Implicit filtering. The Global search methods which are also based on classical Optimization techniques based upon the

construct and optimization of a function which underestimates their original one. Various algorithms such as Lipschitzian based partitioning, Branch-and-Bound method where it partitions the search space and will find the lower and upper bound values and the inferior 1 are eliminated. The branch and bound method will not be practically feasible to apply since it eliminates the inferior function which cannot be done in the proposed work since we are concerned with not eliminating any player who is bought by a club or part of a college team, but it on how we could improvise the player in the best possible way. The surrogate management framework is a classical Global search Optimization method in which the accuracy of the surrogate model is done by stepwise search to produce optimal points. It evaluates the candidate function along with the surrogate model and does the measurement. The branch and fit method also deploy the concept of surrogate models and randomization which are combined. Optimizing the models based on the evaluation of other points versus candidate points is used. The above algorithms will not be suitable for our research work since we are experimenting with a dynamic set of parameters with constant self-reference to the methods and functions and peer references in the data set.

Simulated annealing is a metaheuristic trajectory-based method, each iteration the algorithm generates the latest trial, which is compared with the incumbent and is accepted with probability function. Simulated annealing was initially proposed for handling problems based on Optimization and has been experimented and results obtained for global optimum have been recorded. But the guaranteed solution is not sustained and a good solution cannot be obtained within a finite set of iteration. Simulated annealing is not practically feasible for the research work carried out.

We have used the population-based evolutionary technique of the Gravitational Search Algorithm (GSA) which has features such as non-deterministic function as what is required in player selection from the given data set. The GSA algorithm is agent-based as the direct methods which cannot be practically created to. The position of every agent is important towards the system state and hence can be directly co-related to every player position which affects the team. The Other important aspect in GSA is that the agent's absolute fitness is directly dependent on system state as we can correlate the fitness of every player which is important towards the overall aspect of a club or a team. This aspect is not available at direct methods. An important and valuable aspect is that as compared to the direct method is completely inform the method as contradicting to direct methods that have very limited information. The rigorous solution of Agent dynamics which is feasible in GSA is not practically feasible in direct method and hence the selection of GSA towards initial simple population-based to completely normalize on what needs to obtain as result towards estimating the influential parameter.

4. Gravitational Search Algorithm (GSA)

4.1. Introduction to GSA

Gravitational search algorithm (GSA) which is based on soft computing technique is exhilarated on a conceptual skeleton and

more relevantly derived from nature and has its origins in gravitational kinematics Kelton et al. in [9]. The model is based on the motion of masses which moves under the influence of gravity Abhinav Sharma et al. in [10]. The preeminent concept of GSA is in the collection of objects which internally interact with each other because of various influences of objects under the Newtonian gravity and principles of motion Damodar Reddy et al. in [11]. Masses measure the performances of objects. When there are plenty of objects in a particular region the gravity, force, and masses of all these objects play an important role as the influence of the objects and by the law of attraction and forces a global movement of all objects toward other objects of heavier masses P.B. De Oliveria et al. in [12]. The position of every object corresponds to a solution to the problem. During iteration, the position of the objects is updated and the best value along with the corresponding object is stored P.B. De Oliveria et al. in [12]. As per the law of physics, heavier masses tend to move more calmly or lazily a more calm or lazy way than the lighter ones Tongzaing et al. in [13]. After a specific number of iterations, the algorithm terminates before which the best fitness value transforms to global fitness value for estimating the influential parameter of players' goal-scoring capabilities and the positions of each corresponding object yield a global solution. Rajendra Kumar et al. in [14].

Heuristics and optimization play a very important role in any of the research fields of work which needs to be done Mohammad Massoud et al. in [15]. It is required for enabling to infer new knowledge out of the existing ones Ehsanol et al. in [16]. Without heuristics, it would be hard to conclude that none could learn many things that are not possible. Optimization is an entity where the quality of any work just doesn't depend on finishing it alone but in making the best and effective use of all the available resources Ehsanol et al. in [16]. Optimization completes the task but finishes the task quickly and accurately Tolba et al. in [17]. Hence concerning both heuristics and optimization, many algorithms are designed and developed using heuristics and optimization concepts and Gravitational Search Algorithm very much ascertains the fact of it. GSA because of its practical application as a nature-based algorithm has gained tremendous significance amidst the research and scientific community.

4.2. Working of SPEGSA

SPEGSA, as the name suggests, is very much taken all its inspiration from Mother Nature as it is established on one of Newton's laws of gravity and law of motion Xing et al. in [18]. Most of the optimization algorithms have a high usage of Animals or insects and their different techniques in optimization since they make use of the best of the resources when it comes to optimization. Whether it is Ant Colony Optimization, Particle Swarm Optimization, Whale optimization, and many others, have seen the same techniques in technology as it can reflect on a real-world problem to an optimized level Alli A et al. in [19]. SPEGSA has more than enough credibility in providing a very accurate, effectual, and powerful high-quality solution for any such optimization-based problems Gonzalez et al. in [20].

SPEGSA is categorized under a population-based method due to various objects involved. SPEGSA is made up of certain things termed as particles, which are primarily comprised of four major parameters. That is the position, inertial mass, active gravitational mass, and passive gravitational mass Kyuchang Kang et al. in [21].

One of the unique features of SPEGSA is that it is a memory-less algorithm that does not compromise on efficiency as compared with any other memory-based algorithms R. Priyadarshini et al. in [22]. Since SPEGSA is in its infancy stage, more studies will be based on it. There is a high potential of the algorithm to provide optimal solutions for various problems in any possible domain.

G, which is gravitational constant computed in iteration t, is computed as follows,

$$G(t) = G_0 e^{-\alpha/T} \tag{1}$$

Where G_0 and α need to be loaded before search. Their values will be gradually decremented during the search Esmat et al. in [23]. T is the total number of iterations. The masses of every object will abide by the law of gravity as mentioned in the equation "(2)".

$$F = G \frac{M_1 M_2}{R^2} \tag{2}$$

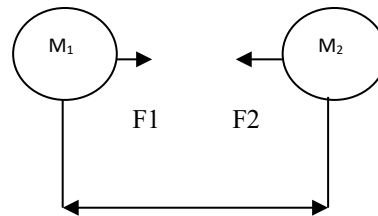


Figure 2: Force Acting between Mass M1 and M2.

Equation "(2)" represents Newton's law of gravity, where, F is a magnitude of gravitational force acting upon various objects. G is gravitational constant. M_1 and M_2 masses of first and second respectively objects Seyed et al. in [24]. R is the distance between two objects M_1 and M_2 . Newton's second law states that, when force F is applied to an object, the object moves with an acceleration 'a' depending on the applied force and the object mass 'M' as in the equation "(3)".

$$a = \frac{F}{M} \tag{3}$$

Masses are of three kinds, Active gravitational mass M_a , Passive gravitational mass M_p , and Inertial mass M_i . The gravitational force F_{ij} that acts on mass i by mass j is defined by,

$$F_{ij} = \frac{GM_{aj} \times M_{pi}}{R^2} \tag{4}$$

M_{aj} and M_{pi} are active and passive masses of objects j and i respectively.

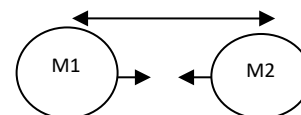


Figure 3: Gravitational Mass.

The acceleration of the object will be calculated as below.

$$a_i = \frac{F_{ij}}{M_{ii}} \quad (5)$$

Where M_{ii} is inertia mass of the particle. While searching the particles update their velocities and positions as below equations.

$$V_i(t+1) = rand_i \times V_i(t) + a_i(t) \quad (6)$$

$$X_i(t+1) = rand_i \times V_i(t) + a_i(t) \quad (7)$$

4.3. Pseudocode for SPEGSA

Step 1: Initialize Player Gravitational constant G_0 , a , ϵ along with iteration counter t .

Step 2: Randomly generate the initial population which consists of N particles, the position of each particle defined by,

$$X_i = (x_{i1}(t), x_{i2}(t), \dots, x_{id}(t), \dots, x_{in}(t)) \text{ for } i=1,2..n$$

Step 3: Iterate until termination criteria are satisfied.

Step 4: Evaluation of all particles in the population is done and the best, worst particles are assigned.

Step 5: Updating of player gravitational constant by $G(t) = G_0 e^{-\alpha/t}$

Step 6: The force applied when particle j acts upon particle I at a specific time (t) is calculated.

$$F_{ijd}(t) = \frac{G(t)M_{pi}(t) \times M_{aj}(t)}{R_{ij}(t) + \epsilon} (X_{jd}(t) - X_{id}(t))$$

Step 7: Total force acting upon particle i at iteration t is calculated.

$$F_{id}(t) = \sum rand_j F_{ijd}(t), j \neq i$$

Step 8: Calculate the inertial mass.

Step 9: The acceleration of the particle 'i' is calculated.

Step 10: Computer the velocity and position of the particle 'i'.

Step 11: Loop steps until termination criteria are satisfied.

Step 12: Optimal Solution attained in the process.

5. Proposed Methodology

5.1. Experimental Setup

This study makes use of the SPEGSA model and proposes an efficient way of finalizing the players required to be fielded for a game in soccer. This takes into consideration various attributes as variables of players, Jersey Number (JN), Age, Goals Assisted (GA), Goals Scored (GS), and Goals Saved (GSV). The model helps the coaches and managers in helping them to select the best 11 players for the next game in a league based on various parameters. The data set is passed through different numbers of particles and iteration in variation so that the results obtained are varied and recorded. The parameter selection for goal scoring forecasting in which four input parameters were used for estimating the players to be shortlisted and in forecasting which was for a period of one season, such as Jersey Number, Age, GS, GA, GSV since the forecasting and evaluating of current players have been made for the next possible match.

5.2. Dataset

The dataset available for the current season for a particular team playing in 'A' division football tournament and the College football team is considered and the range of values for a particular

attribute is given. The parameters which are used in the present study are:

i. GA (Range: 0 to 8), ii. GS (Range: 0 to 8), iii. GSV (Range: 0 to 9), JN(Range: 1 to 23).

In the training process, the Dataset is passed through different values of NOP (Number of Particles) to calculate the mass and variations in the number of iterations. As with GSA, the data set will result in different approximations of best mass, and hence calculating it further will estimate the amount of goal scored and the team to be selected. The dataset has been trained for the following values of NOP and iteration, NOP=70 and iteration=900, NOP=100 and iteration=1000, NOP=70, and iteration=700, for which the best mass values were recorded and evaluated further. The GSA model is proposed which takes effect into all the variables until the current season data is defined in the equation "(8)".

$$F_t = D_{a0} + D_{a1} * JN + D_{a2} * Age + D_{a3} * GA + D_{a4} * GS + D_{a5} * SPEGSA \quad (8)$$

Where D_{a0} , D_{a1} , D_{a2} , D_{a3} , D_{a4} , and D_{a5} are various coefficients that will be evaluated using SPEGSA and F_t determines the fitness function to find the number goals to be scored by a player.

5.3. Estimating of Coefficients by SPEGSA

The main purpose behind this implementation is to enhance upon the aforesaid coefficients based SPEGSA algorithm from the player's record on historical grounds of Goals Scored data. The parameters in total used in this model show the effect of the above-mentioned variables. The model parameter can be evaluated to reduce the error between the current goal scored by a player and the simulated goal-scoring estimation output using the fitness function f , for SPEGSA is defined in the equation "(9)"

$$f = \min (\sum [GS \text{ Estimation} - \text{Actual GS}]^2_{ni} = 1) \quad (9)$$

Where n represents the experimental data set values.

5.4. Metrics and Performance

Our main objective is to compare the actual goal scored by a player in addition to other attribute values such as Goal Saved and Goal Assisted which will influence the coach or the manager to take unbiased decisions on team selection for optimal performance. Hence the error between the actual goal scored and estimation of goal scored is forecasted using i) Mean Absolute Error (MAE) ii) Mean Absolute Percentage Error (MAPE) iii) Mean Square Error (MSE) which are defined as follows.

$$MAE = \sum (X_t - F_t) / n = \sum et * n \quad (10)$$

$$MAPE = \sum |X_t - F_t| / X_t * 100 = \sum |et| / X_t * 100 \quad (11)$$

$$MSE = \sum |X_t - F_t|^2 / 2n \quad (12)$$

X_t is original data at period t , F_t is the estimation at period t , et is predicted forecast error at period t , while n is the number of observations.

6. Experimental Results

Estimation of goal scored by a player along with goal saved and assisted will optimize upon the best 14 players to be fielded at the time when maximum required and form the historical data of a player in the current season. These experimental evaluations of players might not only be pivotal for the manager or coach to select a player for the rest of the matches in the league but indeed will provide added information for selecting.

The particles in the problem space were assigned with the initial random values ranges between 0 and 1. Each particle will have its initial random position and velocity. Each particle will move place to place based on the amount of velocity and its position. The particle will have its own memory space to share their knowledge to reach the closest local best and hence reaches the global best. So, the minimizing function equation “(3)” reaches its global minima.

Table 1. Parameter Estimation Player Selection data model using SPEGSA

NOP/Iteration	D _{a0}	D _{a1}	D _{a2}
	20/ 100	-0.5209	0.4979
50/ 400	0.72929	-0.0724	0.02214
70/ 700	0.53006	0.15263	0.00455
70/ 900	0.34065	-0.6285	0.18529
100/ 1000	-0.0592	0.04748	0.25852

In SPEGSA Algorithm, every particle will have the following, its location, inertial mass, active gravitational mass, and passive gravitational mass. The SPEGSA parameters used in the Goal scoring model uses the number of mass as 70, initial position and velocity are random values between 0 and 1, gravitational constant (G) is 1.

Table 2. Parameter Estimation Player Selection data model using SPEGSA

NOP/Iteration	D _{a3}	D _{a4}	D _{a5}
	20/ 100	1.5004	0.1972
50/ 400	-0.7272	-0.6975	0.4576
70/ 700	0.03521	0.25772	0.3956
70/ 900	0.64474	1.07111	-0.1896
100/ 1000	0.19907	1.27980	-0.2068

Because of the random values generated at every point of execution have always got the best mass deterred from the previous one, which is required in this implementation. SPEGSA algorithm concept helps to explore and exploit the optimal values. The exploitation feature helps the particles in a particular region or environment to move towards another particle which has the heavier mass within the same. The values for G and M were calculated to update the position for each iteration until it reaches the optimal solution. The obtained results were tested using evaluation metrics, from different parameters that were available in our dataset such as, the actual goal scored, estimation from SPEGSA, and Fitness function.

The estimation of coefficients in “Table 1” and “Table 2” is for different values of NOP and Iterations respectively.

- NOP=20 iteration=100
- NOP=50 iteration=400.
- NOP=70 iteration=700.
- NOP=70 and iteration=900.
- NOP=100 and iteration=1000.

Players for a club or team at the time of bidding. When a player is bought for 100 million pounds, what will be the impact of such players for the club to improve upon will be a rather difficult scenario when bidding of a player happens? Hence the iterative method of multiple linear regressions using GSA was proposed to find the coefficient for Da0, Da1, Da2, Da3, Da4, and Da5 which is the current player record for the season.

Using these coefficients in the model equation “(4)”, the players selected on the number of goals scored is estimated for the known 23 data set of players, results once obtained were plotted in “Figure 4”, “Figure 5”, “Figure 6”, “Figure 7”, and “Figure 8”, respectively. The performances MAE, MAPE, MSE for the validated data of both the models were tabulated in “Table 1” and “Table 2”.

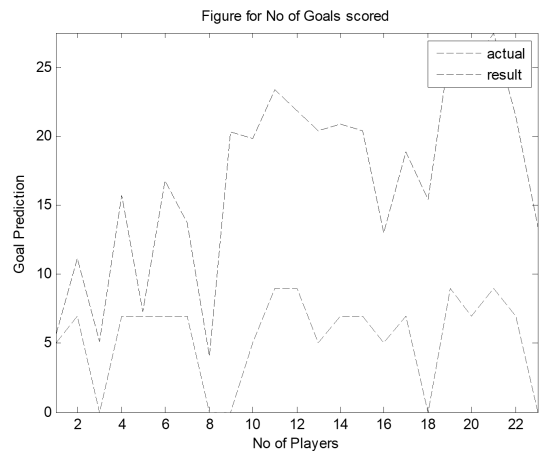


Figure 4: Goal Scored Estimation for NOP=20 and Iteration=100.

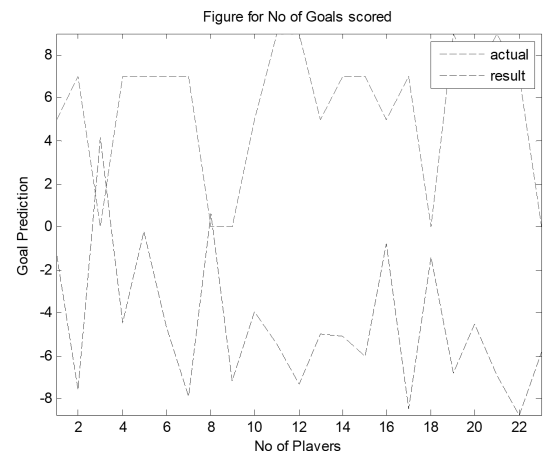


Figure 5: Goal Scored Estimation for NOP=50 and Iteration=400.

In “Figure 4”, the number of particles was considered as 20 and the iteration was 100 and it was observed that the i) Mean Absolute Error (MAE) was recorded with value 1.222073 ii) Mean Absolute Percentage Error (MAPE) was 22.30633 and iii) Mean Square

Error (MSE) was 8.812746. As the NOP and iteration decreased considerably, the actual goal scored value for each player, and the estimated goal while plotted on the graph showed significant drift with relation to each other. And at any point in time, they had a resemblance to each other.

In “Figure 5”, the number of particles was considered as 50 and the iteration respectively was 100 and it was observed that, the i) Mean Absolute Error (MAE) was recorded with value 13.31048 ii) Mean Absolute Percentage Error (MAPE) was 175.0847 and iii) Mean Square Error (MSE) was 218.3298. At this instance, the MAE, MAPE, and MSE experienced an exponential change in their values and it was conclusive that the actual and predicted results of goal scored had some instances of players matched but had a larger variance with each other.

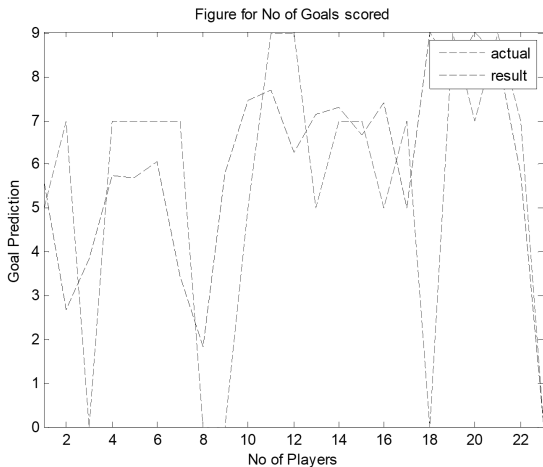


Figure 6: Goal Scored Estimation for NOP=70 and Iteration=700.

In “Figure 6”, the number of particles was considered as 70 and the iteration respectively was 700 and it was observed that, the i) Mean Absolute Error (MAE) was recorded with value -2.72485 ii) Mean Absolute Percentage Error (MAPE) was 30.82959 and iii) Mean Square Error (MSE) was 9.821473. The actual goal scored value for each player and the estimated goal while plotted on the graph showed that for every player the results obtained for actual and estimated run parallel to each other and did not drift away at many points of time.

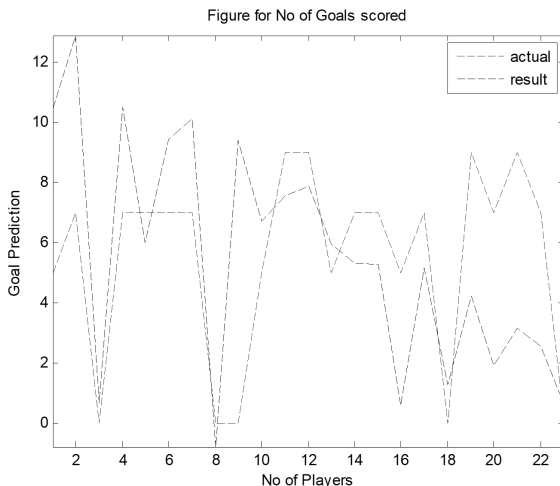


Figure 7: Goal Scored Estimation for NOP=70 and Iteration=900.

In “Figure 7”, the number of particles was considered as 70 and the iteration respectively was 900 and it was observed that the i) Mean Absolute Error (MAE) was recorded with value -0.03328 ii) Mean Absolute Percentage Error (MAPE) was 36.41223 and iii) Mean Square Error was (MSE) 14.08006. The actual goal scored value for each player and the estimated goal scored were significantly matching and hence the actual and predicted goal scored values were best for the set of values considered for NOP and number of iterations.

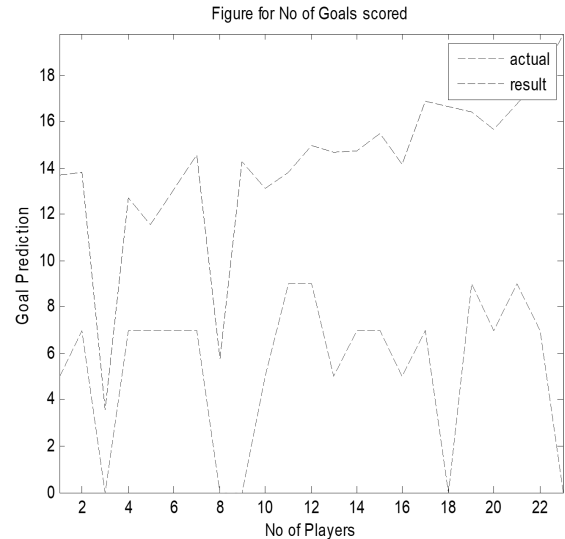


Figure 8: Goal Scored Estimation for NOP=100 and Iteration=1000.

In “Figure 8”, the number of particles was considered as 100 and the iteration respectively was 1000 and it was observed that, the i) Mean Absolute Error (MAE) was recorded with value -1.01699 ii) Mean Absolute Percentage Error (MAPE) was 13.69986 and iii) Mean Square Error (MSE) was 12.14193. The actual goal scored value for each player and the estimated goal while plotted on the graph showed that for every player there was a significant drift in terms of results.

In the above graphs plotted against various values of NOP and Iteration, we have varied with different values for the same to observe the kind of results obtained for MSE, MAE, and MAPE and have observed the approximations what the estimations predict. Observations from “Figure 5”, “Figure 6”, “Figure 7”, and “Figure 8” have estimations which are not close to the actual goal-scoring capability of every player in the team, and hence the lines in each of the graph are escalated above each other and run in parallel. But in “Figure 6” we observe that the lines in the graph intersect at regular intervals indicating a substantial coincidence in the estimation of goals scoring.

The proposed model helps to understand that even though a change in NOP and iterations yield us optimal results, even though in some instances the input parameters in the dataset have better accuracy. As shown in “Figure 6”, the goal scored in actual and estimated is accurate. The efficiency of SPEGSA is enhanced with sufficient parameters of soccer and GSA. The experimental analysis was made by varying the number of iterations and particles with the existing parameters yield better results with an accuracy of 90.18%. Hence the efficiency of the player predicting model performed well was evident from the results of GSA tabulated.

7. Conclusion and Future work

The GSA model was proposed to estimate the coefficients of goal prediction and player predicting model. Based on the models developed, trained, and validated, the estimated results which were observed can be improved and bettered in quite propositions. Hence the dataset can be trained with more statistical methods and Soft computing techniques such as Neural Networks and Backpropagation method.

Conflict of Interest

We know of no conflicts of interests associated with this publication, and there has been no significant financial support for this work.

References

- [1] N. Moosavian a and H. Moosavianb (2017). "Testing Soccer League Competition Algorithm in Comparison with Ten Popular Meta-heuristic Algorithms for Sizing Optimization of Truss Structures", IJE TRANSACTIONS A: Basics Vol. 30, No. 7, pp 926-936.10.5829/idosi.ije.2017.30.07a.01
- [2] Darren J Paul and George P Nassis (2015), "Testing Strength and Power in Soccer Players The Application of Conventional and Traditional Methods of Assessment", *Journal of Strength and Conditioning Research*, pp 1748-1758.
- [3] Miguel Angel Pe´rez-Toledano1, Francisco J. Rodriguez, Javier Garcı´a-Rubio, Sergio Jose ´ Ibañez (2019), "Players' selection for basketball teams, through Performance Index Rating, using multi-objective evolutionary algorithms"
- [4] Muhammad Bukhari Burhanudin, Muhammad Syasman suhaimi, Muhammad Thaqif Isa, Umi Kalsom Yousuf, Mohd Nor Akmal Khalid (2018), "Football Player Selection Using Fuzzy Logic", pp 1-7.
- [5] Bastien Talgorn, Sebastian Le Digabel, Micheal Kokkalaras, "Statistical Surrogate Formulations for Simulation-based design optimization", Vol 137, Issue 2, *Journal of Mechanical Design*, 2015.
- [6] S. Sankaran, Charles Audet, Alison L, "A method for stochastic constrained optimization using derivative-free surrogate pattern search and collection", Vol 229, Issue 12, pp 4664-4682, 2010.
- [7] Pooriya Beghai, Daniel Caragieri, Thomas Bewley, "Delaunay-based derivative-free optimization via global surrogates, Part I: Linear constraints", *Journal of Global optimization*, pp 1-52, Springer, 2015.
- [8] Charles Audet, Christophe Tribes, "Mesh-based Nelder-Mead algorithm for inequality constrained optimization", *Computational Optimization and Applications*. Vol 71, Issue 2, pp 331-352, Springer, 2018.
- [9] Kelton A.P Costa, Luis A.M. Pereira, Rodrigo Y.M, Nakamura, Clayton R. Pereira, Joaa.P.Papa, and Alexandre Xavier Falco, (2015). "A nature inspired approach to speed up optimum path forecast clusterist", Elsevier, *Information Science*, doi:10.1016/j.ins.2014.09.025
- [10] Abhinav Sharma, Sanjay Mathur, and R. Gowri (2018). "Adaptive beamforming for linear antenna using Gravitational Search Algorithm", Springer, AISC Vol 624, doi: 10.1007/978-981-10-5903-2_121
- [11] Damodar Reddy Edla, Diwakar Tripathi1, Ramalingaswamy Cheruku1, and Venkatanaresbhabu Kuppilil (2017). "An Efficient Multi-layer Ensemble Framework with BPSOGSA-Based Feature Selection for Credit Scoring Data Analysis", Springer Link, *Arabian Journal for Science and Engineering*, doi: 10.1007/s13369-017-2905-4
- [12] P. B. de Moura Oliveira, Josenalde Oliveira, and José Boaventura Cunha (2018). "Trends in Gravitational Search Algorithm," Springer International Publishing. doi: 10.1007/978-3-319-62410-5_33
- [13] Tongziang Wang, Xiangling Wei, Jianhua Fan, and Tao Liang (2018). "Adaptive Jammer Localization", Elsevier, Vol 141, doi/10.1016/j.comnet.2018.05.002.
- [14] Rajendra Kumar Khndanga and Jitendriya Ku Satapathy (2015). "A new hybrid GA- GSA algorithm for tuning damping controller parameters for a unified power flow controller", *International Journal of Electrical Power and International Journal of Electrical power and Energy systems*, doi: 10.1016/j.ijepes.2015.07.016
- [15] Mohammad Massoud Javidi and Fatemeh Zarsifi Kermani (2017). "Utilizing the advantages of both global and local search strategies for finding a small subset of features in a two-stage method", Springer Link, doi:10.1007/s10489-018-1159-5.
- [16] Ehsanol ah Assareh and Majtaba Biglari z (2015). "A novel approach to capture the maximum power from variable speed wind turbines using PI controller, RBF neural network and GSA evolutionary algorithm", Elsevier, *Renewable and sustainable energy reviews*, doi:10.1016/j.rser.2015.07.034
- [17] Mohamed A. Tolba1,Ahmed A. Zaki Diab1,Vladimir N. Tulskey1, and Almoataz Y. Abdelaziz (2018). VLSI approach for optimal capacitors allocation in distribution networks based on hybrid PSOGSA optimization algorithm", Springer link, doi: 10.1007/s00521-017-3327-7.
- [18] Bo Xing and Tshidlidzi Marwala (2018). "Introduction to intelligent search Algorithms", Springer Link *Smart Maintenance for Human Robot Interaction*, SSDC, Val 129, pp 33-64.
- [19] Alli A and D. Jhon Aravindhar (2017). "A Comparative Study on Cash Management Models using Soft Computing Techniques", *International Journal of Applied Engineering Research* ISSN 0973-4562 Volume 12.
- [20] Beatriz Gonzalez, Fevrier Valdez, and Patricia Melin (2017). "A Gravitational Search Algorithm using Type – 2 Fuzzy logic for parameter adaption", Springer linkg SCI, Vol 667, pp 127-138.
- [21] Kyuchang Kang, Chang Seaok Bae, Henry Wing fung yeung and yuk ying chung (2018) "Hybrid Gravitational Search Algorithm", Elsevier *Applied soft computing*, Vol 66, pp 319-329.
- [22] R. Priyadarshini, M. R. Panda, and N. Dash (2017)." An Improved Backpropagation Neural Network Model Based on Gravitational Search Algorithm for Multinomial Classification", Springer Link, doi: 10.1007/978-981-10-6890-4_17.
- [23] Esmat Rashedi, Elaheh Rashedi, and Hoosein Nezamabadi pour (2018). "A Comprehensive Survey on Gravitational Search Algorithm", Elsevier, *Swarm and Evolutionary computation* available online.
- [24] Seyed Ali Mirjalli and Amir Gandomi (2017). "Chaotic gravitational constants for Gravitational Search Algorithm", Elsevier, Vol 53, pp 407-419.

Sentence retrieval using Stemming and Lemmatization with Different Length of the Queries

Ivan Boban^{*1}, Alen Doko², Sven Gotovac³

¹Faculty of Mechanical Engineering, Computing and Electrical Engineering, University of Mostar, 88 000, Bosnia and Herzegovina

²Institut for Software Technology, German Aerospace Center, 28199, Germany

³Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, 21 000, Croatia

ARTICLE INFO

Article history:

Received: 25 March, 2020

Accepted: 13 May, 2020

Online: 11 June, 2020

Keywords:

Sentence retrieval

TF-ISF

Data pre-processing

Stemming

Lemmatization

ABSTRACT

In this paper we focus on Sentence retrieval which is similar to Document retrieval but with a smaller unit of retrieval. Using data pre-processing in document retrieval is generally considered useful. When it comes to sentence retrieval the situation is not that clear. In this paper we use TF – ISF (term frequency - inverse sentence frequency) method for sentence retrieval. As pre-processing steps, we use stop word removal and language modeling techniques: stemming and lemmatization. We also experiment with different query lengths. The results show that data pre-processing with stemming and lemmatization is useful with sentences retrieval as it is with document retrieval. Lemmatization produces better results with longer queries, while stemming shows worse results with longer queries. For the experiment we used data of the Text Retrieval Conference (TREC) novelty tracks.

1. Introduction

Sentence retrieval consists of retrieving relevant sentences from a document base in response to a query [1]. The main objective of the research is to present the results of sentence retrieval with TF – ISF (term frequency – inverse sentence frequency) method using data pre-processing consisting of stop word removal and language modeling techniques, stemming and lemmatization. Stemming and lemmatization are data reduction methods [2].

Previous work mentions the usefulness of the pre-processing steps with document retrieval. Contrary to that when it comes to sentence retrieval the usefulness of pre-processing is not clear. Some paper mentions it vaguely without concrete results. Therefore, we will try to clarify the impact of stemming and lemmatization on sentence retrieval and present this through test results. As additional contribution we will test and discuss how pre-processing impacts sentence retrieval with different query lengths. Because sentence retrieval is similar to document retrieval and stemming and lemmatization techniques have shown a positive effect on document retrieval, we expect these procedures to have a beneficial effect on sentence retrieval as well.

In our tests we use the State of The Art TF – ISF method in combination with stemming and lemmatization. For testing and evaluation, data from the TREC novelty tracks [3 - 6], were used.

*Ivan Boban, +38763484395 & ivan.boban@hotmail.com

www.astesj.com

<https://dx.doi.org/10.25046/aj050345>

This paper is organised as follows. Previous work is shown in section 2., an overview of methods and techniques is shown in section 3., in section 4. data set and experiment setup are presented, result and discussion are presented in section 5 and 6, and the conclusion is given in section 7.

2. Previous research

2.1. Sentence retrieval in document retrieval

Sentence retrieval is similar to document retrieval, and document retrieval methods can be adapted for sentence retrieval [7]. When it comes to Document retrieval the State of The Art TF – IDF (term frequency – inverse document frequency) method is commonly combined with preprocessing steps stemming and stop word removal. However, sentences of a document have an important role in retrieval procedures. In the paper [8], research results have shown that the traditional TF – IDF algorithm has been improved with sentence-based processing on keywords helping to improve precision and recall.

2.2. Document retrieval with stemming and lemmatization

Stemming and lemmatization are language modeling techniques used to improve the document retrieval results [9]. In [10] the authors showed the impact of stemming on document retrieval, using short and long queries. The paper [10] proved that

stemming has a positive effect on IR (the ranking of retrieved documents was computed using $TF - IDF$). Paper [11] compares document retrieval precision performances based on language modeling techniques, stemming and lemmatization. In papers [9, 11] it is shown that language modeling techniques (stemming and lemmatization) can improve document retrieval.

2.3. $TF - ISF$ sentence retrieval with stemming and lemmatization

When it comes to stemming and lemmatization and their impact on the $TF - ISF$ method, the results are not clearly presented, unlike the $TF - IDF$ method, where the impact is clear. In paper [12] stemming is mentioned in context of sentence retrieval. The paper states that stemming can improve recall but can hurt precision because words with distinct meanings may be conflated to the same form (such as "army" and "arm"), and that these mistakes are costly when performing sentence retrieval. Furthermore, paper [12] states that terms from queries that are completely clear and unambiguous, can match with sentences that are not even from the same topic after the stop word removal and stemming process.

3. Using $TF - ISF$ method for sentence retrieval in combination with stemming and lemmatization

For sentence retrieval in this paper we use $TF - ISF$ method based on vector space model of information retrieval. $TF - ISF$ was also used in [13, 14].

The ranking function is as follows:

$$R(s|q) = \sum_{t \in q} \log(tf_{t,q} + 1) \log(tf_{t,s} + 1) \log\left(\frac{n + 1}{0.5 + sf_t}\right) \quad (1)$$

Where:

- $tf_{t,q}$ – number of appearances of the term t in a query,
- $tf_{t,s}$ – number of appearances of the term t in a sentence,
- n – is the number of sentences in the collection and
- sf_t – is the number of sentences in which the term t appears,

The search engine for sentence retrieval with ranking function (1) uses data pre-processing consisting of following three steps: stop word removal, stemming and lemmatization. In information retrieval, there are many words that present useless information. Such words are called stop words. Stop words are specific to each language and make the language functional but they do not carry any information (e.g. pronouns, prepositions, links, ...) [15]. For example, there are around 400 to 500 stop words in the English language [15]. Words that often appear at the collection level can be eliminated through some tools like RapidMiner or programmatically, so as not to have an impact on ranking.

There are several different methods for removing stop words presented in [16] like:

- Z-Methods;
- The Mutual Information Method (MI);
- Term Based Random Sampling (TBRS);

In this paper we used the classic method of removing stop words based on a previously compiled list of words.

Part of the list of words to be removed by pre-processing is shown in Figure 1 which is a snippet from our source code:

```
//RapidMiner stopwords
static Dictionary<string, bool> _stop = new Dictionary<string, bool>
{
    { "abaft", true },
    { "aboard", true },
    { "about", true },
    { "above", true },
    { "across", true },
    { "afore", true },
    ...
}
```

Figure 1: Example part of the stop words list

Stemming refers to the process of removing prefixes and suffixes from words. When it comes to stemming, there are many different algorithms. One of them use the so called "bag of words" that contain words that are semantically identical or similar but are written as different morphological variants. By applying stemming algorithms, words are reduced to their root, allowing documents to be represented by the stems of words instead of original words. In information retrieval, stemming is used to avoid mismatches that may undermine recall. If we have an example in English where a user searches for a document entitled "How to write" over which he raises the query "writing", it will happen that the query will not match the terms in the title. However, after the stemming process, the word "writing" will be reduced to its root (stem) "write", after which the term will match the term in the title. We use the Porter's stemmer, which is one of the most commonly used stemmers, which functions on the principle that it applies a set of rules and eliminates suffixes iteratively. Porter's stemmer has a well-documented set of constraints, so if we have the words "fisher", "fishing", "fished", etc., they get reduced to the word "fish" [17]. Porter's stemmer algorithm is divided into five steps that are executed linearly until the final word shape is obtained [18]. In paper [19] it was proposed modified version of the Porter stemmer.

Lemmatization is an important pre-processing step for many applications of text mining, and also used in natural language processing [20]. Lemmatization is similar to stemming as both of them reduce a word variant to its "stem" in stemming and to its "lemma" in lemmatizing [21]. It uses vocabulary and morphological analysis for returning words to their dictionary form [11, 20]. Lemmatization converts each word to its basic form, the lemma [22]. In the English language lemmatization and stemming often produce same results. Sometimes the normalized/basic form of the word may be different than the stem e.g. "computes", "computing", "computed" is stemmed to "comput", but the lemma of that words is "compute" [20]. Stemming and lemmatization have an important role in order to increase the recall capabilities [23, 24].

4. Data set used and experiment setup

Testing was performed on data from the TREC Novelty tracks [3]-[5]. Three Novelty Tracks were used in the experiment: TREC 2002, TREC 2003 and TREC 2004. Each of the three Novelty Tracks has 50 topics. Each topic consisting of "titles", "descriptions" and "narratives".

Figure 2. shows a part of the file related to one topic labeled "N56" from TREC 2004 novelty track with parts "titles", "descriptions" and "narratives" [25]:

```
<num>Number: N56

<title>
Woodstock Music Festival Reunion

<toptype>
Event

<desc>Description:
Woodstock 99 music festival reunion in Rome, NY

<narr>Narrative:
Relevant documents contain opinions on the planning,
location, data, events, participants, results, community
reactions, and problems associated with this festival.
Opinions can be from the public, personal, local government,
and police sources. The first Woodstock event is not
relevant.
```

Figure 2: Example of the topic N56 from TREC 2004 novelty track

Two types of queries were used in the experiment. The short query uses the <title> part and the longer query the <desc> part. To each of 50 topics 25 documents were assigned. Each of the 25 documents contains multiple sentences.

Figure 3 shows a snippet from one of the 25 documents assigned to topic N56, which has multiple sentences, which are in the format: <s docid="xxx" num="x">Content of Sentence </s>.

```
<s docid="NYT19980812.0284" num="9"> This time, Woodstock fans
probably won't be worrying about getting slowed down in the
mud on Max Yasgur's farm in upstate New York.</s>
</P>
<P>

<s docid="NYT19980812.0284" num="10"> This time, it might be
on the bandwidth.</s>
</P>
<P>

<s docid="NYT19980812.0284" num="11"> Beginning tomorrow,
Infoseek Corp. should expect plenty of fiftysomethings to log
on to its simulcast of the 29th anniversary concerts for the
original Woodstock music festival in Bethel, N.Y.</s>
</P>
<P>
```

Figure 3: Example of part within the document from TREC 2004 novelty track

In our experiment we extract single sentences from the collection. During the extraction we assign a docid (document identifier) and num (sentence identifier) to each sentence.

Three data collections were used, (Table 1 and Table 2).

Table 1: Description of three data collections

Name of the collection	Number of topics	Number of queries (title/desc)
TREC 2002	50	50
TREC 2003	50	50
TREC 2004	50	50

Table 2: Overview of number of sentences per document

Name of the collection	Number of documents per topic	Number of sentences
TREC 2002	25	57792
TREC 2003	25	39820
TREC 2004	25	52447

For results evaluation one file is available which contain a list of relevant sentences [25]. Figure 4 shows a snippet from the relevant sentence file.

```
N55 XIE20000523.0098:15
N56 APW19990205.0174:7
N56 APW19990205.0174:8
N56 APW19990205.0174:9
N56 APW19990205.0174:10
N56 APW19990205.0174:13
N56 APW19990205.0174:15
N56 NYT19990409.0104:12
N56 NYT19990409.0104:15
N56 NYT19990409.0104:16
N56 NYT19990409.0104:17
N56 NYT19990409.0104:22
```

Figure 4: File with list of relevant sentences

The format of the list of relevant sentences shown in Figure 4 is: N56 NYT19990409.0104:16.

Where:

- N56 - indicates the topic number,
- NYT19990409.0104 - indicates a specific document,
- 16 - indicates the sentence number - identifier.

N56 NYT19990409.0104: 16 defines sentence "16" of document "NYT19990409.0104" as relevant to topic "N56".

Using the presented TREC data we test at first the *TF – ISF* method without any pre-processing. Then we test the same *TF – ISF* method with stemming and with lemmatization. All three tests we do twice: First time with short queries and second time with long queries. In all of our tests we use stop word removal.

We denote the baseline method as *TF – ISF*, the method with stemming we denote as *TF – ISF_{stem}* and the method with lemmatization we denote as *TF – ISF_{lem}*.

5. Result and discussion

As already mentioned, we wanted to test if data pre-processing steps stemming and lemmatization affect the sentence retrieval. Also, we want to analyse if the effect of pre-processing is different when using different query lengths. For test evaluation we used standard measures: P@10, R-precision and Mean Average Precision (MAP) [26, 27].

Precision at x or P@x can be defined as:

$$P@x(q_j) = \frac{\text{number of relevant sentences within top } x \text{ retrieved}}{x} \quad (2)$$

The P@10 values shown in this paper refer to average P@10 for 50 queries.

R-precision can be defined as [26]:

$$R - \text{precision} = \frac{r}{|Rel|} \quad (3)$$

Where:

- |Rel| is the number of relevant sentences to the query,

- r is the number of relevant sentences in top $|Rel|$ sentences of the result.

As with P@10 we also calculate the average R-precision for 50 queries. Another well-known measure is Mean Average Precision which gives similar results to R-precision.

Mean Average Precision and R-precision is used to test high recall. High recall means: It is more important to find all of relevant sentences even if it means searching through many sentences including many non-relevant. In opposite to that P@10 is used for testing precision.

Precision in terms of information retrieval means: It is more important to get only relevant sentences than finding all of the relevant sentence.

For result comparison we used two tailed paired t -test with significance level $\alpha=0.05$. Statistically significant improvements in relation to the base $TF - ISF$ method (without data pre-processing) are marked with a (*). The results of our tests on different data sets are presented below in tabular form. Table 3 shows the results of our tests on TREC 2002 collection with short queries presented on Figure 2 and labeled with <title>.

Table 3: Test results using TREC 2002 collection with short query

TREC 2002 - title			
	$TF - ISF$	$TF - ISF_{stem}$	$TF - ISF_{lem}$
P@10	0,304	0,328	0,34
MAP	0,1965	*0,2171	*0,2149
R-prec.	0,2457	0,2629	0,2575

From Table 3 we see that the method with stemming $TF - ISF_{stem}$ and the method with lemmatization $TF - ISF_{lem}$ show statistically significant better results in comparison to the baseline method, when it comes to MAP measure.

Table 4 shows the results of our tests on TREC 2002 collection with longer queries presented on Figure 2 and labeled with <desc>.

Table 4: Test results using TREC 2002 collection with longer query

TREC 2002 - description			
	$TF - ISF$	$TF - ISF_{stem}$	$TF - ISF_{lem}$
P@10	0,332	0,296	0,324
MAP	0,2075	0,2157	*0,2176
R-prec.	0,2490	0,2601	0,2570

Only $TF - ISF_{lem}$ provides better results and statistically significant differences in relation to the base $TF - ISF$ method (without data pre-processing), when the MAP measure is used. We can see that stemming performs a little worse when it comes to longer queries in relation to the base $TF - ISF$ method. Table 5 and Table 6 show the results of our tests using TREC 2003 collection with short and longer queries respectively.

Table 5 show that $TF - ISF_{stem}$ and $TF - ISF_{lem}$ provide better results and statistically significant differences in relation to the base $TF - ISF$ method, when the MAP and R-prec. measure are used.

Table 5: Test results using TREC 2003 collection with short query

TREC 2003 - title			
	$TF - ISF$	$TF - ISF_{stem}$	$TF - ISF_{lem}$
P@10	0,692	0,712	0,714
MAP	0,5765	*0,5911	*0,5887
R-prec.	0,5470	*0,5611	*0,5593

Table 6: Test results using TREC 2003 collection with longer query

TREC 2003 - description			
	$TF - ISF$	$TF - ISF_{stem}$	$TF - ISF_{lem}$
P@10	0,734	0,738	0,738
MAP	0,5914	*0,6059	*0,6049
R-prec.	0,5617	0,5699	*0,5750

Table 6 is shows that lemmatization keeps showing statistically significant better results even with long queries, unlike the method with that uses stemming. Table 7 and Table 8 show the results of our tests on TREC 2004 collection with short and longer queries.

Table 7: Test results using TREC 2004 collection with short query

TREC 2004 - title			
	$TF - ISF$	$TF - ISF_{stem}$	$TF - ISF_{lem}$
P@10	0,434	0,448	0,444
MAP	0,3248	*0,3390	0,3331
R-prec.	0,3366	0,3385	0,3387

Table 8: Test results using TREC 2004 collection with longer query

TREC 2004 - description			
	$TF - ISF$	$TF - ISF_{stem}$	$TF - ISF_{lem}$
P@10	0,498	0,49	0,498
MAP	0,3540	*0,3644	*0,3635
R-prec.	0,3583	0,3688	0,3699

Table 7 shows that $TF - ISF_{stem}$ method with short queries, provides better results in comparison to the baseline, when it comes to MAP measure. Table 8 shows that with longer queries both methods shows statistically significant better results. When taking a look at all the tables above we see that stemming and lemmatization often give statistically significant better results when it comes to MAP and R-Prec. Therefore, we can assume that these pre-processing steps have similar positive effect on sentence retrieval as they have on document retrieval. Let us analyse how query length impacts our two methods ($TF - ISF_{stem}$ and $TF - ISF_{lem}$). Table 9 shows an overview of the overall number of statistically significant better results for four pairs (stem - short queries, stem - long queries, lem - short queries, lem - long queries).

As we can see $TF - ISF_{stem}$ seems to go better with short queries and $TF - ISF_{lem}$ seems to go better with long queries. At the moment we do not have enough data to examine this behaviour further. But that will be a topic for further research of us.

Table 9. Performance of $TF - ISF_{stem}$ and $TF - ISF_{lem}$ in regard to query length

Number of statistically significant better results with methods		
	$TF - ISF_{stem}$	$TF - ISF_{lem}$
Short queries	4	3
Long queries	2	4

6. Additional result analysis

To understand the reasons why we have better results using $TF - ISF_{stem}$ and $TF - ISF_{lem}$ methods in relation to the base $TF - ISF$ method, we analysed the positioning of one relevant sentence in the test results of the different methods when using the TREC 2003 collection. Table 10 and 11 analyze one sentence ("Two of John F. Kennedy Jr., 's cousins, David and Michael, both sons of Robert Kennedy, died young, the latter of a drug overdose in 1984, as did four Kennedys of the preceding generation") from topic "N42" in two different scenarios using short and long query and with 3 different methods. One of them is baseline method and the remaining two use stemming and lemmatization. As already said we match the sentence with two different queries: Short query ("John F. Kennedy, Jr. dies") and long query ("John F. Kennedy, Jr. was killed in a plane crash in July 1998").

Table 10 shows the matching of the sentence with short query with resulting sentence position for each of the three methods.

Table 10: Analysis of the sentence and the short query with the different methods

$TF - ISF$	
Short query	"john", "f", "kennedy", "jr", "dies"
Sentence content	"john", "f", "kennedy", "jr", "s", "cousins", "david", "michael", "sons", "robert", "kennedy", "died", "young", "latter", "drug", "overdose", "kennedys", "preceding", "generation"
Sentence position	(24)
$TF - ISF_{stem}$ (stemming)	
Short query	"john", "f", "kennedi", "jr", "die"
Sentence content	"john", "f", "kennedi", "jr", "s", "cousin", "david", "michael", "son", "robert", "kennedi", "die", "young", "latter", "drug", "overdos", "kennedi", "preced", "generat"
Sentence position	(1)
$TF - ISF_{lem}$ (lemmatization)	
Short query	"john", "f", "kennedy", "jr", "die",
Sentence content	"john", "f", "kennedy", "jr", "s", "cousin", "david", "michael", "son", "robert", "kennedy", "die", "young", "latter", "drug", "overdose", "kennedy", "precede", "generation"
Sentence position	(1)

Table 11 shows the same as Table 10 but with long query.

Table 11: Analysis of the sentence and the long query with the different methods

$TF - ISF$	
Long query	"john", "f", "kennedy", "jr", "killed", "plane", "crash", "july", "1998",
Sentence content	"john", "f", "kennedy", "jr", "s", "cousins", "david", "michael", "sons", "robert", "kennedy", "died", "young", "latter", "drug", "overdose", "kennedys", "preceding", "generation"
Sentence position	(67)
$TF - ISF_{stem}$ (stemming)	
Long query	"john", "f", "kennedi", "jr", "kill", "plane", "crash", "juli", "1998"
Sentence content	"john", "f", "kennedi", "jr", "s", "cousin", "david", "michael", "son", "robert", "kennedi", "die", "young", "latter", "drug", "overdos", "kennedi", "preced", "generat"
Sentence position	(63)
$TF - ISF_{lem}$ (lemmatization)	
Long query	"john", "f", "kennedy", "jr", "kill", "plane", "crash", "july", "1998",
Sentence content	"john", "f", "kennedy", "jr", "s", "cousin", "david", "michael", "son", "robert", "kennedy", "die", "young", "latter", "drug", "overdose", "kennedy", "precede", "generation"
Sentence position	(62)

Table 10 and Table 11 shows how stemming and lemmatization help to position relevant sentences closer to the top of search result.

More precisely, every match of words between query and sentence is marked bold. Matches that occurred with stemming or lemmatization but not with the baseline are marked as bold and underlined.

In Table 10 and Table 11 we clearly can see some words that could be matched thanks to stemming and lemmatization. For example, if we look at a short query and a sentence through three different methods shown in Table 10, we can see how the word "dies" and "died", in query and sentence is reduced by the stemming and lemmatization to the word form "die", through which it is possible to overlap between the query and the sentence. Also, the tables show a few more examples that show how some words could be matched thanks to stemming and lemmatization, and why a sentence has a better position in the search result.

7. Conclusion

In this paper we showed through multiple tests that pre-processing steps stemming and lemmatization have clear benefits when it comes to sentence retrieval. In most of our tests we got better results when combining $TF - ISF$ with stemming or lemmatization. However, the positive effects only appeared with the measures MAP and R-prec. which improve recall. At the same time the pre-processing steps did not show any negative effects on sentence retrieval. Therefore, we think that stemming and

lemmatization is generally beneficial to sentence retrieval, we saw that stemming tends to show better result with short queries while lemmatization tends to show better results with longer queries which we will explore in more detail in the future.

References

- [1] Doko, A., Stula, M., & Stipanicev, D. (2013). A recursive tf-idf based sentence retrieval method with local context. *International Journal of Machine Learning and Computing*, 3(2), 195.
- [2] Florijn, W. J. (2019). Information retrieval by semantically grouping search query data (Master's thesis, University of Twente).
- [3] Harman, D. (2002). Overview of the TREC 2002 novelty track. In *Proceedings of the eleventh text retrieval conference (TREC)*.
- [4] Soboroff, I., & Harman, D. (2003). Overview of the TREC 2003 novelty track. In *Proceedings of the twelfth text retrieval conference (TREC)*.
- [5] Soboroff, I. (2004). Overview of the TREC 2004 novelty track. In *Proceedings of the thirteenth text retrieval conference (TREC)*.
- [6] Text REtrieval Conference (TREC) Novelty Track. (2003, March 4). Retrieved March 19, 2020, from <https://trec.nist.gov/data/novelty.html>
- [7] Doko, A., Stula, M., & Seric, L. (2015). Using TF-ISF with Local Context to Generate an Owl Document Representation for Sentence Retrieval. *Computer Science & Engineering: An International Journal*, 5(5), 01–15.
- [8] Vetriselvi, T., Gopalan, N. P., & Kumaresan, G. (2019). Key Term Extraction using a Sentence based Weighted TF-IDF Algorithm. *International Journal of Education and Management Engineering*, 9(4), 11.
- [9] Samir, A., & Lahbib, Z. (2018, April). Stemming and Lemmatization for Information Retrieval Systems in Amazigh Language. In *International Conference on Big Data, Cloud and Applications* (pp. 222-233). Springer, Cham.
- [10] Kantrowitz, M., Mohit, B., & Mittal, V. (2000, July). Stemming and its effects on TFIDF ranking. In *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 357-359).
- [11] Balakrishnan, V., & Lloyd-Yemoh, E. (2014). Stemming and lemmatization: a comparison of retrieval performances.
- [12] Murdock, V. (2006). Aspects of sentence retrieval. Ph.D. thesis, University of Massachusetts.
- [13] Allan, J., Wade, C., & Bolivar, A. (2003, July). Retrieval and novelty detection at the sentence level. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 314-321).
- [14] Losada, D. (2008, July). A study of statistical query expansion strategies for sentence retrieval. In *Proceedings of the SIGIR 2008 Workshop on Focused Retrieval* (pp. 37-44).
- [15] V. Srividhya, R. Anitha, Evaluating Preprocessing Techniques in Text Categorization, *International Journal of Computer Science and Application*, Issue 2010.
- [16] Dr. S. Vijayarani, Ms. J. Ilamathi, Ms. Nithya, Preprocessing Techniques for Text Mining - An Overview, *International Journal of Computer Science & Communication Networks*, vol 5(1), pp. 7-16.
- [17] Sandhya, N., Lalitha, Y. S., Sowmya, V., Anuradha, K., & Govardhan, A. (2011). Analysis of stemming algorithm for text clustering. *International Journal of Computer Science Issues (IJCSI)*, 8(5), 352.
- [18] Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, 14(3), 130–137.
- [19] Joshi, A., Thomas, N., & Dabhade, M. (2016). Modified porter stemming algorithm. *Int. J. Comput. Sci. Inf. Technol*, 7(1), 266-269.
- [20] Plisson, J., Lavrac, N., & Mladenic, D. (2004). A rule based approach to word lemmatization. In *Proceedings of IS (Vol. 3, pp. 83-86)*.
- [21] Ms. Anjali Ganesh Jivani, A Comparative Study of Stemming Algorithms, Anjali Ganesh Jivani et al, *Int. J. Comp. Tech. Appl.*, Vol 2 (6), 1930-1938, ISSN:2229-6093.
- [22] Korenius, T., Laurikkala, J., Järvelin, K., & Juhola, M. (2004, November). Stemming and lemmatization in the clustering of finnish text documents. In *Proceedings of the thirteenth ACM international conference on Information and knowledge management* (pp. 625-633).
- [23] Kettunen, K., Kunttu, T., & Järvelin, K. (2005). To stem or lemmatize a highly inflectional language in a probabilistic IR environment?. *Journal of Documentation*.
- [24] Kanis, J., & Skorkovská, L. (2010, September). Comparison of different lemmatization approaches through the means of information retrieval performance. In *International Conference on Text, Speech and Dialogue* (pp. 93-100). Springer, Berlin, Heidelberg.
- [25] Text REtrieval Conference (TREC) TREC 2004 Novelty Track. (2005, February 4). Retrieved March 19, 2020, from https://trec.nist.gov/data/t13_novelty.html
- [26] Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge university press.
- [27] Fernández, R. T., Losada, D. E., & Azzopardi, L. A. (2011). Extending the language modeling framework for sentence retrieval to include local context. *Information Retrieval*, 14(4), 355-389.

Design and Optimization of Dual-Band Branch-Line Coupler with Stepped-Impedance-Stub for 5G Applications

Ayyoub El Berbri^{1,*}, Adil Saadi¹, Seddik Bri²

¹*Control, Pilotage and Supervision of Systems, National Graduate School of Arts and Crafts, Moulay Ismail University, Meknes, Morocco*

²*Materials and Instrumentation, High School of Technology, Moulay Ismail University, Meknes, Morocco*

Article history:

Received: 21 January, 2020

Accepted: 23 May, 2020

Online: 11 June, 2020

Keywords:

Optimization

Tuning

Dual-Band

Branch-Line Coupler

5G

ABSTRACT

This paper presents a design optimization of a dual-band branch-line coupler with stepped-impedance-stub lines. This coupler operates over 5G NR frequency bands n5 and n2, developed by 3GPP for the 5G (fifth generation) mobile network, and these two bands are centered at 0.85 GHz and 1.9 GHz respectively. To achieve the design specifications an adjusted Tuning Space Mapping method is used. This method of optimization moves the hardship of optimization from high-fidelity electromagnetic models to low-fidelity tuning models. The simulated and measured results of this enhanced coupler show good dual-band performance at the two bands.

1. Introduction

The request for mobile communication is more than ever. For this reason, many researchers have focused on the development of a new generation of communication. 5G NR is a new radio access technology developed by 3GPP for the 5G mobile network. We designed the coupler to function in the two bands n5 and n2 from the 3GPP TS 38.101 [1].

Branch-line couplers (BLC) offer a $\pi/2$ power splitting and phase difference, which is suitable in different microwave circuits like power combined amplifiers, phase shifters, data modulators, and balanced mixers [2]. There are different dual-band BLC design that exists in the literature [3-9]. In this work, we use stepped-impedance-stub branches lines (SISBL) for dual-band operation. [10].

This work presents the design of a dual-band BLC with SISBL, designed for 5G NR frequency bands downlink [n2= (0.869–0.894 GHz) / (n5= (1.930–1.990 GHz)], this bands developed by 3GPP for the 5G (fifth generation) mobile network and centered at 0.85 GHz and 1.9 GHz respectively [1]. 5G NR, like most modern techniques, demands height accuracy. In many situations like ours, due to microwave structure complexity of SISBL dual-band BLC,

the theoretical model only gives the initial design that requires optimization so he can meet the design specifications. In our case, for the optimization, we use an adjusted Tuning Space Mapping, the adjustment is made in the tuning model using our engineering expertise and knowledge of the design problem. The advantage is to reduce time-consuming. The enhanced characteristics of this Dual-Band BLC with SISBL are presented.

The simulated results of this enhanced dual-band BLC show good dual-band performances at n5/n2. The performance of its fabricated prototype has also been validated experimentally.

2. Dual-Band BLC with SISBL

Figure 1.a presents the SISBL containing a signal path (Z_3, θ_3) tapped with (Z_1, θ_1) and (Z_2, θ_2) [10]. By multiplying the ABCD matrices of each cascade component, the ABCD matrix of the SISBL is reached. Since the dual-band branch line behaves as a $\pi/2$ line at center frequencies of the two bands (f_p and f_s), the ABCD matrix is formulated as (1).

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 0 & \pm \frac{j}{J} \\ \pm jJ & 0 \end{bmatrix} \quad (1)$$

*Ayyoub El Berbri, Moulay Ismail University & ayyoub.elberbri@gmail.com

Where J the characteristic admittance of the $\pi/2$ line

$$x_f = \arg \min_x U(R_f(x)) \quad (2)$$

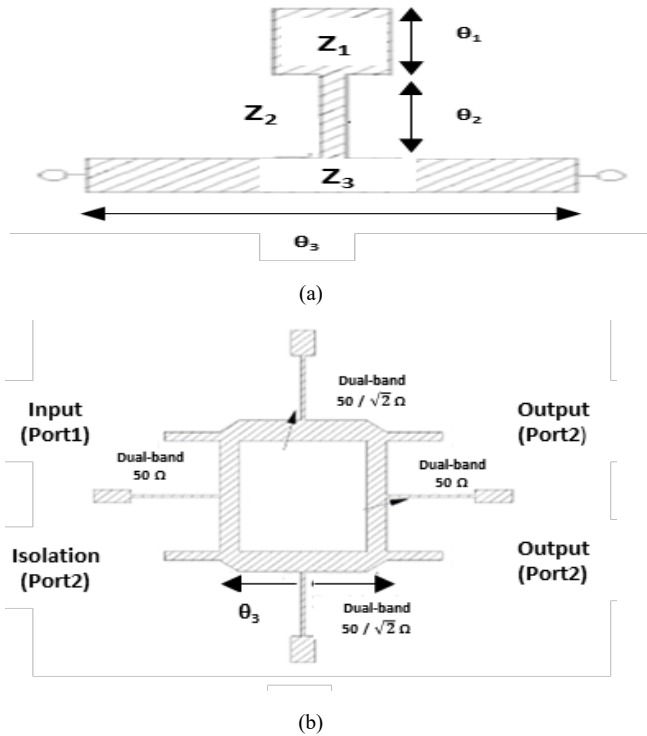


Figure 1. a: Dual-band SISBL; b: The dual-band BLC with SISBL.

The circuit dimensions of the SISBL Figure 1.b can be determined by solving the following equations (1)-(4). Where $(\theta_1, \theta_2, \theta_3)$ the electrical lengths of the lines (1,2,3) respectively, and (Z_1, Z_2, Z_3) the impedances of the lines (1,2,3).

$$\theta_3 = \frac{2n\pi}{(1+r_f)} \quad (3)$$

where $n=1, 2, 3, \dots$; $r_f=fs/fp$ is the frequency ratio.

To get a compact dual-band BLC in (2), n takes the value 1.

$$Z_3 = \frac{1}{J \left| \tan \frac{\theta_3}{2} \right|} \quad (4)$$

$$Z_3 \cot \theta_3 = \frac{2 * Z_2 (Z_1 \cot \theta_1 - Z_2 \tan \theta)}{Z_2 + Z_1 \cot \theta_1 \tan \theta_2} \quad (5)$$

We cast as free variables, the impedance ratio $R=Z_1/Z_2$ and the electrical length ratio $U= \theta_1/\theta_2$. Substituting (2)– (3), $Z_1=R*Z_2$ and $\theta_1=U*\theta_2$ into (4) eliminates Z_2 , leads to (5):

$$\tan [\theta(1+r_f)] = \frac{R [\cot(U\theta_2) + \cot(Ur_f\theta_2)]}{1 - R^2 \cot(U\theta_2) \cot(Ur_f\theta_2)} \quad (6)$$

Equation (4) has only θ_2 for variable since the constants U, R and r_f can be pre-chosen. θ_2 is found using the graph of each side of the equation (4).

Figure 2 presents the variation of the impedances Z_1 and Z_2 versus frequency ratio r_f when R and U are both 0.2. That figure shows that both Z_1 and Z_2 increase as r_f increases. Since only an impedance of 20–120 Ω can be realized using microstrip, the range of r_f is limited to $1.9 < r_f < 2.5$.

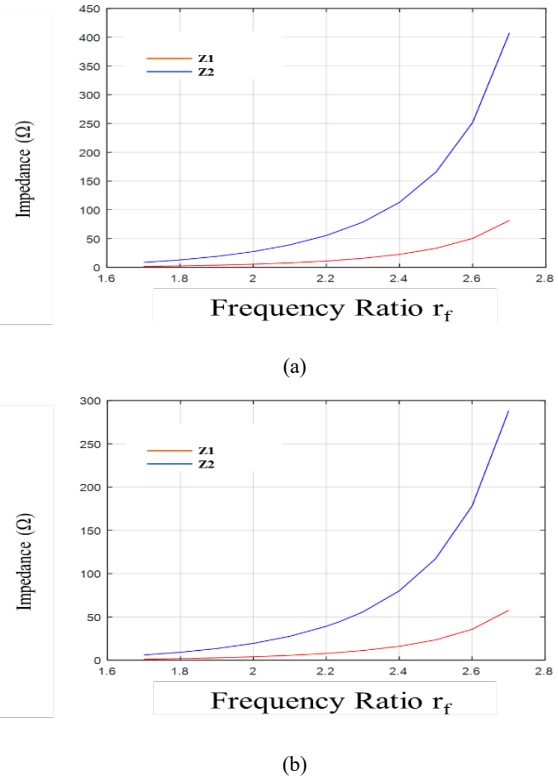


Figure 2: Variation of Z_1 and Z_2 versus r_f . (a) Dual-band 50 Ω branch line. (b) Dual-band 50/ $\sqrt{2}$ Ω branch line

3. Tuning Space Mapping

The utilize of electromagnetic software for design optimization can be problematic, because EM simulations are rigorous. By shifting the optimization load onto a coarse model, for example, a circuit equivalent, space mapping (SM) reduces the computational rate of EM simulation. Over an iterative optimization and updating of the low fidelity models, SM achieves efficient optimization. [11].

Tuning space mapping such as SM belongs to a family of surrogate-based optimization techniques [12-23]. However, the role of the surrogate model is replaced by a so-called tuning model. By introducing tuning components (circuit-theory base components) into the fine model structure, the tuning model is built up. The original optimization problem is as follow [15]:

Where $R_f \in R^m$ is the response vector, U is an objective function, x is the vector of design parameters, and x_f^* is the optimal solution.

At i -th iteration, a tuning model $R_t^{(i)}$ is constructed based on the fine model data. The alignment process to eliminate the gap

between the tuning model response and the response of the fine model at i , it is formulated as:

$$x_{t,0}^{(i)} = \arg \min_{x_t} \|R_f(x^{(i)}) - R_t^{(i)}(x_t)\| \quad (7)$$

To fit the design specifications, $R_t^{(i)}$ needs to be optimized with respect to x_t . This is expressed by:

$$x_{t,1}^{(i)} = \arg \min_{x_t} U(R_t^{(i)}(x_t)) \quad (8)$$

The modifications needed for the design variables are determined using a calibration process. Equation (9) express this process:

$$x^{(i+1)} = C(x^{(i)}, x_{t,1}^{(i)}, x_{t,0}^{(i)}) \quad (9)$$

4. Results and Discussion

We chose $(U, R) = (0.2, 0.2)$, the dimensions of this dual-band BLC have been computed with the use of (2)-(5). Figure 3 illustrates the intersecting points of each side of the equation (4).

The length electric from Figure 3 is $\theta_2 = 72.68^\circ$. The other circuit dimensions $(Z_1, \theta_1, Z_2, Z_3, \theta_3)$ of both the 50Ω and $50\Omega/\sqrt{2}$ branches of the coupler are computed from equations (2), (3) and (4).

The design parameters are $x = [L_0 \ L_{11} \ L_{21} \ L_{31} \ L_{12} \ L_{22} \ L_{32}]^T$ mm. The fine model, as shown in Figure 4, is simulated using a substrate of $\epsilon_r = 4$, height $H = 1.4$ mm and loss tangent = 0.0004.

Figure 6 shows that S_{11} and S_{14} for the second band are below -10dB from 1.76 GHz to 1.88GHz, this is not the entire [1.930 GHz–1.990 GHz] band. And in the same range, the phase difference between output ports (2 and 3) is not $90^\circ \pm 10^\circ$. For better performance, optimization is needed.

To build the tuning model, we first simulate the fine model with the co-calibrated ports Figure 5, the corresponding S44P data file is charged into a 44-port S-parameter file component in the tuning model. After that, an appropriate circuit component is attached to the corresponding ports on the S-parameter component; the tuning parameters are $x_t = [L_{t0} \ L_{t1} \ L_{t2} \ L_{t3} \ L_{t12} \ L_{t22} \ L_{t32}]^T$ mm.

The tuning parameters obtained after optimized the tuning model are $x_{t,1}(0) = [0.632 \ -0.73 \ -2.583 \ 0.343 \ -2.945 \ -1.291 \ -3.884]^T$ mm. With the use of a direct calibration the optimal values of the tuning parameters are converted to the adjustments of the design parameters in a direct manner. After the first iteration, the new design $x^{(1)} = [50.9392 \ 6.75824 \ 38.7104 \ 61.2757 \ 4.45566 \ 38.9794 \ 51.5329]^T$ mm has already satisfied the design specifications Figure 7.

Figure 7 shows that the simulated S_{11} and S_{14} are less than -11dB, within the frequency range of 0.82-0.92 GHz (first band), and within 1.88-2 GHz (second band). Within a similar frequency range, the coupling coefficient is 3 ± 1 dB, the transmission coefficient is 3 ± 1 dB, and the phase difference between output ports (2 and 3) is $90^\circ \pm 10^\circ$. TABLE I summarized the performances of this dual-band BLC.

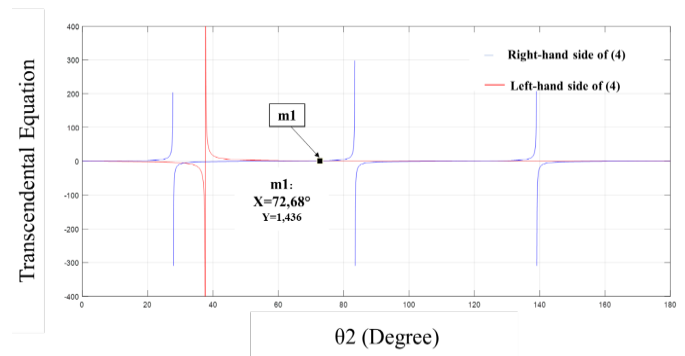


Figure 3: Solutions to (4) for θ_2 of SISBL

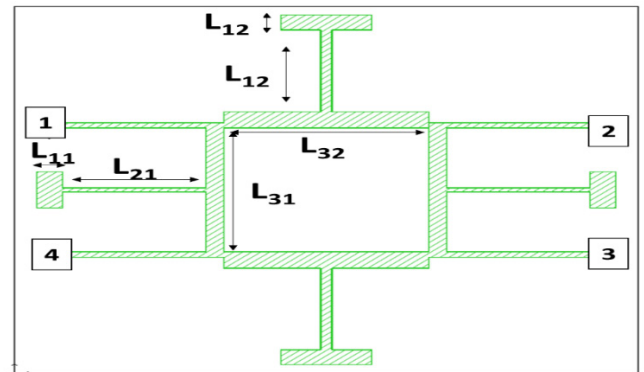


Figure 4: The fine model.

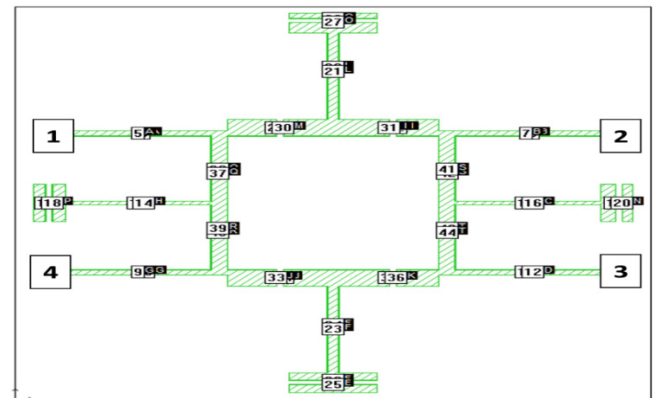


Figure 5: The fine model divided

Figure 8 shows the photograph of the prototype of the enhanced dual-band BLC fabricated using an LPKF machine. As we can see, its structure is relatively simple as it can be fabricated on a single layer printed circuit board with a simple ground plane.

The measurements of the fabricated prototype are done using 1-Port USB Vector Network Analyzer. Figure 9 plots the simulated and measured (S_{11}), it shows that there is a good match between simulated and measured S_{11} of this enhanced dual-band BLC.

TABLE I summarized, the results of the simulation for the initial and the optimized dual-band BLC performance, also the measurements of the fabricated prototype. The optimized dual-band BLC shows better performances for both 5G NR bands, which the initial does not do.

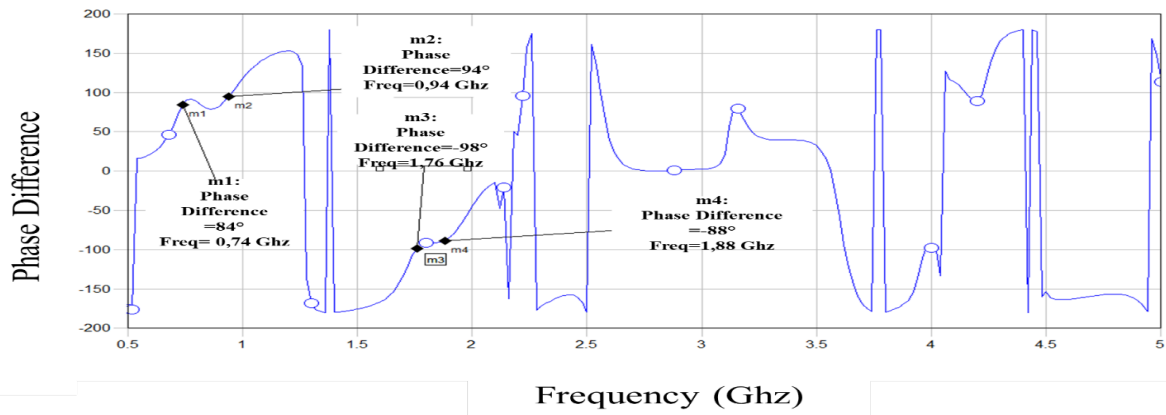
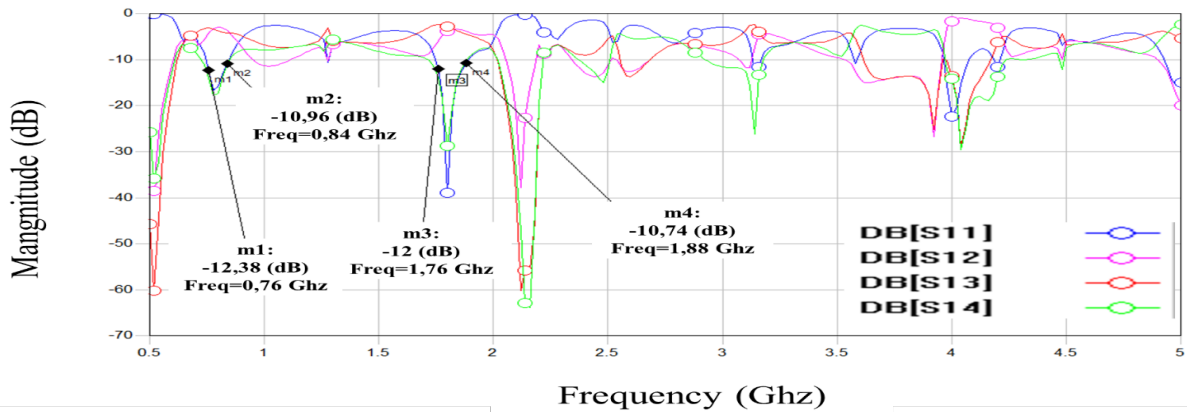


Figure 6. The Initial fine model response; a:S-parametres b: phase difference between Port2 and 3

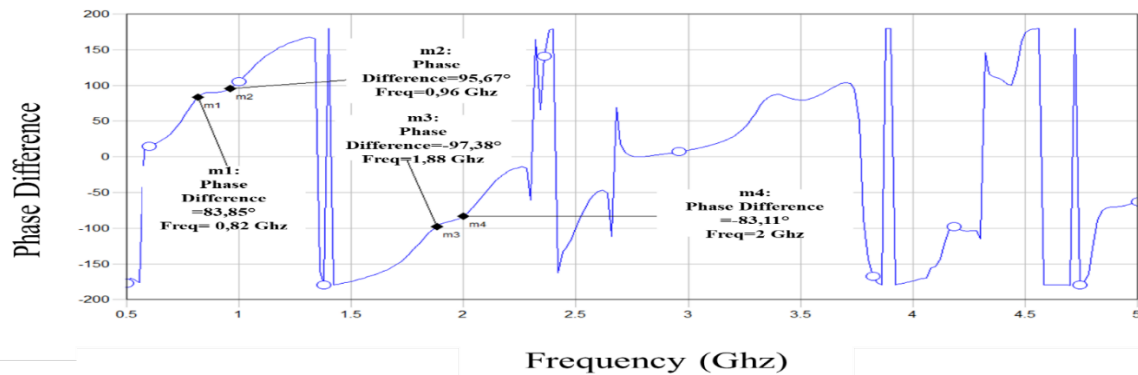
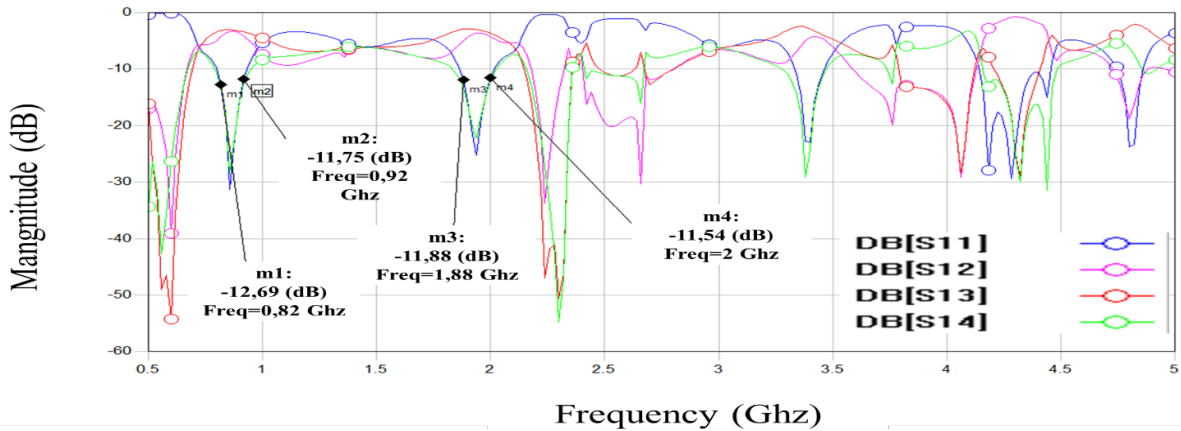


Figure 7: The Optimized fine model response; a:S-parametres b: phase difference between Port2 and 3.

Table 1: The simulated and measured results

Parameter	Initial First band	Initial Second band	Optimized First band	Optimized Second band	Measured First band	Measured Second band
S_{11} (dB)	<-10 dB	<-10 dB	<-10dB	<-10dB	<-10dB	<-10dB
S_{12} (dB)	3 ± 1 dB	3 ± 1 dB	3 ± 1 dB	3 ± 1 dB	-	-
S_{13} (dB)	3 ± 1 dB	3 ± 1 dB	3 ± 1 dB	3 ± 1 dB	-	-
S_{14} (dB)	<-10 dB	<-10 dB	<-10dB	<-10dB	<-10dB	<-10dB
Phase difference	$90^\circ \pm 10^\circ$	$90^\circ \pm 10^\circ$	$90^\circ \pm 10^\circ$	$90^\circ \pm 10^\circ$	-	-
Operating frequency (GHz)	0.76-0.84	1.76-1.88	0.8-0.92	1.88-2	0.8-0.94	1.84-2.04

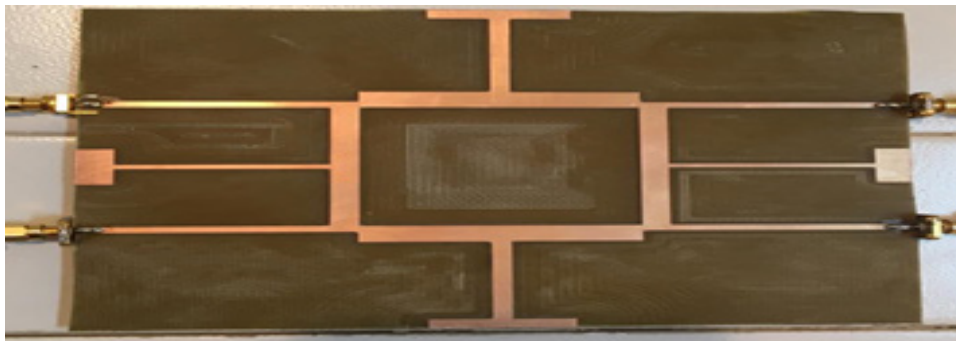


Figure 8: Photograph of the fabricated prototype

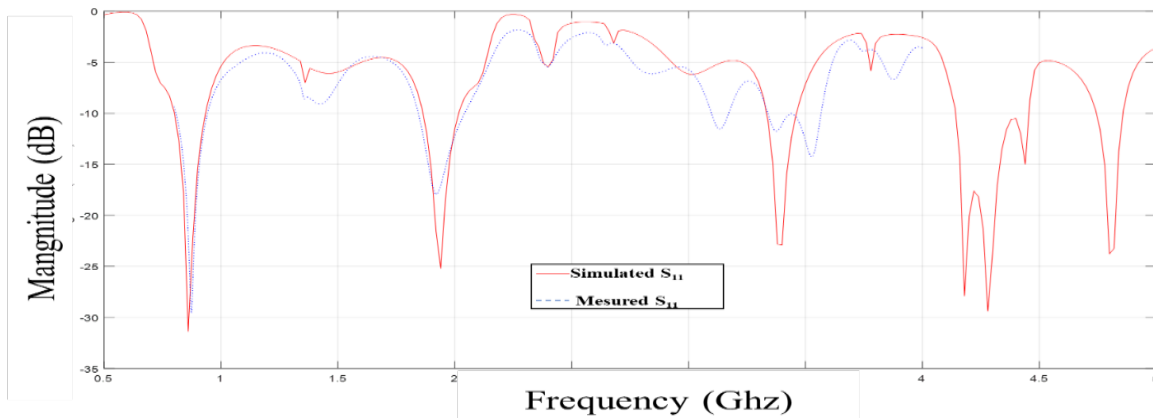


Figure 9: simulated and measured S_{11}

5. Conclusion

A dual-band BLC with SISBL for 5G NR applications has been designed. The coupler performance is further improved through optimization of his physical dimensions, using an adjusted Tuning Space Mapping to work for more complex microwave circuits such as our design. The simulated and measured results show that S_{11} are less than -11dB, within the frequency range of 0.82-0.92 GHz (first band), and within 1.88-2 GHz (second band). The method used permits rapid optimization with accurate results obtained after only one iteration.

References

- [1] 3GPP specification series: 38series.url: <https://www.3gpp.org/DynaReport/38-series.htm> (visited on 20/17/2020)
- [2] D. M. Pozar, Microwave Engineering, 4th ed. New York: Wiley, 2012. www.astesj.com
- [3] M.-J. Park and B. Lee, "Dual-band, cross coupled branch line coupler" IEEE Microw. Wireless Compon. Lett., 15(10), 655–657, 2005. <https://doi.org/10.1109/LMWC.2005.856683>
- [4] I.-H. Lin, M. DeVincentis, C. Caloz, and T. Itoh, "Arbitrary dual-band components using composite right/left-handed transmission lines" IEEE Trans. Microw. Theory Tech., 52(4), 1142–1149, 2004. <https://doi.org/10.1109/TMTT.2004.825747>
- [5] J.-X. Niu and X.-L. Zhou, "A novel dual-band branch line coupler based on strip-shaped complementary split ring resonators" Microw. Opt. Technol. Lett., vol. 49(11), 2859–2862, 2007. <https://doi.org/10.1002/mop.22873>
- [6] K.-K. M. Cheng and F.-L. Wong, "A novel approach to the design and implementation of dual-band compact planar 90 branch-line coupler" IEEE Trans. Microw. Theory Tech., 52(11), 2458–2463, <https://doi.org/10.1109/TMTT.2004.837151>
- [7] C. Collado, A. Grau, and F. D. Flaviis, "Dual-band planar quadrature hybrid with enhanced bandwidth response" IEEE Trans. Microw. Theory Tech., 54(1), 180–188, 2006. <https://doi.org/10.1109/TMTT.2005.860306>

- [8] K.-K. M. Cheng and F.-L. Wong, "Dual-band rat-race coupler design using tri-section branch-line" *Electron. Lett.*, 43(6), 41–42, Mar. 2007. <https://doi.org/10.1049/el:20070018>
- [9] C.-L. Hsu, J.-T. Kuo, and C.-W. Chang, "Miniaturized dual-band hybrid couplers with arbitrary power division ratios" *IEEE Trans. Microw. Theory Tech.*, 57(1), 49–156, 2009. <https://doi.org/10.1109/TMTT.2008.2009036>
- [10] K.-S. Chin, K.-M. Lin, Y.-H. Wei, T.-H. Tseng, Y.-J. Yang "Compact Dual-Band Branch-Line and Rat-Race Couplers with Stepped-Impedance-Stub Lines" *IEEE Transactions on Microwave Theory and Techniques*, 58(5), 2010. <https://doi.org/10.1109/TMTT.2010.2046064>
- [11] J. Meng, S. Koziel, J.W. Bandler, M.H. Bakr, and Q.S. Cheng, "Tuning space mapping: a novel technique for engineering optimization" in 2008 IEEE MTT-S International Microwave Symposium Digest, Atlanta, GA, USA, 2008. <https://doi.org/10.1109/MWSYM.2008.4633001>
- [12] D. Echeverria, P.W. Hemker, "Space mapping and defect correction," *The International Mathematical Journal Computational Methods in Applied Mathematics*, 5(2), 107–136, 2008. https://doi.org/10.1007/978-3-540-78841-6_8
- [13] M.A. Ismail, D. Smith, A. Panariello, Y Wang, and M. Yu, "EM-based design of large-scale dielectric-resonator filters and multiplexers by space mapping" *IEEE Trans. Microwave Theory Tech.*, 52(1), 386–392, 2004. <https://doi.org/10.1109/TMTT.2003.820900>
- [14] T.W. Simpson, J. Peplinski, P.N. Koch, and J.K. Allen, "Metamodels for computer-based engineering design: survey and recommendations" *ENG. COMPUT.*, 17(2), 129–150, 2001. <https://doi.org/10.1007/PL00007198>
- [15] J.W. Bandler, R.M. Biernacki, S.H. Chen, P.A. Grobelny, and R.H. Hemmers, "Space mapping technique for electromagnetic optimization" *IEEE T. MICROW. THEORY.*, 42(12), 2536–2544, 1994. <https://doi.org/10.1109/22.339794>
- [16] N.V. Queipo, R.T. Haftka, W. Shyy, T. Goel, R. Vaidynathan, and P.K. Tucker, "Surrogate-based analysis and optimization," *PROG. AEROSP. SCI.*, 41(1), 1–28, 2005. <https://doi.org/10.1016/j.paerosci.2005.02.001>
- [17] S. Koziel, J.W. Bandler, A.S. Mohamed, and K. Madsen, "Enhanced surrogate models for statistical design exploiting space mapping technology" in IEEE MTT-S International Microwave Symposium Digest, Long Beach, CA, USA, 2005. <https://doi.org/10.1109/MWSYM.2005.1517012>
- [18] S. Koziel, J.W. Bandler, and K. Madsen, "Theoretical justification of spacemapping-based modeling utilizing a data base and on-demand parameter extraction" *IEEE T. MICROW. THEORY.*, 54(12), 4316–4322, 2006. <https://doi.org/10.1109/TMTT.2006.884648>
- [19] S. Koziel and J.W. Bandler, "Microwave device modeling using spacemapping and radial basis functions" in 2007 IEEE/MTT-S International Microwave Symposium, Honolulu, HI, 2007. <https://doi.org/10.1109/MWSYM.2007.380079>
- [20] V.K. Devabhaktuni, B. Chattaraj, M.C.E. Yagoub, and Q.-J. Zhang, "Advanced microwave modeling framework exploiting automatic model generation, knowledge neural networks, and space mapping" *IEEE T. MICROW. THEORY.*, 51(7), 1822–1833, 2003. <https://doi.org/10.1109/MWSYM.2002.1011836>
- [21] A. SAADI, S. BRI, "Non-Linear Optimization of Small Size Microwave Directional Coupler Design Using Implicit Space" *Int. J. Microw. Appl.*, Volume 2(4), 2013. http://warse.org/pdfs/2013/ijma022420_13.pdf
- [22] F. Feng; C. Zhang; W. Na; J. Zhang; W. Zhang; Q.-J. Zhang, "Adaptive Feature Zero Assisted Surrogate-Based EM Optimization for Microwave Filter Design" *IEEE MICROW. WIREL. CO.*, 29(1), 2-4, 2019. <https://doi.org/10.1109/LMWC.2018.2884643>
- [23] B. Liu, V. Grout, A. Nikolaeva, "Efficient Global Optimization of Actuator Based on a Surrogate Model Assisted Hybrid Algorithm" *IEEE T. IND. EL. CON. IN.*, 65(7), 5712-5721, 2018. <https://doi.org/10.1109/TIE.2017.2782203>

A Survey on Image Forgery Detection Using Different Forensic Approaches

Akram Hatem Saber^{1,*}, Mohd Ayyub Khan¹, Basim Galeb Mejbil²

¹Department of Electronics Engineering, Aligarh Muslim University, 202002, India

²Department of Computer Technician Engineering, AL-Esraa University, 10069, Iraq

ARTICLE INFO

Article history:

Received: 18 February, 2020

Accepted: 13 May, 2020

Online: 11 June, 2020

Keywords:

Forgery Detection

Active Forensic Approaches

Passive Forensic Approaches

Tampering Identification

Copy-Move

Digital Watermarking

Digital Signature

Image Cloning

Image Splicing

ABSTRACT

Recently, digital image forgery detection is an emergent and important area of image processing. Digital image plays a vital role in providing evidence for any unusual incident. However, the image forgery may hide evidence and prevents the detection of such criminal cases due to advancement in image processing and availability of sophisticated software tamper of an image can be easily performed. In this paper, we provide a comprehensive review of the work done on various image forgeries and forensic technology. Many techniques have been proposed to detect image forgery in the literature such as digital watermarking, digital signature, copy-move, image retouching, and splicing. The investigation done in this paper may help the researcher to understand the advantage and handles of the available image forensic technology to develop more efficient algorithms of image forgery detection. Moreover, the comparative study surveys the existing forgery detection mechanisms include deep learning and convolution neural networks concerning its benefits and demerits.

1. Introduction

Digital images are the major information source in recent days, due to its availability and sophistication [1]. Also, it is widely used in different fields, detection of digital image forgery is utilized in numerous applications that are linked to media, publication, law, military, medical image science applications, satellite image, and world wide web publications. Because it is very easy to manipulate and edit [2]. For this reason, different types of cameras and the user-friendly software are used to create and edit the digital images [3]. Digital images are frequently used to support the important decision for many situations. Moreover, the digital images are a popular source of information and the reliability of digital image and it becomes an important issue.

For image forensics, the techniques are classified into two, such as the active approach and passive approach. In the case of active approach: in this method, the digital image entails the various types of preprocessing like watermark embedded or signature are added in the original image. Digital watermarking and signature are two different active protection techniques. If the image has tampered, special information is not extracted from the obtained image. Watermarking is one of the methods

of active tampering detection and security structure is embedded into the image but most of the image processing tools are not contained any watermarking or signature module[4].

In recent days different methods are developed for made image reliable and secure that is analogous to watermarking like message authentication code, image checksum, image hash, and image shielding. Passive image forensics is a challenging task in image processing techniques[5]. It is not a particular method for all cases but different methods each can detect the special forgery. The stream of passive tampering detection is to deal with analyzing raw image based on different statistics and semantics of an image content to localize tampering of image[6].

There are several types of image forgery that include image retouching, image splicing, copy and move attack. Image retouching is considered a minimum harmful type of digital image forgery. An original image does not significantly change, but they reduced some features of the original image. This technique is used to edit the image for a popular magazine. This type of image forgery is located in all magazine covers and also it used to improve the specific features of an image[7]. On the other hand, Image splicing or photo montage refers to make a forgery image and it is more aggressive than image

*Akram Hatem Saber, Aligarh, +919515584268, alasmr.2a@gmail.com

www.astesj.com

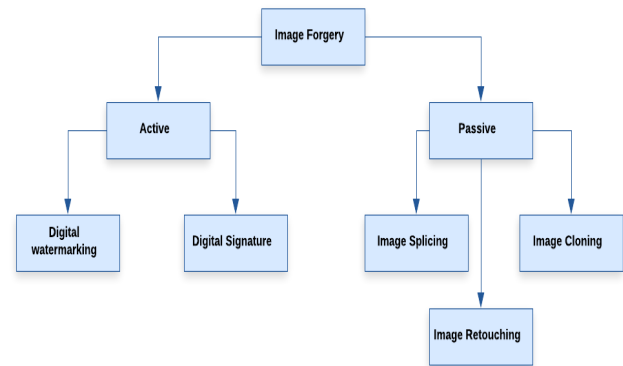
<https://dx.doi.org/10.25046/aj050347>

retouching. Image splicing is an easy process and it pastes the regions from isolated sources. This method is referred to as paste-up formed by sticking together the image by using digital tools like photoshop. This technique is a group of two or more images that are combined to generate fake images[8]. However, copy and move attack is also one of the popular and difficult images tampering technique. It required the cover part of a similar image to add or remove the information. Copy and move attack, the aim is to hide some information in the original image. The detection of the forged image from the original one is very hard. The naked eye is not able to identify the tampered region from a forged image. The image tampering is a general manipulation of digital images. Traditional block-based forgery detection methodologies are categorized as the input images into overlapping and regular image blocks and also tampered regions are identified by matching blocks of pixel or transform coefficient[9].

Normally, the image forgery detection is performed by using the following techniques: JPEG quantization tables, Chromatic Aberration, Lighting, Camera Response Function (CRF), Bi-coherence and higher-order statistics, and Robust matching. The digital cameras encode the images based on JPEG compression [10], which configures the devices at various compression level. Then, the sign of image tampering is evaluated by analyzing the inconsistency of lateral chromatic aberration [11]. In which, the average angular between the local and global parameters is computed for every pixel in the image. If the average value exceeds the threshold, it is stated that the deviation is unpredictable in the image due to the forgery of the image. Then, for each object in the image, the inconsistencies and the illuminating light source is detected to identify the forgery [12]. Typically, different measurements such as infinite, local and multiple are considered for determining the error rate. Then, the CRF is mainly used to expose the image splicing instituted on the geometry invariant of the image. In which, the suspected boundary is identified within each region of the image, and it is validated for identifying the inconsistencies [13]. The bi-coherence features [14] are widely used for detecting the splicing on images that estimate the mean of magnitude and phase entropy for augmenting the images. Moreover, it extracts the features for the authentic counterpart and incorporates it to capture the characteristics of various object interfaces. Finally, the exact replicas are identified by matching the features concerning the block size, which is done by the use of robust matching [15]. But, it requires the human intervention for interpreting the output of replicas detection [16]. Generally, the region duplication is performed on the image based on the geometrical and illumination adjustments. It is a very simple operation in which a continuous portion of pixels is copied and pasted on some other location in the image. This paper is fully focused on the detailed investigation of the image forgery detection mechanisms. The remaining sectors present in the study are arranged as follows: Section II investigates some of the image forgery detection mechanisms used in digital image processing. Section III surveys the forensic approaches and its working procedure for image forgery detection. Section IV presents a detailed investigation of the existing methodologies used for image forgery detection with its advantages and disadvantages. The overall conclusion of the paper is presented in Section V.

2. Digital Image Forgery Detection Methods

Typically, the methodologies used for forgery detection are classified into two types such as active forensics and passive forensics, in which digital watermarking and digital signature are the types of active techniques. Then, the splicing, image retouching, image cloning, and copy-move techniques are the categories of the passive technique [17]. The description of these techniques are investigated in the following sub-sections.



2.1. Digital Watermarking

In this type of image forgery, a digital watermark is added on the photo, which is more or less visible. Then, the appended information is more or less transparent, so it is very difficult to notice the watermark. Ferrara, et al. [18] suggested a new forensic tool for analyzing the original image and forged regions based on the interpolation process. The image splicing can be detected by the use of the conditional Co-occurrence Probability Matrix (CCPM) [19], which uses the third-order statistical features during the forgery detection. Normally, the watermarking schemes are categorized as reversible and irreversible. In which, the image irreversible distortions are avoided based on the original features of the image by using the reversible watermarking techniques. The watermarking can be mainly used to indicate the source or authorized consumer of the image. It is a pattern of bits that is inserted into a digital media for identifying the creator [20]. The watermarking techniques are semi-fragile, fragile, and content based, which are mainly used for image authentication application.

Li, et al. [21] implemented a new method for detecting the copy move forgery, where the Local Binary Pattern (LBP) was utilized to extract the circular blocks. The stages involved in this system are preprocessing, feature extraction, feature matching, and post processing. Here, it is stated that when the region is rotated at different angles, it is highly difficult to detect the forgeries. Hussain, et al. [22] suggested a multi-resolution Weber Local Descriptors (WLD) for detecting the image forgeries based on the features obtained from the chrominance components. Here, the WLD histogram components are calculated and the Support Vector Machine (SVM) classifier is utilized to detect the forgery. In this paper, two different types of forgeries such as splice and copy-move are detected by using the multi-resolution WLD approach.

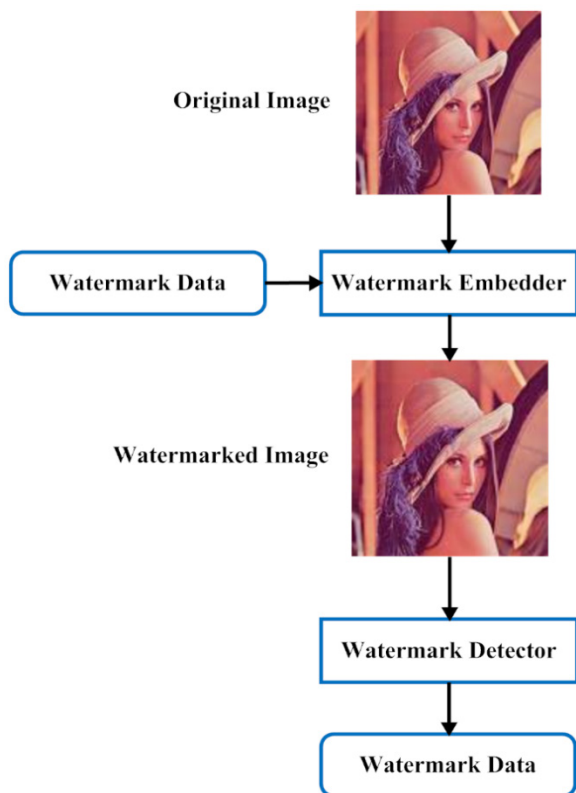


Figure 2. Digital watermarking [18]

2.2. Digital Signature

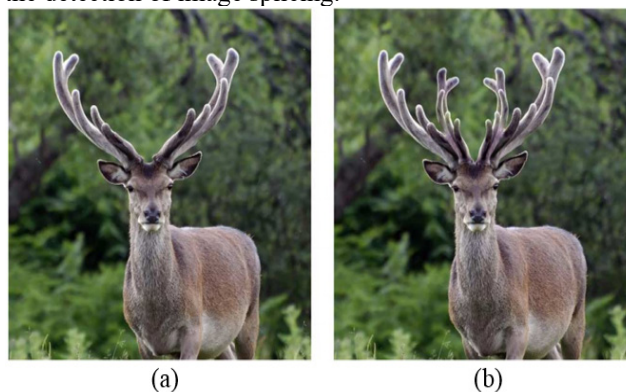
Normally, the authenticity of the digital messages is validated based on the digital signature. Because, based on the valid signature, the recipient can believe that the message is formed by the recognized sender. Thus, the digital signature is widely used in the fields of financial transactions, contract management software, and software distribution [20]. Normally, the digital signature embeds some secondary information, which is obtained from the image. In this method [23], the distinct features are extracted from the image during the initial stage, based on these, the image authenticity is validated. Typically, the digital signature has the following properties:

- Only the sender can sign the image and the receiver can validate the signature
- Unauthenticated users cannot able to forge the signature
- It provides an integrity
- Also, it achieves non-reputation

2.3. Splicing Method

Image splicing is a kind of forgery detection method, in which a single image is created based on the combination of two or more images [24]. It is also termed as image composition, in which various image manipulation operations are performed. Typically, many inconsistencies may be created in the image features due to the splicing operation. In this technique, the composition between the two images is estimated and incorporated for creating a fake image. Based on the image block content, the difference between the illumination and reference illuminate color is estimated. In this digital image

forgery, it is very difficult to extract the exact shape of the image. Typically, the image splicing method [25] is categorized into two types such as boundary-based and region-based. *Alahmadi, et al* [26] suggested a passive splicing forgery detection mechanism for verifying the authenticity of digital images. Here, the features are extracted from the chromatic channel for capturing the tampering artifacts. *Kakar, et al* [27] utilized a forgery detection approach for detecting the splicing in the digital images. Here, the small inconsistencies in the motion blur are detected by analyzing the special characteristics of image gradients [28]. The stages involved in this detection are image subdivision, motion blur estimation, smoothing, blur computation, interpolation and segmentation [29]. The authors of this paper [30] employed a machine learning algorithm for detecting the image splices. The illumination analysis is highly effective for the detection of image splicing [31]. To increase the effect of photorealism, an image splicing operation is performed with the operations of color and brightness adjustment. In this paper [32] the radial distortion from various portions of the image is estimated for the detection of image splicing.



2.4. Image Retouching

Among the other image forgeries, image retouching is considered as the less harmful forgery technique, in which some enhancement can be performed on the image. Also, it is popular in photo editing applications and magazines. *Muhammad, et al* [33] suggested an un-decimated dyadic wavelet transformation technique for detecting the copy-move forgery. Typically, more sophisticated tools are available for making this type of forgery by applying the soft touch on the edges. So, it is very difficult to differentiate the color and texture of the stimulated part with the unoriginal part. Moreover, it makes the forgery detection as highly complicated, because of two or more identical objects in the same image. So, the authors of this paper utilized similarity measurements for detecting this forgery, in which the noisy inconsistency is analyzed between the copied and moved parts. Here, it is stated that the transformation methods such as FMT, Scale Invariant Feature Transform (SIFT), and Discrete Wavelet Transform (DWT) can detect the forgery in a highly compressed image. *Ghorbani, et al* [34] recommended a Discrete Cosine Transform Quantization Coefficients Decomposition (DCT-QCD) for detecting the copy-move forgery. The integrity and authenticity verification of digital

images is a very difficult process, specifically the images used for news items, medical records, and court law. Because the copy-move forgery may be created for those types of images.



Figure 4 (a). Forged image and (b). Real image [33]

2.5. Copy-Move Method

Among the other forgery methods, the copy-move method an extensively used type of image tampering, where the specific portion is copied and pasted on some other region [35]. The main motive of this method is to hide a significant element or highlight a precise object. Bayram, et al [36] implemented a proficient method for detecting the copy-move forgery. The authors stated that the block matching procedure is used to detect this type of forgery by separating the image into overlapping chunks. Also, it identifies the duplicated connected image blocks by finding the distance between the neighbor blocks [37]. For taking the forgery decision, only the duplicate blocks detection is not enough, because the natural images have many similar blocks [38]. Moreover, the Fourier Mellin Transform (FMT) is used to perform the operations like scaling, translation, and rotation for image forgery detection [39].

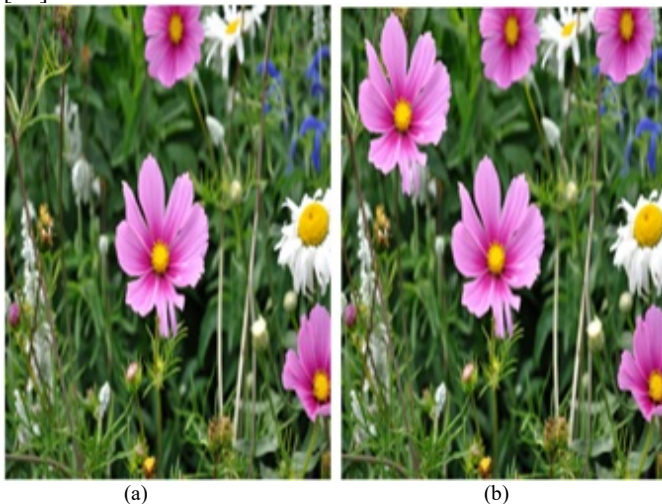


Figure 5 (a). Original image and (b). Tampered image [36].

As shown in figure above copy move image forgery (a) original image and (b) is tampered image Mahdian, et al [40] utilized a detection method for identifying the copy-move forgery based on the blur moment invariants. This detection

methodology can detect blur degradation, noise, and some other arbitrary changes in the duplicate image regions like noise addition and gamma correction gamma is a non-linear adjustment to individual pixel values. The steps involved in this method are image tiling with overlapping, representation blur moment invariants, transformation, similarity analysis, and map creation for duplication region detection. Moreover, the dimensionality of blocks was reduced by using the principle component transformation. Muhammad, et al [41] employed a Dyadic undecorated Wavelet Transformation (Dew) technique for detecting blind copy-move image forgery detection. This transformation technique aimed to extract the low frequency and high-frequency components by estimating the similarity between the blocks [42]. Moreover, the Euclidean distance is computed between every pair of blocks in the image. Then, the match is identified by computing the threshold value between the sorted lists [43]. In the wavelet transformation, the downsampling process is not involved, and the coefficients are not shrunk between the scales. Lynch, et al [44] aimed to detect the copy-move forgery by the use of expanding block algorithms. Also, it intended to identify the duplicated regions in the image by estimating the size and shape [45]. In this paper [46], it is stated that the copy-move forgery is performed for hiding the region of the image by wrapper it with a duplicate image. Still, recognizing the forged region is extremely intricate due to the precise copy of another region [47]. This detection mechanism contains the stages of feature extraction, comparison, and similarity estimation for taking copy decisions [48]. As shown in figure below the procedure of copy move technique first step the input image preprocessed, second step block division, third step feature extracted, last step the blocks which carry same feature triggered and mapped as a forgery.

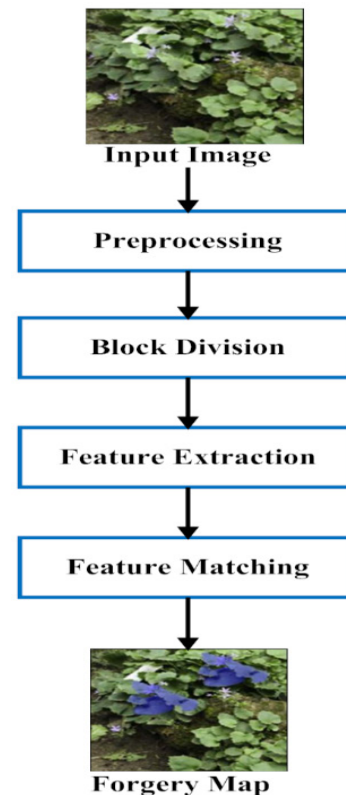


Figure 6. Procedure of Copy Move technique [48]

3. Forensic Approaches

In this section, some of the forensic approaches used for digital image forgery detection are surveyed with its working stages. *Omen, et al* [49] utilized a fractal dimension and Support Vector Decomposition (SVD) techniques to detect and isolate the duplicated regions in the image. In this scheme, the image is classified into various groups based on its fractal dimension, which is used to identify the variations. Then, the copied and pasted regions are identified by using an efficient texture-based classification technique. Here, it is stated that the SVD is one of the widely used robust and reliable matrix factorization methods, which offers algebraic and geometric invariant features for classification. Also, the SVD technique provides maximum energy packaging for exhibiting good stability from distortion. It helps to locate the duplicated regions by adding noise and avoiding the blurred edges. *Chierchia, et al* [50] implemented a Bayesian Markov Random Field (MRF) technique to identify the image forgeries based on the sensor pattern noise. Here, the observed statistics and prior knowledge were balanced by the use of a Bayesian approach. Also, the reliability of forgery detection is improved by using the global optimization algorithm. *Bianchi and Piva* [51] developed a new forensic algorithm for discriminating against the original and forged regions in the image. Here, the effects of cumulation between various DCT coefficients are extracted with the simplified map by using the unified statistical model.

Murali, et al [52] investigated various image forgery detection mechanisms for identifying the forged regions in the forged image. In this paper, it is stated that the copy-move and copy create types are the two kinds of image forgeries, which are implemented at earlier stages. It is detected by using the JPEG compression analysis and filtering algorithms. Here, the algorithms are evaluated based on the factors of image formation, time complexity, multiple forgery detection, and image transformation. *Piva* [53] provided a comprehensive overview of image forensics for determining whether the image content is authenticated or not. The methods investigated in this paper were acquisition-based, coding-based methods, and editing based methods. *Pan, et al* [54] suggested a feature matching technique for identifying the duplicated regions in the digital image.

4. Comparative Study

This section surveys the existing forgery detection mechanisms with respect to its own benefits and demerits. This study is mainly focused on the detection of image forgery by using various forensic approaches. The methods that have been investigated in this analysis are digital signature verification, digital image watermarking, cosine transformation, authentication watermarking, SURF, wavelet transformation, binary pattern extraction, deep learning, block matching, and blind image forgery detection.

Table 1. Comparative analysis of various image forensic approaches

S.No	Paper Title	Methods Used	Tampering Detection Type	Pros/Cons	Publication Year
1.	Research issues and challenges for multiple digital signatures	Digital Signature Verification Schemes [55]	The validity of multiple digital signatures are verified.	Advantage: 1. It provides the clear overview of various signature verification schemes with its specific limitations. Disadvantage: 1. However, it failed to state an efficient and robust signature verification scheme	2005
2.	ROI based tamper detection and recovery for medical images using reversible watermarking technique	Digital image watermarking [56]	It is used to detect the locations of the tampered portion inside the Region of Interest (ROI).	Advantages: 1. Good performance in terms of hiding capacity and visual quality 2. High embedding capacity Disadvantages: 1. Lack of reversibility 2. Limited hiding capacity Induced distortions inside the regions	2010
3.	A comparison study on copy-cover image forgery detection.	Discrete Cosine Transformation (DCT) and Principle Component Analysis [57]	It detected a copy-move image forgery.	Advantages: 1. Energy compaction property 2. Reduced time complexity 3. Increased accuracy Disadvantages: 1. It required to locate the possible inconsistency 2. Increased false positive rate	2010
4.	A chaotic system based fragile watermarking scheme for image tamper detection	Authentication watermarking scheme [58].	It locates the tampered regions for image authentication.	Advantages: 1. High security 2. Superior tamper detection and localization Disadvantages: 1. Increased computational complexity 2. Required to improve the performance	2011
5.	DWT-DCT (QCD) based copy-move image forgery detection	Discrete Wavelet Transformation (DWT) and Discrete	It detected a copy-move image forgery in an accurate manner.	Advantages: 1. Better accuracy 2. Reduced dimensionality of features Disadvantages:	2011

		Cosine Transformation (DCT) [34]		<ol style="list-style-type: none"> 1. Heavy compression 2. It required to remove the position of pasted areas 3. Increased complexity 	
6.	Detection of region duplication forgery in digital images using SURF	Speeded Up Robust Features (SURF) [59]	A copy move forgery is detected with better detection performance.	<p>Advantages:</p> <ol style="list-style-type: none"> 1. Better detection rate 2. It evaluated the image with different angles <p>Disadvantages:</p> <ol style="list-style-type: none"> 1. Required to reduce the false match rate 2. Also, it failed to identify the small copied regions. 	2011
7.	Passive copy move image forgery detection using undecimated dyadic wavelet transform	Undecimated dyadic wavelet transformation [33]	A copy move image forgery is detected efficiently.	<p>Advantages:</p> <ol style="list-style-type: none"> 1. It estimated the methods based on three case studies 2. Better performance results <p>Disadvantages:</p> <ol style="list-style-type: none"> 1. Noise estimation is not robust 2. It is not translation invariant 	2012
8.	A novel video inter-frame forgery model detection scheme based on optical flow consistency	Inter-frame forgery model detection mechanism [60]	It detected the frame insertion and deletion forgery.	<p>Advantages:</p> <ol style="list-style-type: none"> 1. It provides the good performance by efficiently identifying the frame insertion and deletion <p>Disadvantages:</p> <ol style="list-style-type: none"> 1. Reduced precision 2. Increased false detection rate 	2013
9.	Digital image tamper detection techniques-a comprehensive study	Fragile watermark detection technique [61]	Authentication based tampering detection is performed.	<p>Advantages:</p> <ol style="list-style-type: none"> 1. Robust watermark 2. It accurately pinpoint the forgeries <p>Disadvantages:</p> <ol style="list-style-type: none"> 1. It required a digital signature on the images 2. Not highly efficient 	2013
10.	Survey on blind image forgery detection	Blind image forgery detection [62]	It detects the copy-move, splicing, and retouching image forgeries.	<p>Advantages:</p> <ol style="list-style-type: none"> 1. It evaluated different number of matches for forgery identification 2. It efficiently identified the duplicated blocks <p>Disadvantages:</p> <ol style="list-style-type: none"> 1. It required to analyze the quality of image 2. Increased time consumption 	2013
11.	Splicing image forgery detection based on DCT and Local Binary Pattern	Local Binary Pattern (LBP) and Discrete Cosine Transformation (DCT) [26]	Here, an image splicing forgery is detected accurately.	<p>Advantages:</p> <ol style="list-style-type: none"> 1. Better detection performance 2. Increased accuracy <p>Disadvantages:</p> <ol style="list-style-type: none"> 1. Increased complexity 2. Not highly efficient 	2013
12.	A Forensic Method for Detecting Image Forgery Using Codebook	SIFT based feature extraction and codebook generation [63]	Dissimilar types of image tampering are concentrated in this paper that includes enhancing, compositing and copy move.	<p>Advantages:</p> <ol style="list-style-type: none"> 1. Highly efficient 2. Better accuracy <p>Disadvantages:</p> <ol style="list-style-type: none"> 1. Requires more time for detection 2. It distorts the content 3. Inconclusive results 	2013
13.	Region Duplication Forgery Detection using Hybrid Wavelet Transforms	Hybrid wavelet transformation technique [64]	It detected a copy move image forgery and region duplication forgery.	<p>Advantages</p> <ol style="list-style-type: none"> 1. Effective compression 2. It detected the duplicated regions with increased accuracy <p>Disadvantages</p> <ol style="list-style-type: none"> 1. It failed to detect the duplicated regions, when the copied region is rotated or scaled 2. Not highly efficient 	2014
14.	Digital image forgeries and passive image authentication techniques: A survey	Passive image authentication techniques [20]	A copy move image forgery is detected in an efficient way.	<p>Advantages:</p> <ol style="list-style-type: none"> 1. Reduced computational complexity 2. Increased robustness <p>Disadvantages:</p> <ol style="list-style-type: none"> 1. Sharp edge disturbances after splicing 2. Not reliable feature extraction 	2014

15.	Image Forgery Detection using Speed up Robust Feature Transform, Wavelet Transform, Steerable Pyramid Transform and Local Binary Pattern	Discrete Wavelet Transformation (DWT) and dyadic wavelet transformation techniques [65]	Copy move image forgery is detected with better accuracy.	Advantages: 1. Good efficiency 2. Reliable Disadvantages: 1. Not more suitable for noisy image 2. Time complexity is high	2016
16.	An Evaluation of Digital Image Forgery Detection Approaches	Pixel based image forgery detection [19]	Image splicing, copy-move and image resampling forgeries are detected.	Advantages: 1. Better accuracy 2. High reliability Disadvantages: 1. Will not work in the noisy image 2. Time consuming	2017
17.	A Review Paper on Digital Image Forgery Detection Techniques	Brute force, block based and key point based techniques [66]	A generalized schema is developed for detecting a copy move image forgery.	Advantages: 1. Reduced complexity 2. Quite robust Disadvantages: 1. Not efficient for complicated background and texture 2. Less accurate	2017
18.	Boosting Image Forgery Detection using Resampling Features and Copy-move Analysis	Deep learning mechanism [67]	The copy move image features are identified for detecting the forgery.	Advantages: 1. Reduced false positive 2. Highly efficient Disadvantages: 1. Not highly robust 2. Less accurate	2018
19.	Accurate and Efficient Image Forgery Detection Using Lateral Chromatic Aberration	Lateral Chromatic Aberration (LCA) and block matching algorithm [68]	Image forgery is detected by analyzing the hypothesis testing problem.	Advantages: 1. Increased efficiency 2. Reduced complexity Disadvantages: 1. Increased estimation error 2. Not suitable for noisy images	2018
20.	Recent Advances in Passive Digital Image Security Forensics: A Brief Review	Passive digital image forensic approaches [69]	It detected the image forgeries based on the artifacts.	Advantages: 1. Better generalization ability 2. Minimized time consumption Disadvantages: 1. Handling difficulty in most forgery cases 2. Performance degradation	2018
21	Image Splicing Detection using Deep Residual Network	this approach three classifiers Multiclass Model using SVM Learner, K-NN and Naïve Bayes are used to train the classifier model[70]	Spliced image forgery detection using image as input for CNN and processed through various layers	Advantages: 1- Increase the accuracy 2- Localization of spliced forged image efficiently Disadvantages: 1- Not suitable for copy-move forgery detection 2- Required highly performance system to implement the algorithms	2019
22	Image splicing forgery detection combining coarse to refined convolutional neural network and adaptive clustering	paper proposes detection method with two parts: Coarse-to- refined convolutional neural network (C2RNet) and diluted adaptive Clustering, replace patch-level CNN in C2RNet.[71]	Spliced image forgery detection with two parts (C2RNet) and diluted adaptive Clustering.	Advantages: 1- Decrease the computational complexity. 2- Tremendous decrease in the time. Disadvantages: 1- Slightly Poorer in visual performance. 2- Poorer in Recall than that of several of comparison methods.	2019
23	Image Forgery Detection: A Low Computational-Cost and Effective Data-Driven Model	low computational-cost and effective data-driven model as a Modified deep learning-based model [72]	Daubechies wavelet transform is utilized, representing YCrCb patches inside the image, neural network used to classify forged patches.	Advantages: 1- Reduce computational cost. 2- Increase accuracy. Disadvantages: 1- Not highly robust 2- Time complexity is high	2019

24	Morphological Filter Detector for Image Forensics Applications	Mathematical morphological filter detector (considered Gaussian low pass and Median filtering)[73]	operates on grayscale images, propose a non-trivial extension of a deterministic approach originally detecting erosion and dilation of binary images	Advantages: 1- Robust to image compression 2- Very high accuracy Disadvantages 1- Mathematical complexity 2- Time complexity	2020
25	Constrained Image Splicing Detection and Localization With Attention-Aware Encoder-Decoder and Atrous Convolution	Newly methods used AttentionDM for CISDL[74]	Splice forgery detection, and detects whether one image has forged regions pasted from the other	Advantages: 1- Performance improved 2- Computational improved Disadvantages: 1- Equal error rate and detection rate reduced 2- Slightly slower than DMAC	2020
26	Deep Learning Local Descriptor for Image Splicing Detection and Localization	Deep convolution neural network CNN, a two branch CNN used with automatically learn hierarchical [75]	Image splice detection and localization scheme	Advantages: 1- Robustness against JPEG compression 2- Highly detection accuracy Disadvantages: 1- Huge complexity while used 30 linear high pass filter 2- Future fusion is complex	2020

5. Conclusion and Future Work

This paper surveyed various image forensics approaches for identifying the forgeries performed on the digital images. The techniques investigated in this paper are digital signature, digital watermarking, copy-move, image splicing, and image cloning. Most of the authors stated that image forgery detection is a highly complicated process due to the advent of various manipulation and editing tools. The feature is also playing an essential role in forgery detection because the features are highly sensitive to some forgery operations. Moreover, different image processing techniques such as preprocessing, feature extraction, feature selection, and classification are highly useful for detecting the forgeries in an exact manner. The passive methods are highly suitable for forgery detection compared to the active approaches. Because it analyzes the pixel variations and estimates the geometrical illuminations in an efficient manner. Among the other passive methods, the copy-move and image splicing are widely used by many researchers due to its benefits of reduced complexity and increased accuracy.

References

- [1] B. Mahdian and S. Saic, "A bibliography on blind methods for identifying image forgery," *Signal Processing: Image Communication*, vol. 25, pp. 389-399, 2010.
- [2] J. Li, X. Li, B. Yang, and X. Sun, "Segmentation-based image copy-move forgery detection scheme," *IEEE Transactions on Information Forensics and Security*, vol. 10, pp. 507-518, 2015.
- [3] H.-D. Yuan, "Blind forensics of median filtering in digital images," *IEEE Transactions on Information Forensics and Security*, vol. 6, pp. 1335-1345, 2011.
- [4] C. Chen, J. Ni, and J. Huang, "Blind detection of median filtering in digital images: A difference domain based approach," *IEEE Transactions on Image Processing*, vol. 22, pp. 4699-4710, 2013.
- [5] X. Lin, J.-H. Li, S.-L. Wang, A.-W.-C. Liew, F. Cheng, and X.-S. Huang, "Recent Advances in Passive Digital Image Security Forensics: A Brief Review," *Engineering*, 2018/02/17/ 2018.
- [6] M. D. Ansari, S. P. Ghreera, and V. Tyagi, "Pixel-based image forgery detection: A review," *IETE journal of education*, vol. 55, pp. 40-46, 2014.
- [7] C.-M. Pun, X.-C. Yuan, and X.-L. Bi, "Image forgery detection using adaptive oversegmentation and feature point matching," *IEEE Transactions on Information Forensics and Security*, vol. 10, pp. 1705-1716, 2015.
- [8] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra, "Geometric tampering estimation by means of a SIFT-based forensic analysis," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, 2010, pp. 1702-1705.
- [9] W. Li and N. Yu, "Rotation robust detection of copy-move forgery," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*, 2010, pp. 2113-2116.
- [10] M. C. Stamm, S. K. Tjoa, W. S. Lin, and K. R. Liu, "Undetectable image tampering through JPEG compression anti-forensics," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*, 2010, pp. 2109-2112.
- [11] V. Christlein, C. Riess, and E. Angelopoulou, "On rotation invariance in copy-move forgery detection," in *Information Forensics and Security (WIFS), 2010 IEEE International Workshop on*, 2010, pp. 1-6.
- [12] E. Kee and H. Farid, "Exposing digital forgeries from 3-D lighting environments," in *Information Forensics and Security (WIFS), 2010 IEEE International Workshop on*, 2010, pp. 1-6.
- [13] M. Kobayashi, T. Okabe, and Y. Sato, "Detecting forgery from static-scene video based on inconsistency in noise level functions," *IEEE Transactions on Information Forensics and Security*, vol. 5, pp. 883-892, 2010.
- [14] Z. He, W. Sun, W. Lu, and H. Lu, "Digital image splicing detection based on approximate run length," *Pattern Recognition Letters*, vol. 32, pp. 1591-1597, 2011.
- [15] M. Jaber, G. Bebis, M. Hussain, and G. Muhammad, "Accurate and robust localization of duplicated region in copy-move image forgery," *Machine vision and applications*, vol. 25, pp. 451-475, 2014.
- [16] D. Cozzolino, D. Gragnaniello, and L. Verdoliva, "Image forgery detection through residual-based local descriptors and block-matching," in *Image Processing (ICIP), 2014 IEEE International Conference on*, 2014, pp. 5297-5301.
- [17] G. Muhammad, "Multi-scale local texture descriptor for image forgery detection," in *Industrial Technology (ICIT), 2013 IEEE International Conference on*, 2013, pp. 1146-1151.
- [18] P. Ferrara, T. Bianchi, A. De Rosa, and A. Piva, "Image forgery localization via fine-grained analysis of CFA artifacts," *IEEE Transactions on Information Forensics and Security*, vol. 7, pp. 1566-1577, 2012.
- [19] A. Kashyap, R. S. Parmar, M. Agrawal, and H. Gupta, "An Evaluation of Digital Image Forgery Detection Approaches," *arXiv preprint arXiv:1703.09968*, 2017.
- [20] S. Mushtaq and A. H. Mir, "Digital image forgeries and passive image authentication techniques: A survey," *International Journal of Advanced Science and Technology*, vol. 73, pp. 15-32, 2014.
- [21] L. Li, S. Li, H. Zhu, S.-C. Chu, J. F. Roddick, and J.-S. Pan, "An efficient scheme for detecting copy-move forged images by local binary patterns,"

- Journal of Information Hiding and Multimedia Signal Processing*, vol. 4, pp. 46-56, 2013.
- [22] M. Hussain, G. Muhammad, S. Q. Saleh, A. M. Mirza, and G. Bebis, "Image forgery detection using multi-resolution Weber local descriptors," in *2013 IEEE EUROCON*, 2013, pp. 1570-1577.
- [23] T. Mahmood, T. Nawaz, R. Ashraf, M. Shah, Z. Khan, A. Irtaza, et al., "A survey on block based copy move image forgery detection techniques," in *Emerging Technologies (ICET), 2015 International Conference on*, 2015, pp. 1-6.
- [24] D. Chauhan, D. Kasat, S. Jain, and V. Thakare, "Survey on Keypoint Based Copy-move Forgery Detection Methods on Image," *Procedia Computer Science*, vol. 85, pp. 206-212, 2016/01/01/ 2016.
- [25] T. Huynh-Kha, T. Le-Tien, S. Ha-Viet-Uyen, K. Huynh-Van, and M. Luong, "A Robust Algorithm of Forgery Detection in Copy-Move and Spliced Images," *International Journal of advanced Computer Science and Applications*, vol. 7, pp. 1-8, 2016.
- [26] A. A. Alahmadi, M. Hussain, H. Aboalsamh, G. Muhammad, and G. Bebis, "Splicing image forgery detection based on DCT and Local Binary Pattern," in *Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE*, 2013, pp. 253-256.
- [27] P. Kakar, N. Sudha, and W. Ser, "Exposing digital image forgeries by detecting discrepancies in motion blur," *IEEE Transactions on multimedia*, vol. 13, pp. 443-452, 2011.
- [28] A. C. Gallagher and T. Chen, "Image authentication by detecting traces of demosaicing," in *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*, 2008, pp. 1-8.
- [29] J. G. Han, T. H. Park, Y. H. Moon, and I. K. Eom, "Efficient Markov feature extraction method for image splicing detection using maximization and threshold expansion," *Journal of Electronic Imaging*, vol. 25, p. 023031, 2016.
- [30] T. J. De Carvalho, C. Riess, E. Angelopoulou, H. Pedrini, and A. de Rezende Rocha, "Exposing digital image forgeries by illumination color classification," *IEEE Transactions on Information Forensics and Security*, vol. 8, pp. 1182-1194, 2013.
- [31] F. Hakimi and I. M. H. Zanjani, "Image-Splicing Forgery Detection Based On Improved LBP and K-Nearest Neighbors Algorithm," *International Journal Of Electronics Information & Planning*, 2015.
- [32] H. R. Chennamma and L. Rangarajan, "Image splicing detection using inherent lens radial distortion," *arXiv preprint arXiv:1105.4712*, 2011.
- [33] G. Muhammad, M. Hussain, and G. Bebis, "Passive copy move image forgery detection using undecimated dyadic wavelet transform," *Digital Investigation*, vol. 9, pp. 49-57, 2012.
- [34] M. Ghorbani, M. Firouzmand, and A. Faraahi, "DWT-DCT (QCD) based copy-move image forgery detection," in *Systems, Signals and Image Processing (IWSSIP), 2011 18th International Conference on*, 2011, pp. 1-4.
- [35] Z. Mohamadian and A. A. Pouyan, "Detection of duplication forgery in digital images in uniform and non-uniform regions," in *Computer Modelling and Simulation (UKSim), 2013 UKSim 15th International Conference on*, 2013, pp. 455-460.
- [36] S. Bayram, H. T. Sencar, and N. Memon, "An efficient and robust method for detecting copy-move forgery," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, 2009, pp. 1053-1056.
- [37] M. Jaber, G. Bebis, M. Hussain, and G. Muhammad, "Improving the detection and localization of duplicated regions in copy-move image forgery," in *Digital Signal Processing (DSP), 2013 18th International Conference on*, 2013, pp. 1-6.
- [38] N. B. A. Warif, A. W. A. Wahab, M. Y. I. Idris, R. Ramli, R. Salleh, S. Shamshirband, et al., "Copy-move forgery detection: Survey, challenges and future directions," *Journal of Network and Computer Applications*, vol. 75, pp. 259-278, 2016.
- [39] J. ZHENG, W. HAO, and W. ZHU, "Detection of Copy-move Forgery Based on Keypoints' Positional Relationship," *JOURNAL OF INFORMATION & COMPUTATIONAL SCIENCE*, vol. 9, pp. 4729-4735, 2012.
- [40] B. Mahdian and S. Saic, "Detection of copy-move forgery using a method based on blur moment invariants," *Forensic science international*, vol. 171, pp. 180-189, 2007.
- [41] G. Muhammad, M. Hussain, K. Khawaji, and G. Bebis, "Blind copy move image forgery detection using dyadic undecimated wavelet transform," in *Digital Signal Processing (DSP), 2011 17th International Conference on*, 2011, pp. 1-6.
- [42] K. Asghar, Z. Habib, and M. Hussain, "Copy-move and splicing image forgery detection and localization techniques: a review," *Australian Journal of Forensic Sciences*, vol. 49, pp. 281-307, 2017.
- [43] N. Chaitawittanun, "Detection of copy-move forgery by clustering technique," *International Proceedings of Computer Science & Information Technology*, vol. 50, 2012.
- [44] G. Lynch, F. Y. Shih, and H.-Y. M. Liao, "An efficient expanding block algorithm for image copy-move forgery detection," *Information Sciences*, vol. 239, pp. 253-265, 2013.
- [45] L. Chen, W. Lu, and J. Ni, "An image region description method based on step sector statistics and its application in image copy-rotate/flip-move forgery detection," *International Journal of Digital Crime and Forensics (IJDCF)*, vol. 4, pp. 49-62, 2012.
- [46] A. Kaur and R. Sharma, "Optimization of copy-move forgery detection technique," *Computer Engineering and Applications Journal*, vol. 2, 2013.
- [47] M. Srivedi, C. Mala, S. Sandeep, and N. Meghanathan, "Copy-move image forgery detection in a parallel environment," *SIPM, FCST, ITCA, WSE, ACSIT, CS and IT*, vol. 6, pp. 19-29, 2012.
- [48] B. Liu, C.-M. Pun, and X.-C. Yuan, "Digital image forgery detection using JPEG features and local noise discrepancies," *The Scientific World Journal*, vol. 2014, 2014.
- [49] R. S. Oommen, M. Jayamohan, and S. Sruthy, "Using Fractal Dimension and Singular Values for Image Forgery Detection and Localization," *Procedia Technology*, vol. 24, pp. 1452-1459, 2016/01/01/ 2016.
- [50] G. Chierchia, G. Poggi, C. Sansone, and L. Verdoliva, "A Bayesian-MRF approach for PRNU-based image forgery detection," *IEEE Transactions on Information Forensics and Security*, vol. 9, pp. 554-567, 2014.
- [51] T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," *IEEE Transactions on Information Forensics and Security*, vol. 7, pp. 1003-1017, 2012.
- [52] S. Murali, G. B. Chittapur, and B. S. Anami, "Comparision and analysis of photo image forgery detection techniques," *arXiv preprint arXiv:1302.3119*, 2013.
- [53] A. Piva, "An overview on image forensics," *ISRN Signal Processing*, vol. 2013, 2013.
- [54] X. Pan and S. Lyu, "Region duplication detection using image feature matching," *IEEE Transactions on Information Forensics and Security*, vol. 5, pp. 857-867, 2010.
- [55] M.-S. Hwang and C.-C. Lee, "Research Issues and Challenges for Multiple Digital Signatures," *IJ Network Security*, vol. 1, pp. 1-7, 2005.
- [56] O. M. Al-Qershi and B. E. Khoo, "ROI-based tamper detection and recovery for medical images using reversible watermarking technique," in *Information Theory and Information Security (ICITIS), 2010 IEEE International Conference on*, 2010, pp. 151-155.
- [57] F. Y. Shih and Y. Yuan, "16 A Comparison Study on Copy-Cover Image Forgery Detection," *Multimedia Security: Watermarking, Steganography, and Forensics*, p. 297, 2017.
- [58] S. Rawat and B. Raman, "A chaotic system based fragile watermarking scheme for image tamper detection," *AEU-International Journal of Electronics and Communications*, vol. 65, pp. 840-847, 2011.
- [59] B. Shivakumar and L. Baboo, "Detection of region duplication forgery in digital images using SURF," *IJCSI International Journal of Computer Science Issues*, vol. 8, 2011.
- [60] J. Chao, X. Jiang, and T. Sun, "A novel video inter-frame forgery model detection scheme based on optical flow consistency," in *The International Workshop on Digital Forensics and Watermarking 2012*, 2013, pp. 267-281.
- [61] M. Mishra and F. Adhikary, "Digital image tamper detection techniques-a comprehensive study," *arXiv preprint arXiv:1306.6737*, 2013.
- [62] T. Qazi, K. Hayat, S. U. Khan, S. A. Madani, I. A. Khan, J. Kołodziej, et al., "Survey on blind image forgery detection," *IET Image Processing*, vol. 7, pp. 660-670, 2013.
- [63] E. Agnes, S. D. Mahalakshmi, and D. K. Vijayalakshmi, "A Forensic Method for Detecting Image Forgery Using Codebook," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, 2013.
- [64] T. K. Sarode and N. Vaswani, "Region Duplication Forgery Detection using Hybrid Wavelet Transforms," *International Journal of Computer Applications*, vol. 90, 2014.
- [65] R. Kaur and T. Sharma, "Image Forgery Detection using Speed up Robust Feature Transform, Wavelet Transform, Steerable Pyramid Transform and Local Binary Pattern," *International Journal of Modern Computer Science and Applications (IJMCSA)*, vol. 4, 2016.

- [66] N. K. Gill, R. Garg, and E. A. Doegar, "A review paper on digital image forgery detection techniques," in *Computing, Communication and Networking Technologies (ICCCNT), 2017 8th International Conference on*, 2017, pp. 1-7.
- [67] T. M. Mohammed, J. Bunk, L. Nataraj, J. H. Bappy, A. Flenner, B. Manjunath, *et al.*, "Boosting Image Forgery Detection using Resampling Detection and Copy-move analysis," *arXiv preprint arXiv:1802.03154*, 2018.
- [68] O. Mayer and M. C. Stamm, "Accurate and Efficient Image Forgery Detection Using Lateral Chromatic Aberration," *IEEE Transactions on Information Forensics and Security*, 2018.
- [69] X. Lin, J.-H. Li, S.-L. Wang, F. Cheng, and X.-S. Huang, "Recent Advances in Passive Digital Image Security Forensics: A Brief Review," *Engineering*, 2018.
- [70] Ankit Kumar Jaiswal and Rajeev Srivastava, "Image Splicing Detection using Deep Residual Network," 2nd International Conference on Advanced Computing and Software Engineering (ICACSE-2019).
- [71] Bin Xiao, Yang Wei, Xiuli Bi, Weisheng Li, and Jianfeng Ma, "Image splicing forgery detection combining coarse to refined convolutional neural network and adaptive clustering " *Elsevier Information Sciences*, 2019.
- [72] Thuong Le-Tien, Hanh Phan-Xuan, Thuy Nguyen-Chinh, and Thien Do-Tieu, " Image Forgery Detection: A Low Computational-Cost and Effective Data-Driven Model " *International Journal of Machine Learning and Computing*, Vol. 9, No. 2, April 2019.
- [73] Giulia Boato, Duc-Tien Dang-Nguyen, and Francesco G. B. Denatale, " Morphological Filter Detector for Image Forensics Applications" *IEEE Access* 2020.
- [74] Yaqi Liu, and Xianfeng Zhao, "Constrained Image Splicing Detection and Localization With Attention-Aware Encoder-Decoder and Atrous Convolution" *IEEE Access* 2020.
- [75] Yuan Rao, Jiangqun Ni, and Huimin Zhao, "Deep Learning Local Descriptor for Image Splicing Detection and Localization" *IEEE Access* 2020.

A Model for Operationalizing the Information Technology Strategy Based on Structuration View

Thami Batyashe, Tiko Iyamu*

Department of Information Technology, Cape Peninsula University of Technology, South Africa

ARTICLE INFO

Article history:

Received: 10 March, 2020

Accepted: 24 March, 2020

Online: 11 June, 2020

Keywords:

IT strategy implementation

IT strategy operationalization

Structuration theory

ABSTRACT

Many organisations adopt and implement information technology (IT) but fail to operationalise it. As a result, the process of implementation is continually repeated without achieving the goals and objectives, which are often to gain competitive advantage and sustainability. This study employs structuration theory as lens to examine and understand the factors that influence operationalisation of IT strategy in an organisation. The case study approach was employed, and semi-structured interviews technique was used to collect data. The hermeneutics approach was used in the analysis, which was guided by the duality of structure from the perspective of structuration theory. From the analysis six factors were found to primarily influence the operationalisation of IT strategy in an organisation. Based on the factors, a model was developed, which is intended to guide both IT and business managers in the operationalisation of IT strategy.

1. Introduction

The pervasiveness of Information Technology (IT) has compelled organisations in different business spheres to adopt it. Thus, IT divisions are necessary to enable and support innovations in organisations. However, the deployment and use of IT in organisation has never been straightforward or as easy as sometimes proclaimed. Author [1] argues that developing technology is generally viewed as a variable and erratic undertaking. It is often complex to both individuals and organisations at large, with complexities attributable to both human and technology factors. Furthermore, [2] emphasise that interrelated factors, technical, social, and organisational make implementation of information technologies extremely complex.

Based on this some of the complexities on one hand, essentialities on another hand, the use of IT solutions clearly require strategy in fulfilling business needs and requirements over a period of time. The researcher [3] assert that whether or not an organisation intends to strive for any competitive advantage, Information Systems (IS) or IT will still require a strategy to manage it, if only to circumvent being disadvantaged by the conduct of others. Accordingly, [4] affirms that IT enables organisations to implement strategies and to realise objectives.

Additionally, the rapid changes in business and technological environments compel many organisations to adopt strategies in

response to the ever-changing business needs and new opportunities. The objectives are often to increase their capability to escalate competitive and sustainable in line with the organisational vision and strategic intent. Thus, many organisations develop strategies. Some authors argue, though, that most strategic initiatives remain on paper, and are only as good as the paper they are written on [5]-[7].

The need to deliver heightened business value and streamlined processes through IT is greater than ever before. Thus, organisations put emphasis on IT strategy and operationalisation thereof to continually enable and support their processes and activities [8]. Moreover, operationalising IT strategy often assists an organisation to change in a more informed and systematic way, thereby managing challenges such as IT ineffectiveness, an IT approach that is vague or uncertain, business and IT plans that are not aligned, IT being reactive as opposed to proactive and inconsistency of IT practises with best practices. Scholar [9] states the greatest benefits of IT strategy seem to be realised when IT investment is linked with other aligned investments and strategies, and all new business processes seem to be important in realising the maximum benefit of IT.

Based on these and other factors, many organisations attempt to operationalise the IT strategies in order to realise their organisational goals and objectives. However, this has not been easy; instead, some organisations develop IT strategy year-in and year-out. Also, if only some human actors adopt, implement and operationalise the strategy, realising the goals and objectives may

*Tiko Iyamu, Email: connectvilla@yahoo.com

be hampered. Therefore, organisations constantly develop and implement IT strategies, unaware of the numerous challenges hindering the operationalisation of the IT strategy.

The remainder of paper is divided into five main sections. The first and second sections are a review of literature. This is followed by the research methodology that was employed in the study, and analysis of the data. The fifth section presents the results from the analysis, based on which the conclusions were drawn in the last section.

2. Information Technology Strategy and its implementation

Through the innovative use of IT, organisations are able to outperform their competitors [10]. Concurring, [11] are of the view that disruptive innovation is putting some organisations at the lead in a highly competitive environment. Innovations concerning IT have the potential to provide valuable opportunities for organisations [12]. Clearly, IT is not merely a support function. It has become embedded in the systems and processes of many organisations. With the rapid spread of IT and the increasing connectivity of the modern world, relying on an IT strategy is no longer a luxury for organisations; indeed, it has become necessary for survival.

The aim of the IT strategy is often to create a plan that manages investments on IT solutions. Accordingly, [13] assert IT strategy aims to create a medium to long-term plan for introducing information systems and to manage related IT investments. [14] state that IT strategy uses IS to support business strategy; it is the main plan of the IS function, and the collective view of the role of IT within the organisation. [15] affirm that an IT strategy concerns the use of IT to support business operation and strategy. However, [4] maintains that an IT strategy is a phrase that concerns a complex mix of concepts, ideas, visions, experiences, objectives, knowledge, recollections, views and opportunities that provide overall guidance for certain actions in the interest of specific outcomes within the computing environment.

Some studies indicate that while organisations develop comprehensive IT strategy plans, they are unable to implement them successfully, thereby leading to poor overall organisational performance [3, 16, 17]. It is much simpler to reflect on a good strategy than to implement it; thus, the interest in implementing strategies, in practice, has intensified, primarily because good strategies are not necessarily implemented successfully. Authors, [17, 18] articulate a different view, stating that inadequately implementing a strategy may not be bad in an environment where strategies themselves may often be flawed; incorrect implementation may be a valuable source of bottom-up consideration for better strategies. As a result, even after more than a decade of research in the disciplines of information technology strategy, implementation and operationalisation are not fully understood.

A critical challenge within IT strategic implementation is that little has been investigated in terms of how to successfully implement strategic change linked to the use of it [19]. Despite the interest and the vital role of implementing the strategy, most

strategy implementations fail. One challenge organisations experience is that of putting an implementation team in place to execute and operationalise the IT strategy [17].

Hence, [17] emphasises that as a result of the prominent deficiency, a conclusion can be drawn verifying a lack of expertise in implementing strategies in organisations. It is apparent that on one hand the implementation of IT strategy does not happen by default, and on the other hand, after the strategy is implemented, the operationalisation is normally left to happen by itself. A comprehensive, coherent IT strategy and implementation plan alone does not guarantee the success of IT. Authors [3] are of the view that a sustainable, strategic approach to support every aspect of IT is inclusive in the IT strategy. Thus, it is critical to operationalise this strategy in fulfilling the objectives.

Regardless of the type and level of strategy in an organisation in the end management is faced with putting strategy into practice which is described as the implementation of tactics so that the organisation moves in the desired strategic direction [20]. Implementation of the IT strategy enables and ensures the use of systems, rendering IT solutions capable of supporting organisational practices [19]. These authors [21] explain that failure to put the implemented IT strategy to good use manifests into strategy blindness. Hence, [19] define strategy blindness as an organisation's inability to achieve the strategic intent of implementation of available IT abilities. While much attention is paid to the challenge of implementing an IT strategy that aligns organisational strategy to IT investment, there is a dearth of information pertaining to putting IT strategy into practice successfully [19].

3. Structuration View

Structuration theory's (ST) main emphasis is to understand how social practices are structured across time and space. The theory of structuration is a general sociological theory regarded as connecting multiple levels from society down to the individual [22]. Academic [23] postulate that although ST only infrequently refers to IT, it has been extensively used in IS studies because it is regarded as particularly useful to describe unexpected results of IT implementation. Structuration theory plays a significant role in this study in comprehending the social, organisational and personal contexts within which an IT strategy is implemented and operationalised. The theory draws a vital connection to comprehend an IT strategy, which on the one hand is constrained or enabled by the societal context in which it operates, and on the other hand, is a means for sustaining or amending that context. As far back as 2003 [24] explain ST has a significant role to play in the advancement of understanding how technological systems support human interaction in societal, organisational and personal contexts. It has been argued that without social interpretation, technology can be viewed as 'meaningless' [25].

Therefore, in this study, ST serves as a lens to understanding the meaning of the actions, rules and resources associated with the operationalisation of an IT strategy. The interaction between

agency and structure signifies a mutually constituted duality. Explaining structuration [26] advocates that human agents (agency) continuously produce, reproduce and change social societies (structures). Similarly, an IT strategy operationalising recursively produces outcomes that mutually reproduce the social world, because the rules and resources available for formulation, development, implementation and dissemination are distinctive to every organisation. One of the main tenets of structuration is the dual relationship between agency and structure.

Duality of structure is described as the repeated relation between human and structures, whereas structures shape human actions, and in turn, form the structure [4]. Whereas, [27] refers to duality of structure as the relationship between agency and structure which poses one of the most prevalent and challenging issues in social theory, asserting that structure exists only in and through the actions of human agents [28]. These dynamics may adversely affect the thoroughness and validity of the processes, technologies and capabilities required to implement an IT strategy rendering it operational.

The role of structure can therefore be seen as both a constraining and enabling element for human action. Thus, [29] suggest that structures in organisations have these enabling and constraining aspects, enabling because they provide a valuable framework for social dealings, but also constraining as they afford little flexibility for how individuals conduct themselves and interact within the organisation's boundaries. While structuration theory assists in explaining communication practices within organisations and helps in clarifying how employees understand their organisational rules, structures can be useful as well as adverse for organisations and employees. Therefore, structuration theory expresses the power of agency and structure over time in a social system [30].

4. Research Approach

The study employed the qualitative method because views and opinions of participants were required in achieving the objectives. According to [31] qualitative method, assist researchers to get hold of the thoughts and beliefs of participants, which enables comprehending the meaning that people attribute to their experiences. Qualitative method undertakes to enhance the understanding of why things are the way they are in social world and why people act the way that they do [32]. The case study approach was employed primarily because real-life setting was the focus. The approach enables an in-depth exploration of a real-life phenomenon in its natural setting [33]. An organisation from the private sector was selected as a case. Private organisations are companies that are owned by private investors. At the time of the study, the organisation needs to operationalise its IT strategy, therefore making it interesting and appropriate to examine the factors influencing the operationalising of its IT strategy. Triumph Technologies, was selected for the study, to gain an empirical understanding of how IT strategy can be operationalised in a real-life setting.

Triumph Technologies (TT) is a multinational privately-owned organisation in the telecommunication industry. The organisation

is wholly owned by the employees. The organisation operates in over 170 countries and regions, including South Africa. The head office referred, to as 'headquarters' (HQ) is in Asia and the regional office is situated in South Africa. The organisation was selected as a case for study based on specific criteria, including: (i) a developed IT strategy in the organisation; (ii) a willingness to participant in the study; and (iii) previously established distinct foci for the organisation.

According to [34] the purpose of qualitative interviewing is to express and clarify individuals' real-world life as they live it, feel it, experience it and make sense of accomplishments. Semi-structured interview was used for the data collection. [this author [35] describe semi-structured interviews as starting with defined questions, however the interviewer has the autonomy to adapt the questions to a specific directions response in an effort to allow for more spontaneous and instinctive conversations between the interviewer and interviewee. Sixteen (16) participants were interviewed until a point of saturation was reached. Generally, researchers should carry on the interview participants until the field of interest is saturated, meaning until anything new is said by the all participants [35]. Through the hermeneutic approach form the perspective of the interpretivist approach the data was analysed using the lens of the structuration theory as a guide. The results from the case study give a deeper understanding to how IT strategy is operationalised in an organisation, making a case for generalisation. According to [36:1452], "the end product of qualitative analysis is a generalization, regardless of the language used to describe it".

5. Data analysis through duality of structure

The data collected from the case was analysed with a hermeneutic approach from the interpretive paradigm. The analysis was guided by structuration theory employing duality of structure as a lens to guide the analysis. A summary of the analysis is depicted in Table, below. Through duality of structure, the research examined how rules and resources enable and constrain agencies to operationalise the IT strategy, producing and reproducing events, processes and activities in the organisation.

Agencies at TT were divided into two categories: technical and non-technical. Technical agencies comprised proprietary technologies and IT systems; playing an integral part in developing and implementing systems and innovations to operationalise the IT strategy. At TT, the IT specialists and business users are the non-technical agents. Most of the non-technical agents, in particular, the IT specialists and IT management representatives such as the CIO, report into the international structures. Although the regional office has IT specialists, the headquarters IT support team based in Asia remotely supports the regional office in South Africa.

At the organisation, structure was classified under rules as IT policies, and resources were the IT and business people and the processes. These IT policies are the guidelines followed to operationalise the IT strategy. Resources include IT and business people and processes. The IT people are IT specialists and IT management teams that implement and operationalise the IT strategy. The business people are non-IT employees who are participating in operationalising the IT strategy. Processes are used by IT and business employees to perform business activities and actions.

Table 1: Overview of Analysis at Triumph Technologies

	Signification	Domination	Legitimation
Structure	At the organisation, some employees considered the IT strategy and its operationalisation as very critical and significant in achieving efficiency and effectiveness. This consideration was based on factors, such as people, technologies processes, and continuous learning.	Owing to the IT strategy impacting business processes, activities and events, IT staff put emphasises on operationalising it. The IT staff where the main role players driving the implementation and operationalise the IT strategy. Due to their involvement as main role players, some were aware of the IT strategy and had the skills to operationalise it. Because knowledge and skills were critical, some employees used it to dominate the environment.	The organisation had rules, policies, processes, frameworks and controls that guided and managed carrying out of the IT strategy activities and actions that authorised some employees' actions and behaviour, legitimating it.
	Interpretive scheme	Facility	Norm
Modality	At Triumph Technologies, based on different views and interpretations from various employees across the organisation, the IT strategy and operationalisation of it, on the one hand, enabled and on the other hand constrained business activities, events, processes and policies.	In an effort to operationalise the IT strategy different resources (people) such as people in the headquarters in Asia were employed in various roles as enablers and at the same time constrained some processes. An example is the remote support provided by the people in the headquarters. This constrained some business events, as they had to wait for assistance from someone who is based far and remote.	It was the organisational culture to that the people, processes and technologies involved and through which the IT strategy was operationalised were allocated and managed from a central point, the headquarters in Asia. This included skilled people, new processes and advanced technologies and training, roles, responsibilities.
	Communication	Power	Sanction
Interaction	Different means and ways were used to communicate, share knowledge and information about IT strategy and how it can be operationalised in the organisation. This included teleconference, video conference, meetings, workshops and electronic training. The CIOs were accountable and responsible in sharing information about IT strategy its operationalisation with the stakeholders.	The roles some employees occupy in the organisation delegate authority, in particular CIO and the heads of the different IT divisions to manage and control how to operationalise the IT strategy. This type of control, where power played a role, increases the interest and contribution from some of the employees, because they respect the authority bestowed on the person.	The work ethics and culture of some employees in the organisation was to adhere to instructions and go beyond the call of duty to operationalising the IT strategy. This was not only based on their employment agreement with the organisation, also on the work-centric and customer-centric attitude and believes.

The discussion that follows should be read with the Table to gain better understanding of the data analysis.

5.1. Signification/Interpretive scheme/Communication

In the organisation, the IT strategy was the roadmap defining what and how solutions should be deployed. This includes planning, IT systems development and management of telecommunication devices. In addition, through the IT strategy, synergy and consolidation of artefacts and systems were carried out. From this viewpoint, some employees considered the IT strategy significant in that it simplifies the numerous activities that were carried out within the environment. In operationalising the IT strategy, the consolidation approach reduced the numerous systems, some of which were redundant and others duplications. This ensures that the solutions selected and deployed were unified, enhancing consistency, standardisation, and reducing complexity, promoting efficiency and effectiveness, and advancing the organisation's competitiveness.

Another important aspect of the IT strategy was that its operationalisation enabled a seamless link between the branches of the organisation across the world, between the Asian and African continents. This enabled Triumph Technologies to achieve the organisational goals and objectives by reducing operational cost and increasing competitiveness. This was important to both the management of the organisation including some employees. However, many employees did not fully grasp the significance of consolidating and unifying technology

solutions in the organisation, as these employees thought it was fanciful, or nice to have. Others, however, understood the cost implications as well as the efficiency and effectiveness that such initiatives contribute to the environment. This diverse understanding was based on individual and group interpretation of the activities undertaken within the business units as enabled and supported by the IT unit through operationalisation of its strategy. The interpretation was influenced by communication.

In Triumph Technologies, electronic mail (email) and mobile electronics applications (apps) were the primary methods of communication. Video conferencing and teleconferencing were secondary methods of communication within the organisation, because both video conferencing and teleconferencing were often used for meetings and clarifications of subjects that had previously been communicated through the email. Spoken language was a challenge during communication, whether management-to-employee or employee-to-employee. This critically influenced the interpretation of contents during operationalisation of the IT strategy within the organisation. Occasionally language had to be translated for other employees or stakeholders. In the process of language translation, some of the meanings or contexts are misconstrued.

Operationalisation of the IT strategy was influenced, enabled and constrained within Triumph Technologies by the significance associated to it. Moreover, interaction was of mixed feelings because some employees were privileged and others were not in terms of sharing organisational information. Thus, meanings

which individuals and groups make of the technology solutions and artefacts affected the operations. In addition, communication was not always straightforward, which often influenced employee interpretations and the value they associated with the IT strategy.

5.2. Domination/Facility/Power

At Triumph Technologies, there were imbalances from various perspectives, such as allocation of tasks and information sharing. The imbalances enabled and sometimes constrained events, processes and activities, consciously or unconsciously. These were actions that reproduced themselves during the operationalisation of the IT strategy in the organisation. During operationalisation of the IT strategy, various facilities were employed, including processes and spoken languages. The facilities were employed from two viewpoints, personal and organisational. At the organisational front, processes were followed in the operationalisation of the IT strategy towards achieving the goals and objectives of the organisation. From a personal perspective, some employees spoke in the language that friends among colleagues understood, excluding others from participating in discussions.

A South African language (Sesotho) and an Asian language (Hakka) were commonly spoken divisively to exclude colleagues from discussions. In addition, some employees having close or have personal relationships with their managers preferred to speak in the language only both understand instead of the generally accepted language of the environment, which was English at the time of this study. The use of the English language was mainly because they, the promoters, did not have a choice but to employ an inclusive approach for tasks to be carried out. The reliance on a particular spoken language to exclude certain colleagues was at some point a hindrance to the operationalisation of the strategy. This was so because many of the interested employees, or those with the necessary skill-sets, found it difficult to participate in discussions, affecting their overall execution of tasks. This worsened as the exclusion approach was also practiced in formal meetings.

Through the preferred spoken language, employees unconsciously created networks within the organisation, meaning that networks were formed along language lines. Consciously or unconsciously, the networks regulated activities of the IT strategy during operationalisation. This was primarily because some of the employees were more loyal to their networks than the organisational objectives. Another reason for loyalty was attributed to the fact that some employees admitted to receiving more information from their networks than from the formal hierarchical structure within the organisation. In the operationalisation of the IT strategy, there were also factors of power at personal levels and from organisational hierarchical levels (positions). This factor caused imbalance in the organisation during the operationalisation of the IT strategy. At a personal level, the source of power came from knowledge, which some employees acquired through continuous learning and the privilege to information.

Although there was power associated with knowledge, skills and understanding of the IT strategy and its operationalisation in the organisation, there was also power that bestowed on the positions. The staff in the headquarters (HQ) had power to approve or reject activities relating to the operationalisation of the IT strategy in the organisation. The HQ team includes the Chief Information Officer (CIO) and the IT specialists in Asia. Business initiatives were discussed with the HQ team who have the ultimate decision-making power. It is clear that during the operationalisation of the IT strategy, there were imbalances, which means that some employees were dominant over their colleagues. This dominance was based on levels of access to facilities that were sources of power. Power was enacted by the facilities, enabling and simultaneously constraining activities in the operations of the IT strategy

5.3. Legitimation/Norms/Sanction

At Triumph Technologies, operationalisation of the IT strategy entails various activities through different processes, rules and regulations to fulfil organisational requirements, goals and objectives. These actions were assessed and deemed eligible for use within the organisation. Thereafter, actions were executed by humans using facilities such as technology solutions (devices), spoken language and face-to-face meetings to operationalise the IT strategy.

In operationalising the IT strategy, micro and macro approaches were employed at middle management and lower management, respectively. The different management approaches, micro and macro, were employed because of the hierarchical structured nature of the environment. The macro focuses on strategic intent, while the micro was operational. Thus, the approaches were purposely followed to enforce the different types of instructions, rules and regulations through the hierarchy, for different events and activities during operationalisation of the IT strategy.

At both micro and macro levels, long working hours (beyond the prescribed eight working hours) and late-night meetings were held. Although some employees were initially not accustomed to this culture, with time, they became acclimated to this as it gradually became the norm as operationalisation of the IT strategy continued within the organisation. Few other actions, such the use of certain spoken languages for exclusivity, were also norm. This occurred even though they were consciously or unconsciously used to enable or constraint in one way or the other, the activities involved in the operationalising of the IT strategy.

Even though the facilities were approved for organisational purposes, some of the actions that manifested were not entirely geared towards achieving the goals and objectives of the IT strategy. For example, Hakka was spoken for exclusivity purposes. Despite its negative connotation, it became a culture, a way of conducting the business of operationalising the IT strategy, practiced over a period of time within the organisation. Some

employees accepted this practice, not because they liked or agreed with it, but because they felt that had no choice.

This was because the senior organisational management sanctioned the practice. Management and even some employees sanctioned some of the actions, such as the long hours of meetings, meetings at late hours, and the use of the Hakka language for exclusivity. This was not because they wanted to, but because it facilitated productivity in the operationalisation of the IT strategy in the organisation. These actions were practiced, and eventually became the norm, mainly because they were first sanctioned by the management at the HQ, the decision-making authority in Asia.

At Triumph Technologies, as an initiative to educate aspiring IT specialists to address the different spoken language imbalances, learning materials were presented. The intention of this initiative was to make operationalisation of IT strategy easier and more efficient, creating a culture of learning and inclusion. The learning culture was sanctioned by everyone who wanted to acquire skills and knowledge and participate in operationalising the IT strategy. The culture of learning encouraged employee awareness of the IT strategy, learning and understanding why and how to operationalise it.

In operationalising the IT strategy, many of the human actions as well as the technological solutions were reproductive. Even though the actions and technological solutions were eligible (legitimate) within the frame of the organisation, they were not always to promote organisational interest. In addition, some of the actions and activities that were considered as the norm were not generally agreed upon by the many of the employees. For example, only a few of the employees agreed to the abnormal working hours to protect their jobs. The management sanctioned activities and actions intended for the benefit of the organisation, but with little regard for the consequences to the employee.

6. Discussions and findings

Six factors were identified from the analysis that enabled and simultaneously constrained the operationalisation of the IT strategy at Triumph Technologies (TT): hierarchical consciousness; technology solutions; network of people; training and skill-set; exclusivity vs inclusivity; and language differentiation (Figure). The figure needs to be perused with the discussion in mind to ascertain exactly how the factors shape IT strategy and its operationalisation.

The model depicted in Figure 1, present factors that are interrelated. Thus, the factors influence and are being influenced by others. In other words, these factors enable and constrain each other during the IT strategy operationalization process at Triumph Technologies.

6.1. Hierarchical consciousness

Hierarchical levels are necessary in an environment to steer information appropriately [37], such as IT strategy solution. Author [38] suggests that processing information or tasks that involve many behavioural options require consciousness. This is

to avoid potential disintegration of solutions such as the operationalisation of the IT strategy within an environment. Furthermore, [39] explain that consciousness can play a role in enabling tasks within and environment. To the contrary [37] argue that some users often lose consciousness of their tasks as they navigate within hierarchy. But successful integration of artefacts or solutions requires clear consciousness of the people that are involved [39].

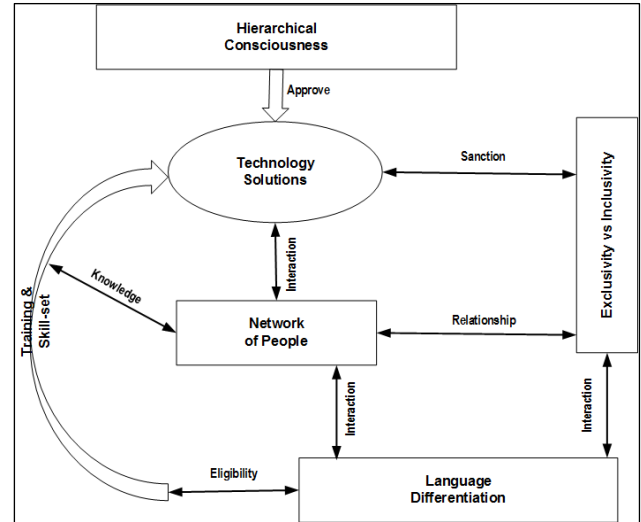


Figure 1: Model for operationalizing an IT strategy

At Triumph Technologies, adherence to organisational structure was considered an important influencing factor in operationalising the IT strategy in the organisation. During the operationalising of IT solutions, approval was sought from senior management and structures in Asia, a practice accepted by both IT specialists and business users, irrespective of whether or not they agreed with the strategy and its processes. This enabled smoothness of the processes and various activities as well as employee inclusiveness in the operationalisation of the IT strategy. Additionally, the approval of the strategy ensures that the solutions operationalised are in alignment with the organisation's universal strategy.

As organisational structure allows strategy and its process to circumvent duplication of IT solutions, promoters of the IT strategy verified and validated each innovation with senior management. The verification and validation processes occurred by way of interaction among stakeholders involved in operationalising the strategy. Without approval through the organisational structures, activities and events involving operationalisation would potentially be delayed, with some activities even facing termination or rejection. Thus, IT specialists and business users were intentionally conscious, aware of the significance of the organisational structures in carrying out their responsibilities related to the operationalisation of the IT strategy.

6.2. Technology solutions

Technological solutions refer to information systems and technological tools or artefacts used to enable and support activities [40]. IT strategy defines the solutions and arranges them in priority perspective for more efficient organisational use. This

evolves over time, gradually addressing the changing needs of an organisation [41]. Technological solutions do not operate in and of themselves, but require human expertise [42].

Technology solutions were defined by the IT strategy, including standard deployment, management and use of the solutions for best organisational purposes. The IT systems, IT infrastructures and telecommunication devices were the main aspects of the IT strategy, with IT systems involving mobile applications, applications, electronic flows (e-flows) and tools. At Triumph Technologies, IT infrastructures consist of servers, laptops, desktops and notebooks used by the employees to manage processes and activities. Telecommunication devices were employed for teleconference and videoconference meetings with the headquarters and other branches globally.

The IT strategy was operationalised to enable deployment of the technology solutions, with the intent of improving organisational efficiencies. During operationalisation of the technology solutions, processes and activities were managed attentively to ensure appropriateness and suitability in accordance with organisational purposes. This was because technology solutions both influence and are influenced by other factors such as hierarchical consciousness, skill-sets and networks of people (Figure). The process of operationalisation required legitimisation that happens through hierarchical consciousness of management. Also required were appropriate skill-sets and the deliberate involvement of various personnel. Above all, interaction and relationships among stakeholders were of critical importance.

6.3. Network of People

Network of people refers to conscious or unconscious groupings of employees within an organisation. According to [43], people engage in networks for various purposes, both personal and organisational. These authors [44] explain how the interaction that occurs within networks of people influence technology deployment within an organisation. The success or failure of operationalisation of IT strategy can be influenced by the interactions and actions within networks of people. Scholar [45] argue that in recent years, traditional hierarchical approaches are struggling against challenges of an emerging relational set-up in which decisions cannot be imposed but must emerge from the interactions among actors.

Alignment of various agencies played a significant role in operationalising the technology solutions as the agencies formed a homogenous network of people, consciously or unconsciously, intended to achieve business objectives of the organisation. The networks, formed based on spoken languages, skills and competencies, were enabling as well as constraining in the operations of personnel. From the enabling front through the networks, deliverables were fostered, primarily because employees were either acquaintances or friends, and based on strength of relationships, they offered various levels of support to each other. From the constraining perspective, collaboration between various networks were challenging because of factors such as language differences, which, while including some, often excluded others.

In the operationalisation of IT strategy, it is important that the different networks of people involved have not only the skill-set and understanding of various processes but work collaboratively with one another to achieve organisational business objectives. Thus, skill-sets and collaboration of various people were significant in operationalising the IT strategy. Skill-set deficiencies and lack of training surrounding various processes involved in operationalisation meant inefficiency and ineffectiveness of the IT strategy.

6.4. Training and skill-set

The roles of employees are not as easily ascertained as believed; otherwise, the operationalisation of technology solutions will be even more complex due to human actions [44]. The different standards and levels of employee actions, based on knowledge and skill, determine the success of activities within an environment [45] so it is critical that organisations involve employees with the right of skill-sets as that is critical for competitive advantage [46]. Thus, it is essential to train and develop employees appropriately about operationalising the IT strategy in the organisation.

Training and development meant that employees in the organisation were equipped with vital knowledge and skills for understanding the processes and activities involved in operationalising technological solutions as defined by the IT strategy. Training and development were often conducted through different methods and mediums such as electronic learning (e-learning), which gave employees the convenience of accessing training material and course participation in the operations of the IT strategy in the organisation. The training enabled some of the employees to carry out their responsibilities from anywhere, and at any time, through their mobile devices.

Through training and development, knowledge about the technology solutions was acquired. Therefore, networks of people had the capabilities and knowledge to operationalise the IT strategy. The importance of training and development during operationalisation was for the network of people to generate a common understanding about the processes and activities when interacting during operationalisation; however, employees were also eligible to interact in different languages, which created a language barrier in the organisation.

6.5. Language Differentiation

Understanding of activities and tasks is mediated by language of instruction and engagement through facilitating communication among team members [47]. Thus, devising effective strategy is necessary to bridge the language barrier and manage significantly negative activity [48]. Even though training programs are carried out, they do not always consider language barriers, an oversight that can engender additional complexities in an environment [49]. This needs more attention in that through language the communicating of thoughts, ideals and knowledge is manifest, so language is clearly an influence in terms of how IT strategy is operationalised.

The spoken language was used, whether consciously or unconsciously, to enable and occasionally constrain operationalisation of the IT strategy in the organisation. On the one

hand, when employees of the same network communicate using a preferred language, such as Hakka, in sharing knowledge and ideas to ease understanding about technology solutions and processes, smooth operationalisation is heightened; but on the other hand, when employees who are unfamiliar with the network language become difficult, communication challenges escalate. This is a constraining barrier during the operationalisation of the IT strategy. This situation was reproduced time and again in operationalising the IT strategy in the organisation.

The language differentiation influenced and was influenced by networks of people, by the exclusivity or inclusivity of employees, and by the use of technology solutions. This was both enabling and constraining in operationalising the IT strategy, as explained above. The most important thing is that language differentiation has been identified as an essential influencing factor when operationalising IT strategy in an environment as it creates division among employees in operationalising the IT strategy in the organisation.

6.6. Exclusivity and Inclusivity

Inclusiveness aims to enrol as many as possible participants while exclusiveness is about access by only a privilege few. According to [50] inclusivity is a process that genuinely and legitimately allows broader participation in an activity. However, deceptive actors tend to use more cognition, inclusivity and exclusivity in words when interacting with groups within environment [51]. Postulated by [52] an understanding of information system continuance for information-oriented mobile applications requires a dramatic shift from exclusivity to inclusivity to influence operationalisation of the IT strategy.

In operationalising the IT strategy in an organisation, exclusivity and inclusivity of employees were both enabling and at the same time constraining. Exclusivity minimises too many opinions and options, mineralising complications inherent in decision-making. However, the same factor of exclusivity deprived certain employees from participating in processes and activities tasked for the execution of IT strategy. The concept of inclusiveness was beneficial to both the business and IT units of the organisation, from an alignment viewpoint, as alignment between business and IT units was instrumental in operationalising the IT strategy in the organisation. Despite the positive aspect of inclusivity, it was also constraining. For example, too many people could not be involved in certain decisions, especially those requiring technical expertise.

In the environment and during operationalisation of the IT strategy, exclusivity or inclusivity of a group of employees was sometimes consciously and sometimes unconsciously created. This happened at various levels, from senior management to technical expertise. Some employees were privileged, granted exclusive access to information pertaining to IT strategy operationalisation. Both exclusivity and inclusivity of employees influenced and was influenced by the relationship and interactions during operationalisation of IT strategy in the organisation, impacting how some employees were nominated for skills development, but not others, how processes were defined, and how

tasks were assigned to certain distinct individuals in the operationalising the IT strategy in the organisation.

7. Conclusion

This paper provides a clear distinction between IT strategy implementation and operationalisation. This is a confusion that has contributed to the misunderstanding, and negativity which the IT strategy has received for many years. The study reveals, and makes it possible to gain better understanding of the factors that influence operationalisation of an IT strategy in an organisation, which were not empirically known. The factors are critical as they assist in achieving business goals and objectives. Thus, the research intended to benefit academics and professionals alike that focus on operationalising IT strategies in organisations. The academics domain gain from this research through its addition to existing literature in the subject areas of information technology strategy, implementation and operationalisation. Professionals in the business sphere, the benefits come from gaining better understanding of the influential factors involve in operationalising IT strategies in organisations.

References

- [1] C. Christensen, *The innovator's dilemma: when new technologies cause great firms to fail*. Harvard Business Review Press, 2013.
- [2] K. Cresswell, A. Sheikh, "Organizational issues in the implementation and adoption of health information technology innovations: an interpretative review" *International journal of medical informatics*, 82(5), e73-e86, 2013. <https://doi.org/10.1016/j.ijmedinf.2012.10.007>
- [3] J. Peppard, J. Ward, *The Strategic Management of Information Systems: Building a Digital Strategy*, 4th ed. Sussex, UK: John Wiley & Sons, 2016.
- [4] T. Iyamu, "Politicking the Information Technology Strategy in Organisations. Strategic Information Technology Governance and Organizational Politics in Modern Business, IGI Global.
- [5] J. Richardson, "The business model: an integrative framework for strategy execution" *Strategic Change*, 17(5-6),133-144, 2008. <https://doi.org/10.1002/jsc.821>
- [6] M. Birshan, M. J. Kar, "Becoming more strategic: Three tips for any executive" *McKinsey Quarterly*, 2012.
- [7] P. Nuntamanop, I. Kauranen, B. Igel, "A new model of strategic thinking competency" *Journal of Strategy and Management*, 6(3), 242-264, 2013, <https://doi.org/10.1108/JSMA-10-2012-0052>
- [8] T. Iyamu. "A Model for Operationalising Influencing Factors in IT Strategy Deployment" *International Journal of Social and Organisational Dynamics in IT*, 1(4), 159-170, 2012. DOI: 10.4018/ijcsodit.2011100104
- [9] Z. Wang, "Strategic Fit Issues in Information System Research: Concept, Operationalization and Future Directions" *International Journal of Hybrid Information Technology*, 7(1), 13-24, 2014, <http://dx.doi.org/10.14257/ijhit.2014.7.1.02>
- [10] J. Kandampully, A. Bilgihan, T. C. Zhang, "Developing a people-technology hybrids model to unleash innovation and creativity: The new hospitality frontier" *Journal of Hospitality and Tourism Management*, 29(2016), 154-164, 2016. <https://doi.org/10.1016/j.jhtm.2016.07.003>
- [11] J. Kroh, H. Luetjen, D. Globocnik, C. Schultz, "Use and efficacy of information technology in innovation processes: The specific role of servitization" *Journal of Product Innovation Management*, 35(5), 720-741, 2018. <https://doi.org/10.1111/jpim.12445>
- [12] R. J. Kauffman, J. Liu, D. Ma, "Technology investment decision-making under uncertainty" *Information Technology and Management*, 16(2), 153-172, 2015. <https://doi.org/10.1007/s10799-014-0212-2>
- [13] C. Yeh, G. G. Lee, J. Pai, "How information system capability affects e-business information technology strategy implementation: An empirical study in Taiwan" *Business Process Management Journal*, 18(2), 197-218, 2012. <https://doi.org/10.1108/14637151211225171>
- [14] D. Q. Chen, M. Mocker, D.S. Preston, A. Teubner, "Information systems strategy: reconceptualization, measurement, and implications" *MIS quarterly*, 34(2), 233-259, 2010.

- [15] T. Cui, Y. Hua, H. Toe, J. Li, "Information technology and open innovation: A strategic alignment perspective" *Information & Management*, 52(2015), 348-358, 2015. <https://doi.org/10.1016/j.im.2014.12.005>
- [16] R. Pérez Estébanez, E. Urquía Grande, C. Muñoz Colomina, "Information technology implementation: evidence in Spanish SMEs" *International Journal of Accounting & Information Management*, 18(1), 39-57, 2010. <https://doi.org/10.1108/18347641011023270>
- [17] J. Bartenschlager, *Implementing IT strategy—Laying a foundation Tagungsband der INFORMATIK*, 2011.
- [18] E. Lee, P. Puranam, "The implementation imperative: Why one should implement even imperfect strategies perfectly" *Strategic Management Journal*, 37(8), 1529-1546, 2015. <https://doi.org/10.1002/smj.2414>
- [19] V. Arvidsson, J. Holmström, K. Lyytinen, "Information systems use as strategy practice: A multi-dimensional view of strategic information system implementation and use" *The Journal of Strategic Information Systems*, 23(1), 45-61, 2014. <https://doi.org/10.1016/j.jsis.2014.01.004>
- [20] A. A. Palladan, K.B. Abdulkadir, Y. W. Chong, "The effect of strategic leadership, organization innovativeness, information technology capability on effective strategy implementation: A study of tertiary institutions in Nigeria" *Journal of Business and Management*, 18(9), 109-115, 2016. DOI: 10.9790/487X-180901109115
- [21] J. Peppard, R. D. Galliers, A. Thorogood, A. "Information systems strategy as practice: Micro strategy and strategizing for IS" *Journal of Strategic Information Systems*, 23(1):1-10, 2014, <https://doi.org/10.1016/j.jsis.2014.01.002>
- [22] A. F. V. Veenstra, Melin, U. K. Axelsson, "Theoretical and practical implications from the use of structuration theory in public sector information systems research" In 2014 the European Conference on Information Systems (ECIS) 2014, Tel Aviv, Israel, 2014.
- [23] G. Walsham, Cross-cultural software production and use: a structural analysis. *MIS Quarterly*, 359-380, 2002.
- [24] J. Brodie, J. Evans, L. Brooks, M. Perry, "Supporting human interaction through digital technology: theory and practice" in 2003 1st international symposium on Information and communication technologies (pp. 403-408). Trinity College Dublin, 2003.
- [25] J. Rose, R. Hackney, "Towards a structural theory of information systems: A substantive case analysis. In 2003 IEEE 36th Annual Hawaii International Conference on System Sciences, Hawaii, 2003.
- [26] B. E. Wiggins, G. B. Bowers, "Memes as genre: A structural analysis of the memescape" *New media & society*, 17(11), 1886-1906, 2015. <https://doi.org/10.1177/1461444814535194>
- [27] D. M. Mezzanotte, "Planning Enterprise Architecture: Creating organizational knowledge using the Theory of Structuration to build Information Technology" In 2016 IEEE Software Engineering Research, Management and Applications (SERA), 14th International Conference.
- [28] D. M. Mezzanotte, J. Dehlinger, "Building Information Technology Based on a Human Behavior-Oriented Approach to Enterprise Architecture" In 2013. International Conference on Software Engineering Research and Practice (SERP), The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).
- [29] C. J. Liberman, M. L. Doerfel, "Structuring Organizational Communication: Employees' Role and Network Position as Predictive of Institutional Talk About the Adoption of Technology" In 2012 IEEE System Science (HICSS), 45th Hawaii International Conference, 2012.
- [30] R. Whittington, "Giddens, structuration theory and Strategy as Practice" In Damon Golsorkhi, Linda Rouleau, David Seidl, & Eero Vaara (Eds.), *Cambridge handbook of strategy as practice* (pp. 109–126). Cambridge: Cambridge University Press, 2010.
- [31] J. Sutton, J. Z. Austin, 2015. *Qualitative research: Data collection, analysis, and management*. The Canadian journal of hospital pharmacy, 68(3), 226-231, 2015. DOI: 10.4212/cjhp.v68i3.1456
- [32] F. Tuli, "The basis of distinction between qualitative and quantitative research in social science: reflection on ontological, epistemological and methodological perspectives" *Ethiopian Journal of Education and Sciences*, 6(1), 97-108, 2011. DOI: 10.4314/ejesc.v6i1.65384
- [33] R. K. Yin, 2017. *Case study research and applications: Design and methods*. 6th ed. Thousand Oakes, CA: Sage Publications, 2017.
- [34] U. Schultze, M. Avital, "Designing interviews to generate rich data for information systems research. *Information and Organization*" 21(1), 1-16, 2011.
- [35] Z. Austin, J. Sutton, "Qualitative research: Getting started" *The Canadian journal of hospital pharmacy*, 67(6), 436-440, 2014. doi: 10.4212/cjhp.v67i6.1406
- [36] D. F. Pilot, C. T. Beck, "Generalization in Quantitative and Qualitative Research: Myths and Strategies" *International journal of nursing studies*, 47(11), 1451-1458, 2010. <https://doi.org/10.1016/j.ijnurstu.2010.06.004>
- [37] J. Rekimoto, M. Green, "The information cube: Proceedings of the Third Annual Workshop" in 1993 on Information Technologies & Systems (WITS'93) (pp. 125-132). 1993.
- [38] A. Keller, "The evolutionary function of conscious information processing is revealed by its task-dependency in the olfactory system" *Frontiers in Psychology*, 5(62),1-7, 2014. <https://doi.org/10.3389/fpsyg.2014.00062>
- [39] L. Mudrik, N. Faivre, C. Koch. "Information integration without awareness" *Trends in Cognitive Sciences*, 18(9), 488-496, 2014. <https://doi.org/10.1016/j.tics.2014.04.009>
- [40] R. Adams, R. S. Jeanrenaud, J. Bessant, D. Denyer, P. Overy. "Sustainability-oriented innovation: A systematic review" *International Journal of Management Reviews*, 18(2):180-205, 2016. <https://doi.org/10.1111/ijmr.12068>
- [41] A. Garcia-Holgado, F. J. Garcia-Peñalvo, "Preliminary validation of the metamodel for developing learning ecosystems" in 2017 ACM 5th International Conference on Technological Ecosystems for Enhancing Multiculturality, 2017. <https://doi.org/10.1145/3144826.3145439>
- [42] R. Welford, *Corporate environmental management 3: Towards sustainable development*. London: Routledge, 2016.
- [43] K. Y. Lin, H. P. Lu, "Why people use social networking sites: An empirical study integrating network externalities and motivation theory" 2011. <https://doi.org/10.1016/j.chb.2010.12.009>
- [44] S. S. Mkhomazi, T. Iyamu, T. "Human Interaction in the Regulatory of Telecommunications Infrastructure Deployment" in 2014 International Working Conference on Transfer and Diffusion of IT (pp. 324-333). Springer, Berlin, Heidelberg, 2014.
- [45] M. Del Giudice, F. Caputo, F. Evangelista. "How are decision systems changing? The contribution of social media to the management of decisional liquefaction" *Journal of Decision systems*, 25(3), 214-226, 2016. <https://doi.org/10.1080/12460125.2016.1187546>
- [46] R. A. Noe, J. R. Hollenbeck, B. Gerhart, P. M. Wright, P. M. Human resource management: Gaining a competitive advantage. New York, NY: McGraw-Hill Education, 2017.
- [47] M. Barak, A. Watted, H Haick. 2016. "Motivation to learn in massive open online courses: Examining aspects of language and social engagement" *Computers & Education*, 94, 49-60, 2016. <https://doi.org/10.1016/j.compedu.2015.11.010>
- [48] C. L. Timmins. "The impact of language barriers on the health care of Latinos in the United States: a review of the literature and guidelines for practice" *The Journal of Midwifery & Women's Health*, 47(2),80-96, 2002. [https://doi.org/10.1016/S1526-9523\(02\)00218-0](https://doi.org/10.1016/S1526-9523(02)00218-0)
- [49] De Jesus-Rivas, M. H. A. Conlon, C. Burns, "The impact of language and culture diversity in occupational safety" *Workplace health & safety*, 64(1), 24-27, 2016. DOI: 10.1177/2165079915607872
- [50] D. Loveridge, O. Saritas, O. Reducing the democratic deficit in institutional foresight programmes: a case for critical systems thinking in nanotechnology. *Technological Forecasting and Social Change*, 76(9), 1208-1221, 2009. <https://doi.org/10.1016/j.techfore.2009.07.013>
- [51] S. M. H.O, J. T. Hancock, C. Booth, M. Burmester, X. Liu, S.S. Timmarajus, "Demystifying insider threat: Language-action cues in group dynamics" in 2016 IEEE 49th Hawaii International Conference on System Sciences (HICSS) (pp. 2729-2738), 2016.
- [52] C. T. Chisita, F. Chinyemba, "Utilising ICTs for Resource Sharing Initiatives in Academic Institutions in Zimbabwe: Towards a New Trajectory" In *Managing Knowledge and Scholarly Assets in Academic Libraries* (IGI Global, 2017).

Applications of Causal Modeling in Cybersecurity: An Exploratory Approach

Suchitra Abel*, Yenchih Tang, Jake Singh, Ethan Paek

Department of Computer Science and Engineering, Santa Clara University, 95053, United States

ARTICLE INFO

Article history:

Received: 16 January, 2020

Accepted: 15 May, 2020

Online: 11 June, 2020

Keywords:

Causal Modeling

Cybersecurity

Data Breach

ABSTRACT

Our research investigates the use of causal modeling and its application towards mapping out cybersecurity threat patterns. We test the strength of various methods of data breaches over its impact on the breach's discovery time as well as the number of records lost. Utilizing a Causal Modeling framework, we simulate the isolation of confounding variables while testing the robustness of varying estimators. The motivation is to shed a unique insight provided by the usage of Causal Modeling in cybersecurity.

1. Introduction

The purpose of this paper is to demonstrate the application of Causal Modeling in the domain of Cybersecurity. We engage in the scientific inquiry into the underlying causes of data breaches. Using methods of causal analysis that link concepts to observations, and a rationale connecting concepts to practice. The notion of causality, as used in Computer Science, provides principles that guide the problem specification, elaboration of the procedures, and interpretation of datasets. We employ causal modeling for the purpose of providing a computable measurement of a certain group of data breaches.

We tackle a variety of data breach problems that affect our industries and have an impact on the overall economy. In our work, we demonstrate how the usage of Causal Modeling can help us locate such data breach problems. Statistical analysis is enough for identifying associative relationships. While this is useful for general analysis, Causal Modeling provides a different structure with interventions included in it. Interventions tell us what would have happened if events other than the ones we are currently observing had happened. Such interventions allow us to avoid unnecessary steps and come directly to the point. It can also provide justification as to why and how the desired step or conclusion is arrived and provide defense for potential future cases. In order to intervene, we needed to estimate the effect of changing an input from its current value, for which no data exists. Such questions, involving estimating a counterfactual, are common in decision-making scenarios.

Statistical Prediction is the estimation of an outcome based on the observed association between a set of independent variables and a set of dependent variables. Its main application is forecasting.

Causality is the identification of the mechanisms and processes through which a certain outcome is produced. It can be used in predicting future events that are similarly connected via mechanisms and processes. Causal relations are not features that can be directly read off from the data, but have to be inferred. The field of causal discovery is concerned with this inference and the assumptions that support it.

Our research focuses on two aspects of Causal Modeling: Causal Discovery and Causal Inference. Causal Discovery algorithms try to derive causal relations from observational data. Given a set of data, a causal discovery algorithm returns a set of statements regarding the causal interactions between the measured variables.

Causal Inference is the process of drawing a conclusion about a causal connection based on the conditions of the occurrence of an effect. The main difference between causal inference and statistical inference of association is that the former analyzes the response of the effect variable when the cause is changed. The process shows causal direction, which is rarely found by statistical correlation alone. For example, a question that causal reasoning can answer is: Is there a causal link between the distribution across values of a certain variable X and values of another variable Y?

Causal inference process solves causal problems systematically, by methods such as counterfactual analysis,

*Suchitra Abel, Santa Clara University Santa Clara, CA 95053, USA.

sabel@scu.edu

www.astesj.com

<https://dx.doi.org/10.25046/aj050349>

graphical models, and the association between counterfactual and graphical methods.

Causal modeling resolves questions about possible causes by providing explanation of phenomena as the result of previous events. One can generate a plausible explanation for gaps within cybersecurity infrastructure. The usage of Causal Modeling can help us locate a set of data breach problems and help provide a solution for such problems.

The objective of this research is to evaluate the risks of data breach of cybersecurity incidents with the overall aim to identify patterns of importance amongst the dataset, accomplished by noting causes and effects in the modeling process. This is achieved by studying the characteristics of the VERIS Community Database (VCDB) of cybersecurity incidents. VCDB is a widely used open-source dataset containing a breadth of information regarding data breaches.

2. Background

We offer a scientific method based on the notion of causation. Following are the motivations behind the use of Causal Modeling for Cybersecurity:

One can draw from past experiences, and try to build a probability distribution [1]. Standard probability theory has been productive in these problems and similar ones, when the past experiences are readily available for analysis. But there are instances where it fails to provide adequate concepts and mathematical methods, particularly when the past experiences are either not available, or are not relevant.

A context like breach of data can interact with the phenomena of interest in ways that standard probability theory does not productively capture; that is, in ways that standard probability theory does not provide insights and methods for useful modeling and fails to capture key concepts. Some of these key concepts are the necessary and sufficient conditions that produce the essential model of the cause-effect relationships involved.

A necessary condition is one that is required if a certain effect is to follow. A sufficient condition, on the other hand, is enough for certain effects to follow.

Some of the usage of the necessary and sufficient conditions are as follows: we have to look for causes that are common in the cases where the effect also occurs. Thus, some event is not a necessary condition if it happens without the effect occurring [2,3].

We can explore causal modeling on observational data. In general, to determine whether or not an uncertain variable x_k (the supposed effect) is responsive or unresponsive to decision d we have to answer the query "Would the outcome of x_k have been the same had we chosen a different alternative for d ?" Questions of this form are counterfactual queries [4,5].

We define the Counterfactual World as follows. there are some uncertain variables, X (of which x_k is an instance), such as data leakage (including some uncertainty as to why, and are we sure about the leakage?) in the scenario; there is also the set of potential causes C [6]. Possible candidates for the causes in C are:

- malware in the system

- hacking
- human error

Let U be the total set of possible effects pertaining to some scenario S . These are possibilities that should be determined correctly. There are variables $X \subseteq U$, which are uncertain variables. We also have a set of decisions D (for example, the decision that the data leakage is, indeed, there, and that it is there because of the bugs). Given these notions, the concept of counterfactual world can be defined. A counterfactual world of X and D is any instance of such world retained by $X \cup D$, after the decision maker selects a particular instance of D [7].

Definitions of unresponsiveness and responsiveness are to be understood next. Suppose that we have some uncertain variables, which form a set X . Also, suppose that we have a set of decisions D . There can be counterfactual worlds D that can form union with the set X . D is the set of scenarios where there is a list of counterfactual decisions (and the outcomes associated with the decisions), which may never take place in the real world as we encounter it. X is unresponsive to D , denoted as $X \nleftrightarrow D$, if X assumes the same instance in all counterfactual worlds of $X \cup D$ [8]. That is, instances of X do not affect the status of $X \cup D$. In the case of Cybersecurity, an example of a counterfactual world can be one in which no cybersecurity compromise is ever reported. These counterfactual variables are not observed, and, most probably, will never be observed. Examples of X can be concerns about Cybersecurity. These two can form a union, but X is unresponsive to D , since the instances of such concern do not affect the union. In contrast, one can think of a set X as being responsive to a set D . In this case, let the set X be the same as before, namely, the set of concerns about Cybersecurity, for example, concerns about data leakage as an element of X . The counterfactual world D can be one where Cybersecurity compromise is supposed to be reported to computer users, but ignored.

If concerns about the data leakage problem is an example of X , then it can assume different instances in different counterfactual worlds of $X \cup D$.

For example, "If one had this concern about data leakage, then one may or may not have ignored the Cybersecurity compromise report".

This shows that some instances of X can belong to some counterfactual world of $X \cup D$. Therefore, X is responsive to D .

X refers to the collections of events (indicating, for example, different states of data leakage) some of which occur after decision(s) D have been made. Given decisions D , the variables in the set C are causes of x with respect to D if all the following three conditions are met:

- Condition 1: x is not a member of C .
- Condition 2: x is responsive to D .
- Condition 3: C' is a minimal set of variables such that x is unresponsive to D in worlds limited by C' (that is, $x \nleftrightarrow D$, and C' is a minimal set such that $x \nleftrightarrow c' D$).

The third condition is saying that C has a definite influence on x being responsive to D. The influence is that the relevant cause (or causes) must be included in whichever set of variables that also necessarily differ (being responsive) in accord with x being responsive to D. So, the set C' that limits the relation of x with D (regarding responsiveness) is a minimal set.

The following are the brief explanations with regard to the system discussed here.

- Condition 1 affirms that the effect (X) is not a member of the set of causes.
- Condition 2 affirms that for x (data leakage) to be caused with respect to decision D (data leakage must have been caused by the bugs in the system), it must be responsive to that decision.
- Condition 3 states the following: suppose that one can find a set of variables Y such that X, data leakage, can be different in different counterfactual worlds only when Y is different. In that case, Y must contain a set of causes.

Our approach in this paper is showing the effects of intervention. Causal modeling helps us ask the right questions about causation and helps us devise a way to emulate it by means that are not intrusive. Our emulation of interventions are based on observational studies and using data to find causal relation between them.

Causal relations are not features that can be directly read off from the data, but have to be inferred. The field of causal discovery is concerned with the inference and the assumptions that support it. Instrumental variable method ensures that we obtain the close-to-correct causal effect, even if there are unobserved conditions. Combining propensity-based and regression-based methods provides us with a causal estimate that is accurate whenever the model is correctly specified.

The potential outcomes framework can be detailed as follows: counterfactual variables such as “knowledge and action of a person P had he received the information that the cybersecurity of his computer system has been compromised” and “knowledge and action of a person P had he not received the information that the cybersecurity of his computer system has been compromised” are as appropriate as traditional variables such as “knowledge and action of a person P” – though one of these counterfactual variables is not observed, and most probably, will never be observed, in the case of this person P.

2.1. *Common Cause, Confounding, Control, and Instrumental Variables.*

Common causes explain the fact that there are concepts related to causation that are more important than correlation.

Suppose that a person has received a “Compromised Host” notice from some authorities, and also his computer-savvy friend (who may or may not know about the notice) has checked this person’s computer and is confident that attackers have gained unauthorized access to this person’s computer. Therefore, this person is worried about cybersecurity, and would like to take steps.

What are the causes of receiving such as notice? What are the causes of this computer-savvy friend being confident that attackers have gained unauthorized access to this person’s computer? If there is some disaster, it could cause the “Compromised Host” notice to go out. It could also cause one’s computer-savvy friend being confident that attackers have gained unauthorized access to this person’s computer.

If a disaster happens, both of these are likely. This means in a data set one can find a correlation between the two.

We know there is no causal effect of receiving a “Compromised Host” notice on one’s computer-savvy friend (who may or may not know about the notice) being confident that attackers have gained unauthorized access to this person’s computer, or vice versa. This is the essence of “correlation does not imply causation”.

When there is a common cause between two variables, then the variables will be correlated. This is part of the reasoning behind the phrase, “There is no correlation without causation”.

Suppose that we are dealing with two concepts, named A and B. If neither A nor B has been definitely known to cause the other, and the two are correlated, there must be some common cause of the two. It may not be a direct cause of each of them, but it is there somewhere. This implies that we need to control for common causes if we are trying to estimate a causal effect of A on B.

Common cause variation is fluctuation caused by unknown factors resulting in a steady but random distribution of output around the average of the data.

Suppose that we take the average of the data, and do a steady but random distribution of output around the average. There will be unknown factors that will result in that distribution. This will cause a source of variation called common cause variability. This is a measure of the potential of the process – which includes how well the process can perform, if and when special cause variation is removed. Common cause variation is also called random variation, or non-controllable variation.

If we do not include hidden common causes in our model, we will estimate causal effects incorrectly. This is similar to the notion of confounders (in this particular case, some cybersecurity disaster has happened).

Confounding variables are to be understood in terms of data generating model. Pearl defines the concept of confounding as follows: Let X denote some independent variable (for example, the “Compromised system “notice), and Y some dependent variable (the person is worried and wants to take action). We might want to estimate what effect X has on Y, without regard to other potential factors; for example, if the person is, at the same time, not feeling well. We say that X and Y are confounded by some other variable Z whenever Z is a cause of both X and Y. In our case, Z is that some cybersecurity disaster has happened.

One can state that X and Y are not confounded whenever the observationally witnessed association between them is the same as the association that would be measured in a controlled experiment, with x randomized.

An equality here can be stated as $P(y | do(x)) = P(y | x)$; this can be verified from the data generating model provided that we have all the equations and probabilities associated with the model. This is done by simulating an intervention $do(X = x)$ and checking whether the resulting probability of Y equals the conditional probability $P(y | x)$.

Control is a concept related to confounders. Suppose that we are attempting to assess the effectiveness of the notice being given, from population data. The data shows that prior knowledge about such incidents (Z) influences the state of mind .e.g. worry and wanting to take action (Y). In this scenario, Z confounds the relation between X (his computer-savvy friend takes the action of telling him) and Y since Z is a cause of both X and Y .

We hope to obtain an unbiased estimate $P(y | do(x)) = P(y | x)$. In cases where only observational data are available, an unbiased estimate can only be obtained by "adjusting" for all confounding factors, which means conditioning on their various values and averaging the result.

This gives an unbiased estimate for the causal effect of X on Y . The same adjustment formula works when there are multiple confounders except, in this case, the choice of a set Z of variables that would guarantee unbiased estimates must be done with care. One can view cause-effect relationships via directed acyclic graphs; one should also link causal parameters and observed data, such as information about the subjects studied, as well estimation of the resulting parameters.

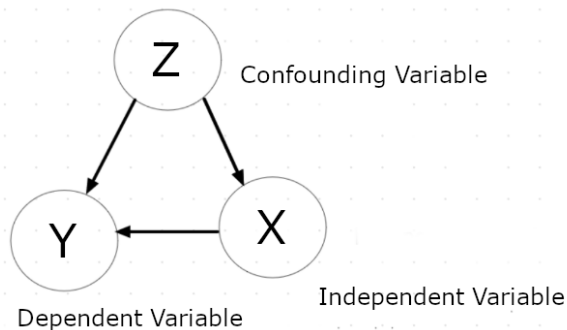


Figure 1. Example of a Causal Model

3. Overview

Since experimentation is not feasible for simulating real world data breaches, the analysis relies solely on observational data. In this regard, Judea Pearl's theory of Counterfactual World theory is extended with the use of propensity scores to calculate causal inference. The main issue to tackle regarding the use of observational data is the bias within the data caused by confounding variables, both known and unfounded. These include the previously mentioned common causes, instrumental variable, and any other covariates.

We thus present the use of causal modeling as a tool for gaining insight into how data breaches occur, and the degree to which certain associations behind these breaches can be seen as causal. We present a subset of open-sourced data offered by Verizon Communication. We then apply principles of Pearl's Causal

inference through the software library DoWhy in order to understand the causal effects of our interventions.

3.1. Methodology

We concluded that DoWhy, a Microsoft open source Causal Modeling framework, was most appropriate for this current project, for its ease of use and abundant resources. It also provided an intuitive method to implement the Model -> Identify -> Estimate -> Refute structure of the analysis. All of these were readily provided by DoWhy and were thus implemented with DoWhy's built-in functions. Due to the limitation on data availability regarding data breaches, we believe these provided enough for an exploratory analysis on the subject [9,10].

DoWhy also provides a principled way of modeling a given problem as a causal graph so that all assumptions are unequivocal and explicit. It provides an integrated interface for causal inference methods, combining the two major frameworks of graphical models and potential outcomes. It also automatically tests for the validity of assumptions if possible and assesses the robustness of the estimate to violations.

It is important to note that DoWhy builds on two of the most powerful frameworks for causal inference: graphical models and potential outcomes. It uses graph-based criteria and do-calculus for modeling assumptions and identifying a non-parametric causal effect. For estimation, it switches to methods based primarily on potential outcomes.

In the following paragraphs we will describe the techniques to use for our analysis: Propensity Score Matching, Propensity Score Stratification, and Linear Regression Estimator. These techniques can all be founded within the DoWhy framework.

Linear Regression Estimator provides a baseline analysis assuming an evenly distributed dataset. It provides a foundation to compare results with the other methods. As linear regression only describes a correlation between the treatment and outcome, Propensity Score Matching and Propensity Score Stratification both use linear regression while adding additional processes in order to account for confounding variables and properly compartmentalize each data entry to find a causal relationship between the treatment and outcome.

Propensity Score Stratification takes the propensity scores of each entry and classifies them into equal sub-groups. These subgroups are classified by the similarity of the covariates. The aim is to have each sub-group represent a distribution that accurately represents a non-biased dataset to the best of its ability.

Propensity Score Matching instead takes the propensity scores of each entry and finds the entries with the highest propensity scores within the treatment group and finds the entries within the control group with covariates that most closely match each treatment group entry. This attempts to establish parity between the covariates of the treatment group and the control group.

Both Matching and Stratification work to remove bias from high-dimensional datasets. They do so by balancing out the treated and control groups with processes that emulate a random distribution in an experiment. This is done by evaluating the propensity score of each group. The propensity score represents the probability of the treatment on each sample in the treatment

group and is calculated by mapping the outcomes to a linear regression line. The difference in the methods in how they use the propensity score to balance out the treatment and control group.

In addition, refuters are necessary in the causal analysis process in order to verify the robustness of the results. The following methods all check the effects of confounding variables and compare them with the treatment to concretely establish a causal relationship. Generally, the refuters all involve rerunning the same causal analysis methods with the following changes to the dataset:

- Placebo Refuter: Replaces the treatment variable with a placebo variable with random values
- Data Subset Refuter: Runs the program over a randomly chosen subset of the original data
- Common Cause Refuter: Generates a random confounding variable

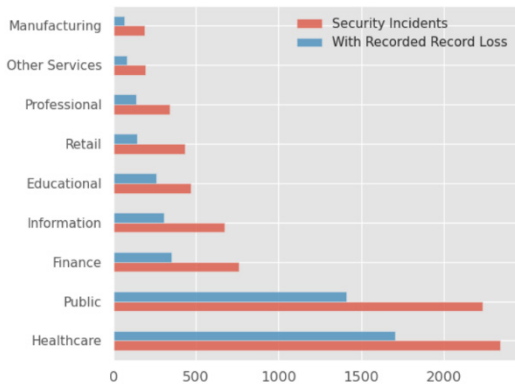


Figure 2: Number of Data Breaches by Industry

3.2. Data Acquisition

We note that high-quality information on real-world cybersecurity incidents through academic or otherwise publicly accessible channels is likely to be unrepresentative of the nature in which breaches occur on a broader scale. As a result, we focus on analyzing healthcare privacy breach data, which generally enjoys stringent reporting standards. Our reasoning is as follows.

For the private sector, disclosure of breaches can negatively impact short term company value as well as consumer trust. A report by IBM’s Ponemon Institute in 2019 estimates the global average impact of having a data breach to an organization to be 3.9 million US dollars, representing a rise of 12% over the course of five years. For organizations with fewer than 500 employees, this cost averages to 2.5 million dollars [11]. Voluntary disclosure of data breaches may be unpalatable in light of this [12].

In contrast, government and healthcare institutions are generally under greater legal pressure to disclose similar incidents. For instance, the Department of Human and Health Care Services (HHS) in the United States mandates that information regarding data breaches involving over 500 individuals be disclosed to the media within 60 days of discovery. Structured collection of such breaches is made publicly available through the HHS website [13].

Initial exploration was performed on the VERIS Community Database (VCDB). The VCDB is an open dataset covering a broad spectrum of security incidents occurring throughout both public

and private sectors. Data available through this channel represents a small portion of data contained in a more comprehensive report presented in Verizon’s annual Data Breach Investigation Report. The VCDB is attractive as there are few publicly available repositories containing annotated security breach information [14].

The VCDB follows the Vocabulary for Event Recording and Incident Sharing (VERIS) framework. Generally, information surrounding security incidents is divided into four categories: Actor, Attribute, Asset, and Action. Actor pertains to the entity or entities responsible for the data breach. Asset characterizes the type of information lost, as well as how accessible said information was. Attribute refers to the degree which the asset in question was affected, as well as the severity of the incident, the medium of transmission, and if said data was exposed to the public. Finally, Action describes how the security breach was carried out; such as if the breach was a result of malware, or simply negligence. Additional data on affected industry and incident timeline are included as well.

To accommodate the wide variety in reporting standards, VERIS uses a fine-grained approach for characterizing security incidents, using a nested key-value store to accommodate some 173 attributes.

We will take the “actor” category as an example. For any given incident, the individual or individuals responsible could be categorized as either external, internal (affiliated with the organization), a partner (associate, but not directly affiliated), or simply unknown or not available. Within each type of actor lies a different subcategory. For instance, the “external actor” label can represent a criminal organization, foreign government, former employee, or a combination thereof. As a result, many of the keys contain lists as values, as represented in (1).

$$\begin{aligned}
 \text{“actor”} &: \{ \\
 & \quad \text{[“external”} : \{ \\
 & \quad \quad \text{“variety”} : [\text{“Mother Nature”}, \\
 & \quad \quad \text{“Criminal Organization”}], \\
 & \quad \quad \text{“motive”} : [\text{“NA”}, \\
 & \quad \quad \text{“Espionage”}, \\
 & \quad \quad \text{“Ideology”}] \\
 & \quad \dots \\
 & \} \dots
 \end{aligned}
 \tag{1}$$

3.3. Why Healthcare?

While the VCDB contains many features describing the companies that were victims of data breach, little information seems to be provided regarding the situation preceding and during the data breach. Therefore, to maintain a degree of uniformity of each company, narrowing down to one industry like healthcare would mitigate discrepancies within the dataset.

Furthermore, the VCDB utilizes a JSON-formatted, hierarchical data structure presented as a list of key-value pairs.

While each record adheres to the same general schema, sparsity arises as a result of how much data is disclosed by each entity, or simply what information is relevant to which sector. Referring back to US Healthcare data breach disclosure law, we can expect a baseline of data to be provided, such as the number of individuals affected, the type of breach, and the vector of attack.

The next logical step was to transform the data from a hierarchical format into a two-dimensional, tabular structure. A strong motivation for this was to make the data both more comprehensible and consistent.

Transforming the database for VCDB was straightforward thanks to the open-source library Verispy. Verispy converts the deeply nested structure of the original VCDB dataset into a two-dimensional grid-like format. by performing “one-hot encoding” on each of the categorical variables.

This leads to a relatively consistent dataset with the caveat of vastly increasing the (perceived) dimensionality. The final table consists of 2,347 columns, containing 2108 (89%) Boolean entries, 147 (6%) string or string-like entries, and the remaining 92 (5%) numerical entries.

1839 entries within VCDB are related to healthcare out of 8363 data breach entries. In order to further narrow down VCDB into healthcare, we further take out all irrelevant variables to our causal model (described in the next section) as well as drop all entries that have empty values in any of those variables. This drops the final dataset entry count to 106 entries, a mere 1.3% of the original VCDB size. This demonstrates sparseness of the VCDB dataset, despite the breadth of information available within.

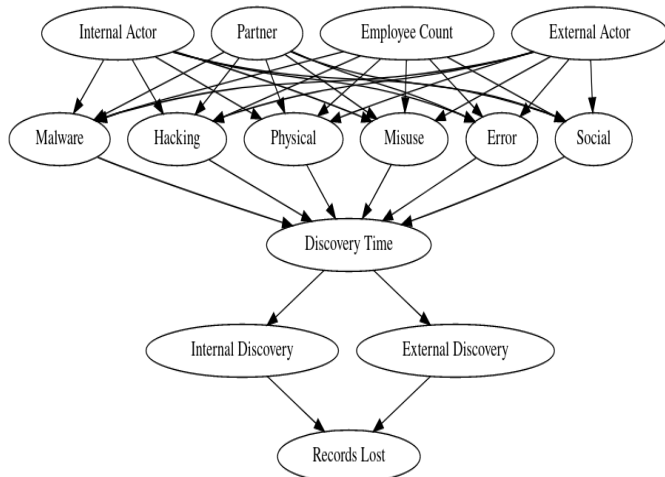


Figure 3. Causal Model for Cybersecurity (on VCDB)

3.4. Data Breach Model

Figure 3 represents the causal graph used as the basis for our causal analysis. The variables are all taken from VCDB and were decided on how accurately they could be mapped to a timeline of the data breach. Since all observational data given to us are all post-data breach, the way to approximate a causal effect for this analysis is to generate a model that shows a progression of events. Many of these variables and their sequencing were derived from personal interpretation than any logical standpoint. We will take a look at each variable type with their justifications.

- Actor
- Employee Count
- Action
- Discovery Time
- Discovery Method
- Records Lost

‘Actor’ is referring to the one who instigates the action against the victim. This could be a single person, a group of people, or even a natural disaster. In the causal model, the actor is spread amongst three categories: Internal, External, Partner. Internal actors are those who work within the company that is affected by the breach. External actors are those with no affiliation whatsoever with the company. Finally, partners do have or are part of an organization that has an affiliation with the company but are not from the company themselves. This variable represents a general categorizable description regarding the perpetrator of the data breach and is put near the top of the causal graph because the ‘actor’ is the one that will begin this data breach event.

‘Employee Count’ represents the general size of the victim company, which is represented by an integer value. Employee count was chosen as it is a variable that conveys a simple, but ordinal description of healthcare organizations.

Each of the types of data breaches (‘Malware’, ‘Hacking’, ‘Physical’, ‘Misuse’, ‘Error’, ‘Social’) are labeled as ‘Actions’ within VCDB. These are the treatment variables in which the analysis will be performed on. Each action is a binary state, and while there are a few rare cases that have multiple ‘Actions’ at once, these are still considered one data breach.

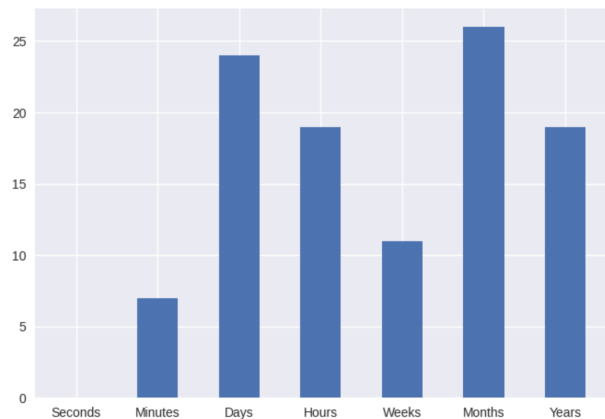


Figure 4. Distribution of ‘Discovery Time’

‘Discovery Time’ is the unit of time it took for the data breach to be discovered. VCDB does not have discovery time as an integer number. Instead, the variable is categorized as six different ranges of numbers, getting subsequently larger. Going under the assumption that the larger unit means that the actual discovery time was longer, the units were combined into one variable from 1-6, each representing a greater scale of time. The unit of time represents the general time frame of the data breach being discovered. Discovery time is one of the outcomes that is used to measure causality of data breaches. ‘Containment Time’ and ‘Exfiltration Time’ were considered as well. However, a high

proportion of these entries remain unfilled in the VCDB dataset and are therefore not of much use.

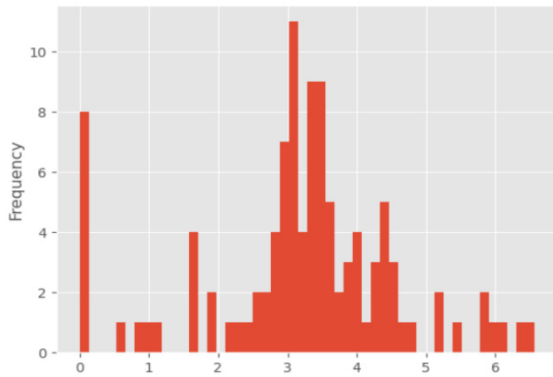


Figure 5. Distribution of Records Lost (log-scaled)

‘Discovery Method’ is the method by which the victim was first able to discover the data breach. Like ‘Actor’, this is also split into External, Internal, and Partner, which represents the relationship of the individual or group that discovered the breach to the victim company. External meant those unrelated to the company, Internal part of the company, and Partner are those affiliated with but not directly part of the company.

‘Records Lost’ is the second outcome we will be using as an outcome to test the causality of the causal model. Similar to Discovery Time, Records Lost is not an integer value, but ranges of values of subsequently greater number. This variable is also similarly combined into one variable ranging from 0-6. One major caveat is that this variable doesn’t have a defined unit and thus the scale of a unit of record is determined by each individual company. Part of the decision to focus on healthcare companies only was to mitigate this ambiguity.

4. Results and Analysis

Causal estimate calculations were run across all six ‘Actions’ (Social, Physical, Misuse, Malware, Hacking, and Error) and two outcomes (Discovery Time, Records Lost). This means multiple runs using the same causal model and dataset but changing the ‘Action’ and ‘Outcome’ input for each run until all permutations of each variable was covered. This was then repeated across all refutations. The causal estimate results for Propensity Score Matching on Discovery Time and Records Lost are shown in Figure 6 and 7, respectively.

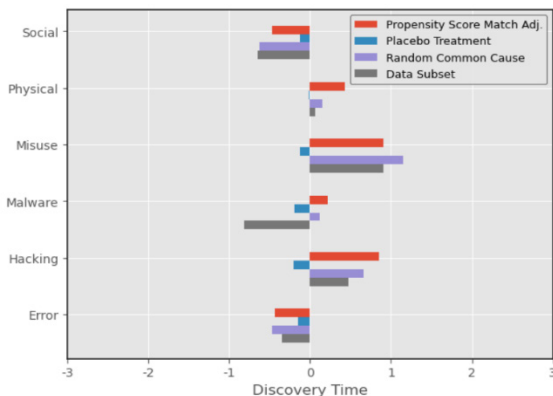


Figure 6. Causal Estimates for Discovery Time

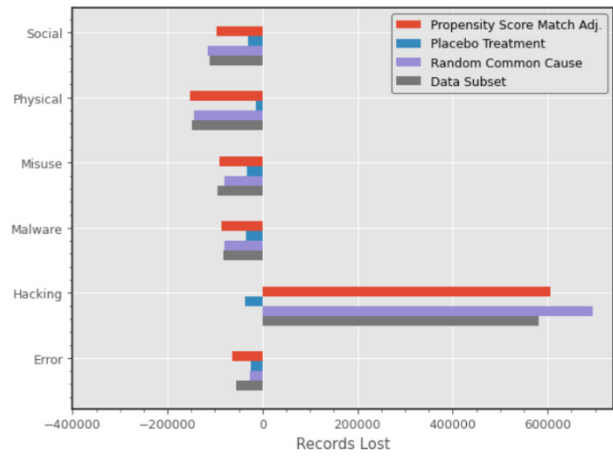


Figure 7. Causal Estimates of Records Lost

As seen in Figure 6 shows, the causal estimate of each action on Discovery Time shows quite a range of values and distributions across all actions. 3 actions (Physical, Misuse, Hacking) have positive causal estimates, indicating a potential strong causal relationship between the actions and lengthy discovery times. On the other hand, Social and Error turn out negative causal estimates, meaning that the impact of those two variables on discovery time is minimal. Lastly, Malware has a unique scenario where there is a split between the Propensity Score and its refuters.

For records lost (Figure 7), Hacking returns an overwhelming higher causal estimate compared to all other actions. In fact, all the other actions return a negative causal estimate. This does not necessarily mean the lack of causal effect of the other actions on records lost. However, it does provide strong indication that the greatest impact when it comes to records lost during a data breach is most likely the result of hacking as opposed to all other methods. Interestingly, this is backed up by both the Random Common Cause and Data Subset, but not the Placebo Refuter. In the Placebo case, the causal estimate returns a comparable negative value to the other actions. A possible explanation can be traced back to the nature of the dataset. While our causal model brings into consideration other causes of data breach, the distribution of the effect of each cause can be hard to separate. This is exacerbated when the Placebo Refuter randomizes the treatment variable, setting it so that every single entry in the dataset can also be considered part of the method of hacking.

This Placebo Refuter discrepancy is reflected across all the actions, which each return strongly diminished causal estimates. However, Hacking remains the only variable where the causal estimate goes from a positive to a negative value.

Another quirk to note is the large value of the causal estimate of Hacking on Records Lost. The reason for this exaggerated value is likely due to a lack of a solid control group within our data. The dataset provides us with a large selection of data breaches in a wide variety of companies. What the dataset lacks are scenarios where no data breach has occurred, generating an inherent bias within the dataset. This bias makes it so that the data do not fit well into linear regression, hence providing an overly large value as the result.

The most unexpected outcome was that propensity score stratification gave inconclusive results when ran on DoWhy, hence the lack of data on this portion of the analysis. After some analysis,

we come to the conclusion that, due to the binary nature of each action, the distribution of the linear regression is not clear enough for stratification to be able to quantify and compartmentalize the dataset into groups. Hence, the resulting value outputs an inconclusive value due to a lack of substantial strata. This applies to stratification and not matching because matching disregards parts of the dataset with low propensity score; in stratification they still have an impact due to those data entries being assigned into strata.

Overall, in this specific scenario and dataset, Hacking would prove to be the most impactful amongst all methods of data breach. However, the refuters give strong indication on where this impact is limited regarding not only the action itself but the dataset as a whole.

5. Conclusion and Future Work

The principal findings of this paper demonstrate the unique perspective of the causal modeling approach. Because we cannot realistically set up an experiment on data breach incidents, particularly in which all factors are readily provided, DoWhy and Causal Modeling allow us to simulate such experiments and make inferences with a degree of robustness based on events that would otherwise be difficult to duplicate.

We identify a subset of factors in the Verizon Community Database and create a hypothesis based on the theory that malware and hacking are the most prominent causes of data breaches. Through propensity score matching and stratification, we measure the strength of the action behind data breaches. By running refutation tests, we are able to verify how well these metrics hold up, similar to how traditional experiments employ control groups or utilize a placebo treatment.

Ample room remains for the use of causal modeling in cybersecurity. We limit the scope of the factors considered in the Verizon Dataset to Actions in order to emphasize the results of the exploratory approach. A larger and denser dataset could utilize the causal model better.

Other fields of cybersecurity lend themselves well to causal modeling. In particular, the use of Directed Acyclic Graphs to model vectors of attack in a network intrusion scenario could lead to different approaches into how such cases are handled.

The study is important to the readers in the scientific community since it is relevant to formulating policies in industry and government, in order to avoid such problems in the near future. Given the context of the work, exhibited in the paper, our findings are worthy of note.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgments

We thank the Data Science Discovery Research Program of UC Berkeley and their participant, Rubina Aujla of UC Berkeley, 2018-2019, for contributions to causal modeling.

References

- [1] P. Zornig, Probability Theory and Statistical Applications, De Gruyter. ISBN-13: 978-3110363197
- [2] B. McLaughlin, On the Logic of Ordinary Conditionals, Buffalo, NY: SUNY Press, 1990.
- [3] J. Pearl, Causality: Models, Reasoning, and Inference, Cambridge: Cambridge University Press, 2000.
- [4] J. Pearl, Causality, 2nd edition, Cambridge University Press, 2009.
- [5] S. Thornley, R.J. Marshall, S. Wells, R. Jackson, "Using Directed Acyclic Graphs for Investigating Causal Paths for Cardiovascular Disease", Journal of Biometrics & Biostatistics, 2013, 4:182. doi:10.4172/2155-6180.1000182
- [6] P. Menzies, H. Beebee. "Counterfactual theories of causation.", 2001.
- [7] A. Chesher, A. Rosen, Counterfactual worlds. No. CWP22/15. Centre for Microdata Methods and Practice, Institute for Fiscal Studies, 2015.
- [8] A. Agresti, An Introduction to Categorical Data Analysis, 3rd Edition, Wiley Series in Probability and Statistics.
- [9] A. Sharma, E. Kiciman, 2020. Causal Inference and Counterfactual Reasoning. In 7th ACM IKDD CoDS and 25th COMAD (CoDS COMAD2020), January 5–7, 2020, Hyderabad, India. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3371158.3371231>
- [10] A. Sharma, E. Kiciman, et al. DoWhy: A Python package for causal inference. 2019.
- [11] IBM Security 2019 Cost of a Data Breach Study: Global Overview
- [12] R. Anderson, 2001. Why Information Security is Hard-An Economic Perspective. In Proceedings of the 17th Annual Computer Security Applications Conference (ACSAC '01). IEEE Computer Society, USA, 358.
- [13] Department of Health and Human Services, 2013. Modifications to the HIPAA Privacy, Security, Enforcement, and Breach Notification Rules Under the HITECH Act and the GINA Act; other Modifications to the HIPAA Rules (78 FR 5565), pp. 5565-5702
- [14] VERIZON. Data Breach Investigations Reports Overview, 2019 (DBIR).

Racial Categorization Methods: A Survey

Krina B. Gabani*, Mayuri A. Mehta, Stephanie Noronha

Department of Computer Engineering, Sarvajani College of Engineering and Technology, Surat, 395001, Gujarat, India

ARTICLE INFO

Article history:

Received: 30 January, 2020

Accepted: 25 May, 2020

Online: 11 June, 2020

Keywords:

Ethnicity

Ethnic group

Race

Racial classification

Anthropometry

Facial features

Race identification

Face recognition

Machine learning

ABSTRACT

Face explicitly provides the direct and quick way to evaluate human soft biometric information such as race, age and gender. Race is a group of human beings who differ from human beings of other races with respect to physical or social attributes. Race identification plays a significant role in applications such as criminal judgment and forensic art, human computer interface, and psychology science-based applications as it provides crucial information about the person. However, categorizing a person into respective race category is a challenging task because human faces comprise of complex and uncertain facial features. Several racial categorization methods are available in literature to identify race groups of humans. In this paper, we present a comprehensive and comparative review of these racial categorization methods. Our review covers survey of the important concepts, comparative analysis of single model as well as multi model racial categorization methods, applications, and challenges in racial categorization. Our review provides state-of-the-art technical information concerning racial categorization and hence, will be useful to the research community for development of efficient and robust racial categorization methods.

1. Introduction

Human face expresses social information that is highly useful in automated systems. It provides soft biometric information of human such as race, gender, age, identity and emotions [1-7]. This information is significant in interdisciplinary research areas such as psychology science, computer vision science, neuroscience, anthropological science as well as in the social security and forensic art department. Amongst the various types of biometric information, race information is crucial and is required for a wide range of applications. Race is a group of human beings differentiated based on physical or social attributes. Race conveys social and cultural traits of different communities. Facial features such as eyes, eyebrow, ear, nose, cheek, mouth, chin, forehead area and jaw differ from human to human and are highly dependent on racial category [8-14]. Figure 1 shows the difference in facial features of different racial groups.

Race analysis is essential in contemporary applications such as criminal judgment and forensic art [15-20], aesthetic surgery [21], healthcare [22-26], medico legal [27-29], video security surveillance and public safety [30], human computer interface [31-33] and face recognition [34]. In such applications, race analysis is required for identification of individuals. Several racial

categorization methods have been proposed in literature. They are either single model racial categorization methods or multi model racial categorization methods. Single model racial categorization method uses facial features to recognize race [35-45]. Conversely, multi model racial categorization method considers fusion of physical characteristics such as gait pattern and audio clues in addition to facial features [5, 7, 46-50]. The majority of the practical applications involve single model racial categorization because facial data is available in large quantities compared to gait pattern and audio clues. However, there are applications that consider gait pattern and audio cues in absence of facial image. The single model racial categorization methods mainly differ from each other with respect to classification approach such as Support Vector Machine (SVM) [51-54], Convolutional Neural Network (CNN) [38, 44, 55-59], Artificial Neural Network (ANN) [60], Local Binary Pattern (LBP) [61] and Local Circular Pattern (LCP) [62]. In literature, participants based racial categorization methods such as diffusion model [63] and implicit racial attitude [64] are also available. The multi model racial categorization methods involve classification approaches such as SVM [65-67], logistic regression [66], Adaboost [66], random forest [66], CNN [68] and Haar-LBP histogram [69].

*Krina B. Gabani, Email: krinagabani007@gmail.com

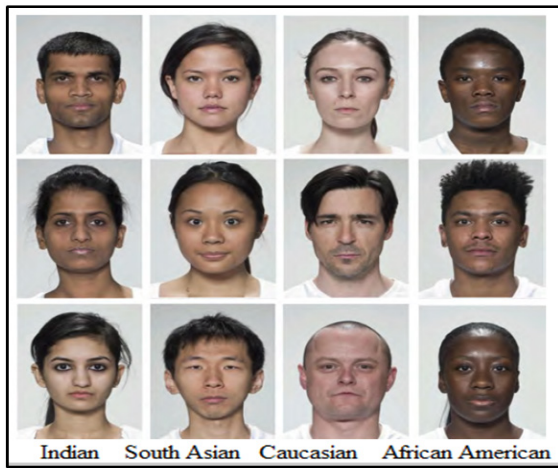


Figure 1: Illustrative facial characteristic variance of human races. (Figure source: <http://www.faceresearch.org/>)

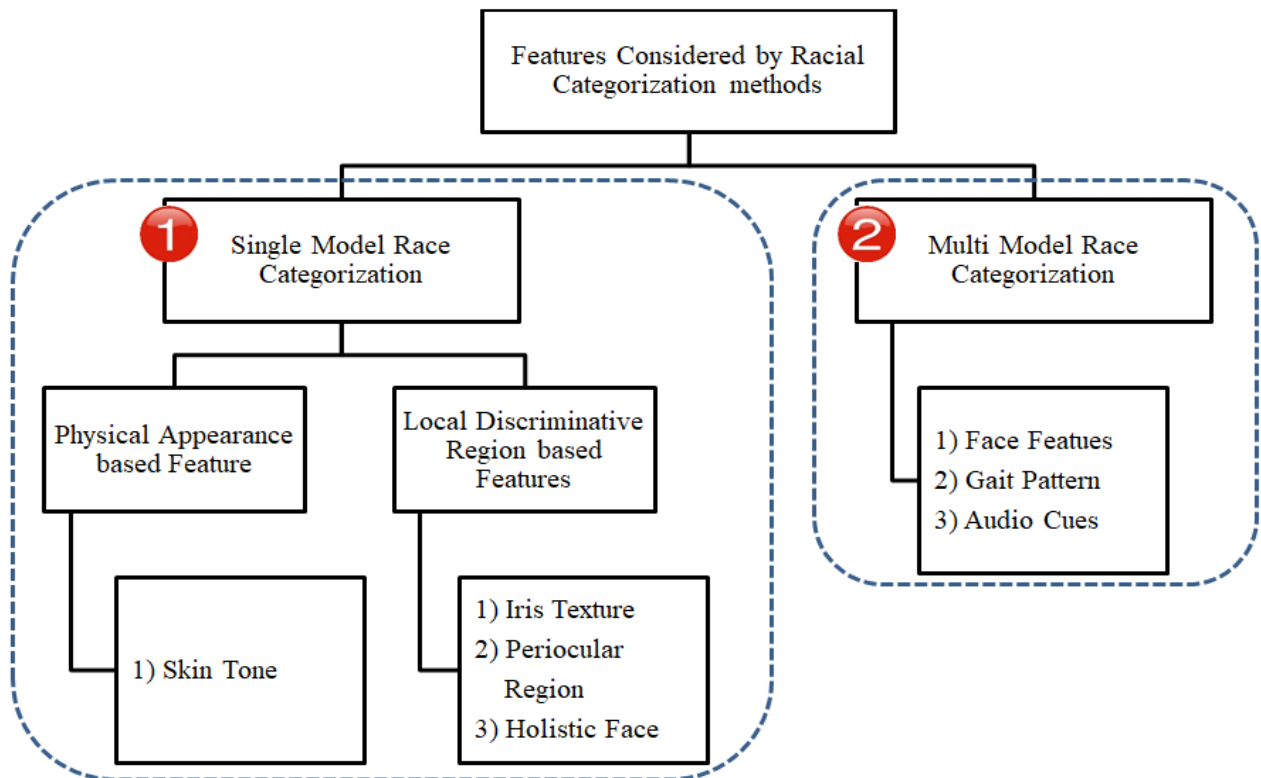
In this paper, we present a thorough and extensive study of various racial categorization methods. First, we present a taxonomy of available racial categorization methods. Then we describe several single model and multi model racial categorization methods. Based on our study, we identify several parameters to evaluate them. Subsequently, we present parametric evaluation of single model racial categorization methods and multi model racial categorization methods separately based on identified parameters. Next, we illustrate applications of racial categorization and future research direction in the field of racial categorization. Our comprehensive and comparative survey will serve as a catalogue to researchers in this area.

The rest of the paper is structured as follows: In section 2, we describe the taxonomy of racial categorization methods and features considered by different racial categorization methods. In section 3, we illustrate various single model racial categorization methods and their parametric evaluation. In section 4, we illustrate various multi model racial categorization methods and their parametric evaluation. Section 5 describes the major applications of racial categorization. In section 6, we list key challenges in the field of racial categorization. Finally, section 7 specifies conclusion and feature scope in the field of racial categorization.

2. Classification of Racial Categorization Methods

Racial categorization methods are broadly categorized into two categories: single model and multi model. As shown in Figure 2, the features used by single model and multi model methods to classify humans into features or local discriminative region based features or combination of both. Amongst the discriminative region based features, iris texture, periocular region or/and holistic face are used for racial categorization [70-77]. Multi model racial categorization takes into consideration face features, gait pattern and audio cues to classify humans into various race categories. Gait pattern is also useful to recognize the biometric information of humans [6, 78-80].

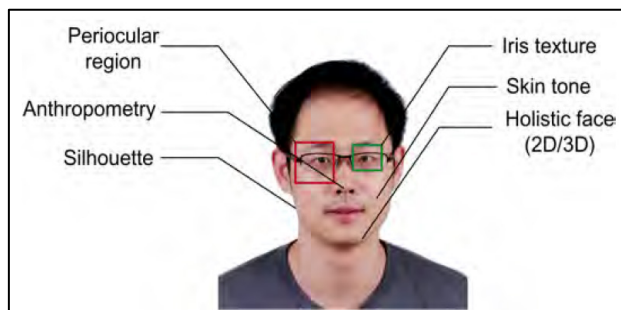
At this juncture, we clarify that at the top of all categorizations, a human is mainly divided into two categories, namely race and ethnicity, on the basis of his/her physical appearance and social appearance respectively. However, some researchers use words race and ethnicity interchangeably [24, 35, 81].



: Features considered for racial categorization

2.1. Features Considered by Single Model Racial Categorization

Figure 3 shows the facial features such as periocular region, anthropometry distance, silhouette, iris texture, skin tone and holistic face.



Skin tone differs mainly due to geographical location of humans. African-American, South-Asian, East-Asian, Caucasian, Indian and Arabian have different skin tones. Skin tone plays a minor role in identification of racial groups because skin color may also differ due to varying lighting conditions during the image capturing process [56, 64].

Like fingerprints, iris texture is a significant biometric characteristic of humans because it is unique for every human being [82]. It is highly useful for racial categorization because different race groups such as American, Indian and so on have different iris texture [59, 74-76, 82-84]. The key limitation of this feature is that it cannot be considered if race is to be identified from video because video may be of low quality and hence, may not give precise iris texture information [85-86].

Periocular region is defined as the region surrounded by eye. It is a region that overlays eyebrows, eyelid, eyelash and canthus [52]. It gives rich texture and biometric information as compared to iris texture [30]. Some facial features get influenced due to different facial poses and expressions. However, the periocular region does not get affected due to facial poses and expressions. Hence, it is considered as the most reliable feature for racial categorization [52, 62, 77].

Holistic face provides the texture information of various facial features such as eyes, nose, mouth, cheek, chin, skin color and jaw line [43, 51, 63, 87-89]. Extra frontal face features such as hairline and hair color in combination with cropped aligned face features ease racial categorization process [54].

2.2. Features Considered by Multi Model Racial Categorization

Multi model racial categorization improves accuracy of racial categorization via fusion of facial features with other human features such as gait pattern and audio cues [90] (Figure 4). Below we discuss the features considered by multi model racial categorization.

Gait pattern, also known as the walking pattern, is a prominent biometric feature that varies from human to human and is used for identification of a person [91-92]. Advanced racial categorization methods use gait pattern fused with facial features for overall

effective racial categorization [92-97]. For videos in which humans at near to moderate distance have been captured, facial features are sufficient to identify the race. However, for videos in which humans at far distance have been captured, gait pattern is highly useful to identify the race of human because facial features of humans at far distance are not clearly visible. Thus, fusion of gait pattern with facial features improves overall racial categorization accuracy [65].

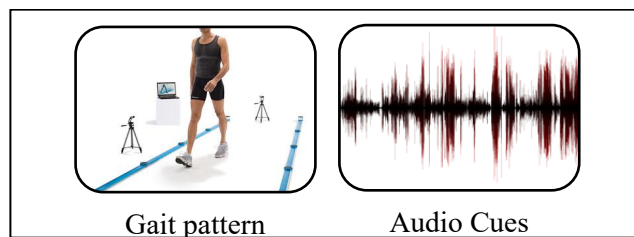


Figure 4: Features considered by multi model racial categorization

Audio pattern differs from race to race [98]. It is useful to identify race in case a video sequence or image of a person is not available. For instance, it is useful to identify race from phone calls.

3. Single Model Racial Categorization Methods

Race depends on physical and social characteristics of humans. As geographic distance increases, variation in facial features of inter-races become visible. As facial data is easily available compared to gait pattern, the majority of the racial categorization applications use a single model racial categorization method. Moreover, it has been revealed in literature that facial features are more prominent for race categorization [99-100]. Below we discuss various single model racial categorization methods available in literature.

3.1. Multi Ethnical Categorization using Manifold Learning

In [51], authors have proposed a method for intra-racial categorization based on facial landmarking. This method classifies eight intra-races residing in China based on facial landmarking concept. It includes Active Shape Model (ASM) to locate 77 facial landmarks. The landmarks are used to calculate three types of geometric facial features: distance, angle, and ratio. These features are provided to different classifiers such as Bayesian Net, Naive Bayesian, SMO, J4.8, RBF Network and LibSVM to identify the race category. The dimensionality reduction process carried out by manifold learning approach is useful to reduce the complexity. Though this method is efficient, it is not useful to identify race from a person's profile face images.

3.2. GWT and Ratina Sampling based Ethnicity Categorization

Multiclass SVM based ethnicity categorization method is proposed in [52]. Figure 5 shows the key steps involved in this method.

As shown in figure, first image is normalized via applying rotation operation and changing resolution. The resolution of the image is changed in such a manner that it maintains distance of 28 pixels between two inner corners of eyes. Subsequently, eye and mouth facial features are extracted by fusion of Gabor

Wavelet Transfer (GWT) and retina sampling for efficient categorization. GWT is used to extract accurate orientation and frequency of facial features. Retina sampling method is used to set facial feature points. The features are fed to multiclass SVM classifier for ethnicity identification. Typically eye is considered as a most prominent feature for racial categorization due to its pose invariant characteristic. On the other hand, uncertainty is introduced by mouth region due to its pose variant characteristic. The disadvantage of this method is that GWT provides erroneous features in case of hollow around the eyes. Moreover, Gabor features reflect error due to variation in frequency of eyelashes.

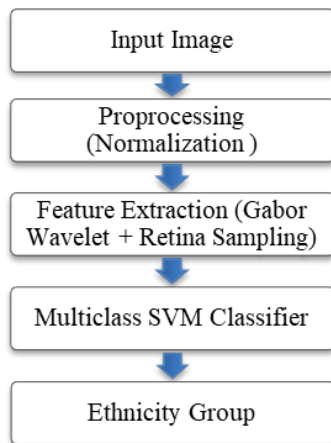


Figure 5: Steps for ethnicity categorization using GWT and retina sampling

3.3. Real Time Racial Categorization

A new method for racial categorization by fusion of Principal Component Analysis (PCA) and Independent Component Analysis (ICA) is introduced in [53]. It consists of major two steps: feature extraction and classification. During the feature extraction step, facial features are extracted using PCA. Subsequently, ICA is used to map and generate new facial features from facial features generated by PCA. New facial features are more suitable for efficient racial categorization. During the classification step, SVM classifier is applied in conjunction with ‘321’ algorithm to classify races. ‘321’ algorithm is inspired by the bootstrap approach for real time racial categorization from video streams. The categorization accuracy of this method can further be enhanced by including pre-processing step to diminish noise from the image and for face alignment.

3.4. Binary Tree based SVM for Ethnicity Detection

In this method, fusions of texture and shape facial features have been considered for better ethnicity categorization [54]. Figure 6 shows the functioning of this method. The first step pre-processing involves the operations such as image resize, image enhancement and image conversion. Then texture features are extracted using Gabor filter and shape features are extracted using Histogram Oriented Gradient (HOG). Subsequently, texture features and shape features are fused together. The fused feature vector is large and it requires more computational time. Hence, Kernel Principle Component Analysis (KPCA) algorithm is applied to reduce dimensionality and complexity. Fused facial features are given as input to binary tree based SVM for ethnicity detection.

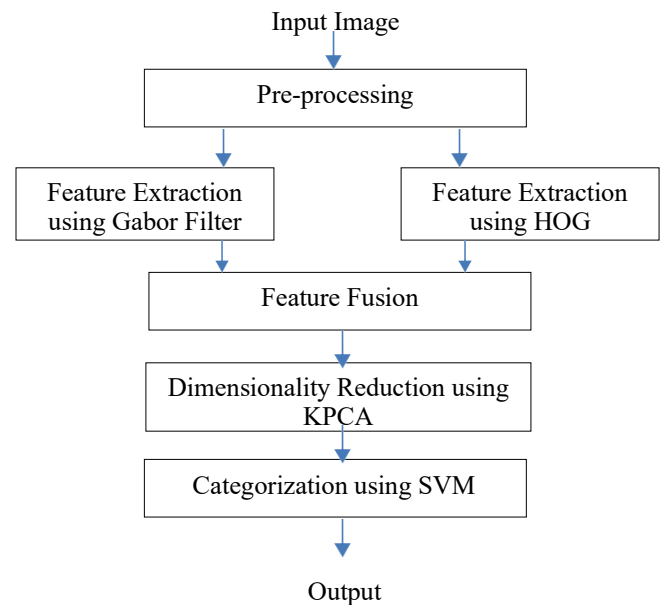


Figure 6: Steps for ethnicity detection using binary tree based SVM [54]

3.5. Racial Categorization using CNN

In [55], a hybrid supervised deep learning based racial categorization method has been proposed. It uses VGG 16 convolution neural network for facial feature extraction and categorization. 224 X 224 face image is given as an input to VGG-16 network for race prediction. Any CNN requires millions of images for training from scratch, which is critical a situation for the medical domain. Hence, to overcome an issue of small dataset, authors have used hybrid approach via fusing VGG 16 with image ranking engine to improve race prediction. It has been shown that image ranking engines work efficiently with CNN based classifiers even for small dataset. The fused feature information extracted by CNN and image ranking engine is used by SVM to learn racial class labels. This hybrid method provides better categorization accuracy.

Begin

1. Face detection using cascade classifier
2. Mark different facial features like nose, eyes and mouth.
3. Calculate distance and ratio between the marked facial features.
4. Detect different geometric facial features.
5. Detect skin color using YCbCr color model.
6. Detect forehead area using Sobel edge detection.
7. Normalize the forehead area considering coordinates of face and eyes and by applying following equation.

$$\text{Normalized Forehead Area} = \frac{\text{ForeheadArea}}{\text{TotalFaceArea}}$$

8. Create normalized feature matrix.
9. Train and validate neural network using feature matrix.
10. Test neural network for racial categorization.

End

Figure 7: Steps involved in racial categorization using ANN

3.6. Neural Network based Racial Categorization

In [56], skin color, forehead area, sobel edge and geometric features are fused for efficient race estimation. Authors have proposed two methods: 1) using Artificial Neural Network (ANN) and 2) using convolution neural network. The steps involved in racial categorization using ANN are shown in Figure 7.

CNN based racial categorization method uses pre-trained VGGNet for racial categorization. It has been observed by authors that CNN based method gives more accurate racial prediction for the given image as compared to ANN based method.

3.7. Local Circular Pattern for Race Identification

Local circular pattern for race identification method works on texture and shape features extracted from 2D face and 3D face respectively. A local circular pattern is an advanced version of a local binary pattern produced for feature extraction. LCP improves the widely utilized LBP and its variants by replacing binary quantization with clustering approach. As compared to LBP, LCP provides higher accuracy even for noisy data. Moreover, AdaBoost algorithm is used for selection of better features and thereby to improve the categorization accuracy. Experimental results have revealed that this method is time efficient and memory efficient.

3.8. Biometric based Machine Learning Method

In [62], authors have proposed a method that focuses on eye region features for racial categorization. The method comprises of five major steps shown in Figure 8. First facial coordinates are located using the DLib library. Subsequently, the region of interest (eye region) is extracted. Then features extracted by LBP and HOG are integrated for efficient racial categorization. LBP and HOG both are individually useful to extract features for categorization. However, fusion of LBP features with HOG features gives higher accuracy compared to other feature fusion approaches. The performance is tested utilizing different classifiers such as SVM, Multi-Layer Perceptron (MLP) and Quadratic Discriminant Analysis (QDA).

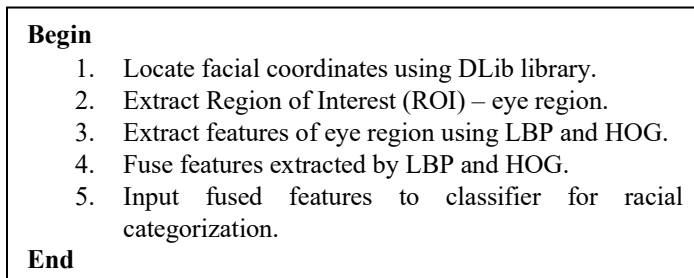


Figure 8: Steps of biometric based machine learning method for racial categorization

3.9. Diffusion Model and Implicit Racial Attitude for Racial Group Identification

In [63-64, 101], manual racial categorization method is proposed. Race is identified by performing several tasks with participants. Diffusion model is used to identify response time boundaries of different participants. This method takes into consideration the visualization of participants for their own race

and other races. Skin color is a less effective feature for automated racial categorization methods due to lightning conditions. However, it is a prominent feature for manual race prediction.

3.10. Performance Evaluation of Single Model Racial Categorization Methods

The above discussed race categorization methods are automated except the last one which is manual race categorization method. With increasing technology, manual race categorization is less effective and less useful as compared to automated racial categorization. Amongst the various automated racial categorization methods, CNN based methods produce more accurate results [55-56] as it considers deep facial features for racial categorization. We observed that the intra-race categorization is not much focused by the researchers in their study.

Based on our study on aforementioned single model racial categorization methods, we have identified the following parameters to compare them: dataset used, racial/ethnic class considered, region of interest, feature extraction operator/s and classifiers used. Dataset refers to the source of data. It is either available online in the form of a standard dataset or it is self-generated. HOIP Database, FERET (Facial Recognition Technology), FRGC v2.0 and BU-3DFE are examples of standard dataset. Self-generated dataset is created by researchers if their predefined requirements are not satisfied with a standard dataset. Racial/ethnic class represents the racial group targeted for study such as Indian, Chinese, Asian, European, African, African American, Caucasian, Bangladeshi, Mongolian, Caucasian, Negro, Hispanic and Pacific Islander. Region of interest specifies the area of face considered for racial categorization. Different face areas include eyes, eyebrow, ear, nose, cheek, mouth, chin, forehead area and jaw line. Some methods also consider skin color for race prediction. Feature extraction operator/s extracts the facial feature from the image. GWT, ICA, PCA, HOG, LBP, LCP are the features extraction operators used by different methods. Classifier specifies the classification algorithm used by the racial categorization method. Different classifiers such as SVM, CNN, ANN, kernel PCA, MLP, LDA, QDA and Kernel based Neural Network (KNN) have been used in different racial categorization methods. Table 1 presents the assessment of aforementioned single model racial categorization methods based these identified parameters.

4. Multi Model Racial Categorization Methods

Multi model racial categorization is highly useful when we do not have human's facial image information. It has been shown in literature that gait pattern and audio cues are amongst the prominent features for biometric information identification. Hence, multi model racial categorization methods use gait pattern, audio cues or fusion of facial features with gait pattern/audio cues to identify race. However, a smaller number of multi model racial categorization methods are available in literature because race data that includes gait pattern or audio cues is not available easily.

4.1. Multi-view Fused Gait based Ethnicity Classification

In [102], authors have proposed a method that identifies ethnicity from seven gait patterns captured from seven different

Table 1: Parametric Evaluation of Single Model Racial Categorization Methods

Method	Dataset used	Racial/ethnic class considered	Region of interest	Feature extraction Operator/s	Classifier/s used
Multi Ethnical Categorization using Manifold Learning [51]	Self-Generated	Chinese (8-Subgroups)	Face	-	Bayesian Net, Naive Bayesian, J4.8, RBF (Radial Basis Function) Network, LibSVM
GWT and Ratina Sampling based Ethnicity Categorization [52]	HOIP Database + Self-Generated	Asian, European, African	Eyes, Mouth	GWT, Ratina Sampling	SVM
Real Time Race Categorization [53]	FERET	Asian, Non-Asian	Face	ICA, PCA	SVM
Binary Tree based SVM for Ethnicity Detection [54]	FERET	Caucasian, African, Asian	Face	Gabor Filter, HOG	SVM, Kernel PCA
Racial Categorization using CNN [55]	Self-Generated	Bangladeshi, Chinese, Indian	Face	-	SVM
Neural Network based Racial Categorization [56]	FERET	Mongolian, Caucasian, Negro	Skin Colour, Forehead Area	-	ANN, CNN
Local Circular Pattern for Race Identification [61]	FRGC v2.0 and BU-3DFE	Whites and East-Asians	Face	LCP	Adaboost
Biometric based Machine Learning Method [62]	FERET	Asian, White, Black or African American, Hispanic, Pacific Islander, Native American	Eyes, Eyebrows, Periocular Region	LBP, HOG	SVM, MLP, LDA, QDA
Diffusion Model (Binary Classifier) [63]	Race Morph Sequence	Asian, Caucasian	Face	Manually	Manually
Implicit Racial Attitude [64]	Facial stimuli used in current research	African American, Caucasian	Skin Colour, Facial Physiognomy	Manually	Manually

angles. Figure 9 shows the key steps involved in the ethnicity identification process. First, all seven gait patterns are converted into corresponding Gait Energy Image (GEI). Next, seven GEIs are fused using three different fusion methods: score fusion, feature fusion and decision fusion. The goal of using three fusion methods is to accurately identify the ethnicity (race) of the person. Subsequently, features are extracted from the fused image using Multi-linear Principal Component Analysis (MPCA) and fed to classifier for ethnicity classification.

4.2. Hierarchical Fusion for Ethnicity Identification

In this method [65], gait pattern and facial features are fused for better ethnicity categorization. Figure 10 shows the two level processing involved in this method. First level involves gait pattern evolution. It takes gait video as an input. It includes the intermediate steps such as gait cycle estimation and GEI

generation. First level also includes SVM for classification. Second level takes face video as an input. It comprises of three

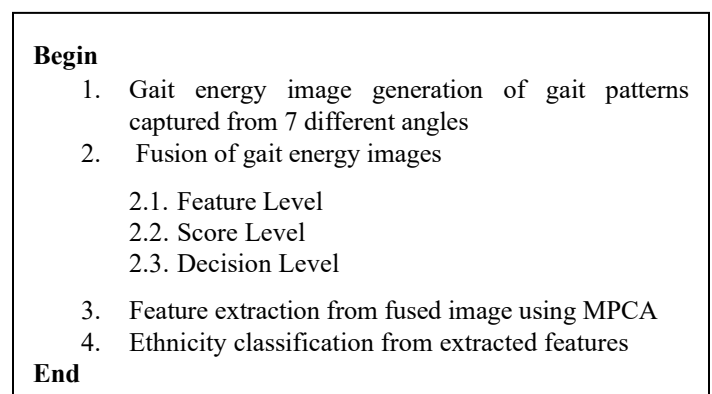


Figure 9: Steps involved in multi-view fused gait based ethnicity classification

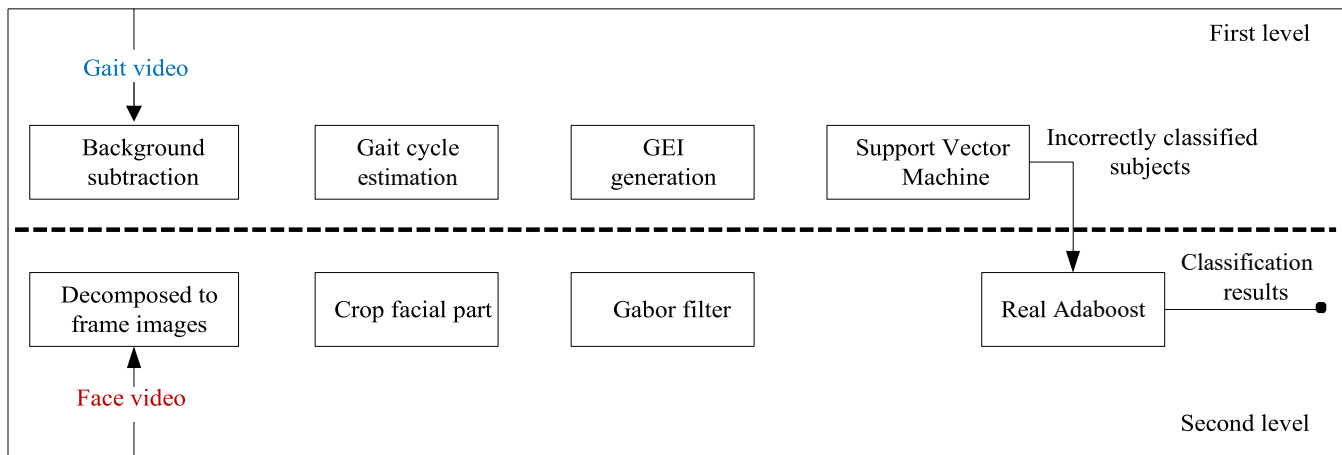


Figure 10: Steps of ethnicity categorization using hierarchical fusion system [65]

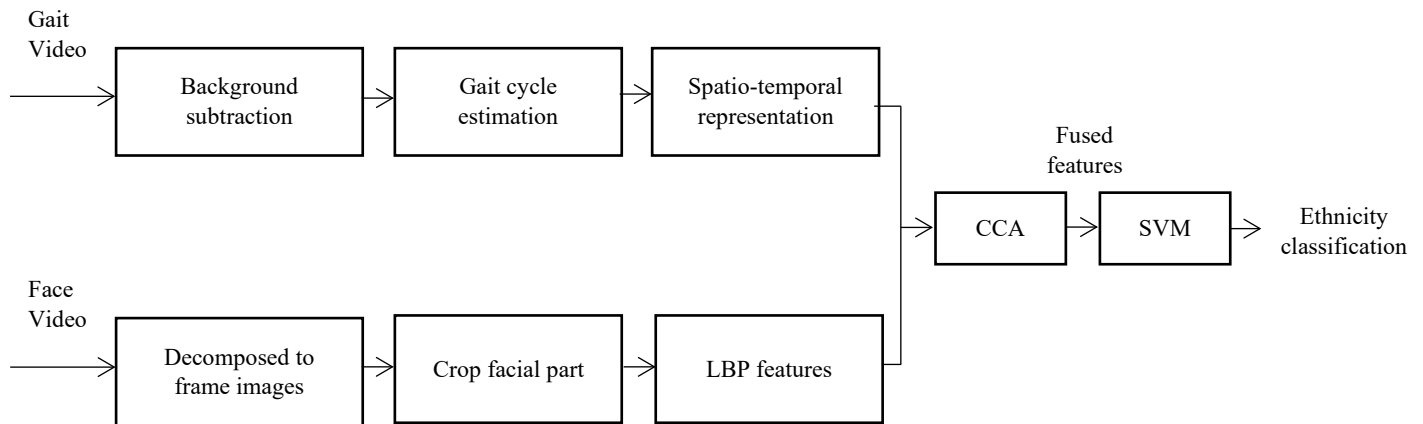


Figure 11: Block diagram of ethnicity classification system [67]

major steps: frame extraction from video, face detection and feature extraction using Gabor filter. Features extracted in first and second level are fused together to get accurate classification. The fused features are given to SVM and then to Adaboost to identify ethnicity.

4.3. Dialogue based Biometric Information Classification

In [66], a method for biometric information classification and deception detection has been proposed. It identifies gender, personality and ethnicity from the audio (dialogue). Lexical and acoustic-prosodic features are extracted from the dialogue. Lexical features are extracted using Linguistic Inquiry and Word Count (LIWC). Acoustic-prosodic features are extracted using Praat. Both types of features are given to different machine learning classifiers such as SVM, logistic regression, Adaboost and random forest for classification.

4.4. Cross-Model Biometric Matching

A new method for cross biometric matching by fusion of voice and facial image is introduced in [68]. The cross model is used for inferring the two types of information: 1) voice from human face and 2) human face from voice. This method involves two key steps: feature extraction and classification. The features are extracted from image as well as voice. The extracted features are fused and given as input to CNN for biometric matching and classification.

4.5. Gait and Face Fusion for Ethnicity Classification

Ethnicity classification system is proposed in [67]. It considers fusion of facial features and gait pattern for better classification. As shown in Figure 11, inputs to this system are gait video and facial video. Both videos are processed in parallel. During gait video processing, first background is subtracted from the video and subsequently each gait cycle pattern is estimated. Next all gait cycle patterns are represented using spatio-temporal representation for gait pattern characterization. During face video processing, frames are extracted from the face video and subsequently the facial part is cropped from the face image. Facial features are extracted from each frame using LBP. Features extracted from gait and face videos are fused together using Canonical Correlation Analysis (CCA). Fused features are given as input to SVM for ethnicity identification.

4.6. Performance Evaluation of Multi Model Racial Categorization Methods

As illustrated in Table 2, we identified the same set of parameters for comparison of multi model racial categorization methods as we identified for single model racial categorization methods. As defined and discussed previously in section 3, they are dataset used, biometric information considered for classification, region of interest, feature extraction operator/s and classifiers used in the method.

Table 2: Parametric Evaluation of Multi Model based Racial Categorization Methods

Method	Dataset used	Biometric information considered	Region of interest	Feature extraction operator/s	Classifier/s used
Multi-view Fused Gait based Ethnicity Classification [102]	Self- Generated	Ethnicity (East-Asian and South-American)	Gait Pattern	Multi-linear Principal Component Analysis	-
Hierarchical Fusion for Ethnicity Identification [65]	Self- Generated	Ethnicity (East-Asian and South-American)	Gait Pattern and Face	Gabor Filter	SVM
Dialogue based Classification [66]	NEO-FFI	Gender, Ethnicity and Personality	Dialogue (Audio)	LIWC and Praat	SVM, Logistic Regression, AdaBoost and Random Forest
Cross-Modal Biometric Matching [68]	VGGFace (Face data) and VoxCeleb (Audio Data)	Gender, Age, Ethnicity and Identity	Audio and Face	-	CNN
Gait and Face Fusion for Ethnicity Classification [67]	Self- Generated	Ethnicity	Gait Pattern and Face	LBP	SVM

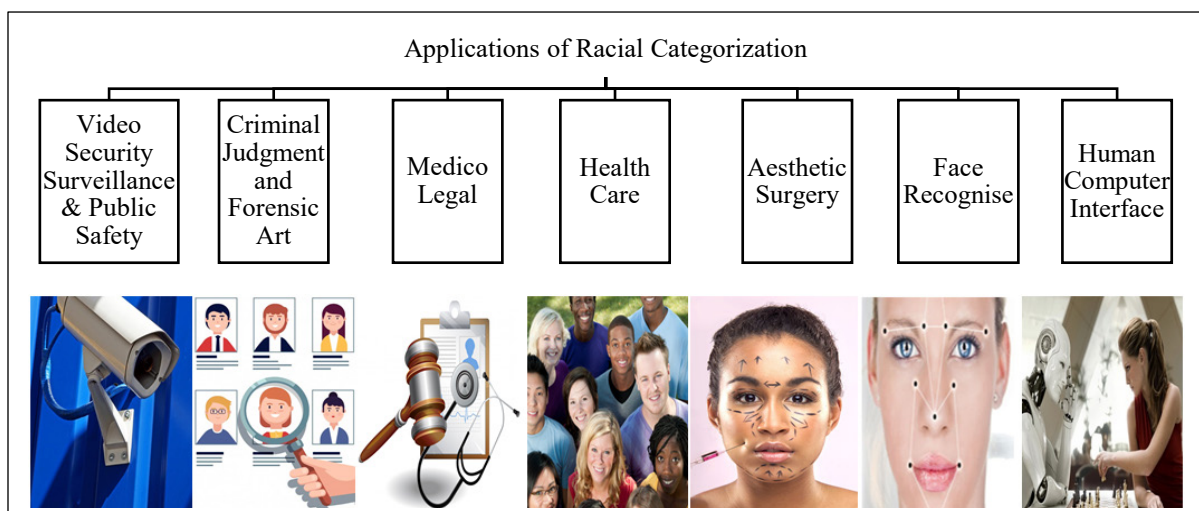


Figure 12: Application areas of racial categorization

5. Application of Racial Categorization

Racial categorization has a high impact on our social life. Race defines common physical characteristics of humans to represent his existence. Physical characteristics of humans of different races differ from each other. Racial categorization is significant for several applications. As shown in Figure 12, the major application areas are video security surveillance and public safety, criminal judgment and forensic art, medico legal, healthcare, aesthetic surgery, face recognition, and human computer interface.

5.1. Video Security Surveillance and Public Safety

Race identification from the subject’s face plays a crucial role in video security surveillance. Video security surveillance system assists in identifying criminals by comparing the detected subject’s image with the existing criminal database. Automated

race identification system fused with video security surveillance system provides quick information about the subject [53]. Such a fused system is already in use at several airports and public places. Moreover, it has been proven useful for applications such as maritime, aviation, mass transformation, government office building, recreational centers, stadium and large retail malls.

5.2. Criminal Judgment and Forensic Art

Crime related investigation requires crucial information related to criminals including cross-country evidence (if any) [103-108]. Race/ethnicity of criminals provides such crucial information. Face is typically considered for criminal investigation because face conveys important information. Particularly, it conveys age, race and gender that are needed for criminal investigation. This information makes the investigation process easy for the government to find the right criminal [104,

109-111]. Moreover, such information is useful to prevent innocent people and provides justice to minority community groups.

Normally the forensic department has the subject's image captured using a public camera. However, it is difficult for the forensic department to manually extract the crucial information from the image. Conversely, racial categorization method can be used to identify the race from the image which assists in further investigation targeting a particular race community [106].

5.3. Medico Legal

Medico legal case is defined as a case of suffering or injury in which examination by the police is essential to determine the cause of suffering or injury. Suffering or injury may be due to several unnatural conditions such as accidents, burning and death. Race provides patient's information that is useful to law enforcing agency for further investigation [112]. By evaluating the race information of the medico legal case, law enforcing agency can obtain history of medico legal cases in that particular racial group [113-115]. Such information eases the investigation process and assists medico legal department in decision making.

5.4. Healthcare

Disease and healthcare issues are conflicting for different geographical areas due to their weather conditions, living sense and food. Healthcare treatment differs for different racial groups [116]. Thus, racial categorization is useful to solve the healthcare issues and to provide quick treatment [117-121]. Moreover, ethnic information is useful to provide appropriate services and special advantages to minority ethnic groups which are defined by the government for the minority and economically low conditions [122].

The center to Eliminate Health Disparities (CEHD) of the University of Texas Medical Branch (UTMB) has implemented the Information System for the health of people of UTMB to

reduce disparities in health. Their information system is also known as REAL (race, ethnicity and language) [122]. Figure 13 illustrates the role of CEHD in the health system of UTMB and Galveston County as a whole. UTMB is a university health center that welcomes patients from diverse backgrounds. It provides services to different racial groups whose income level is below the poverty line. However, the main objectives of this REAL project are (1) to improve the UTMB's health information system for better diagnostics and stratified quality measures by race, ethnicity, language and status (2) to develop and disseminate contingency plans to address disparities through effective partnerships with relevant stakeholders.

5.5. Aesthetic Surgery

Aesthetic surgery is described as a facial plastic surgery either for the beautification of face or to create an attractive face [123]. Anthropometric measurement is the distance between two facial points. It has been revealed in literature that anthropological measurements such as ratio, geometric distance and Euclidean distance are different for different racial groups [124-126]. Geometric and Euclidean distances are the distances between primary or secondary facial landmarks on the frontal face/profile face. Depending on the race of patient, anthropological measurements are derived and used in aesthetic surgery [127-129]. For example, aesthetic surgery for Chinese people and Indian people is different as both groups have different facial features and thereby different anthropological measurements.

5.6. Face Recognition

Racial difference in humans is useful for biometric illustration and human identification [130]. Race cues and race wise anthropological measurements make it easy for face recognition systems to recognize the person [131-134]. Moreover, integration of race information with face recognition makes the face recognition system more intelligent and quick for accurate face recognition [135-136].

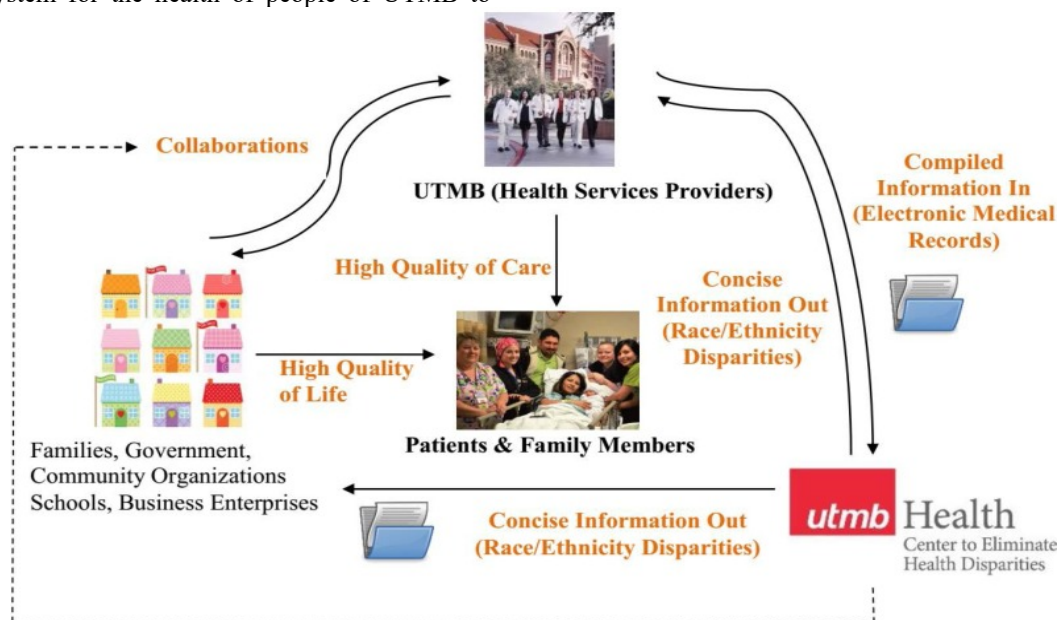


Figure 13: Role of the health center system to eliminate health disparities for different ethnicity [122]

5.7. Human Computer Interface

Nowadays, several systems are automated using robots [137-140]. Consumers of such robotic systems need to interact with robots frequently. In human-robot communication, racial cues play an important role [141-145]. Specifically, by recognizing the race of a human from his face, behavior and expression, robots can deliver the relevant services to humans. Such robotic systems are useful for an open service atmosphere where robots work as humans [146]. In particular, they are useful in hospitals, malls, stadiums, hotels, gaming zones and intelligent HCI organizations for easy communication with humans.

6. Challenges in Racial Categorization

Several challenges are faced to get correct and accurate racial categorization. Below we mention the major challenges faced in racial categorization. These challenges create new opportunities for researchers in this field to carry out further research. An efficient racial categorization method can be developed by overcoming these challenges and higher classification accuracy can be achieved.

6.1. Intra-race Categorization

To the best of our knowledge, intra-race categorization has not been focused much in literature. 85% of the worldwide population is divided into major 7-racial groups, namely African-American, South-Asian, East-Asian, Caucasian, Indian, Arabian and Latino race [30]. Intra-race categorization for aforementioned racial groups is challenging due to severe similarity in facial features and in physical appearance of humans belonging to a particular group [147-148]. It is difficult to infer different clues for intra-race categorization.

6.2. Anthropometry Measurements

Facial landmarking technique is used to measure anthropometry measurements. It has been shown in literature that accuracy of anthropometry measurements and thereby accuracy of racial categorization method varies with respect to the number of facial landmarks [21, 51, 56, 62, 147-148]. Thus, existing landmarking methods can be further improved by increasing the number of landmarks. Moreover, landmarks on forehead area, hairline and earlobe can be additionally considered to increase accuracy further [56]. In addition, anatomists have revealed that Ear pinna and Iannarelli's measures differ for different racial groups. Like finger print, ear pinna is unique for each individual.

6.3. Real-Time Data

It is required to process real-time video streams at public places such as airports, hospitals, health care centers, malls and stadiums for public safety and security systems. However, existing racial categorization methods are not applicable and reliable for processing real-time video stream [53]. They are applicable to only off-line image dataset.

6.4. Manual Racial Categorization

The aforementioned issues are related to automated racial categorization methods. However, the issues faced by manual racial categorization methods are different. The major issue

related to manual racial categorization which involves participants is that the number of stimulus levels for race prediction is limited. Stimulus level is defined as the number of tasks performed by participants for race prediction [56, 61].

7. Conclusion

In this paper, we have presented in-depth review on various single model and multi model racial categorization methods. Moreover, parametric evaluation of racial categorization methods based on identified set of parameters is presented. It has been observed that fusion of facial features and physical appearance provides accurate race categorization. Moreover, it has also been observed that CNN based racial categorization model gives substantially higher accuracy because it extracts deep features. Our rigorous review on racial categorization methods will provide researchers state-of-the-art advancements related to racial categorization methods. Furthermore, the applications and challenges of racial categorization discussed herein will help researchers to develop an efficient and competent racial categorization method.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] S. Gutta, H. Wechsler and P. Phillips, "Gender and ethnic classification of face images", *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*. <https://doi.org/10.1109/afgr.1998.670948>.
- [2] B. Xia, B. Ben Amor and M. Daoudi, "Joint gender, ethnicity and age estimation from 3D faces", *Image and Vision Computing*, vol. 64, pp. 90-102, 2017. <https://doi.org/10.1016/j.imavis.2017.06.004>.
- [3] S. Azzakhni, L. Ballihi and D. Aboutajdine, "Combining Facial Parts For Learning Gender, Ethnicity, and Emotional State Based on RGB-D Information", *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 14, no. 1, pp. 1-14, 2018. <https://doi.org/10.1145/3152125>.
- [4] P. Mukudi and P. Hills, "The combined influence of the own-age, -gender, and -ethnicity biases on face recognition", *Acta Psychologica*, vol. 194, pp. 1-6, 2019. <https://doi.org/10.1016/j.actpsy.2019.01.009>.
- [5] Xiaoli Zhou and B. Bhanu, "Integrating Face and Gait for Human Recognition", 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06). <https://doi.org/10.1109/CVPRW.2006.103>.
- [6] J. Yoo, D. Hwang and M. Nixon, "Gender Classification in Human Gait Using Support Vector Machine", *Advanced Concepts for Intelligent Vision Systems*, pp. 138-145, 2005. https://doi.org/10.1007/11558484_18.
- [7] Ju Han and B. Bhanu, "Statistical feature fusion for gait-based human recognition", *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004. CVPR 2004. <https://doi.org/10.1109/CVPR.2004.1315252>.
- [8] O. Çeliktutan, S. Ulukaya and B. Sankur, "A comparative study of face landmarking methods," *EURASIP Journal on Image and Video Processing*, vol. 2013, no. 1, 2013. <https://doi.org/10.1186/1687-5281-2013-13>.
- [9] Y. Wu and Q. Ji, "Facial Landmark Detection: A Literature Survey," *International Journal of Computer Vision*, 2018. <https://doi.org/10.1007/s11263-018-1097-z>.
- [10] H. Akakin and B. Sankur, "Analysis of Head and Facial Gestures Using Facial Landmark Trajectories," *Biometric ID Management and Multimodal Communication*, pp. 105-113, 2009. https://doi.org/10.1007/978-3-642-04391-8_14.
- [11] L. Farkas, M. Katic and C. Forrest, "International Anthropometric Study of Facial Morphology in Various Ethnic Groups/Races," *Journal of Craniofacial Surgery*, vol. 16, no. 4, pp. 615-646, 2005. <https://doi.org/10.1097/01.scs.0000171847.58031.9e>.

- [12] H. Dibeklioglu, A. Salah and T. Gevers, "A Statistical Method for 2-D Facial Landmarking," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 844-858, 2012. <https://doi.org/10.1109/TIP.2011.2163162>.
- [13] B. Martinez, M. Valstar, X. Binefa and M. Pantic, "Local Evidence Aggregation for Regression-Based Facial Point Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 5, pp. 1149-1163, 2013. <https://doi.org/10.1109/TPAMI.2012.205>.
- [14] M. Taner Eskil and K. Benli, "Facial expression recognition based on anatomy," *Computer Vision and Image Understanding*, vol. 119, pp. 1-14, 2014. <https://doi.org/10.1016/j.cviu.2013.11.002>.
- [15] E. Lloyd, K. Hugenberg, A. McConnell, J. Kunstman and J. Deska, "Black and White Lies: Race-Based Biases in Deception Judgments", *Psychological Science*, vol. 28, no. 8, pp. 1125-1136, 2017. <https://doi.org/10.1177/0956797617705399>.
- [16] G. Breetzke, "The concentration of urban crime in space by race: evidence from South Africa", *Urban Geography*, vol. 39, no. 8, pp. 1195-1220, 2018. <https://doi.org/10.1080/02723638.2018.1440127>.
- [17] W. Threadcraft-Walker and H. Henderson, "Reflections on race, personality, and crime", *Journal of Criminal Justice*, vol. 59, pp. 38-41, 2018. <https://doi.org/10.1016/j.jcrimjus.2018.05.005>.
- [18] C. Lehman, T. Hammond and G. Agyemang, "Accounting for crime in the US: Race, class and the spectacle of fear", *Critical Perspectives on Accounting*, vol. 56, pp. 63-75, 2018. <https://doi.org/10.1016/j.cpa.2018.01.002>.
- [19] B. Jefferson, "Predictable Policing: Predictive Crime Mapping and Geographies of Policing and Race", *Annals of the American Association of Geographers*, vol. 108, no. 1, pp. 1-16, 2017. <https://doi.org/10.1080/24694452.2017.1293500>.
- [20] C. Petsko and G. Bodenhausen, "Race-Crime Congruency Effects Revisited: Do We Take Defendants' Sexual Orientation into Account?," *Social Psychological and Personality Science*, vol. 10, no. 1, pp. 73-81, 2017. <https://doi.org/10.1177/1948550617736111>.
- [21] M. Wai et al., "Nasofacial Anthropometric Study among University Students of Three Races in Malaysia", 2019. <https://doi.org/10.1155/2015/780756>.
- [22] E. Rata and C. Zubaran, "Ethnic Classification in the New Zealand Health Care System", *Journal of Medicine and Philosophy*, vol. 41, no. 2, pp. 192-209, 2016. <https://doi.org/10.1093/jmp/jhv065>.
- [23] J. Reid, D. Cormack and M. Crowe, "The significance of socially-assigned ethnicity for self-identified Māori accessing and engaging with primary healthcare in New Zealand", *Health: An Interdisciplinary Journal for the Social Study of Health, Illness and Medicine*, vol. 20, no. 2, pp. 143-160, 2015. <https://doi.org/10.1177/1363459315568918>.
- [24] S. Mehta et al., "Race/Ethnicity and Adoption of a Population Health Management Approach to Colorectal Cancer Screening in a Community-Based Healthcare System", *Journal of General Internal Medicine*, vol. 31, no. 11, pp. 1323-1330, 2016. <https://doi.org/10.1007/s11606-016-3792-1>.
- [25] L. Culley, "Transcending transculturalism? Race, ethnicity and health-care", *Nursing Inquiry*, vol. 13, no. 2, pp. 144-153, 2006. <https://doi.org/10.1111/j.1440-1800.2006.00311.x>.
- [26] M. Gibbons, "Use of Health Information Technology among Racial and Ethnic Underserved Communities," *Perspect Health Inf Manag*, vol. 8, 2011.
- [27] A. Memarian, S. Mostafavi, E. Zarei, S. Esfahani and B. Ghorbanzadeh, "The Credibility of Cephalogram Parameters in Gender Identification From Medico-Legal Relevance Among the Iranian Populatio," *International Journal of Medical Toxicology and Forensic Medicine*, vil. 9, 2019.
- [28] M. Ibrahim, A. Khalifa, H. Hassan, H. Tamam and A. Hagra, "Estimation of stature from hand dimensions in North Saudi population, medicolegal view," *The Saudi Journal of Forensic Medicine and Sciences*, vol. 1, pp. 19-27, 2018. https://doi.org/10.4103/sjfm.sjfm.10_17.
- [29] B. Bukovitz, J. Meiman, H. Anderson and E. Brooks, "Silicosis: Diagnosis and Medicolegal Implications", *Journal of Forensic Sciences*, 2019. <https://doi.org/10.1111/1556-4029.14048>.
- [30] S. Fu, H. He and Z. Hou, "Learning Race from Face: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 12, pp. 2483-2509, 2014. <https://doi.org/10.1109/TPAMI.2014.2321570>.
- [31] J. Carroll, "Evaluation, Description and Invention: Paradigms for Human-Computer Interaction", *Advances in Computers*, pp. 47-77, 1989. [https://doi.org/10.1016/S0065-2458\(08\)60532-X](https://doi.org/10.1016/S0065-2458(08)60532-X).
- [32] M. Mackal, "The effects of ethnicity and message content on affective outcomes in a computer-based learning environment," 2009.
- [33] V. Evers, "Human - Computer Interfaces: Designing for Culture," 1997.
- [34] B. Marsh, K. Pezdek and D. Ozery, "The cross-race effect in face recognition memory by bicultural individuals", *Acta Psychologica*, vol. 169, pp. 38-44, 2016. <https://doi.org/10.1016/j.actpsy.2016.05.003>.
- [35] L. Farkas, M. Katic and C. Forrest, "International Anthropometric Study of Facial Morphology in Various Ethnic Groups/Races", *Journal of Craniofacial Surgery*, vol. 16, no. 4, pp. 615-646, 2005. <https://doi.org/10.1097/01.scs.0000171847.58031.9e>.
- [36] S. Ozdemir, D. Sigirli, I. Ercan and N. Cankur, "Photographic Facial Soft Tissue Analysis of Healthy Turkish Young Adults: Anthropometric Measurements", *Aesthetic Plastic Surgery*, vol. 33, no. 2, pp. 175-184, 2008. <https://doi.org/10.1007/s00266-008-9274-z>.
- [37] E. Boutellaa, A. Hadid, M. Bengherabi and S. Ait-Aoudia, "On the use of Kinect depth data for identity, gender and ethnicity classification from facial images", *Pattern Recognition Letters*, vol. 68, pp. 270-277, 2015. <https://doi.org/10.1016/j.patrec.2015.06.027>.
- [38] N. Narang and T. Bourlai, "Gender and ethnicity classification using deep learning in heterogeneous face recognition", *2016 International Conference on Biometrics (ICB)*, 2016. <https://doi.org/10.1109/ICB.2016.7550082>.
- [39] Y. Wang, X. Duan, X. Liu, C. Wang and S. Li, "Semantic description method for face features of larger Chinese ethnic groups based on improved WM method", *Neurocomputing*, vol. 175, pp. 515-528, 2016. <https://doi.org/10.1016/j.neucom.2015.10.089>.
- [40] L. Zhao and S. Bentin, "Own- and other-race categorization of faces by race, gender, and age", *Psychonomic Bulletin & Review*, vol. 15, no. 6, pp. 1093-1099, 2008. <https://doi.org/10.3758/PBR.15.6.1093>.
- [41] X. Lu, H. Chen and A. Jain, "Multimodal Facial Gender and Ethnicity Identification", *Advances in Biometrics*, pp. 554-561, 2005. https://doi.org/10.1007/11608288_74.
- [42] G. Guo and G. Mu, "A framework for joint estimation of age, gender and ethnicity on a large database", *Image and Vision Computing*, vol. 32, no. 10, pp. 761-770, 2014. <https://doi.org/10.1016/j.imavis.2014.04.011>.
- [43] A. da Silva, L. Magri, L. Andrade and M. da Silva, "Three-dimensional analysis of facial morphology in Brazilian population with Caucasian, Asian, and Black ethnicity", *Journal of Oral Research and Review*, vol. 9, no. 1, p. 1, 2017. <http://dx.doi.org/10.1590/0103-6440201802027>.
- [44] T. Vo, T. Nguyen and C. Le, "Race Recognition Using Deep Convolutional Neural Networks", *Symmetry*, vol. 10, no. 11, p. 564, 2018. <https://doi.org/10.3390/sym10110564>.
- [45] D. Bobeldyk and A. Ross, "Analyzing Covariate Influence on Gender and Race Prediction From Near-Infrared Ocular Images", *IEEE Access*, vol. 7, pp. 7905-7919, 2019. <https://doi.org/10.1109/ACCESS.2018.2886275>.
- [46] S. Samangooei and M. Nixon, "Performing content-based retrieval of humans using gait biometrics", *Multimedia Tools and Applications*, vol. 49, no. 1, pp. 195-212, 2009. <https://doi.org/10.1007/s11042-009-0391-8>.
- [47] A. Kale, A. Roychowdhury and R. Chellappa, "Fusion of gait and face for human identification", *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*. <https://doi.org/10.1109/ICASSP.2004.1327257>.
- [48] S. Sarkar, P. Phillips, Z. Liu, I. Vega, P. Grother and K. Bowyer, "The humanID gait challenge problem: data sets, performance, and analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162-177, 2005. <https://doi.org/10.1109/TPAMI.2005.39>.
- [49] Ju Han and Bir Bhanu, "Individual recognition using gait energy image", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 316-322, 2006. <https://doi.org/10.1109/TPAMI.2006.38>.
- [50] J. Gomez-Ezeiza, J. Torres-Unda, N. Tam, J. Irazusta, C. Granados and J. Santos-Concejero, "Race walking gait and its influence on race walking economy in world-class race walkers", *Journal of Sports Sciences*, vol. 36, no. 19, pp. 2235-2241, 2018. <https://doi.org/10.1080/02640414.2018.1449086>.
- [51] C. Wang, Q. Zhang, X. Duan and J. Gan, "Multi-ethnic Chinese facial characterization and analysis," *Multimedia Tools and Applications*, 2018. <https://doi.org/10.1007/s11042-018-6018-1>.
- [52] S. Hosoi, E. Takikawa and M. Kawade, "Ethnicity estimation with facial images," *Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, 2004. <https://doi.org/10.1109/AFGR.2004.1301530>.
- [53] O. Yongsheng, W. Xinyu, Q. Huihuan and Yangsheng Xu, "A real time race categorization system," *IEEE International Conference on Information Acquisition*, 2005.
- [54] M. Uddin and S. Chowdhury, "An integrated approach to classify gender and ethnicity," *2016 International Conference on Innovations in Science, Engineering and Technology (ICISSET)*, 2016. <https://doi.org/10.1109/ICISSET.2016.7856480>.

- [55] Z. Heng, M. Dipu and K. Yap, "Hybrid Supervised Deep Learning for Ethnicity Categorization using Face Images," 2018 IEEE International Symposium on Circuits and Systems (ISCAS), 2018.
- [56] S. Masood, S. Gupta, A. Wajid, S. Gupta and M. Ahmed, "Prediction of Human Ethnicity from Facial Images Using Neural Networks," *Advances in Intelligent Systems and Computing*, pp. 217-226, 2017. https://doi.org/10.1007/978-981-10-3223-3_20.
- [57] A. Das, A. Dantcheva and F. Bremond, "Mitigating Bias in Gender, Age and Ethnicity Classification: a Multi-Task Convolution Neural Network Approach," *ECCV*, 2018. <https://link.springer.com/conference/eccv>.
- [58] N. Srinivas, H. Atwal, D. Rose, G. Mahalingam, K. Ricanek and D. Bolme, "Age, Gender, and Fine-Grained Ethnicity Prediction Using Convolutional Neural Networks for the East Asian Face Dataset", 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), 2017. <https://doi.org/10.1109/FG.2017.118>.
- [59] A. Mohammad and J. Al-Ani, "Convolutional Neural Network for Ethnicity Classification using Ocular Region in Mobile Environment", 2018 10th Computer Science and Electronic Engineering (CEEC), 2018. <https://doi.org/10.1109/CEEC.2018.8674194>.
- [60] C. Bagchi, D. Geraldine Bessie Amali and M. Dinakaran, "Accurate Facial Ethnicity Classification Using Artificial Neural Networks Trained with Galactic Swarm Optimization Algorithm", *Advances in Intelligent Systems and Computing*, pp. 123-132, 2018. https://doi.org/10.1007/978-981-13-3329-3_12.
- [61] D. Huang, H. Ding, C. Wang, Y. Wang, G. Zhang and L. Chen, "Local circular patterns for multi-modal facial gender and ethnicity categorization," 2018. <https://doi.org/10.1016/j.imavis.2014.06.009>.
- [62] A. Mohammad and J. Al-Ani, "Towards ethnicity detection using learning based classifiers," 2017 9th Computer Science and Electronic Engineering (CEEC), 2017. <https://doi.org/10.1109/CEEC.2017.8101628>.
- [63] C. Benton and A. Skinner, "Deciding on race: A diffusion model analysis of race-categorisation," 2018. <https://doi.org/10.1016/j.cognition.2015.02.011>.
- [64] E. Stepanova and M. Strube, "The role of skin color and facial physiognomy in racial categorization: Moderation by implicit racial attitudes," *Journal of Experimental Social Psychology*, vol. 48, no. 4, pp. 867-878, 2012. <https://doi.org/10.1016/j.jesp.2012.02.019>.
- [65] D. Zhang, Y. Wang and Z. Zhang, "Ethnicity Classification Based on a Hierarchical Fusion", *Biometric Recognition*, pp. 300-307, 2012. https://doi.org/10.1007/978-3-642-35136-5_36.
- [66] S. Levitan, Y. Levitan, G. An, M. Levine, R. Levitan, A. Rosenberg and J. Hirschberg, "Identifying Individual Differences in Gender, Ethnicity, and Personality from Dialogue for Deception Detection," *NAACL-HLT*, pp. 40-44, 2016.
- [67] D. Zhang, Y. Wang, Z. Zhang and M. Hu, "Ethnicity classification based on fusion of face and gait", 2012 5th IAPR International Conference on Biometrics (ICB), 2012. <https://doi.org/10.1109/ICB.2012.6199781>.
- [68] A. Nagrani, S. Albanie and A. Zisserman, "Seeing Voices and Hearing Faces: Cross-Modal Biometric Matching," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [69] Hengliang Tang, Yanfeng Sun, Baocai Yin and Yun Ge, "Face recognition based on Haar LBP histogram", 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), 2010. <https://doi.org/10.1109/ICACTE.2010.5579370>.
- [70] R. Sáenz and M. Morales, "6 Demography of Race and Ethnicity", *Handbooks of Sociology and Social Research*, pp. 163-207, 2019. https://doi.org/10.1007/978-3-030-10910-3_7.
- [71] H. Qiu, S. Fang and K. Song, "Method and apparatus for face classification," *Google patent*, 2018.
- [72] C. Lv, Z. Wu, D. Zhang, X. Wang and M. Zhou, "3D Nose shape net for human gender and ethnicity classification", *Pattern Recognition Letters*, 2018. <https://doi.org/10.1016/j.patrec.2018.11.010>.
- [73] M. Singh, S. Nagpal, M. Vatsa, R. Singh, A. Noore and A. Majumdar, "Gender and ethnicity classification of Iris images using deep classifier", 2017 IEEE International Joint Conference on Biometrics (IJCB), 2017.
- [74] H. Jang, J. Yoon, Y. Kim and Y. Park, "Classification of Iris Colors and Patterns in Koreans", *Healthcare Informatics Research*, vol. 24, no. 3, p. 227, 2018. <https://doi.org/10.4258/hir.2018.24.3.227>.
- [75] G. Mabuza-Hocquet, F. Nelwamondo and T. Marwala, "Ethnicity Distinctiveness Through Iris Texture Features Using Gabor Filters", *Intelligent Information and Database Systems*, pp. 551-560, 2017. https://doi.org/10.1007/978-3-319-54430-4_53.
- [76] D. Bobeldyk and A. Ross, "Analyzing Covariate Influence on Gender and Race Prediction From Near-Infrared Ocular Images", *IEEE Access*, vol. 7, pp. 7905-7919, 2019. <https://doi.org/10.1109/ACCESS.2018.2886275>.
- [77] H. Chen, M. Gao, K. Ricanek, W. Xu and B. Fang, "A Novel Race Classification Method Based on Periocular Features Fusion", *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 31, no. 08, p. 1750026, 2017. <https://doi.org/10.1142/S0218001417500264>.
- [78] S. Arseev, A. Konushin and V. Liutov, "Human Recognition by Appearance and Gait", *Programming and Computer Software*, vol. 44, no. 4, pp. 258-265, 2018. <https://doi.org/10.1134/S0361768818040035>.
- [79] W. Chi, J. Wang and M. Meng, "A Gait Recognition Method for Human Following in Service Robots", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 9, pp. 1429-1440, 2018. <https://doi.org/10.1109/TSMC.2017.2660547>.
- [80] I. Rida, A. Bouridane, G. Marcialis and P. Tuveri, "Improved Human Gait Recognition", *Image Analysis and Processing ICIAP 2015*, pp. 119-129, 2015. https://doi.org/10.1007/978-3-319-23234-8_12.
- [81] P. Shah and N. Davis, "Comparing Three Methods of Measuring Race/Ethnicity", *The Journal of Race, Ethnicity, and Politics*, vol. 2, no. 1, pp. 124-139, 2017. <https://doi.org/10.1017/rep.2016.27>.
- [82] R. Lee, B. Chon, S. Lin, M. He and S. Lin, "Association of Ocular Conditions With Narrow Angles in Different Ethnicities", *American Journal of Ophthalmology*, vol. 160, no. 3, pp. 506-515.e1, 2015. <https://doi.org/10.1016/j.ajo.2015.06.002>.
- [83] B. Latinwo, A. Falohun, E. Omidiora and B. Makinde, "Iris Texture Analysis for Ethnicity Classification Using Self-Organizing Feature Maps", *Journal of Advances in Mathematics and Computer Science*, vol. 25, no. 6, pp. 1-10, 2018. <https://doi.org/10.9734/JAMCS/2017/29634>.
- [84] G. Mabuza-Hocquet, F. Nelwamondo and T. Marwala, "Ethnicity Prediction and Classification from Iris Texture Patterns: A Survey on Recent Advances", 2016 International Conference on Computational Science and Computational Intelligence (CSCI), 2016. <https://doi.org/10.1109/CSCI.2016.0159>.
- [85] M. Singh, S. Nagpal, M. Vatsa, R. Singh, A. Noore and A. Majumdar, "Gender and ethnicity categorization of Iris images using deep classifier", 2017 IEEE International Joint Conference on Biometrics (IJCB), 2017.
- [86] S. Lagree and K. Bowyer, "Predicting ethnicity and gender from iris texture," 2011 IEEE International Conference on Technologies for Homeland Security (HST), 2011. <https://doi.org/10.1109/THS.2011.6107909>.
- [87] K. Huri, E. (Omid) David and N. Netanyahu, "DeepEthnic: Multi-label Ethnic Classification from Face Images", *Artificial Neural Networks and Machine Learning - ICANN 2018*, pp. 604-612, 2018. https://doi.org/10.1007/978-3-030-01424-7_59.
- [88] C. Wang, Q. Zhang, W. Liu, Y. Liu and L. Miao, "Facial feature discovery for ethnicity recognition", *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 9, no. 1, p. e1278, 2018. <https://doi.org/10.1002/widm.1278>.
- [89] M. Sajid et al., "Facial Asymmetry-Based Anthropometric Differences between Gender and Ethnicity", *Symmetry*, vol. 10, no. 7, p. 232, 2018. <https://doi.org/10.3390/sym10070232>.
- [90] G. Shakhnarovich, L. Lee and T. Darrell, "Integrated face and gait recognition from multiple views", *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. <https://doi.org/10.1109/CVPR.2001.990508>.
- [91] D. Zhang, Y. Wang and Z. Zhang, "Ethnicity Categorization Based on a Hierarchical Fusion," *Biometric Recognition*, pp. 300-307, 2012. https://doi.org/10.1007/978-3-642-35136-5_36.
- [92] A. Mecca, "Impact of Gait Stabilization: A Study on How to Exploit it for User Recognition", 2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), 2018. <https://doi.org/10.1109/SITIS.2018.00090>.
- [93] Y. Liu, K. Lu, S. Yan, M. Sun, D. Lester and K. Zhang, "Gait phase varies over velocities", *Gait & Posture*, vol. 39, no. 2, pp. 756-760, 2014. <https://doi.org/10.1016/j.gaitpost.2013.10.009>.
- [94] Y. Wang, S. Yu, Y. Wang and T. Tan, "Gait Recognition Based on Fusion of Multi-view Gait Sequences", *Advances in Biometrics*, pp. 605-611, 2005. https://doi.org/10.1007/11608288_80.
- [95] Xiayi Huang and N. Boulgouris, "Gait Recognition using Multiple Views", 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, 2008. <https://doi.org/10.1109/ICASSP.2008.4517957>.
- [96] A. Ross and A. Jain, "Information fusion in biometrics", *Pattern Recognition Letters*, vol. 24, no. 13, pp. 2115-2125, 2003. [https://doi.org/10.1016/S0167-8655\(03\)00079-5](https://doi.org/10.1016/S0167-8655(03)00079-5).

- [97] M. Deng and C. Wang, "Human gait recognition based on deterministic learning and data stream of Microsoft Kinect", *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1-1, 2018. <https://doi.org/10.1109/TCSVT.2018.2883449>.
- [98] L. Sturm, K. Donahue, M. Kasting, A. Kulkarni, N. Brewer and G. Zimet, "Pediatrician-Parent Conversations About Human Papillomavirus Vaccination: An Analysis of Audio Recordings," *Journal of Adolescent Health*, vol. 61, no. 2, pp. 246-251, 2017. <https://doi.org/10.1016/j.jadohealth.2017.02.006>.
- [99] A. Rehman, G. Khan, A. Siddiqi, A. Khan and U. Khan, "Modified Texture Features from Histogram and Gray Level Co-occurrence Matrix of Facial Data for Ethnicity Detection", 2018 5th International Multi-Topic ICT Conference (IMTIC), 2018. <https://doi.org/10.1109/IMTIC.2018.8467231>.
- [100] E. Loo, T. Lim, L. Ong and C. Lim, "The influence of ethnicity in facial gender estimation", 2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA), 2018. <https://doi.org/10.1109/CSPA.2018.8368710>.
- [101] Young, D., Sanchez, D. and Wilton, L. (2019). Biracial perception in black and white: How Black and White perceivers respond to phenotype and racial identity cues. <https://doi.org/10.1037/cdp0000103>.
- [102] D. Zhang, Y. Wang and B. Bhanu, "Ethnicity classification based on gait using multi-view fusion", 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 2010. <https://doi.org/10.1109/CVPRW.2010.5544614>.
- [103] L. Krivo, M. Véléz, C. Lyons, J. Phillips and E. Sabbath, "Race, Crime, And The Changing Fortunes of Urban Neighborhoods, 1999–2013", *Du Bois Review: Social Science Research on Race*, vol. 15, no. 1, pp. 47-68, 2018. <https://doi.org/10.1017/S1742058X18000103>.
- [104] W. Billy Huang, "Immigrant crime in Taiwan: perspectives from Eastern Asia", *Forensic Research & Criminology International Journal*, vol. 6, no. 3, 2018. <https://doi.org/10.15406/frcij.2018.06.00210>.
- [105] M. Hollis and R. Martínez, "Theoretical Approaches to the Study of Race, Ethnicity, Crime, and Criminal Justice", *The Handbook of Race, Ethnicity, Crime, and Justice*, pp. 203-207, 2018. <https://doi.org/10.1002/9781119113799.part2>.
- [106] S. Gabbidon and H. Greene, "Race and Crime," SAGE Publication, 2019.
- [107] S. Awaworyi Churchill and E. Laryea, "Crime and Ethnic Diversity: Cross-Country Evidence", *Crime & Delinquency*, vol. 65, no. 2, pp. 239-269, 2017. <https://doi.org/10.1177/0011128717732036>.
- [108] R. Martínez, M. Hollis and J. Stowell, "The Handbook of Race, Ethnicity, Crime, and Justice", 2018. <https://doi.org/10.1002/9781119113799.part1>.
- [109] A. Fernandes and R. Crutchfield, "Race, Crime, and Criminal Justice", *Criminology & Public Policy*, vol. 17, no. 2, pp. 397-417, 2018. <https://doi.org/10.1111/1745-9133.12361>.
- [110] R. Martínez and M. Hollis, "An Overview of Race, Ethnicity, Crime, and Justice", *The Handbook of Race, Ethnicity, Crime, and Justice*, pp. 11-16, 2018. <https://doi.org/10.1002/9781119113799.part1>.
- [111] J. Unnever, "Ethnicity and Crime in the Netherlands", *International Criminal Justice Review*, vol. 29, no. 2, pp. 187-204, 2018. <https://doi.org/10.1177/1057567717752218>.
- [112] T. A. LaVeist, "Beyond dummy variables and sample selection: what health services researchers ought to know about race as a variable," *Health Serv Res*, pp. 1-16, 1994.
- [113] T. Stewart, "Medico-legal aspects of the skeleton. I. Age, sex, race and stature," *American Journal of Physical Anthropology*, vol. 6, no. 3, pp. 315-322, 1948. <https://doi.org/10.1002/ajpa.1330060306>.
- [114] A. Dubey, S. Roy and S. Verma, "Significance of Sacral Index in Sex Determination and Its Comparative Study in Different Races", 2019. <https://doi.org/10.16965/ijar.2016.153>.
- [115] R. Kumar, "Forensic odontology: A medico legal guide for police personnel", *International Journal of Forensic Odontology*, vol. 2, no. 2, p. 80, 2017. https://doi.org/10.4103/ijfo.ijfo_13_17.
- [116] J. Betancourt, "Defining Cultural Competence: A Practical Framework for Addressing Racial/Ethnic Disparities in Health and Health Care", *Public Health Reports*, vol. 118, no. 4, pp. 293-302, 2003. <https://doi.org/10.1093/phr/118.4.293>.
- [117] J. Williams, R. Walker and L. Egede, "Achieving Equity in an Evolving Healthcare System: Opportunities and Challenges", *The American Journal of the Medical Sciences*, vol. 351, no. 1, pp. 33-43, 2016. <https://doi.org/10.1016/j.amjms.2015.10.012>.
- [118] C. Jones, B. Truman, L. Elam-Evans, "Using Socially Assigned Race to Probe White Advantages in Health Status," *Race, Ethnicity, and Health: A Public Health Reader*, 2013.
- [119] R. Harris, D. Cormack and J. Stanley, "The relationship between socially-assigned ethnicity, health and experience of racial discrimination for Māori: analysis of the 2006/07 New Zealand Health Survey", *BMC Public Health*, vol. 13, no. 1, 2013. <https://doi.org/10.1186/1471-2458-13-844>.
- [120] D. Cormack, R. Harris and J. Stanley, "Investigating the Relationship between Socially-Assigned Ethnicity, Racial Discrimination and Health Advantage in New Zealand", *PLoS ONE*, vol. 8, no. 12, p. e84039, 2013. <https://doi.org/10.1371/journal.pone.0084039>.
- [121] R. Harris, D. Cormack, J. Stanley and R. Rameka, "Investigating the Relationship between Ethnic Consciousness, Racial Discrimination and Self-Rated Health in New Zealand", *PLOS ONE*, vol. 10, no. 2, p. e0117343, 2015. <https://doi.org/10.1371/journal.pone.0117343>.
- [122] W. Lee, S. Veeranki, H. Serag, K. Eschbach and K. Smith, "Improving the Collection of Race, Ethnicity, and Language Data to Reduce Healthcare Disparities: A Case Study from an Academic Medical Center," *Perspect Health Inf Manag*, vol. 13, 2016.
- [123] R. Holliday, "Aesthetic surgery as false beauty", *Feminist Theory*, vol. 7, no. 2, pp. 179-195, 2006. <https://doi.org/10.1177/1464700106064418>.
- [124] T. A. Elsamny, A. N. Rabie, A. N. Abdelhamid and E. A. Sobhi, "Anthropometric Analysis of the External Nose of the Egyptian Males," *Aesthetic Plastic Surgery*, vol. 42, no. 5, pp. 1343-1356, 2018. <https://doi.org/10.1007/s00266-018-1197-8>.
- [125] I. Mohammed, T. Mokhtari, S. Ijaz, A. Omotosho, A. Ngaski, M. Milanifard and G. Hassanzadeh, "Anthropometric study of nasal index in Hausa ethnic population of northwestern Nigeria," *Journal of contemporary medical science*, vol. 4, 2018.
- [126] A. Sa, O. B, A. Gt, F. Op And I. Ag, "Anthropometric Characterization Of Nasal Parameters In Adults Oyemekun Ethnic Group In Akure Southwest Nigeria", *International Journal of Anatomy and Research*, vol. 6, no. 22, pp. 5272-5279, 2018. <https://dx.doi.org/10.16965/ijar.2018.178>.
- [127] S. Ozdemir, D. Sigirli, I. Ercan and N. Cankur, "Photographic Facial Soft Tissue Analysis of Healthy Turkish Young Adults: Anthropometric Measurements," *Aesthetic Plastic Surgery*, vol. 33, no. 2, pp. 175-184, 2008. <https://doi.org/10.1007/s00266-008-9274-z>.
- [128] E. Plemons, "Gender, Ethnicity, and Transgender Embodiment: Interrogating Classification in Facial Feminization Surgery", *Body & Society*, vol. 25, no. 1, pp. 3-28, 2018. <https://doi.org/10.1177/1357034X18812942>.
- [129] T. Elsamny, A. Rabie, A. Abdelhamid and E. Sobhi, "Anthropometric Analysis of the External Nose of the Egyptian Males", *Aesthetic Plastic Surgery*, vol. 42, no. 5, pp. 1343-1356, 2018. <https://doi.org/10.1007/s00266-018-1197-8>.
- [130] B. Marsh, K. Pezdek and D. Ozery, "The cross-race effect in face recognition memory by bicultural individuals," *Acta Psychologica*, vol. 169, pp. 38-44, 2016. <https://doi.org/10.1016/j.actpsy.2016.05.003>.
- [131] T. Tong, T. Key, J. Sobiecki and K. Bradbury, "Anthropometric and physiologic characteristics in white and British Indian vegetarians and nonvegetarians in the UK Biobank", *The American Journal of Clinical Nutrition*, vol. 107, no. 6, pp. 909-920, 2018. <https://doi.org/10.1093/ajcn/nqy042>.
- [132] N. Mehta and R. Srivastava, "The Indian nose: An anthropometric analysis", *Journal of Plastic, Reconstructive & Aesthetic Surgery*, vol. 70, no. 10, pp. 1472-1482, 2017. <https://doi.org/10.1016/j.bjps.2017.05.042>.
- [133] E. Ozsahn, E. Kızılkant, N. Boyan, R. Soames and O. Oguz, "Evaluation of Face Shape in Turkish Individuals," *Int. J. Morphol.*, 2016. <https://scielo.conicyt.cl/pdf/ijmorphol/v34n3/art15.pdf>.
- [134] D. Chettri, D. Sinha, D. Chakraborty and D. Jain, "Naso-facial anthropometric study of Female Sikkimese University Students", *IOSR Journal of Dental and Medical Sciences*, vol. 16, no. 03, pp. 49-54, 2017. <https://doi.org/10.9790/0853-1603074954>.
- [135] M. Ho and K. Pezdek, "Postencoding cognitive processes in the cross-race effect: Categorization and individuation during face recognition", *Psychonomic Bulletin & Review*, vol. 23, no. 3, pp. 771-780, 2015. <https://doi.org/10.3758/s13423-015-0945-x>.
- [136] J. Cavazos, E. Noyes and A. O'Toole, "Learning context and the other-race effect: Strategies for improving face recognition", *Vision Research*, 2018. <https://doi.org/10.1016/j.visres.2018.03.003>.
- [137] J. Anacleto and A. Carvalho, "Improving Human-Computer Interaction by Developing Culture-sensitive Applications based on Common Sense Knowledge," *Human-Computer Interaction*.
- [138] A. Cowell and K. Stanney, "Manipulation of non-verbal interaction style and demographic embodiment to increase anthropomorphic computer character credibility", *International Journal of Human-Computer Studies*, vol. 62, no. 2, pp. 281-306, 2005. <https://doi.org/10.1016/j.ijhcs.2004.11.008>.

- [139] A. Schlesinger, W. Edwards and R. Grinter, "Intersectional HCI", Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17, 2017. <https://doi.org/10.1145/3025453.3025766>.
- [140] C. Munteanu, H. Molyneaux, W. Moncur, M. Romero, S. O'Donnell and J. Vines, "Situational Ethics", Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15, 2015. <https://doi.org/10.1145/2702123.2702481>.
- [141] A. Schlesinger, W. Edwards and R. Grinter, "Intersectional HCI," Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17, 2017. <https://doi.org/10.1145/3025453.3025766>.
- [142] R. Bellini et al., "Feminist HCI," Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18, 2018. <https://doi.org/10.1145/3170427.3185370>.
- [143] S. Breslin and B. Wadhwa, "Towards a Gender HCI Curriculum," Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15, 2015. <https://doi.org/10.1145/2702613.2732923>.
- [144] S. Yardi and A. Bruckman, "Income, race and class," Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12, 2012. <https://doi.org/10.1145/2212776>.
- [145] M. Allison and L. Kendrick, "Towards an Expressive Embodied Conversational Agent Utilizing Multi-Ethnicity to Augment Solution Focused Therapy," International Florida Artificial Intelligence Research Society Conference, pp. 332-337, 2013.
- [146] J. Waycott et al., "Ethical Encounters in Human-Computer Interaction", Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16, 2016. <https://doi.org/10.1145/2851581.2856498>.
- [147] P. Kumar, M. Bashour and V. Packiriswamy, "Anthropometric and Anthroposcopic Analysis of Periorbital Features in Malaysian Population: An Inter-racial Study", Facial Plastic Surgery, vol. 34, no. 04, pp. 400-406, 2018. <https://doi.org/10.1055/s-0038-1648224>.
- [148] S. Jilani, H. Ugail and A. Logan, "Inter-Ethnic and Demic-Group Variations in Craniofacial Anthropometry: A Review," PSM Biological Research, vol. 1, 2019.

Enhancing Decision Making Capabilities in Humanitarian Logistics by Integrating Serious Gaming and Computer Modelling

Za'aba Bin Abdul Rahim^{1,*}, Giuseppe Timperio¹, Robert de Souza¹, Linda William²

¹The Logistics Institute Asia Pacific, National University of Singapore, 21 Heng Mui Keng Terrace #04-01, Singapore 119613

²School of Informatics & IT, Temasek Polytechnic, 21 Tampines Avenue 1, Singapore 529757

ARTICLE INFO

Article history:

Received: 30 January, 2020

Accepted: 25 May, 2020

Online: 11 June, 2020

Keywords:

Supply Chain Management

Humanitarian Logistics

Serious Gaming

Computer Modelling

Enhanced Decision Making

ABSTRACT

The field of humanitarian logistics has in recent times gained an increasing attention from both academics and practitioners communities alike. Although various research groups have addressed theoretical and technical developments in humanitarian logistics using conventional research tools, applied research appears to be often dependent on practitioners' inputs. This paper is an attempt to fill the existing gaps between academic research and practitioners' needs and proposes an integrated framework that consists of serious games and computer modelling. The serious games component aims to raise awareness on humanitarian logistics issues as well as provide a platform to facilitate the acquisition of inputs from humanitarian practitioners. Based on these inputs, a computer model will be developed. To test the framework, a real-life case study about the prepositioning of strategic stockpiles in Indonesia, one of the countries with the highest disaster risk exposure on a global scale, was used. Findings of this work highlight the role of serious games as risk-free environments for players to design strategies enhancing disaster preparedness in conjunction with broadly used research methodologies such as computer modelling.

1. Introduction

This paper is an extension of work originally presented in 2018 Winter Simulation Conference (WSC), Gothenburg, Sweden by de Souza, et al. by [1]

ReliefWeb, a specialized digital service of the UN Office for the Coordination of Humanitarian Affairs (OCHA), reported that between the years 1994 and 2013, more than 6,500 natural calamities occurred globally. On average, 218 million lives were affected, and 1.35 million deaths were recorded annually [2]. In the year 2017 alone, for instance, more than 300 catastrophes occurred, of which more than half were natural [3]. These disasters claimed more than 11,000 people's lives, and it's estimated that they have caused an economic loss of over 300 billion US dollars. It is almost two times as many as the USD 180 billion economic losses estimated for the precedent year, and well above the average USD 190 billion monetary losses for the period 2005-2015 [3]. In addition, the combination of climate change, political and social crises, and progressive urbanization of

disaster-prone areas seem to push even further the frequency and scale of humanitarian crises globally.

Therefore, in a context of increased number and scale of humanitarian crises in conjunction with reduced resources, there is a need for greater efficiency in the deployment of relief assets. This is particularly relevant for logistics (purchasing, transportation, and storage), that is by far the biggest contributor to humanitarian expenditures (60-80%) [4].

In humanitarian context, logistics can be defined as the activity of "planning, implementing and controlling the efficient, cost-effective flow of and storage of goods and materials as well as related information, from point of origin to point of consumption for the purpose of alleviating the suffering of vulnerable people" [5]. Compared to the commercial context, humanitarian logisticians are required to be cost-effective, agile, and time responsive.

Generally, disaster relief operations are run in a highly chaotic environment where the needs of affected population are in a fast and continuous development, and resources such as

*Za'aba Bin Abdul Rahim, Email: zaaba@nus.edu.sg

www.astesj.com

<https://dx.doi.org/10.25046/aj050351>

supplies, people, funds, and technologies are scarce [6] [7]. Hence, the logistics processes and disaster relief chains need to be streamlined to speed up the response time for the pressing need of critical supplies. Given these requirements, humanitarian supply chains require to be designed proactively, well ahead of the time when disasters occur [8] [9].

To address supply chain issues, computer models are conventionally used. The main benefit of computer models is that they allow to replicate (to a certain extent) real-life operations in virtual environments. In turns, this allows to identify possible solutions, to test their feasibility in a virtual environment, and to mitigate the risk of making costly mistakes prior to the implementation in a real-world setting [10]. Amongst others, computer simulation appears to be the most popular decision support tool for supply chain design decisions as well as for the assessment of supply chain policies. A number of simulation techniques were used by other scholars to address research issues in the humanitarian logistics space including System Dynamics (SD) [11] [12], Discrete Event Simulation (DES) [13] [14] and Agent Based Simulation (ABS) [15] [16].

However, practical experience suggests that while the aforementioned research techniques are indeed relevant, practitioners' inputs are also a key component for research inquiries related with humanitarian logistics. In fact, while methodologies such as network optimization or simulation can be used to generalize and extend the humanitarian logistics preparedness through scenario-based analysis, it appears that to date there is no clear method facilitating the acquisition of inputs from practitioners.

In another front, there has also been an increase in the use of serious games to increase awareness and provide an avenue for wider audience to learn and experience humanitarian logistics, as well as to facilitate the acquisition of experiential inputs from practitioners. Serious games have been around for many years and have recently gained increased popularity [17]. Playing games increases interactivity among players and provide time-flexibility to pace their learning in their own time [18]. Serious games in humanitarian logistics provide a safe environment for the players to plan, implement and evaluate their humanitarian logistics strategies without the dangers of real-world consequences. It provides different sets of scenarios or environments compared to computer simulation models [19].

In this paper, we focus on developing a decision support framework integrating the use of serious games and computer models to address complex supply chain issues in the domain of disaster relief. We apply a role-playing simulation-based board game, titled THINKLog: Humanitarian Logistics Gameplay to gather inputs from practitioners. THINKLog is an interactive board game designed that can be scaled to include different scenarios for logistics and supply chain management including for humanitarian logistics. THINKLog is a learning framework where different scenarios can be generated by combining different concepts and challenges in logistics. One of the scenarios that is available is the Humanitarian Logistics Gameplay. This gameplay helps to introduce the importance and complexity of supply chain management and logistics in humanitarian context and deepen the understanding of humanitarian logistics for the players. The

gameplay focuses on the warehouse location (Preparedness), inventory management and deliveries (Response).

This game has been played both locally (Singapore) and regionally (Southeast Asia). In this paper, the results from one of the game workshops are presented as an evidence that the game provided an interactive method to apply and experiment with supply chain concepts. The workshop was conducted with an official operating in disaster preparedness and response from multiple ASEAN countries at the ASEAN Coordinating Centre for Humanitarian Assistance on Disaster Management (AHA) Centre Executive (ACE) Programme in November 2018 in Jakarta, Indonesia.

Using the gathered inputs from the practitioners during and after the game session, a computer model is developed. As a case study, we applied it to a real-life case study about prepositioning of life-saving kits in Indonesia. Subsequently, a simulation model using AnyLogistix software was developed.

The rest of the manuscript is organized as follows. Section 2 debates the academic literature previously developed in this domain and focuses its attention on simulation-based serious games applied to supply chain and logistics and computer models in the niche of stockpile prepositioning. Section 3 describes the problem statement and methodology used in this paper. Section 4 describes the THINKLog – Humanitarian Logistics Gameplay game, a real-life simulation model and key findings of both. Section 5 summarizes the key findings of this work alongside its limitations and future extensions.

2. Literature Review

In this section, we review relevant academic literature in the areas of simulation-based serious games and computer simulation applied to the domain of humanitarian logistics and stockpile prepositioning.

2.1. Simulation-Based Serious Gaming

Recently, there has been an upward trend towards the use of serious games in creating awareness in specific areas and domains. They are widely used in areas like education [20], military [21], healthcare [22] and city planning [23]. Games help players grasp the learning objectives by providing hands-on engaging and motivating experiences. This manuscript will emphasize humanitarian simulation-based games.

Simulation-based games embed the process of hypothesizing, probing, and reflecting upon the simulated world within the game to promote learning and create awareness [24]. By using specific storylines in the game, players try to solve a particular task or challenge to learn specific concepts. Players are able to perform repeated experimentations and understand the consequences of their decisions, thus promoting players' incentive and learning transfer to the real-life context. [19]. Some other studies have even shown that these kind of games advanced players' awareness and understanding of scientific knowledge [19] [25].

There is a fair number of supply chain simulation-based games available in the market that are used as a learning tool to teach certain topics of supply chain management. In this paper, we looked into simulation-based games with humanitarian

storylines. Two of such games are AFTERSHOCK [26] and STOP DISASTERS! [27]. AFTERSHOCK is a board game where players assume roles of different agencies involved in the response of an emergency and its early phase of the recovery. STOP DISASTERS! is an online game whereby the player is needed to prepare for the impending disaster by improving the infrastructure in order to save the lives.

These games do offer a certain level of experiential experience or learning to the players. However, they do not provide an avenue to elicit current issues, especially faced by humanitarian logistics domain experts, practitioners and decision makers. There may be valuable information/data that may be captured, analysed and derived from the games that are not necessarily captured.

Table 1 shows the summary and comparison of the two aforementioned games against our proposed one namely THINKLog.

2.2. Computer Modelling in Humanitarian Logistics

Logistics system design with focus on storage and transportation systems have received significant attention in the domain of commercial supply chains [28], [29]. Nevertheless, the importance of humanitarian logistics applications has been recognized only in recent times. One of the areas receiving the greatest attention in the space of humanitarian logistics is stockpile prepositioning and network design.

Balcik & Beamon (2008) are considered pioneers in this field. Their study analyses the facility location problem in consideration of constraints such as budget and capacity of facilities. The overarching goal of their work is to optimize the flow of relief supplies all the way to disaster affected zones. McCoy & Brandeau [30] looked at storage and shipping policies and the main outcomes of their work include effective distribution strategies and trade-off between stock size and service level. Roh, et al. [31] analyzed the supply prepositioning location problem from both macro and micro perspectives using inputs from high-level decision-makers and operational experts.

Simulation literature studies humanitarian logistics problems using different methods including system dynamics, discrete-event simulation, and agent-based modeling.

System dynamics (SD) is a modelling technique often used to evaluate the nonlinear behavior of complex systems. Basic elements of SD modelling are known as stocks, flows, internal feedback loops, table functions and time delays [32]. One of the first applications of SD in the humanitarian logistics space was by Besiou et al. [11]. Their work applies SD to the issue of fleet management. Costa et al. [12] instead used SD with the aim of addressing the challenge of coordination in disaster response.

Discrete-event simulation (DES) is a modelling technique used to sequence discrete events over time. Once a particular event occurs, the state of the system will change. No changes will occur between two consecutive events [33]. Iakovou, et al. [13] developed a DES model to assess the impact of emergency sourcing on supply chain performances. Noreña, et al. [14] used DES to understand the robustness of the logistics of medical supply in Bogotá (Colombia).

Table 1. Games Comparison

	STOP DISASTERS!	AFTERSHOCK	THINKLog (humanitarian logistics version)
Topic	Pre-disaster buildings capabilities and defences	Interagency cooperation in response to emergency and early recovery phase of a humanitarian crisis	Logistics and facility location identification in humanitarian context
Type of game	Simulation	Simulation	Simulation
Objective	To build the defence of an existing structures for the population in face of an upcoming disaster	To address the urgent humanitarian needs of the local population.	To build a warehouse based in the concept of Multi Criteria Decision Making (MCDM) to deliver an uninterrupted and continuous supply of the required life-saving kits to the disaster area
Type of tasks	Operation activities/tasks	Operation activities/tasks	Strategic planning (in preparation stage) and operation activities/tasks (in response stage)
Playing mode	Computerized	Board game	Board game with digital companion app

Agent-based modeling (ABM) refers to a series of computer simulation techniques whereby the actions of autonomous pre-defined agents are simulated [15]. Horner & Widener [16] used an integrated approach ABM and GIS to foresee the damage caused by heavy storms on the transportation network in use case of Florida City.

2.3. Research Gaps

Despite the significant body of literature available in the domain of stockpile prepositioning, to our knowledge it appears that most research groups have tackled humanitarian logistics issues, and facility location problems in particular, from a purely conceptual standpoint with limited understanding of practitioners' viewpoints. Although these approaches provide fine solutions to this class of problems, it is also evident that there is a need for supplementing them with practitioner perspectives. It is intended that, decision-makers will be able to gain insights on the following matter:

- How to leverage on serious games to acquire inputs from decision makers and practitioners?
- How to integrate serious-games and computer models to assure the robustness of the computer model and its relevance to the practice?

2.4. The Contribution of The Current Research

This paper demonstrates the applicability of the proposed framework for network design in the use case of Indonesia. The work presented in this manuscript will apply THINKLog and simulation to demonstrate the relevance of the proposed solution approach. The contribution of our study is on the integration of serious games and computer modelling with the aim of generating insights into the domain of network design using practitioner’s inputs.

3. Problem Statement and Methodology

Indonesia is the 14th largest nation by size, spans across three time zones and counts over 260 million people living across more than seventeen thousand islands [9]. The country sits on the edges of four tectonic plates and is part of the so called “Ring of Fire”. According to Statista [34], between 1900 and 2016, 113 earthquakes were recorded in the Indonesian archipelago, which claimed nearly 200,000 human lives. According to the Centre for Research and Epidemiology of Disasters, Indonesia ranks in the ‘top 5 countries most frequently hit by natural disasters’ after China, India, the Philippines, and the United States [35]

3.1. Problem Statement

When disasters hit in remote areas of the archipelago, existing response capacities are significantly stretched. Despite the efforts of the national authorities to boost inter-island and intra-island connectivity through infrastructure development [36], the existence of ad-hoc logistics hubs to store those critical inventories to be deployed in times of emergency, would be highly beneficial for the effective provision of humanitarian assistance. Hence, the case study will focus on the pre-positioning of relief supplies at strategic locations across Indonesia to improve country’s logistics capabilities to respond to natural disasters.

3.2. Methodology

Because the problem statement requires the embedding of inputs from practitioners, decision makers, and humanitarian logistics domain experts, the integration of serious games for the collection of inputs, and computer simulation as an analytical closed form of analysis is proposed.

In particular, to address the identified problem statement, the following multi-phase framework encompassing two main steps was conceptualized and tested:

- (1) Phase 1. Socializing the problem at hand with practitioners, decision makers, and humanitarian logistics domain experts, and gathering of inputs using serious games.
- (2) Phase 2. Building the computer model and running scenario-based analysis for the identification of the solution to the problem at hand.

Figure 1 is an outline of the proposed framework which was tested in the use case of stockpile prepositioning in the use case of Indonesia using THINKLog game in integration and computer simulation in AnyLogistix.

4. THINKLOG Game

4.1. THINKLOG Game, An Overview

ThinkLog is an expandable interactive logistics and supply chain management board game that was designed specifically to help players in learning about SCM concepts. The game aims to complement teaching and learning activities in classrooms and workshops by introducing SCM concepts through roleplaying and simulation. The initial design of the game was for a pure board game play: main board, demand cards, gameplay/rules and game master [37]. This game also includes a digital companion application [38]. With the inclusion of this app, more SCM concepts and different scenarios of game play can be added, such as the Humanitarian Logistics game play.

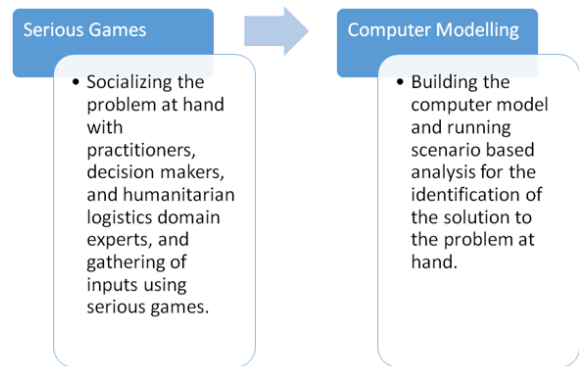


Figure 1. Proposed Framework



Figure 2: THINK:Log game components

4.2. Humanitarian Logistics Gameplay

The main learning objectives for this gameplay are to introduce the notions of Multi Criteria Decision Making (MCDM), and particularly its applications in warehouse location selection problems for disaster preparedness, while providing a high level

view on the challenges related with information flow and coordination in logistics and supply chain management.

MCDM, that is a structured framework able to analyze decisions problems involving complex objectives concepts, is adapted in this gameplay to help the identification of possible locations for the warehouses [9], [39], [40].

There are two stages in this gameplay - Preparation and Response. In the Preparation, the players input their weightage to the criteria defined in the MCDM framework directly through the companion app. Figure 3 shows the input screen for the players. There are 5 criteria for consideration, including Distance, Congestion, Cost, Coverage and Risk. Figure 4 shows one of the criteria and its definition. The companion app will then generate recommended locations based on the weightage as input by the player. In the second stage namely Response, players will have to deliver the requested relief items to the affected population while taking into consideration the possible random events that may occur and disrupt their distribution operations. Examples of events include congestions and /or further calamities that may impede their logistics plan. Figure 5 shows the basic game flow of this game play.

Players assume the role of a Humanitarian Agency Officer in charge of designing and coordinating an uninterrupted and continuous flow of relief items to disaster affected areas. The player has to strategize the most optimal way to deliver such relief items. In the event where the player fails to deliver the required number of items within the stipulated time, the player will ‘earn’ a failure token as a penalty. The player with the least number of failure tokens wins the game.

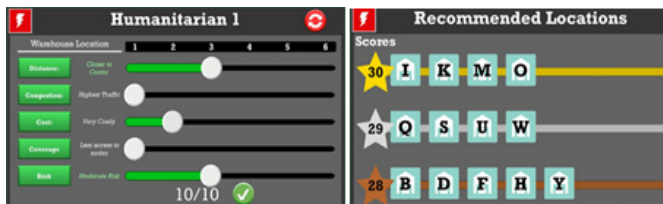


Figure 3: Screenshots of the Warehouse Location selection screen using the digital companion application



Figure 4: Sample screenshot of the MCDM criteria in the digital companion application

4.3. Game Experience

In November 2018, we conducted a workshop using the THINKLog game in Jakarta, Indonesia. This workshop was part of the ASEAN Coordinating Centre for Humanitarian Assistance on Disaster Management (AHA) Centre Executive (ACE)

Programme conducted by the ASEAN Coordinating Centre for Humanitarian Assistance (AHA Centre). The exercise involved 16 senior government officials operating in disaster preparedness and response from multiple ASEAN countries, whom had limited knowledge in humanitarian logistics.

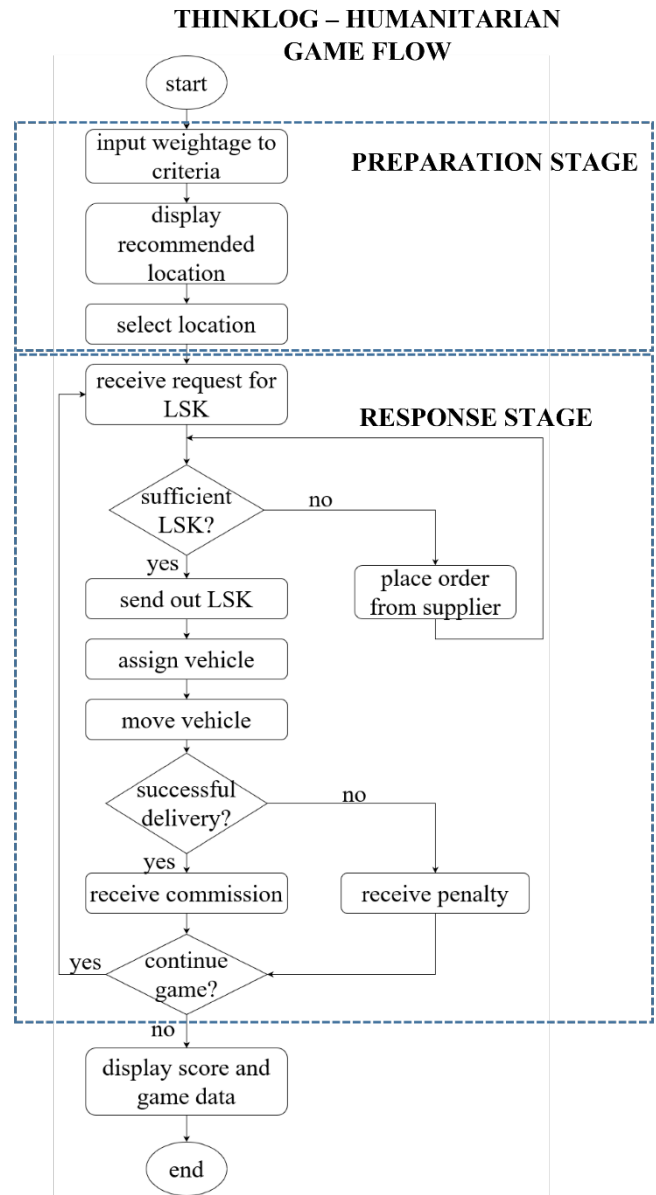


Figure 5: Basic game flow of THINKLog Humanitarian play

We conducted a brief presentation on humanitarian logistics and related supply chain concepts prior to the game play session. After the game play, participants were asked to fill up feedback forms. We then used those data to evaluate the impact of the game to increase the participants’ understanding about humanitarian logistics and the overall game experience. We also used the workshop as a platform to gather inputs from the participants to be used in the simulation model.

The session allowed us to gather two separate sets of inputs. The first set is to evaluate participants’ game experiences and it used a questionnaire of 14 questions. The sample of questions are shown in Table 2 below.

Table 2: Sample of feedback questions

0: Strongly disagree 1: Disagree 2: Neutral 3: Agree 4: Strongly agree

Your Opinion	Rate (0-4)
1. I feel that the whole session is interesting	0 – 1 – 2 – 3 – 4
2. I would like to play this game again in future	0 – 1 – 2 – 3 – 4
3. I encounter difficulties in understanding the game rules	0 – 1 – 2 – 3 – 4
4. I know exactly what I need to do in the game	0 – 1 – 2 – 3 – 4
5. I feel that every steps/decisions made in this game has meaningful meaning	0 – 1 – 2 – 3 – 4
6. I feel very involved in this game	0 – 1 – 2 – 3 – 4
7. Interaction with other players motivates me to understand this game better	0 – 1 – 2 – 3 – 4
8. Overall session is too long/boring	0 – 1 – 2 – 3 – 4
9. The game motivates me to ask questions/discuss	0 – 1 – 2 – 3 – 4
10. I obtained interesting information through this session	0 – 1 – 2 – 3 – 4
11. I am interested in the content provided in this game	0 – 1 – 2 – 3 – 4
12. I did not learn anything from this session	0 – 1 – 2 – 3 – 4
13. The Companion App helps me understand the game	0 – 1 – 2 – 3 – 4
14. The Companion App helps me learn from the game	0 – 1 – 2 – 3 – 4

Participants were asked to provide answers using the 5-point Likert scale with 1 as strongly disagrees and 5 as strongly agree. There were sixteen valid responses received. The positive experience’s average score is above 3 and the negative experience’s is below 2, as shown in Figure 6. We may conclude that the participants had a good overall game experience.

The second feedback asked the participants to list down the key learning points that they have gathered from the session. We then compared the responses with the game’s intended learning objectives. We received ten valid responses. We extracted the responses and identified the topics that they mentioned in their

responses. The topics and the connection between the topics are shown in Figure 7.

From the analyzed feedback forms, we identified three main topics, namely: Plan, Warehouse and Location. Three out of six responses (3/6) that mentioned about Plan also mentioned about Warehouse and Location. It can be further breakdown to important considerations for planning the warehouse location, such as distance, congestion, cost, coverage and risk.

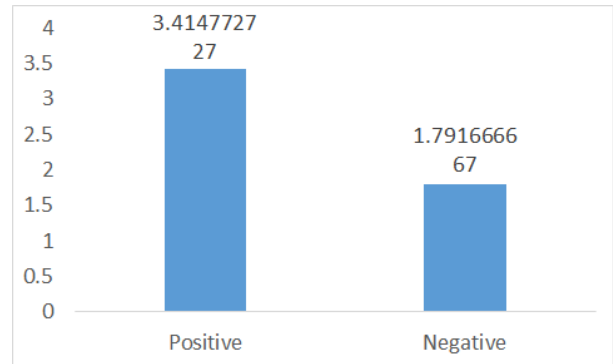


Figure 6: Overall experience

With the results, we are able to conclude that:

- The participants’ learning outcome is aligned with the game’s learning objectives.
- The game exposes the participants to the humanitarian logistics complexity.
- The feedback confirmed the need of embedding key elements such as distance (closeness to affected areas), congestion (the lower, the better), cost (the lower, the better), coverage (closer to highly populated disaster prone areas) and risk (facility to be located in a disaster free zone).

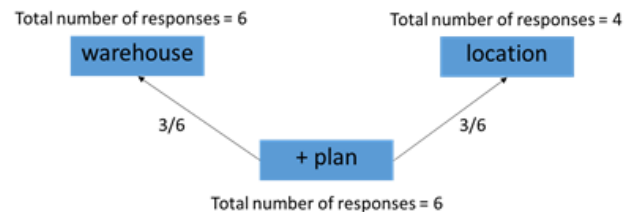


Figure 7: Learning Outcome

In addition, we also recorded the practitioners’ inputs during and after the game session. Their inputs include the critical criteria to use for determining the location of a warehouse, and the importance of each criterion.

The feedback shows that the game enhances the players’ awareness to the importance of decision making, especially in facility location. Players spent a substantial amount of time inputting the weightage for the criteria, deliberating which criteria should be of a higher importance. As the game provides 5 selected decision criteria for the MCDM, the players are more eager to

experiment these using real-data simulation model that incorporates even more criteria for consideration, which will also involve the framework of selecting those criteria. Table 3 shows the criteria that was selected for the THINKLog game as compared to the criteria used for the model in Section 4.4. [41]

During the game play session, it was also noted that most, if not all, players placed higher weightage to Risk and Access to affected zones criteria. All players have the same number of vehicles (2 trucks) and warehouse capacity (30 LSK) in the game. As the game has only 1 disaster area (demand point), this may be sufficient. Whereas in real-life situation, there may be more than 1 demand point at any given point of time during a disaster.

Also, for this game play configuration, we intentionally excluded the transportation cost that may occur to deliver the LSK to the disaster area, for ease of play purpose. As such, the player may not be able to optimize its relief operation.

Table 3: The criteria used in THINKLog game

Criteria	Description used in THINKLog game	THINKLog game
Coverage	Refers to the number of nodes connected.	✓
Access to affected zones	Refers to the time taken from when the order is placed till the order is delivered. The closer one is to the affected area, the shorter lead time it will have,	✓
Risk	Refers to the chance of a disaster striking the node.	✓
Access to infrastructure	Not used in the game	✗
Access to corridor	Not used in the game	✗
Congestion	Refers to the frequency of traffic congestions happening around that area	✓
Cost	Refers to the maintenance cost of a warehouse in a node. The area around the downtown costs higher than those at the outskirts	✓
National development plan (NDP)	Not used in the game	✗

Section 4.4 describes the model we have developed to help locate the most appropriate alternative locations for the warehouse (facilities) to serve the different demand points, taking into consideration the transportation cost.

4.4. Simulation Model for Humanitarian Logistics

Based on the practitioners’ input and the game session’s outcomes of the game play, we developed a computer simulation model to delve the understanding on network design for stockpile prepositioning in the use case of Indonesia. The aim of this model was to determine, out of 9 alternative locations, which were the

most appropriate 6 sites to use for the positioning of emergency response facilities.

There were 186 demand points identified across the archipelago. Each demand points is served by a single facility. The model’s overarching objective is to identify the possible locations for the facilities to serve these demand points, while keeping the transportation cost at the minimum. We used the actual distance and land fuel cost in this model to calculate the transportation cost between the facility and demand points. For those demand points without land connections (e.g. small islands), we set the fuel cost to extremely high values as to replicate intermodal transportation.

The results of the model are summarized in Figure 8 while the entire case study can be found in the paper by The Logistics Institute – Asia Pacific [42]. Figure 8(a) and 8(b) show possible network configurations identified by the computer model whereas figure 8(c) shows the network structure prospected by discussions with practitioners.

By inputting these 3 alternative network configurations into the model, findings show that the network structure identified using network optimization has the potential to reduce the transportation cost by 15%.



8a) Optimum configuration from Simulation tool



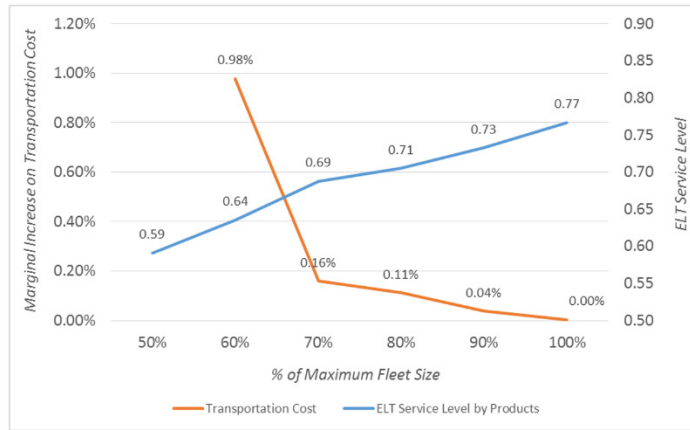
8b) Second best configuration from Simulation tool



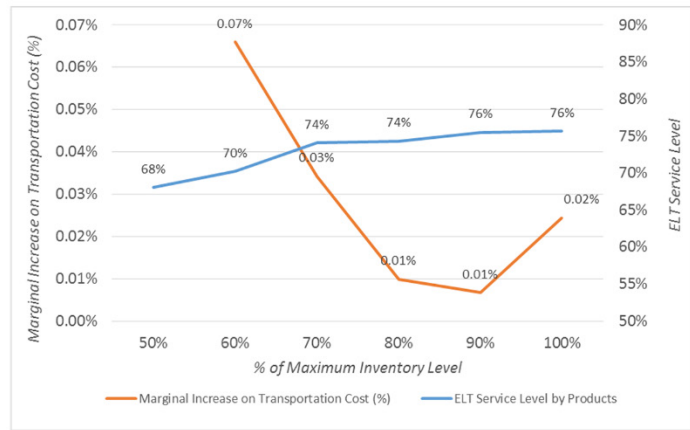
8c) Configuration from preliminary discussions with several experts in Indonesia

Figure 8: Disaster Relief Model

Subsequently, we stress-tested the network configuration upon two key logistics parameters namely fleet size and inventory. Figure 9 shows the results of the computer simulation model in regard to service level at when fleet size and inventory are set at different values.



9a) Impact of Fleet size on Transportation Cost and service level



9b) Impact of Inventory Level on Transportation Cost and service level

Figure 9: Simulation Result for Fine-tuning the Network Configuration on Two Parameters

5. Conclusion

In this paper, the focus was on integrating serious games and computer modelling as a novel way to raise awareness on humanitarian logistics issues as well as to facilitate the acquisition of inputs from humanitarian practitioners. A real-case simulation model was developed with the aim of selecting the most suitable locations for establishing a network of facilities for prepositioning stocks of life-saving goods in Indonesia. We also had preliminary discussions with several experts on a prospected network configuration. When we compared these two configurations, the simulation model configurations allow for an estimated cost saving of 15%.

A role-based board game, THINKLog, is proposed as an easy-to-use tool to raise awareness in humanitarian logistics for non-technical personnel. The game has two stages, the Preparation and Response stage. The feedbacks gathered from the game play session shows that the participants have a good overall game experience with a substantial positive influence on the players. This is supported by an evaluation of the learning

objectives by the participants whom reported that they were able to grasp them during the game play.

This work has few limitations. First, the sample size of senior government officials operating in disaster preparedness and response from multiple ASEAN nations is fairly small. Playing the game with a larger group of practitioners would allow to gather deeper insights on the game play and inputs to the computer model. Secondly, on the simulation model, the dataset on small and medium scale disasters is confined to the biennium 2014-2015. An extension of this database with the inclusion of a greater number of disasters would provide a more accurate estimation of demand.

For future extension, we see an opportunity to add more immersive game features to the game, like augmented reality (AR). This may enhance the game's visualization and overall gameplay experience. Secondly, we would like to evaluate the THINKLOG game using a broader group of individuals with different age groups, educational and professional backgrounds, and level of understanding of supply chain and humanitarian logistics concepts. This would enable us to gather a more comprehensive view of game's ability to raise awareness.

Conflict of Interest

The authors declare no conflict of interest related to this paper.

Acknowledgment

The THINKLog board game was first developed as part of "Temasek Foundation International –National University of Singapore Urban Transportation Management Programme in Indonesia" programme. It is supported by Temasek Foundation International and Coordinating Ministry for Economic Affairs of Indonesia.

References

- [1] R. D. de Souza, L. William, G. Timperio and Z. B. A. Rahim, "Simulation M And Simulation-Based Serious Gaming In Humanitarian Logistics," in *Winter Simulation Conference*, 2018.
- [2] ReliefWeb, "The human cost of natural disasters 2015: a global perspective," 6 3 2015. [Online]. Available: <https://reliefweb.int/report/world/human-cost-natural-disasters-2015-global-perspective>. [Accessed 5 12 2017].
- [3] Swiss Re Institute, "Natural catastrophes and man-made disasters in 2017: a year of record-breaking losses," Swiss Re Institute, 2018.
- [4] L. Van Wassenhove, "Humanitarian Aid Logistics: Supply Chain Management in High Gear," *Journal of Operational Research Society*, vol. 57, no. 5, pp. 475-489, 2006.
- [5] A. S. Thomas and L. R. Kocpczak, "Fritz Institute," 2005. [Online]. Available: <http://www.fritzinstitute.org/pdfs/whitepaper/fromlogisticsto.pdf>. [Accessed 25 March 2016].
- [6] B. Balcik and M. Beamon, "Facility Location in humanitarian relief," *International Journal of Logistics: Research and Applications*, vol. 11, no. 2, pp. 101-121, 2008.
- [7] Y. Nahleh, A. Kumar and F. Daver, "Facility Location Problem in Emergency," *International Journal of Mechanical, Aerospace, Mechatronic, and Manufacturing Engineering*, vol. 7, no. 10, pp. 2113-2118, 2013.
- [8] C. L'Hermitte, B. Brooks, M. Bowles and P. Tatham, "Investigating the strategic antecedents of agility in humanitarian logistics," *Disasters*, vol. 41, no. 4, p. 672-695, 2016.

- [9] G. Timperio, G. Panchal, A. Samvedi, M. Goh and R. De Souza, "Decision support framework for location selection and disaster relief," *Journal of Humanitarian Logistics and Supply Chain Management*, vol. 7, no. 3, pp. 222-245, 2017.
- [10] C. Thierry, G. Bel and A. Thomas, "Supply Chain Management Simulation: An Overview," in *Supply Chain Management Simulation*, Wiley-ISTE; 1 edition, 2008, pp. 1-39.
- [11] M. Besiou, O. Stapleton and L. N. Van Wassenhove, "System dynamics for humanitarian operations," *Journal of Humanitarian Logistics and Supply Chain Management*, vol. 1, no. 1, pp. 78-103, 2011.
- [12] O. Costa, J. Santos, M. Martins and U. Yoshizaki, "A system dynamics analysis of humanitarian logistics coordination," Delft, 2015.
- [13] E. Iakovou, D. Vlachos, C. Keramydas and D. Partsch, "Dual sourcing for mitigating humanitarian supply chain disruptions," *Journal of Humanitarian Logistics and Supply Chain Management*, vol. 4, no. 2, pp. 245-264, 2014.
- [14] D. Noreña, R. Akhavan-Tabatabaei and L. Yamín, "Using discrete event simulation to evaluate the logistics of medical attention during the relief operations in an earthquake in Bogota," Phoenix, AZ, USA, 2011.
- [15] V. Grimm and S. F. Railsback, "Individual-based Modeling and Ecology," 17 May 2004. [Online]. Available: <http://www2.humboldt.edu/ecomodel/documents/Grimm-Railsback05.pdf>. [Accessed 20 August 2017].
- [16] M. W. Horner and M. J. Widener, "The effects of transportation network failure on people's accessibility to hurricane disaster relief goods: a modeling approach and application to a Florida case study," *Natural Hazards*, vol. 59, no. 3, p. 1619-1634, 2011.
- [17] U. Ritterfeld, M. Cody and P. Vorderer, "Serious games: Mechanisms and effects," *International Journal of Gaming and Computer-Mediated Simulations*, vol. 1, no. 3, pp. 89-94, 2009.
- [18] W. William, "Current trends in educational technology research: The study of learning environments," *Educational Psychology Review*, vol. 14, no. 3, p. 331-351, 2002.
- [19] H.-T. Hou, "Integrating cluster and sequential analysis to explore learners' flow and behavioral patterns in a simulation game with situated-learning context for science courses: A video-based process exploration," *Computers in Human Behavior*, vol. 48, no. 1, pp. 424-435, 2015.
- [20] M. Graafland, J. M. Schraagen and P. Schijven. M., "Systematic review of serious games for medical education and surgical skills training," *British Journal of Surgery*, vol. 99, no. 10, pp. 1322-1330, 2012.
- [21] C. Lim and H. Jung, "A study on the military Serious Game," *Advanced Science and Technology Letters*, vol. 39, pp. 73-77, 2013.
- [22] M. A. Garcia-Ruiz, J. Tashiro, B. Kapralos and M. V. Martin, "Crouching Tangents, Hidden Danger: Assessing Development of Dangerous Misconceptions within Serious Games for Healthcare Education," in *Virtual Immersive and 3D Learning Spaces: Emerging Technologies and Trends*, Kansas State University, USA, IGI Global, 2011, p. 412.
- [23] A. Gómez-Rodríguez, J. González-Moreno, D. Ramos-Valcárcel and L. Vázquez-López, "Modeling serious games using AOSE methodologies," in *11th International Conference on Intelligent Systems Design and Applications*, 2011.
- [24] J. Hamari, D. Shernoff, E. Rowe, B. Coller, J. Asbell-Clarke and T. Edwards, "Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning," *Computers in Human Behavior*, vol. 54, no. 1, pp. 170-179, 2016.
- [25] M. Ma, A. Oikonomou and L. Jain, *Innovations in Serious Games for Future Learning*. In *Serious Games and Edutainment Applications*, London: Springer, 2011.
- [26] PAXsims, "AFTERSHOCK," 4 2015. [Online]. Available: <https://paxsims.wordpress.com/aftershock/>.
- [27] UN Office for Disaster Risk Reduction, "STOP DISASTERS!," 2019. [Online]. Available: https://www.stopdisastersgame.org/stop_disasters/.
- [28] S. Conceição, S. Pedrosa, A. Neto, M. Vinagre and E. Wolff, "The facility location problem in the steel industry: a case study in Latin America," *Production Planning & Control*, vol. 23, no. 1, pp. 26-46, 2010.
- [29] H. Steenhuis and E. De Bruijn, "Production Planning & Control," *Assessing manufacturing location*, vol. 15, no. 8, pp. 786-795, 2007.
- [30] J. McCoy and M. Brandeau, "Efficient stockpiling and shipping policies for humanitarian relief: UNHCR's inventory challenge," *Operations Research-Spektrum*, vol. 33, no. 3, pp. 673-698, 2011.
- [31] S. Roh, S. Pettit, I. Harris and A. Beresford, "The pre-positioning of warehouses at regional and local levels for a humanitarian relief organisation," *International Journal of Production Economics*, vol. 170, no. Part B, pp. 616-628, 2015.
- [32] System Dynamics Society, "Introduction to System Dynamics," 2010. [Online]. Available: <http://lm.systemdynamics.org/what-is-s/>. [Accessed 9 January 2018].
- [33] S. Robinson, *Simulation: The Practice of Model Development and Use*, 1st ed., Chichester: John Wiley & Sons Ltd, 2004.
- [34] Statista, "Countries with the most earthquake fatalities 1900-2016," 2017. [Online]. Available: <https://www.statista.com/statistics/269649/earthquake-deaths-by-country/>. [Accessed 5 January 2018].
- [35] AIR, "Five Countries Most Frequently Hit by Natural Disasters," 24 September 2015. [Online]. Available: <http://www.air-worldwide.com/Blog/Five-Countries-Most-Frequently-Hit-by-Natural-Disasters/>. [Accessed 21 February 2018].
- [36] Indonesia-investments, "Indonesia-investments," 6 June 2017. [Online]. Available: <https://www.indonesia-investments.com/news/todays-headlines/business-investment-climate-of-indonesia-improving-competitiveness/item7878>. [Accessed 20 February 2018].
- [37] Lindawati, E. Nugroho, R. Fredericco, Z. B. A. Rahim and R. de Souza, "ThinkLog: Interactive learning for supply chain management," in *2017 IEEE 6th International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 2017.
- [38] L. William, A. Rahim, R. de Souza, E. Nugroho and R. Fredericco, "Extendable Board Game to Facilitate Learning in Supply Chain Management," *Science, Technology and Engineering Systems Journal*, vol. 3, no. 4, pp. 99-111, 2018.
- [39] M. Köksalan, J. Wallenius and S. Zionts, "An early history of multiple criteria decision making," *Journal of Multi-Criteria Decision Analysis*, vol. 1, no. 2, pp. 87-94, 2013.
- [40] J. Brans, P. Vincke and B. Mareschal, "How to select and how to rank projects: The PROMETHEE method," *European Journal of Operational Research*, vol. 24, no. 2, pp. 228-238, 1986.
- [41] The Logistics Institute - Asia Pacific, "How to identify the most appropriate locations for establishing an efficient network of emergency facilities? - A Discussion Paper," The Logistics Institute - Asia Pacific, Singapore, 2016.
- [42] The Logistics Institute - Asia Pacific, "Integrated Decision Support Framework for Enhancing Disaster Preparedness: A Pilot Application in Indonesia," The Logistics Institute - Asia Pacific, Singapore, 2018.

A Framework for Measuring Workforce Agility: Fuzzy Logic Approach Applied in a Moroccan Manufacturing Company

Fadoua Tamtam*, Amina Tourabi

National School of Applied Sciences, Systems Engineering and Decision Support Laboratory, 80000, Morocco

ARTICLE INFO

Article history:

Received: 03 May, 2020

Accepted: 22 May, 2020

Online: 11 June, 2020

Keywords:

Workforce agility

Fuzzy logic

Agile enablers

ABSTRACT

In today's Moroccan business environment, companies need to implement organization agility by developing an agile workforce that is able to deal with the environment volatility. Thus, the agile workforce concept has been appeared as a necessary and sufficient condition to achieve agility. Focusing on agile enablers influencing workforce agility is an important area but currently there is limited literature available. Acknowledging its importance, we continued our literature exploration in order to identify the enablers of workforce agility. Then, we describe a list of four enablers with different criteria and attributes. This paper further proposed fuzzy logic approach to evaluate different measures of the workforce agility. The results suggest that engagement, knowledge sharing, acceptance of changes and self-motivation are the most important attributes of agile workforce. Apart from that, different agile workforce attributes need to be improved in order to achieve the extremely agile level of the workforce.

1. Introduction

This paper is an extension of work originally presented in 4th World Conference on Complex Systems (WCCS) [1].

For many years, unexpected and dynamic changes [2] represent a common reality facing many organizations from different sizes and sectors. This volatility makes traditional approaches useless for fulfilling current organizational goals. Under this pressure [3], enterprises need to implement agility in order to achieve profitability [4]. The agility means the ability to quickly respond and adapt to volatile market environments. According to the different definitions of "agility" proposed in the literature [2], agility includes different competitive criteria like speed, flexibility, innovation, adaptability, proactivity, quality, productivity, profitability, customization, and knowledge [3] which help to focus on products and services driven by customer instead of those driven by the company [5].

The successful implementation of agility requires adapting all enterprise elements such as goals, technology and people to the unexpected changes [2]. Thus, agility concept has been extended to cover different organization areas. Within this paper [5], the term workforce agility appeared as the employee ability to adapt and evolve quickly by providing innovative solutions to different

problems during any phase of the company project/program. Therefore, agile workforce needs to exhibit sufficient skills [6] which have been implicitly mentioned in Zhang and Sharifi [7] definition: Agile workforce is knowledge worker [3] with a broad vision and who is able to deal with environmental turbulence by capturing the advantageous side of these changes [8]. The same definition was presented by others researchers like Breu, et al. [5] who defined the concept as environmental responsiveness to market volatility [9], also Patil and Suresh [6] defined it as the ability to respond to customer needs and uncertain changes within the stipulated time.

From the literature review, the agile workforce is an essential facet of the overall agility level of the organization [3] since it allows to achieve different organizational benefits [5]. However, there was less focus on the theoretical and empirical validation of workforce agility enablers [5]. The main aim of the present paper is to identify the crucial factors influencing the workforce agility [6]. The following section develops a list of four workforce agility enablers by identifying their criteria and attributes from the literature [5].

2. Literature review: Enablers of the agile workforce

After reviewing the literature, four enablers have been identified as important factors playing a crucial role in improving the workforce agility.

* Fadoua Tamtam, Email: fadoua.tamtam@gmail.com

Individually, to be agile, the workforce needs to be adaptable or flexible. Then, an agile workforce needs to be proactive [6], which means self-anticipating the activities influencing positively the changes [2]. Also, workforce with innovative behavior is an agile workforce who can identify the need for a new product/service/process/technology or improving those existing.

From what was mentioned previously, the workforce handles multiple tasks or programs simultaneously, which means the workforce needs to focus on the most important ones for his work. If the working conditions are stressful, the workforce should be resilient. Other behaviors should be demonstrated by the agile workforce as getting knowledge of marketplace, business environment, organization operations and future priorities. These behaviors are grouped under the name business orientation. Also,

it is important for agile workforce to achieve multiple competencies as those related to management, business process change, technical [6], information technology or software which were identified by Breu et al. [5,6]. Other researchers have identified responsiveness and intelligence as the main attributes of agility. Also, highly motivated and informed workforce is beneficial for the organization success.

Collectively, agile workforce needs to cooperate and share knowledge with internal or external groups which ensures fluid information, communication and knowledge flow across these groups which is important for agile team [6].

Table 1 summarizes enablers, criteria and attributes discussed in the workforce agility literature [5].

Table 1: Workforce agility enablers (Adapted from [2,3,6])

Workforce agility enablers	Workforce agility criteria	Workforce agility attributes		
Workforce status (E ₁)	Adaptability/ Flexibility (E ₁₁)	Simultaneously work on multiple work assignments within program or across different groups in organization (E ₁₁₁)		
		Move quickly from role to role, to new tasks and responsibilities (E ₁₁₂)		
		Engage and disengage often and easily with others with a singular focus on task accomplishment as per requirement (E ₁₁₃)		
		Having interpersonal and cultural adaptability (E ₁₁₄)		
		Acquire skills of professional flexibility (E ₁₁₅)		
		Constant and quick learning of new skills, technologies, and procedures (E ₁₁₆)		
		Acceptance rate of the workforce to altered working time or work locations (E ₁₁₇)		
	Competency/ Self-awareness (E ₁₂)	Developing new skills or competencies within a short span (E ₁₂₁)		
		Developing new skills in business process change (E ₁₂₂)		
		Developing new management skills (E ₁₂₃)		
		Developing new technical skills (E ₁₂₄)		
		Developing new skills in information technology (E ₁₂₅)		
		Developing new software skills (E ₁₂₆)		
		Comprehension of new ideas, knowledge, or technologies (E ₁₂₇)		
		Creativity and innovation in problem-solving (E ₁₂₈)		
		Workforce autonomy (E ₂)	Proactivity (E ₂₁)	Dynamically explore new opportunities (E ₂₁₁)
				Identify and anticipate problems related to change (E ₂₁₂)
Accomplish the promising goals (E ₂₁₃)				
Personal initiative (E ₂₁₄)				
Solution of the change related problems (E ₂₁₅)				

	Innovation (E ₂₂)	Continuously work towards gaining proficiency in multiple competency areas (E ₂₂₁)	
		Share the gained knowledge with its partners and collaborators (E ₂₂₂)	
		Exploring into new markets (E ₂₂₃)	
		Generates new ideas to tackle the unidentified change requests made by customers (E ₂₂₄)	
		Generates new ideas to come out of the ambiguous situations faster (E ₂₂₅)	
	Resiliency (E ₂₃)	Taking calculated risks (E ₂₃₁)	
		Coming out comfortably from ambiguous situations quickly (E ₂₃₂)	
		Positive attitude to change, to new ideas and technology (E ₂₃₃)	
		Tolerance to stressful and uncertain situations (E ₂₃₄)	
	Intelligence/ Responsiveness [6] (E ₂₄)	Quick response to changing customer needs and market conditions (E ₂₄₁)	
		Sensing changes (E ₂₄₂)	
		Execute things smoothly (E ₂₄₃)	
	Quickness/ Speed (E ₂₅)	Shorter transition or recovery time (E ₂₅₁)	
		Faster completion time (E ₂₅₂)	
		Products or services delivery speed (E ₂₅₃)	
		Problem-solving speed (E ₂₅₄)	
Workforce job (E ₃)	Focus (E ₃₁)	Set priorities while handling multiple programs (E ₃₁₁)	
		Set priorities to drive towards solutions (E ₃₁₂)	
		Demonstrate a strong sense of urgency to deliver a set goal (E ₃₁₃)	
		Demonstrate right focus (E ₃₁₄)	
	Business Orientation (E ₃₂)	Get knowledge of marketplace (E ₃₂₁)	
		Get knowledge of business environment (E ₃₂₂)	
		Get knowledge of the organization's operations (E ₃₂₃)	
		Get knowledge of future priorities (E ₃₂₄)	
		Get aligned with the organizational values (E ₃₂₅)	
	Informative (E ₃₃)	Serious information seekers (E ₃₃₁)	
		Keep informed in order to achieve the objectives or clarify problems (E ₃₃₂)	
		Personally undertake research, analysis or investigation (E ₃₃₃)	
		Use contacts or information networks to obtain useful information about technologies and processes related to programs they are working on (E ₃₃₄)	
	Workforce involvement (E ₄)	Collaboration (E ₄₁)	Ability to collaborate with other teams, functions and organizations (E ₄₁₁)
			Avoid duplication of efforts (E ₄₁₂)
Use decision-makers from different domains (E ₄₁₃)			
Smooth flow of knowledge and information across the boundaries of groups (E ₄₁₄)			
Willingness to enter unexpected collaborations (E ₄₁₅)			

Self-motivation/ Ownership (E ₄₂)	Tolerance to different or new opinions of people from other disciplines (E ₄₁₆)
	Tolerance to different or new approach in collaboration (E ₄₁₇)
	Ease of communication (E ₄₁₈)
	Autonomy in collaboration (E ₄₁₉)
	Self-motivation for seeking solutions to change issues (E ₄₂₁)
	Workforce needs possibility of growth (E ₄₂₂)
	Workforce needs recognition (E ₄₂₃)
	Workforce needs advancement (E ₄₂₄)
	Workforce needs technical supervision (E ₄₂₅)
	Pursue goals in the face of setbacks (E ₄₂₆)

3. Fuzzy logic approach to measure workforce agility of a Moroccan company

Acknowledging the importance of workforce agility, it has been assessed through several means such as [10] fuzzy logic, exploratory methodology, descriptive statistics, mathematical modeling and discriminate analysis [11].

Due to the ill-defined and vague indicators which exist within agility assessment [12], the decision-makers are unable to make a significant measurement when available information is scarce [13]. Thus, fuzzy logic has been widely used to handle real-life problems which are subjective, vague, and imprecise in nature [14]. These phenomena can be assessed only by linguistic values rather than numerical terms [13]. Using fuzzy concepts, each linguistic term can be associated with membership function [12]. When it is easy to determine an exact membership function we can use only type-1 fuzzy logic system, which means in our circumstances we don't need to evaluate the workforce agility measures by type-2 fuzzy logic system.

4. Fuzzy logic application

4.1. About the company

The proposed assessment has been done for an industrial company (hereafter referred as Indus_Comp) at southern part of Morocco [15]. Indus_Comp was started in 1999 and it designs and develops products dedicated to individual homes and professional projects. Currently Indus_Comp faces the challenges of modernization and competition and it should be able to offer an unlimited choice of products and service. In order to sustain in this agile environment, Indus_Comp needs to identify the weaker enablers [12] which can affect its competitive positioning.

In this context, and according to the hierarchical composition of the company, a questionnaire has been distributed to five decision-makers of Indus_Comp. The collected data are provided below [15].

4.2. Results and discussion

Each decision-maker has been instructed to express the suitable linguistic variables (Table 2) in terms of importance weight and

performance rating against each agile workforce attribute. The linguistic scale has been furnished in Table 3. Then, these linguistic variables have been converted into fuzzy intervals (Table 4). The next step is to calculate the aggregated importance weights and aggregated performance ratings of the agile workforce criteria and enablers (Table 5). Finally, we obtained the fuzzy workforce agility index and fuzzy performance importance index of all the attributes (FPII) [15]. A detailed description of fuzzy logic application is given as follows [16]:

- Step 1: Identification of agile workforce enablers, criteria and attributes [17]: From the literature, four enablers, twelve criteria and sixty-four attributes were identified (Table 1).
- Step 2: Collection of performance rating and importance weights of agile workforce attributes: Five decision-makers (DM1, DM2,..., DM5) from Indus_Comp were asked to provide the required detail [15] in terms of linguistic variables (Table 2).

Table 2: Fuzzy calculation of agile workforce enablers

Agile Workforce	Importance weight					Performance rating				
	DM1	DM2	DM3	DM4	DM5	DM1	DM2	DM3	DM4	DM5
E ₁₁₁	M	H	M	VL	VL	G	F	F	F	VP
E ₁₁₂	H	H	FH	FH	H	G	F	F	F	VG
E ₁₁₃	M	H	FL	M	M	VG	E	F	G	G
E ₁₁₄	M	M	M	M	M	F	F	F	G	G
E ₁₁₅	H	M	FH	M	FH	F	F	G	VG	VG
E ₁₁₆	H	FH	M	H	FH	P	F	P	G	F
E ₁₁₇	FH	H	FH	M	FH	W	W	P	F	P
E ₁₂₁	H	H	M	FH	H	W	W	VP	F	F
E ₁₂₂	M	H	H	FH	H	G	VG	E	E	G
E ₁₂₃	VH	H	H	VH	H	G	F	G	G	VG
E ₁₂₄	H	FH	H	FH	FH	E	E	G	F	E
E ₁₂₅	H	FH	H	FH	H	G	G	G	VG	G
E ₁₂₆	FH	M	FH	M	FH	F	G	F	F	G
E ₁₂₇	H	M	M	M	H	G	G	G	VG	F
E ₁₂₈	VH	H	H	VH	VH	VP	F	P	P	G
E ₂₁₁	H	FH	M	FH	H	W	F	G	F	G
E ₂₁₂	H	H	H	FH	H	G	VG	G	F	F

E ₂₁₃	FH	M	FH	FH	M	W	W	F	G	VP
E ₂₁₄	FH	FH	FH	FH	FH	VP	P	P	F	P
E ₂₁₅	VH	H	H	VH	H	F	G	E	VG	E
E ₂₂₁	H	FH	H	FH	FH	VP	P	W	F	F
E ₂₂₂	FL	M	FL	FL	M	VG	F	E	F	G
E ₂₂₃	H	FH	H	M	FH	W	VP	G	E	G
E ₂₂₄	H	M	M	FH	H	VP	F	F	W	E
E ₂₂₅	FH	H	H	H	FH	F	G	E	VG	G
E ₂₃₁	H	FL	H	M	H	F	F	VG	F	G
E ₂₃₂	M	M	FH	H	FL	G	E	F	G	F
E ₂₃₃	H	VH	FH	FL	FL	E	VG	G	E	E
E ₂₃₄	VH	M	VH	H	VH	VP	P	P	W	G
E ₂₄₁	H	FH	VH	H	H	VP	G	E	G	G
E ₂₄₂	H	FH	M	H	M	F	VG	VG	G	VG
E ₂₄₃	FH	FH	FH	FH	FH	VP	VG	F	G	F
E ₂₅₁	FH	M	FH	FH	M	W	W	VP	F	VP
E ₂₅₂	H	FH	H	FH	M	G	F	G	VG	E
E ₂₅₃	H	M	FH	H	H	W	W	G	F	G
E ₂₅₄	FH	M	H	FH	FH	VP	P	W	P	F
E ₃₁₁	H	FH	H	FH	M	G	F	P	VP	G
E ₃₁₂	FH	FH	H	FH	H	P	W	F	G	P
E ₃₁₃	H	H	H	H	H	W	W	F	F	P
E ₃₁₄	VH	H	VH	VH	H	F	G	P	P	P
E ₃₂₁	FH	H	FH	M	FH	E	VG	G	E	F
E ₃₂₂	H	H	H	H	H	E	E	VG	F	F
E ₃₂₃	FH	M	H	H	H	P	F	F	G	VG
E ₃₂₄	H	FH	H	H	M	E	VG	G	F	G
E ₃₂₅	H	H	H	H	H	E	F	G	F	F
E ₃₃₁	H	M	M	M	M	W	W	F	G	F
E ₃₃₂	FH	FH	H	H	FH	W	P	VP	G	F
E ₃₃₃	H	FH	FH	H	H	E	E	E	G	E
E ₃₃₄	H	M	FH	M	H	VP	P	F	G	F
E ₄₁₁	H	FH	FH	FH	M	F	G	P	F	G
E ₄₁₂	H	FH	H	FH	M	F	E	VG	G	F
E ₄₁₃	H	FH	FH	H	H	G	F	F	E	G
E ₄₁₄	H	H	FH	H	M	VP	F	F	F	F
E ₄₁₅	FH	M	M	FH	FH	E	G	G	F	F
E ₄₁₆	H	FH	H	FH	FH	G	E	F	F	VG
E ₄₁₇	M	M	M	FH	FH	W	W	W	F	P
E ₄₁₈	H	M	FH	FH	H	E	E	E	E	VG
E ₄₁₉	H	FH	H	H	H	E	VG	F	P	F
E ₄₂₁	M	H	M	M	FH	VG	G	F	E	E
E ₄₂₂	FH	FH	M	M	H	W	G	P	F	F
E ₄₂₃	H	M	M	M	FH	G	F	G	E	F
E ₄₂₄	FH	H	H	H	H	G	F	F	F	VG
E ₄₂₅	M	H	FH	H	H	E	G	F	P	VP
E ₄₂₆	FH	H	M	H	FH	W	VP	P	VP	F

- Step 3: Matching the linguistic terms with the corresponding fuzzy intervals [16,17]: Table 3 presents linguistic terms and its appropriate fuzzy numbers for weight and ratings assignment [15] chosen from literature [10,16].

Table 3: Fuzzy intervals for approximating linguistic terms (Adapted from [16])

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)

- Step 4: Application of fuzzy calculations on the weight and rating of each agile attribute, criteria and enabler. First, we

calculate the average fuzzy weight and average fuzzy performance rating of each agile attribute [10] as shown in the following example.

$$E_{111} \text{ average fuzzy weight} = [M+H+M+VL+VL]/5 = (0.3, 0.5, 0.7)/5, (0.7, 0.8, 0.9)/5, (0.3, 0.5, 0.7)/5, (0, 0.05, 0.15)/5, (0, 0.05, 0.15)/5 = (0.26, 0.38, 0.52)$$

$$E_{111} \text{ average fuzzy performance rating} = [G+F+F+F+VP]/5 = (5, 6.5, 8)/5, (3, 5, 7)/5, (3, 5, 7)/5, (3, 5, 7)/5, (1, 2, 3)/5 = (3.0, 4.7, 6.4)$$

Then, we calculate the rating of each agile criterion [10]. An example of this calculation is demonstrated below.

Example: Rating of the criterion E₁₁=

$$[(3.0, 4.7, 6.4) \otimes (0.26, 0.38, 0.52) \oplus (4.2, 5.9, 7.6) \otimes (0.62, 0.74, 0.86) \oplus (5.7, 7.1, 8.4) \otimes (0.36, 0.53, 0.70) \oplus (3.8, 5.6, 7.4) \otimes (0.3, 0.5, 0.7) \oplus (5.0, 6.5, 8.0) \otimes (0.46, 0.62, 0.78) \oplus (3.0, 4.7, 6.4) \otimes (0.54, 0.68, 0.82) \oplus (1.4, 2.6, 4.0) \otimes (0.50, 0.65, 0.80)] / [(0.26, 0.38, 0.52) \oplus (0.62, 0.74, 0.86) \oplus (0.36, 0.53, 0.70) \oplus (0.3, 0.5, 0.7) \oplus (0.46, 0.62, 0.78) \oplus (0.54, 0.68, 0.82) \oplus (0.50, 0.65, 0.80)] = (3.68, 5.27, 6.87)$$

Table 4: Fuzzy calculation of the criteria rating

Agile workforce	Agile workforce attributes	Average fuzzy performance rating	Average fuzzy weight	Criteria rating
E ₁₁	E ₁₁₁	(3.0, 4.7, 6.4)	(0.26, 0.38, 0.52)	(3.68, 5.27, 6.87)
	E ₁₁₂	(4.2, 5.9, 7.6)	(0.62, 0.74, 0.86)	
	E ₁₁₃	(5.7, 7.1, 8.4)	(0.36, 0.53, 0.70)	
	E ₁₁₄	(3.8, 5.6, 7.4)	(0.3, 0.5, 0.7)	
	E ₁₁₅	(5.0, 6.5, 8.0)	(0.46, 0.62, 0.78)	
	E ₁₁₆	(3.0, 4.7, 6.4)	(0.54, 0.68, 0.82)	
	E ₁₁₇	(1.4, 2.6, 4.0)	(0.50, 0.65, 0.80)	
E ₁₂	E ₁₂₁	(1.4, 2.6, 4.0)	(0.58, 0.71, 0.84)	(4.54, 5.97, 7.37)
	E ₁₂₂	(6.8, 8.0, 9.0)	(0.58, 0.71, 0.84)	
	E ₁₂₃	(5.0, 6.5, 8.0)	(0.76, 0.86, 0.94)	
	E ₁₂₄	(6.7, 8.0, 9.0)	(0.58, 0.71, 0.84)	
	E ₁₂₅	(5.4, 6.8, 8.2)	(0.62, 0.74, 0.86)	
	E ₁₂₆	(3.8, 5.6, 7.4)	(0.42, 0.59, 0.76)	
	E ₁₂₇	(5.0, 6.5, 8.0)	(0.46, 0.62, 0.78)	
	E ₁₂₈	(2.6, 4.1, 5.6)	(0.79, 0.89, 0.96)	
E ₂₁	E ₂₁₁	(3.2, 4.7, 6.3)	(0.54, 0.68, 0.82)	(2.46, 3.58, 4.79)
	E ₂₁₂	(4.6, 6.2, 7.8)	(0.66, 0.77, 0.88)	
	E ₂₁₃	(1.8, 2.9, 4.2)	(0.42, 0.59, 0.76)	
	E ₂₁₄	(2.0, 3.5, 5.0)	(0.50, 0.65, 0.80)	
	E ₂₁₅	(6.4, 7.7, 8.8)	(0.76, 0.86, 0.94)	
E ₂₂	E ₂₂₁	(1.8, 3.2, 4.7)	(0.58, 0.71, 0.84)	(3.83, 5.19, 6.53)
	E ₂₂₂	(5.3, 6.8, 8.2)	(0.24, 0.41, 0.58)	
	E ₂₂₃	(3.9, 5.0, 6.1)	(0.54, 0.68, 0.82)	
	E ₂₂₄	(3.1, 4.4, 5.7)	(0.50, 0.65, 0.80)	
	E ₂₂₅	(5.7, 7.1, 8.4)	(0.62, 0.74, 0.86)	
E ₂₃	E ₂₃₁	(4.2, 5.9, 7.6)	(0.52, 0.65, 0.78)	(4.49, 5.90, 7.25)
	E ₂₃₂	(4.9, 6.5, 8.0)	(0.40, 0.56, 0.72)	
	E ₂₃₃	(7.5, 8.6, 9.4)	(0.49, 0.62, 0.74)	
	E ₂₃₄	(2.0, 3.2, 4.5)	(0.71, 0.83, 0.92)	
E ₂₄	E ₂₄₁	(4.9, 6.2, 7.4)	(0.69, 0.80, 0.90)	(4.84, 6.20, 7.53)
	E ₂₄₂	(5.8, 7.1, 8.4)	(0.50, 0.65, 0.80)	
	E ₂₄₃	(3.8, 5.3, 6.8)	(0.50, 0.65, 0.80)	
E ₂₅	E ₂₅₁	(1.0, 2.0, 3.2)	(0.42, 0.59, 0.76)	(2.85, 4.03, 5.32)
	E ₂₅₂	(5.7, 7.1, 8.4)	(0.54, 0.68, 0.82)	
	E ₂₅₃	(2.6, 3.8, 5.2)	(0.58, 0.71, 0.84)	
	E ₂₅₄	(1.6, 2.9, 4.3)	(0.50, 0.65, 0.80)	
E ₃₁	E ₃₁₁	(3.2, 4.7, 6.2)	(0.54, 0.68, 0.82)	

E ₃₁	E ₃₁₂	(2.4, 3.8, 5.3)	(0.58, 0.71, 0.84)	(2.47, 3.94, 5.47)
	E ₃₁₃	(1.6, 2.9, 4.4)	(0.7, 0.8, 0.9)	
	E ₃₁₄	(2.8, 4.4, 6.0)	(0.79, 0.89, 0.96)	
E ₃₂	E ₃₂₁	(6.4, 7.7, 8.8)	(0.50, 0.65, 0.80)	(5.29, 6.78, 8.15)
	E ₃₂₂	(6.0, 7.4, 8.6)	(0.7, 0.8, 0.9)	
	E ₃₂₃	(4.0, 5.6, 7.2)	(0.58, 0.71, 0.84)	
	E ₃₂₄	(5.7, 7.1, 8.4)	(0.58, 0.71, 0.84)	
	E ₃₂₅	(4.5, 6.2, 7.8)	(0.7, 0.8, 0.9)	
E ₃₃	E ₃₃₁	(2.2, 3.5, 5.0)	(0.38, 0.56, 0.74)	(4.01, 5.22, 6.44)
	E ₃₃₂	(2.2, 3.5, 4.9)	(0.58, 0.71, 0.84)	
	E ₃₃₃	(7.8, 8.9, 9.6)	(0.62, 0.74, 0.86)	
	E ₃₃₄	(2.8, 4.4, 6.0)	(0.50, 0.65, 0.80)	
E ₄₁	E ₄₁₁	(3.6, 5.3, 7.0)	(0.50, 0.65, 0.80)	(4.62, 6.04, 7.41)
	E ₄₁₂	(5.3, 6.8, 8.2)	(0.54, 0.68, 0.82)	
	E ₄₁₃	(4.9, 6.5, 8.0)	(0.62, 0.74, 0.86)	
	E ₄₁₄	(2.6, 4.4, 6.2)	(0.58, 0.71, 0.84)	
	E ₄₁₅	(4.9, 6.5, 8.0)	(0.42, 0.59, 0.76)	
	E ₄₁₆	(5.3, 6.8, 8.2)	(0.58, 0.71, 0.84)	
	E ₄₁₇	(1.0, 2.0, 3.3)	(0.38, 0.56, 0.74)	
	E ₄₁₈	(8.2, 9.2, 9.8)	(0.54, 0.68, 0.82)	
E ₄₂	E ₄₁₉	(4.7, 6.2, 7.6)	(0.66, 0.77, 0.88)	(3.81, 5.30, 6.74)
	E ₄₂₁	(6.4, 7.7, 8.8)	(0.42, 0.59, 0.76)	
	E ₄₂₂	(2.6, 4.1, 5.7)	(0.46, 0.62, 0.78)	
	E ₄₂₃	(4.9, 6.5, 8.0)	(0.42, 0.59, 0.76)	
	E ₄₂₄	(4.2, 5.9, 7.6)	(0.66, 0.77, 0.88)	
	E ₄₂₅	(3.9, 5.3, 6.6)	(0.58, 0.71, 0.84)	
	E ₄₂₆	(1.4, 2.6, 3.9)	(0.54, 0.68, 0.82)	

After that, we calculate the rating of each agile enabler by using the same calculation method (Table 5). For example, the rating of the enabler E₁ is calculated as

$$E_1 = [(3.68, 5.27, 6.87) \otimes (0.5, 0.65, 0.8) \oplus (4.54, 5.97, 7.37) \otimes (0.7, 0.8, 0.9)] / [(0.5, 0.65, 0.8) \oplus (0.7, 0.8, 0.9)] = (4.18, 5.66, 7.13)$$

Table 5: Fuzzy calculation of the enablers rating

Agile workforce enablers	Agile workforce criteria	Criteria rating	Fuzzy importance weight of the agile workforce criteria	Enabler rating	Fuzzy importance weight of the agile workforce enablers
E ₁	E ₁₁	(3.68, 5.27, 6.87)	(0.5, 0.65, 0.8)	(4.18, 5.66, 7.13)	(0.5, 0.65, 0.8)
	E ₁₂	(4.54, 5.97, 7.37)	(0.7, 0.8, 0.9)		
E ₂	E ₂₁	(4.54, 5.97, 7.37)	(0.5, 0.65, 0.8)	(4.11, 5.46, 6.80)	(0.5, 0.65, 0.8)
	E ₂₂	(3.83, 5.19, 6.53)	(0.5, 0.65, 0.8)		
	E ₂₃	(4.49, 5.90, 7.25)	(0.5, 0.65, 0.8)		
	E ₂₄	(4.84, 6.20, 7.53)	(0.5, 0.65, 0.8)		
	E ₂₅	(2.85, 4.03, 5.32)	(0.5, 0.65, 0.8)		
E ₃	E ₃₁	(2.47, 3.94, 5.47)	(0.5, 0.65, 0.8)	(4.08, 5.42, 6.74)	(0.5, 0.65, 0.8)
	E ₃₂	(5.29, 6.78, 8.15)	(0.7, 0.8, 0.9)		
	E ₃₃	(4.01, 5.22, 6.44)	(0.5, 0.65, 0.8)		
E ₄	E ₄₁	(4.62, 6.04, 7.41)	(0.5, 0.65, 0.8)	(4.21, 5.67, 7.07)	(0.5, 0.65, 0.8)
	E ₄₂	(3.81, 5.30, 6.74)	(0.5, 0.65, 0.8)		

- Step 5: Calculation of the fuzzy workforce agility index of Indus_Comp [17]: Using the same calculation method, we obtained

$$\text{Fuzzy workforce agility index} = [(4.18, 5.66, 7.13) \otimes (0.5, 0.65, 0.8) \oplus (4.11, 5.46, 6.80) \otimes (0.5, 0.65, 0.8) \oplus (4.08, 5.42, 6.74) \otimes (0.5, 0.65, 0.8) \oplus (4.21, 5.67, 7.07) \otimes (0.5, 0.65, 0.8)] / [(0.5, 0.65, 0.8) \oplus (0.5, 0.65, 0.8) \oplus (0.5, 0.65, 0.8) \oplus (0.5, 0.65, 0.8)] = (4.14, 5.55, 6.93)$$

- Step 6: Matching the fuzzy workforce agility index with the appropriate linguistic label [16,17]: From the literature, we selected five linguistic labels (slowly agile, fairly agile, agile, very agile, extremely agile) (Table 6) [16]. Then, we convert the index into the suitable label using Euclidean distance method [10]. This method calculate the distance from the index to each agility level as shown below [16]:

Table 6: Linguistic labels and the corresponding fuzzy intervals (Adapted from [16])

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)

$$D(\text{FAI, Slowly Agile}) = \{(4.14-0)^2 + (5.55-1.5)^2 + (6.93-3)^2\}^{1/2} = 7.00$$

$$D(\text{FAI, Fairly Agile}) = \{(4.14-1.5)^2 + (5.55-3)^2 + (6.93-4.5)^2\}^{1/2} = 4.40$$

$$D(\text{FAI, Agile}) = \{(4.14-3.5)^2 + (5.55-5.0)^2 + (6.93-6.5)^2\}^{1/2} = 0.95$$

$$D(\text{FAI, Very Agile}) = \{(4.14-5.5)^2 + (5.55-7)^2 + (6.93-8.5)^2\}^{1/2} = 2.53$$

$$D(\text{FAI, Extremely Agile}) = \{(4.14-7)^2 + (5.55-8.5)^2 + (6.93-10)^2\}^{1/2} = 5.13$$

The minimum distance between the index and the linguistic label was obtained for the “agile” label [10], as it is showed in figure 1.

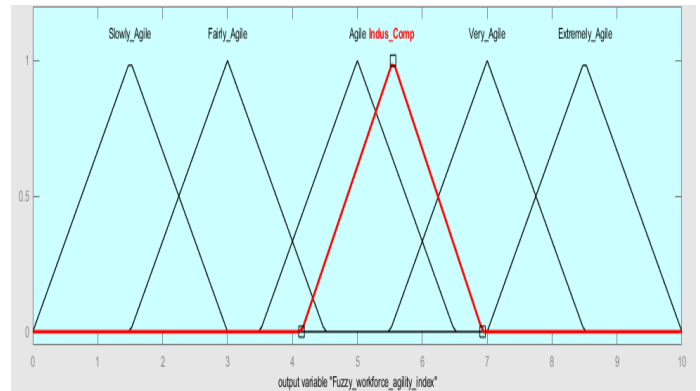


Figure 1: Fuzzy workforce agility index of Indus-Comp

- Step 7: Calculation of fuzzy performance importance index (FPII) and its score [10]: In order to improve the

agility level, we calculate the FPII of each agile workforce attribute. Then, the score of each FPII represents its contribution to the agility of workforce. The score of the 64 FPIIs are presented in table 7. If the score is weak it means that the attribute doesn't contribute to the overall workforce agility [16] and vice versa.

An example of FPII₁₁₁ and its score are calculated as:

$$FPII_{111} = [(1, 1, 1) - (0.26, 0.38, 0.52)] \otimes (3.0, 4.7, 6.4) = (2.22, 2.91, 3.07)$$

$$\text{Score of } FPII_{111} = (2.22 + 4 \times 2.91 + 3.07) / 6 = 2.82$$

Table 7: FPII and ranking score of the agile workforce attributes

Agile workforce attributes	FPII	Ranking score
E ₁₁₁	(2.22, 2.91, 3.07)	2.82
E ₁₁₂	(1.60, 1.53, 1.06)	1.46
E ₁₁₃	(3.65, 3.34, 2.52)	3.25
E ₁₁₄	(2.66, 2.80, 2.22)	2.68
E ₁₁₅	(2.70, 2.47, 1.76)	2.39
E ₁₁₆	(1.38, 1.50, 1.15)	1.42
E ₁₁₇	(0.70, 0.91, 0.80)	0.86
E ₁₂₁	(0.59, 0.75, 0.64)	0.70
E ₁₂₂	(2.86, 2.32, 1.44)	2.26
E ₁₂₃	(1.20, 0.91, 0.48)	0.89
E ₁₂₄	(2.81, 2.32, 1.44)	2.25
E ₁₂₅	(2.05, 1.77, 1.15)	1.71
E ₁₂₆	(2.20, 2.30, 1.78)	2.20
E ₁₂₇	(2.70, 2.47, 1.76)	2.39
E ₁₂₈	(0.55, 0.45, 0.22)	0.43
E ₂₁₁	(1.47, 1.50, 1.13)	1.43
E ₂₁₂	(1.56, 1.43, 0.94)	1.37
E ₂₁₃	(1.04, 1.19, 1.01)	1.13
E ₂₁₄	(1.00, 1.22, 1.00)	1.15
E ₂₁₅	(1.54, 1.08, 0.53)	1.06
E ₂₂₁	(0.76, 0.93, 0.75)	0.87
E ₂₂₂	(4.03, 4.01, 3.44)	3.92
E ₂₂₃	(1.79, 1.60, 1.10)	1.55
E ₂₂₄	(1.55, 1.54, 1.14)	1.47
E ₂₂₅	(2.17, 1.85, 1.18)	1.79
E ₂₃₁	(2.02, 2.06, 1.67)	1.99
E ₂₃₂	(2.94, 2.86, 2.24)	2.77
E ₂₃₃	(3.82, 3.27, 2.44)	3.22
E ₂₃₄	(0.58, 0.54, 0.36)	0.52
E ₂₄₁	(1.52, 1.24, 0.74)	1.20
E ₂₄₂	(2.90, 2.48, 1.68)	2.42
E ₂₄₃	(1.90, 1.85, 1.36)	1.78
E ₂₅₁	(0.58, 0.82, 0.77)	0.77
E ₂₅₂	(2.62, 2.27, 1.51)	2.20
E ₂₅₃	(1.09, 1.10, 0.83)	1.05
E ₂₅₄	(0.80, 1.01, 0.86)	0.95
E ₃₁₁	(1.47, 1.50, 1.12)	1.43
E ₃₁₂	(1.01, 1.10, 0.85)	1.04
E ₃₁₃	(0.48, 0.58, 0.44)	0.54
E ₃₁₄	(0.59, 0.48, 0.24)	0.46
E ₃₂₁	(3.20, 2.69, 1.76)	2.62
E ₃₂₂	(1.80, 1.48, 0.86)	1.43
E ₃₂₃	(1.68, 1.62, 1.15)	1.55
E ₃₂₄	(2.39, 2.06, 1.34)	1.99
E ₃₂₅	(1.35, 1.24, 0.78)	1.18
E ₃₃₁	(1.36, 1.54, 1.30)	1.47
E ₃₃₂	(0.92, 1.01, 0.78)	0.96
E ₃₃₃	(2.96, 2.31, 1.34)	2.26
E ₃₃₄	(1.40, 1.54, 1.20)	1.46

E ₄₁₁	(1.80, 1.85, 1.40)	1.77
E ₄₁₂	(2.44, 2.18, 1.48)	2.11
E ₄₁₃	(1.86, 1.69, 1.12)	1.62
E ₄₁₄	(1.09, 1.28, 0.99)	1.20
E ₄₁₅	(2.84, 2.66, 1.92)	2.57
E ₄₁₆	(2.23, 1.97, 1.31)	1.90
E ₄₁₇	(0.62, 0.88, 0.86)	0.83
E ₄₁₈	(3.77, 2.94, 1.76)	2.88
E ₄₁₉	(1.60, 1.43, 0.91)	1.37
E ₄₂₁	(3.71, 3.15, 2.11)	3.07
E ₄₂₂	(1.40, 1.56, 1.25)	1.48
E ₄₂₃	(2.84, 2.66, 1.92)	2.57
E ₄₂₄	(1.43, 1.36, 0.91)	1.30
E ₄₂₅	(1.64, 1.54, 1.06)	1.48
E ₄₂₆	(0.64, 0.83, 0.70)	0.78

After calculating the scores of the 64 FPIIs, the Indus_Comp management was consulted to decide the appropriate threshold in order to rank the attributes. Subsequently, threshold 1.1 was identified as the management scale [10,16], and sixteen attributes have a lower score than 1.1 (Table 7):

- Acceptance rate of the workforce to altered working time or work locations ;
- Tolerance to different or new approach in collaboration ;
- Creativity and innovation in problem-solving [3].
- Developing new skills or competencies within a short span ;
- Developing new management skills ;
- Continuously work towards gaining proficiency in multiple competency areas ;
- Shorter transition or recovery time ;
- Products or services delivery speed ;
- Problem-solving speed ;
- Set priorities to drive towards solutions ;
- Demonstrate a strong sense of urgency to deliver a set goal ;
- Demonstrate right focus ;
- Keep informed in order to achieve the objectives or clarify problems ;
- Pursue goals in the face of setbacks [6].
- Solution of the change related problems ;
- Tolerance to stressful and uncertain situations [2].

5. Conclusions

Our purpose was evaluating workforce agility. In the theoretical part we presented a conceptual model including four enablers, namely workforce status, workforce autonomy, workforce job and workforce involvement. In the empirical part we used the fuzzy logic. First, we collected data from decision-makers, then selected the appropriate fuzzy numbers for interpreting the linguistic variables, and calculated the fuzzy workforce agility index. Based on Euclidean distance computation, the workforce of Indus_Comp was agile. However,

by calculating FPII and ranking different scores we identified the attributes needing some improvement so that the workforce of Indus_Comp could be extremely agile.

The contribution of this work is to help the companies to evaluate their existing level of the workforce agility by using the fuzzy logic. This method has taken into account the ambiguity of each workforce agility enabler in order to calculate the overall workforce agility index. Then, the company could compare this index with that of competitors in order to identify its workforce attributes strengths and weaknesses.

Despite the above benefits, there are some limitations: In our case, the fuzzy logic depends mainly on the decision-maker only (one parameter) which excludes other employees perception who will have more detail on the technological or technical workforce attributes. Thus, taking into account other parameters can give a better tuning of results. Also, this method needs to be computerized in order to reduce time and errors that may be done in the calculation [15] or even to integrate the results provided by the fuzzy logic with other computing schemes like QFD (Quality Function Deployment), TISM (Total Interpretive Structural Modeling)...etc. For further research, this study should be conducted in different organizations of different sizes and sectors in order to carry out a comparative study. Apart from this, our conceptual model for workforce agility evaluation could also be extended in order to include other workforce aspects.

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] F. Tamtam, A. Tourabi, "Agile capabilities in Moroccan companies: Criteria and practices," in 2019 4th World Conference on Complex Systems (WCCS), Ouarzazate, Morocco, 2019. <https://doi.org/10.1109/ICoCS.2019.8930721>
- [2] B. Sherehiy, W. Karwowski, "The relationship between work organization and workforce agility in small manufacturing enterprises," *International Journal of Industrial Ergonomics*, **44**(3), 466–473, 2014. <https://doi.org/10.1016/j.ergon.2014.01.002>
- [3] R. Qin, D. A. Nembhard, "Workforce agility in operations management," *Surveys in Operations Research and Management Science*, **20**(2), 55–69, 2015. <https://doi.org/10.1016/j.sorms.2015.11.001>
- [4] R. Qin, D. A. Nembhard, "Workforce agility for stochastically diffused conditions—A real options perspective," *International Journal of Production Economics*, **125**(2), 324–334, 2010. <https://doi.org/10.1016/j.ijpe.2010.01.006>
- [5] K. Breu, C. J. Hemingway, M. Strathern, D. Bridger, "Workforce agility: the new employee strategy for the knowledge economy," *Journal of Information Technology*, **17**(1), 21–31, 2002. <https://doi.org/10.1080/02683960110132070>
- [6] M. Patil, M. Suresh, "Modelling the enablers of workforce agility in IoT projects: a TISM approach," *Global Journal of Flexible Systems Management*, **20**(2), 157–175, 2019. <https://doi.org/10.1007/s40171-019-00208-7>
- [7] 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>
- [8] A. Muduli, "Exploring the facilitators and mediators of workforce agility: an empirical study," *Management Research Review*, **39**(12), 1567–1586, 2016. <https://doi.org/10.1108/MRR-10-2015-0236>

- [9] C. L. Bosco, "The relationship between environmental turbulence, workforce agility and patient outcomes," Ph.D Thesis, The University of Arizona, 2007.
- [10] M. Suresh, R. Patri, "Agility assessment using fuzzy logic approach: a case of healthcare dispensary," *BMC Health Services Research*, **17**(1), 394, 2017. <https://doi.org/10.1186/s12913-017-2332-y>
- [11] S. Alavi, D. Abd. Wahab, N. Muhamad, B. Arbab Shirani, "Organic structure and organisational learning as the main antecedents of workforce agility," *International Journal of Production Research*, **52**(21), 6273–6295, 2014. <https://doi.org/10.1080/00207543.2014.919420>
- [12] 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>
- [13] C.-S. Tsai, C.-W. Chen, C.-T. Li, "Align agile drivers, capabilities and providers to achieve agility: a fuzzy-logic QFD approach," in *Supply Chain*, V. Kordic, Ed. I-Tech Education and Publishing, 2008.
- [14] C. K. M. Lee, C. T. Y. Ru, C. L. Yeung, K. L. Choy, W. H. Ip, "Analyze the healthcare service requirement using fuzzy QFD," *Computers in Industry*, **74**, 1–15, 2015. <https://doi.org/10.1016/j.compind.2015.08.005>
- [15] S. S. K. Sahu, "Agility evaluation in fuzzy environment," B.Tech Thesis, National Institute of Technology, 2013.
- [16] 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>
- [17] M. Deksnys, "Organizational agility in high growth companies," Ph.D Thesis, Mykolas Romeris University, 2018.

The Design of an Experimental Model for Deploying Home Area Network in Smart Grid

Fatima Lakrami^{*1}, Najib El Kamoun¹, Hind Sounni¹, Ouidad Albouidya¹, Khalid Zine-Dine²

¹STIC Laboratory, Department of Physics, Faculty of science, Chouaib Doukkali University, El Jadida, Morocco

²Computer Science Department, FSR, Mohammed V University, Rabat, Morocco

ARTICLE INFO

Article history:

Received: 03 January, 2020

Accepted: 10 March, 2020

Online: 18 June, 2020

Keywords:

Smart Grid

Wireless Sensor Network

Mobile Network

QoS

HAN

ABSTRACT

In the smart power grid, designing an efficient and reliable communication architecture has an important role in improving efficiency, and maintaining the connectivity of different network components. The home area network (HAN) provides an energy management system in houses since it enables home energy control and monitoring. So, it is imperative to determine a HAN configuration that provides low cost and better performance. We propose in this paper a conceptual and experimental model for designing a wireless sensor network to monitor a HAN in a Smart Grid. The novelty of this study is that it considers real object and network configuration. The use of ZigBee standard for wireless communication requires a meticulous physical and logical design to overcome limitations related to nodes density, traffic load by PAN, network topology and motion pattern. The current study allows us to characterize the best configuration of the intelligent studied system and to provide a perspective on addressing the problem of a node failure.

1 Introduction

Today, the proliferation of intelligent home appliances makes houses one of the most energy-consuming sectors in the world. So, optimizing consumption in terms of electrical energy has become a necessity, especially with the current depletion of non-renewable natural energy sources and the provision prohibiting polluting sources. The transition from the conventional electricity grid to the Smart Grid is dependent on the deployment of control and communication infrastructures, in parallel with electricity infrastructure from production sites to consumers via transmission and distribution networks. For this reason, the new electrical network uses computer technologies to improve the production, storage, and distribution of electricity. This allows these tools to be used along the entire chain of the power generation architecture, and in the concise term, makes it possible to minimize losses [1].

A Smart Grid is defined as an electrical system capable of integrating, intelligently, the actions of users, consumers, and producers to ensure efficient, economical, and secure electricity supply. With the integration of sophisticated power system electronics, communication technologies and networking, the Smart Grid is envisioned to considerably improve the existing electric grid. All elements of the production and distribution chain are affected by the performance of Smart Grids. Smart Grids consist of Power Generation System

(PGS), appliances, meters, transmission utilities, sensing devices and information gateways that operate in near real-time. Smart Grids are based on the integration of Renewable Energy Systems (RES) and distributed energy storage. Also, Smart Meter (SM) [2] is one of the key components in Smart Grids. The use of smart meters to collect metadata on electricity consumption across the entire network, using sensors, enable suppliers to provide an energy amount according to customer needs. It is more about a communicating electricity network that integrates NTICs (New Technologies of Information and Communication) in its operation, which allows interactions between electricity grid and buildings. The use of NTICs allows distributed and bidirectional management. In fact, with such integration, many optimization options become available to ensure an efficient electricity delivery, economically viable.

The wireless sensor networks (WSNs) have been recently widely recognized as a promising technology, deployed to improve various aspects of Smart Grids. WSNs collect different data (e.g., power production, consumption, weather conditions), these data are analyzed to control the power flow, to balance the power demand/response, to avoid problems of a power blackout and to build intelligent features into the existing electric network. The communication infrastructure of Smart Grid is a composite of three main hierarchical structures which include home area network (HAN), neighborhood area network (NAN) and wide area networks (WAN). A HAN manages

*Fatima Lakrami, Morocco, fatima.lakrami@gmail.com

principally consumers appliances, renewable energy resources, and storage systems. NAN is deployed to establish interaction between data concentrators and smart meters while WAN serves as a backbone for communication between control centers and bulk power system players.

A HAN enable to connect smart electronic devices, that can operate interactively and autonomously, via different wireless/wired protocols. Smart devices (sensors) are generally implemented (embedded) for monitoring and controlling the energy consumption of different forms of smart devices, such as smart cars, smart microwaves, smart refrigerators, or any smart home appliance.

This paper focuses on designing an experimental scheme for deploying a sensor network to monitor smart objects in a Home Area Network (HAN). It is a multidimensional model where several configurations are considered. The proposed scheme is evaluated according to: the type of logical and physical network topology, the traffic load depending on packet size and traffic distribution, the density of nodes and motion pattern.

The rest of the paper is organized as follows: In section II, related works along with their limitations are presented. In section III, we describe the architecture and functioning of the Smart Grid. Section IV studies the deployment of sensor networks for Home Area Networks in Smart Grid. In section V, we discuss the experimental design and evaluation of the performance of the proposed model. Finally, section VII concludes the paper.

2 Related Works

Research works on the use of sensor networks in the context of Smart Grid networks started a few years ago [3]-[6]. Authors in [7] and [8] present a tutorial on deploying sensors for data collection in Smart Grid. A detailed description of the monitoring process in Smart Grid from power generation to power storage along with the communication network architecture and its components is given by [9]. Authors in [10] present an interesting research that covers several important topics in sensing, communicating and managing a Smart Grid. Authors in [11] give a state-of-the-art of home area communications and networking technologies for energy management, authors discuss the challenges risen with the multitude of wireless standards that are currently used in HANs, causing a fragmentation in HAN market. However, Authors precise that ZigBee is still the most popular standard for wireless communication in HANs, due to several reasons: Low cost, low energy consumption etc. In [12] and [13] focus on studying analytically different communication protocols in HAN. Authors give a review of characteristics, limitations and challenges of ZigBee, Bluetooth, Wi-Fi, RFID, SPI and Home Plug. The comparison is performed according to : Data rate, frequency, coverage range, latency, battery life, security and others. Similarly, authors in [14] have demonstrated that ZigBee responds precisely to HAN requirement. However, several wireless technologies are needed to be combined to achieve data transmission in HAN, Authors give the example of ZigBee for household appliances to control and monitor systems and Wi-Fi for connecting digital products to communication equipment. To deal with the problem of heterogeneity of devices, authors in [15] propose the cognitive radio based communications infrastructure for the Smart

Grid. In [16] gives an overview of communication infrastructure in Smart Grid based on cognitive radio for spectrum regulations since it operates in the license-free bands and so, can be applied in the HAN, NAN and WAN.

Authors in [17] and [18] present a framework of Smart Grid with different challenges related to wireless WSN deployed to monitor smart grid. In [19], a review of wired and wireless communication technologies used in Smart Grid is presented, the authors compare the performance of different technologies and their applicability to different types of smart grid networks, the comparison was analytically based on the properties and characteristics of each technology.

In [20], a sequential scheme that controls the transmission sequence of wireless sensor nodes to achieve greater energy efficiency is proposed. The optimal combination of subsequent control level and transmission power is calculated according to different delay requirements. In [21] presents a survey of the challenges related to the deployment of Smart Grids. it proposes a literature review of the properties and constraints of Smart Grid simulation platforms; authors address eventually the new limitations of the deployment of WSN in Smart Grid applications. A Similar work is presented in [22] and [23], where authors study different communication schemes of sensors with smart meters integrated into a Smart Grid. The presented work focused on describing and discussing the technical feasibility of communication in the field of low-voltage consumer devices. For this purpose, two network models have been adopted: Flat and hierarchical. Authors have shown that the hierarchical model reduces packet loss for a high nodes density, although the transmission time is considerably increased, compared to the flat one.

In many research studies, different authors prove that the use of mobility can bring more benefits in improving the performance of a wireless network. In this context, several network management algorithms have been developed, taking advantage of the mobility of one or more particular network nodes. In [24] proposes a method to increase the connectivity time of nodes with the coordinator in IEEE 802.15.4 network, based on beacon exchanges. The proposed solution is based on measuring the signal strength of the received beacons to anticipate if the node moves towards the coordinator or is leaving its coverage area. Thus, the connectivity time is increased by choosing the coordinator with good signal strength, located in the area to which the mobile node is moving. Authors in [25] study different patterns of coordinator mobility in a ZigBee network. The mobile coordinator path can be one of the following trajectories : outer periphery, inner periphery, along a main diagonal, circular, and random path. Obtained results indicate that keeping the coordinator static gives the best performance, or if a trajectory must be chosen, then it should provide sufficient time for each segment of the network to collect all exchanged data. In [26], the authors implement a network node control module into its functional system. If needed, the module reconfigures the sensor to preserve its energy. Based on the same principle, [27] proposes to explore the mobility of the coordinator role to control routing processes for energy balancing between nodes, which provides opportunities for dynamic network reconfiguration. In [28], authors present a review of different approaches related to mobility in sensor networks, offering a comparative study of the various mobility protocols in WSN: MBC-based, MDC-based, they propose also an algorithm

to optimize the trajectory in the MDC. Another study was given by [29] where the performance sensor network is evaluated, for mesh and tree topologies, in the presence of mobility constraint and node failure. A more detailed study of the impact of mobility is given by [30] considering a high nodes density and two mobility schemes: octagonal and random. In [31] investigates the challenges and future perspectives of IoT (Internet of Everything) research in the Smart Grid, and the way it can contribute to improving various network functions for power generation, transmission, distribution, and final utilization. In the same context, a conceptual model for the Smart Grid within the Internet of Things is proposed by [32]. The proposed model is based on IPV6 as the backbone of the Smart Grid communications layer. Authors assume that every object must be assigned an IPV6 address while using a 6Lowpan communication protocol, however, the conceptual model proposed is not yet validated. In [33] provides a cross-layer data communication scheme target-aware based on cross-layer technique to enhance reliability and to reduce the latency in wireless sensor networks for IOT object communication in Smart Grid. The technique called TACT builds a set of nodes known as connected dominating set and each sensor node sends the data to the nodes present in the connected dominating set. TACT and FTACT approaches presented in this paper, outperform other solutions proposed by the literature like Default IEEE 802.15.4, DRX and FDRX, in terms of packet delivery ratio, latency, and data throughput.

3 Smart Grid Architecture and Functioning

The deployment of Smart Grids requires a diversity of technologies. This field of activity is located at the crossing of several sectors: Energy: producers, transporters, distributors and electricity suppliers, Electrical network equipment: manufacturers, electrical, and digital equipment.

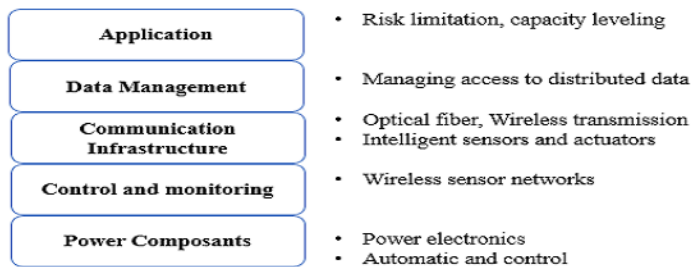


Figure 1: Smart Grid communication plans

Computer networks have a significant role in the interconnection of units in Smart Grid (e. g. Generators, substations, power plants, energy, etc... storage systems, smart meters...) exchanges concern no longer only energy but also information. This information can be a command for control, monitoring, or even maintenance of the consumption chain. In the case of homes (smart houses), grid voltage adjustment is generally managed by control modules equipped with wireless communication interfaces (sensors) and remotely controlled by a control center. Recently, there have been

some perspectives of integrating electric vehicles into the house electrical grid. With the development of Smart Grids, the concept of energy consumption in houses is no longer the same. Actually, the supply and demand of electrical energy are no longer centralized. A house can both consume and sell electricity to a neighbor house using a supply and demand procedure. For this reason, conceiving a distribution system to manage supply and demand is necessary. Several proposals have been reviewed in the literature: IDAPS [34] ADAPS [35] and [36]. At the network level, the data infrastructure, the generic model generally consists of a set of sensors that communicate with the smart meter via a home gateway. The information collected can carry several physical quantities. This information is essential to give an idea of the status of the house in the Smart Grid chain (consumer or consumer/producer) and also on the appropriate energy source to use in real-time. The communication architecture in the Smart Grid network is presented by a hierarchy of two primary levels: the central network (CN) and the access network (AN). The central system (CN) connects control centers, power plants, and the distribution network (DAN), also known as the wide-area network (WAN). Control centers take preventive measures when the electricity grid is in crisis, and, during regular hours, they monitor consumption and generate bills. The most favorable communication technology for communication in a core network is the optical fiber. The access network consists of a home network (HAN) and a neighborhood network (NAN). Home Area Network (HAN) is a multi-vendor network of home appliances that collects data on electricity consumption and transmits it. There is also the Advanced Metering Infrastructure (AMI), which supports fault recovery practices, power monitoring, and other miscellaneous supervision functionality.

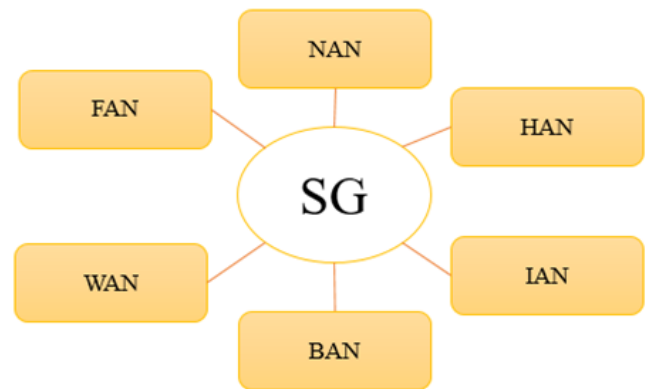


Figure 2: Smart Grid network types

NAN is the network that interconnects AMI HANs located in homes, industries, and businesses. It is also deployed for monitoring support and distribution substation networks. BAN and IAN networks refer to parallel HAN networks when they are implemented in companies/buildings or industrial areas, respectively. The most suitable communication in HAN/BAN/IAN networks is wireless technology. These networks are known for their easy implementation for a large number of users as well as their configuration, and therefore make it possible to supervise electricity consumption at home in a very efficient way. Due to the degree of similarity

between HAN, BAN, and IAN, we only present in the next sections HAN, NAN, and WAN [37].

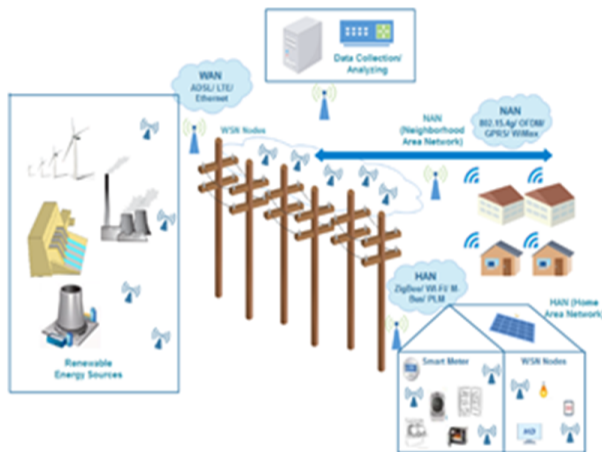


Figure 3: Smart Grid application

3.1 Home Area Network (HAN)

Home Area Network (HAN) connects electrical devices in the home, such as monitors, home appliances, or any intelligent equipment that uses electricity. HAN facilitates communication and interoperability among digital equipment in the house, it also provides better access to digital entertainment tools and increases productivity in plus of organizing the home security. The main challenge in implementing a HAN solution is to simultaneously connect objects inside the home/building while offering smart interoperability features, and to connect the entire home/building network to the outside world for remote monitoring and control. The HAN allows consumers to collect information on their consumption behavior [38] and the electricity usage costs via in-home monitoring devices.

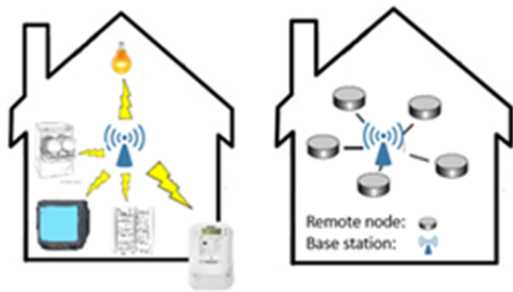


Figure 4: Example of a HAN

Due to the low-bandwidth requirements of HAN applications, it needs cost-effective communication technologies such as Bluetooth, WiFi, Zigbee, PLM. The energy gateway is needed for measurement, management, and communication of energy systems for smart homes and buildings (figure 4). This example design is a bridge between different communication interfaces, such as WiFi, Ethernet, ZigBee, or Bluetooth, ... In demand-side, power consumption is a significant parameter. The Smart Grid enables the two-way flow of electricity and information between supplier and consumer by using

the smart meters and the Advanced Metering Infrastructure (AMI) [39].

3.2 Neighborhood Area Network (NAN)

NAN is established between data collectors and smart meters in a neighborhood area. It is represented by a Wired/wireless connection from meter to the monitored device, either directly or via a data aggregator. To this end, short-range communication technologies can be used to collect the data measured by smart meters to be transmitted to the data concentrator. NAN is the network that interconnects the HAN AMIs that are in homes, industries, and businesses to the WAN. NAN is used for meter reading, demand response, remote disconnect for load control, local command messages, etc. and connects access network, monitoring support, and distribution substation networks. Wireless NAN can be based on Zigbee, Wifi, Wireless HART. One of the primary uses of the neighborhood area network is to communicate with smart meters. Often, we prefer to use meters from a single vendor to simplify communication between the service provider network and the house meter. However, by standardizing communication through the usage of an open and standardized protocols, meters from different vendors can now be interoperable [40].

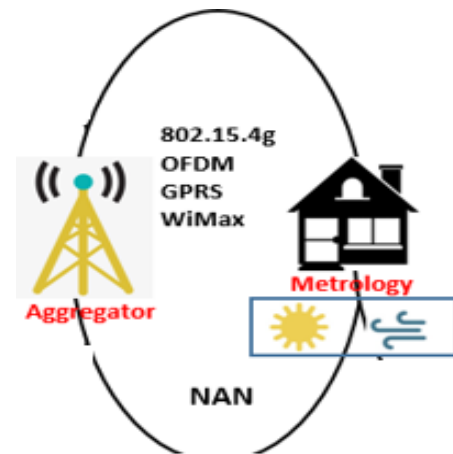


Figure 5: Example of a NAN

The next section discusses the advantages of deploying sensors in smart grid.

3.3 Wide Area Network (WAN)

The WAN creates a communication path between the service provider's data center and data concentrators. It is a high bandwidth and robust two-way communication network, that can handle long-distance data transmissions for Smart Grid monitoring and applications control. In general, WAN networks adopts mainly the communication technologies that provide the best coverage with the lowest cost, such as LTE, cellular networks (2G/3G systems), fiber, power line communication networks. However, other technologies can also be deployed, such as WiMAX, LoRaWAN, NB-IoT, Sigfox.

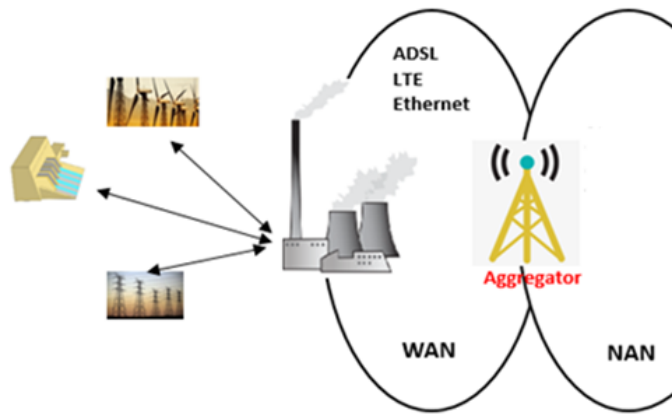


Figure 6: Example of WAN

4 Key Benefits of Deploying WNS in Smart Grid

The market for residential energy management is now increasing dramatically due the growing of Wireless sensor network technology field. So, more and more applications are using this technology. In fact, today's electronic and computer advances made it possible to develop tiny sensors that can perform variant tasks. WSN can now sense the environment, calculate outputs based on the collected data and send results through to the network to a gateway or a smart meter. The IEEE 802.15.4 standard provides three topologies: star, P2P. The system formed is called PAN (Personal Area Network), each PAN has a unique ID and one or several coordinators. Above 802.15.4 [41], the ZigBee network layer allows the creation of more complex networks such as mesh or tree networks through automatic level 3 routing, AODV [42] for example, is used to route data in WSN mesh network.

HAN networks represent a typical deployment of sensor networks in the Smart Grid field. With the arrival of the IoT, home or domestic appliances are now equipped with integrated sensors for monitoring their energy consumption, temperature, or any other data. The development of the embedded system supports such perspectives, especially with the proliferation of home appliances that dominate a large fragment of the electronics market. Smart houses will contain a large number of devices, and therefore will require continuous monitoring of their power consumption for the proper functioning of a Smart Grid system.

The deployment of HAN come is in response to several communication requirements, the network must transport the sensed data/control signal from/to the sensor or the gateway or any device within the network while considering the requirements that should be met when a service is provided. The performance of the network must achieve to enable energy management applications to monitor and control the devices on the home network. In a similar way to LANs, HANs design a term used generally to describe all the intelligence and activity that occurs in HEM systems [43], the local area in LAN is a home for a HAN [44]. the HAN connects devices that are capable of sending and receiving signals from a meter, a

gateway or other wireless device. A typical HAN may consist of the following basic functional components: Node controller (a coordinator), gateway, end devices (sensor), smart meter. Objects in HAN can be connected through Wired or wireless technologies, there are trade-offs that involve power consumption, signaling distance, sensitivity to interference, and security. The main point here is that HANs are not energy management applications, but represent a communication infrastructure for exchanging sensed data (sensors) and commands (actuators) to ensure a real-time monitoring and controlling of energy consumption within a house.

Recent smart homes are equipped with smart appliances and each appliance is considered as a thing (object), each object can be assigned a unique ID and can communicate with the grid to upload and download data and commands from different distant devices. WSN is the most suitable option for HAN, NAN, WAN and smart micro grid applications for integration and operation of renewable energy sources. The Internet of things (IoT) applications are becoming one of the emerging Smart Grid enabling technologies. Several works are now considering HAN component as IoT assets (Referred to Related works section), and develop solutions for enhancing communication and data transmission over the Smart Grid. The ZigBee [45] wireless technology is classified a low cost technology that allow wide deployment in wireless control and monitoring applications. It has been proven that it offers a low power usage and then a longer life with smaller batteries, compared with other wireless standards. In plus, the mesh networking provides high reliability, redundancy and broader range. ZigBee Home and Building Automation focus on enabling smart homes and buildings that can control appliances, lighting, environment, energy management, and security as well as expand to connect with other ZigBee networks. This is the reason that make this technology fit well into the HAN area.

5 Methodology and Proposed Model

If users aim to manage the energy they consume efficiently, they will require a precise and accurate measurements of their energy consumption. Monitoring objects in HAN, implies deploying a WSN architecture. The most popular standard for communication is ZigBee. In ZigBee standard, the real coverage can reach 50 meters. Each end device can communicate with its neighbor (sensor/Coordinator/Router...) within this distance. The disposition of objects in HAN can impact the degree of reachability according to the covered area and the presence of obstacles. Data converges toward a residential gateway that can serve as a smart meter energy gateway if it supports the same PHYs as handled devices and smart meters. HAN must deal also with other issues, like packet loss due to either failure of coordinator or unreachable destination due to the depletion or a node mobility. In plus, consumers are now aware about the depletion of the natural resources, and the need of deploying all the possible effort for the energy conservation, this is why they need to get a real-time visibility of their energy consumption. Therefore, electrical power consumption monitoring on a real-time basis is very essential to avoid exceeding the critical demand level and also to reduce the latency in detecting accidents in the electric power systems or control systems. So, the network must employ its

resources efficiently, to deliver packets in a timely manner.

Designing a HAN is an evolving and changing task. The first step consists on planning the communications architecture to be deployed. The current emphasis on mesh radio technology and the availability of different wireless protocols (ZigBee, RFID, Z-wave, and so on) with different transmission proprieties creates both opportunity and a certain complexity to define the appropriate one (s) to be deployed in a HAN. In fact, designing a communication backbone may depend on several system and data requirements. In this work, we consider further the cost and ease of implementation along with the failure management and the level of QoS provided by the network, for this reason, the proposed model is based on Zigbee.

ZigBee is a bidirectional radio frequency (RF) wireless network standard. The standard operates on the IEEE 802.15 for physical radio specification, it is specially designed for sensors and control device. It is designed for devices requiring low-latency, low-energy consumption and lower bandwidth, it operates in the license-free frequency for short range [46]. However, the multitude of wireless standards that are currently used in HANs cause a fragmentation in HAN market; ZigBee is the most suitable and most implemented protocol in HANs. It is used in most home appliances. ZigBee defined recently an application layer standard for smart energy for HANs with intention of low-cost devices and low energy usage [47].

In order to handle the sensibility of the transported data, the HAN, must provide different levels of data delivery criticality depending on the needs of the Application and supports varied latency requirements messages communicated between various points

within the Smart Grid.

The model designed in this work, considers different dimensions: density, mobility, traffic load/distribution and network topology. Different type of sensors are used for collecting (Energy, Temperature, Humidity..). We design the network according to a real house scheme. The network studied in this work uses ZigBee standard for wireless communication, the design of the ZigBee supports star, mesh, and tree topology, so the tree networks are simulated and evaluated. Each room presents a separate PAN, with a supervised area that covers 1500m2. The density of nodes varies according to the nature of the room, so, there is a specific density per PAN. Two types of Traffic distribution are configured: Uniform and personalized. A detailed description of the experiment details is presented by figure 7.

Several performance metrics are considered for evaluating the experiment. Different measures are:

E2E Delay: Represents the duration between the time a message is queued for transmission at the physical layer until the last bit is received by the destination node.

Throughput: In this study, it represents the amount of data transmitted correctly from the source to the destination in an interval. It is quantified with various factors, including packet collisions, obstacles between nodes, and the type of the topology.

Loss Rate : The packet loss rate is the percentage of packets lost during data transmission. It is a criterion for measuring the quality of a network's services. Is calculated by the following equation (1):

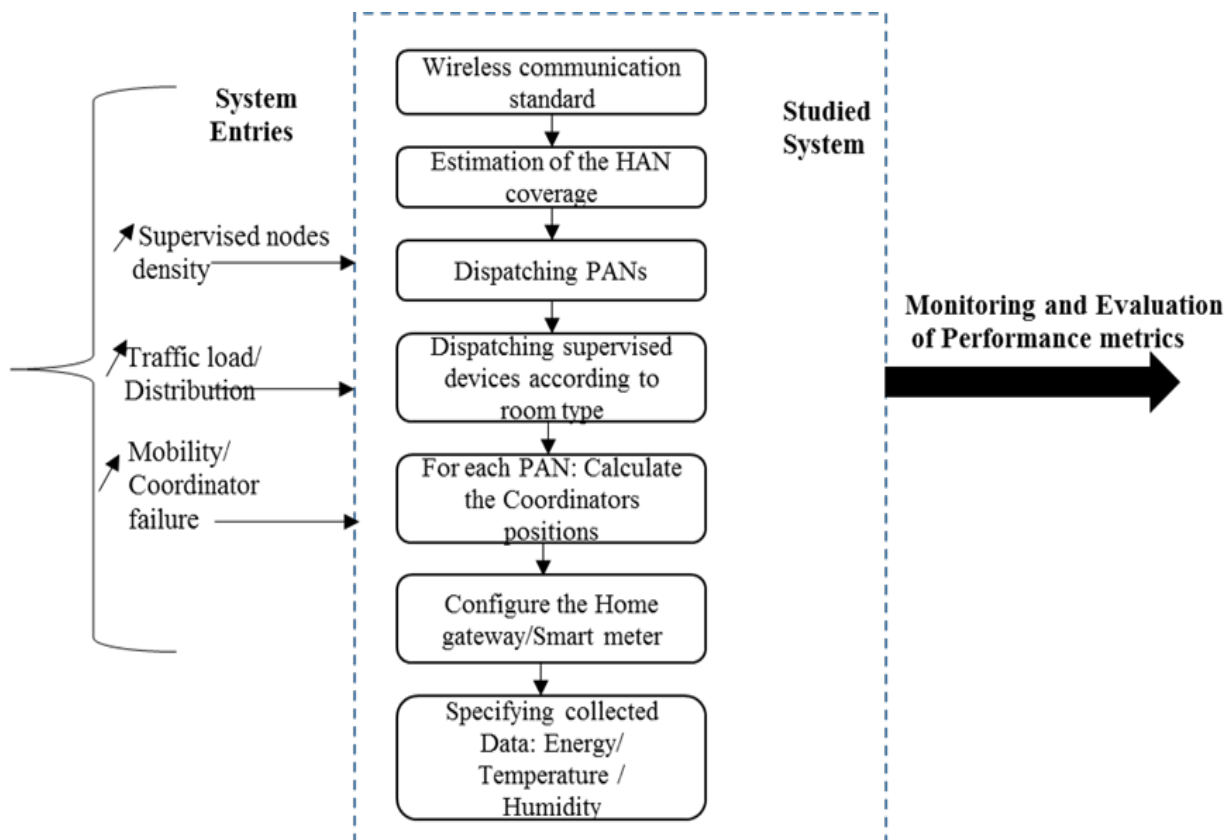


Figure 7: The experimental Model architecture

$$T(\%) = \frac{(T_e - T_r)}{T_e} \times 100 \quad (1)$$

$$RE_i = E_0 - \left(\sum_t E_{tx} + \sum_t E_{rx} \right) \quad (2)$$

$$E_{tx} = P_{cs} \times Packet_size \quad (3)$$

$$E_{rx} = P_{cr} \times Packet_size \quad (4)$$

$L_f = \min(L_0, \dots, L_m)$ with $L_m = E_i / Energy_consumption_per_time_unit$

P_{cs} and P_{cr} are weighting coefficients

We implement in this work a light application that enable nodes to calculate their residual energy to be communicated to a controller charged of collecting data. The controller is also responsible of monitoring and evaluating obtained data. The table below summarizes the configuration parameters:

Table 1: Network Configuration

Recording Time	900s
Traffic Destination	Random
Traffic Distribution	Uniform/ Personalized
Number of nodes	from 2 to 25
Mobility Speed	from 1 to 6m/sec
Other sensors type	Humidity, Electricity, Temperature
Wireless Technology	Zigbee

For each experiment, we fixed the topology and increase the impact of density, mobility and traffic load. First, and while varying density and motion pattern in the network, we fixed a constant traffic load of 6 packets per second with fixed size to be routed between randomly selected end-devices (figure 8). To increase the traffic load, we increase the number of packets sent per second, the traffic distribution is varied according to the specific need of the experiment. All sensor nodes are working in the 2.4 GHz band.

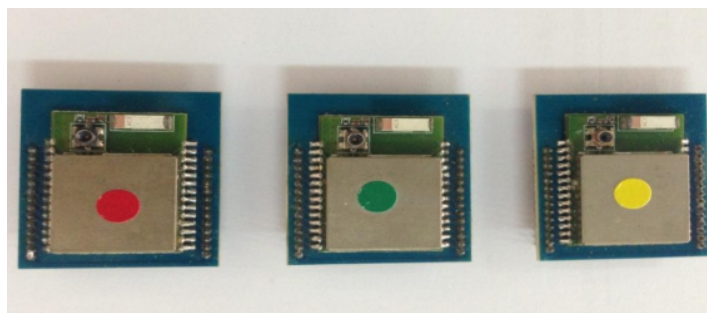


Figure 8: Zigbee modules for Coordinator, sensor and router



Figure 9: The Zigbee platform of the deployed WSN

The very sensitive role of HANs in the connectivity of objects in the smart grid requires prior planning of the deployed network for better performance. In the second part of the experiment, we study the case of mobility, firstly as a network optimization scheme, and secondly as a solution for nodes failure. Mobility can provide significant benefits to the studied network if it is well planned in accordance with the disposition of the nodes in a home network, as well as the nature of the deployed devices. The HAN presented by figure 14 is made of 3 separated PANs: PAN1, PAN2 and PAN3.

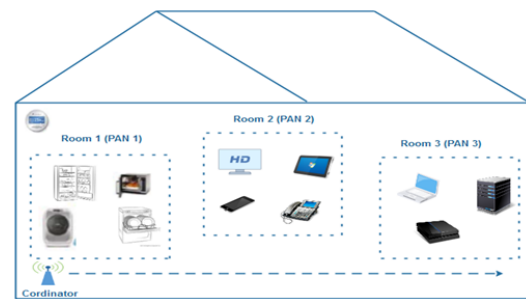


Figure 10: Network topology for studying the impact of mobility

Devices are able to move while having an auto-assigned PAN. To give more flexibility to the network, we define two types of coordinators mobility:

- **Determinist travel:** the coordinator has a cyclical operating mode. The trajectory as well as the pause time are fixed.
- **Random displacement:** the coordinator's position is defined by the user, or generated randomly.

The purpose of this part is to study the behavior of the HAN network, in the case of the presence of one or more mobile coordinators and end devices. The network consists of 3 distinct PANs; each PAN represents a room of a house with an approximate dimension of 200m2. It is imperative to ensure that the 3 PANs coverage does not interfere. In the first scenario, we consider that at a precise time of the day, the coordinator moves to collect data from different nodes (sensors/routers) with different speeds, when the coordinator reaches the nearest node/router of a given PAN it continues its trajectory with no pause time. In the second case of study, the coordinator stops for a pause time, to collect information on the electrical consumption of the supervised devices; the pause time varies as the coordinator speed increases. We consider four mobility patterns:

Circular, Octagonal, pre-defined trajectory, and Random. The pre-defined trajectory consists of moving the coordinator directly from one PAN to another, the start point represent, for example, the center of the supervised area of the house. Each PAN is composed of 5 sensors and 3 routers. While the coordinator moves, two devices from PAN1 also leave their positions towards other PANs with a moderated speed that do not exceed for all scenarios 0.5m/sec. In real context, sensors can be generally embedded in the electronic circuit of the smart home appliance or added manually. They can be static or mobile. For example, for a vacuum cleaner or even an electric vehicle or any device with the capacity to roam over different PANs of the same HAN network, the mobility scheme can vary over time. Here, the coordinator moves according to four patterns already defined above, we use a play car with remote control to transport the coordinator through different PANs and according to different trajectories. We plan to determine the speed and the corresponding pause time, depending on the distribution of the traffic sent, and also the best mobility scheme according to sensors distribution in a HAN. The main objective is to guarantee a reliable network service. The traffic distribution function is also configured to match the coordinator's crossover, this implies that each node is set according to a specific traffic pattern function, based on the coordinator traveling.

6 Results and discussion

6.1 Results related to varying supervised devices density and Traffic Impact for different topologies

We present in this section results related to varying supervised nodes density and traffic load for different topologies. Figures 11, 12, 13 and 14 present successively Loss Rate, Network Lifetime, throughput and delay:

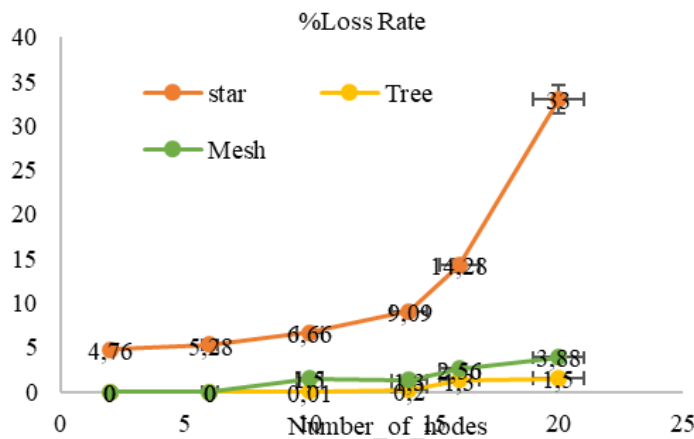


Figure 11: Loss Rate in function of number of Nodes

By examining loss rate results, we can notice that the lowest loss rate is obtained successively for the tree and the mesh network. All nodes in the network are configured to send the collected data at a specific time. In a star topology, the network is controlled by a single coordinator that is responsible for initiating and main-

taining all other devices. As the number of nodes in the network increases, the coordinator becomes saturated with the amount of data received. when the coordinator's buffer is saturated, it will proceed to delete the new incoming packets. In a tree or mesh topology, the coordinator is responsible for starting up the network, but communications can go through intermediate routers, which will create a differential arrival time that will provide additional time to the coordinator to process packets, and then the loss rate (dropped packets) is significantly reduced.

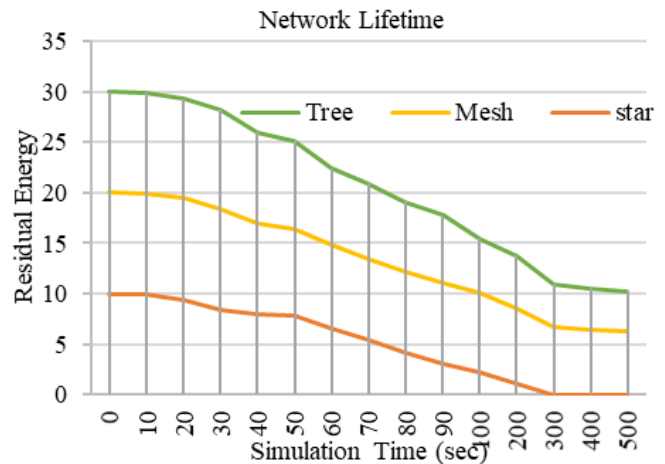


Figure 12: Network Lifetime in function of number of Nodes

Network lifetime is defined as the duration from the time the network is operational until the first node of the WSN depletes its energy completely. Therefore, the lifetime of a sensor node in the WSN depends essentially on the energy consumption of the node with time. At the beginning of each experiment, we set the energy level of the (chargeable) batteries of different nodes in the network to the same initial value. By observing the results of the figure 11, we notice that the first node that runs out is the coordinator of the star topology. A solution to solve this problem is to set up different intervals for sending the data. The mesh network exhausts second. Indeed, when configuring the mesh network, we have placed the coordinators in such a way that they are accessible via several routes. By default, the shortest path is used, which leads to the rapid depletion of one of the sensors.

The throughput (figure 13) represents the quantity of useful data received by all the nodes of the network. The tree topology offers the highest throughput, which is in accordance with the result of the loss rate. In fact, the physical architecture of the network and routing scheme provides a more reliable data transmission. For a reduced number of nodes, star topology outperforms mesh topology. The data routing mechanism in a star topology is simple and efficient, and remains the most appropriate for a HAN network. However, this mechanism becomes less efficient when the number of managed nodes becomes very high.

When the number of nodes is small, the shortest delay (figure 14) is achieved for the star topology, while the mesh topology will result in the highest delay. When the number of nodes increases, the tree topology responds better to the evolution of the network. We can deduce, that in the practical case of a HAN deployed on

a large scale, it is the tree topology that is the best adapted. The star topology remains the best solution for the interconnection of a reduced number of nodes. The deployment of sensor networks in other contexts may give better performance using a mesh routing.

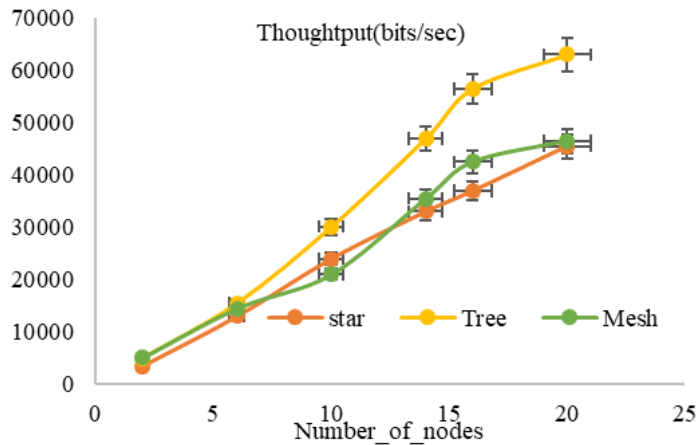


Figure 13: Throughput in function of number of Nodes

When the number of nodes is small, the shortest delay (figure 14) is achieved for the star topology, while the mesh topology will result in the highest delay. When the number of nodes increases, the tree topology responds better to the evolution of the network. We can deduce, that in the practical case of a HAN deployed on a large scale, it is the tree topology that is the best adapted. The star topology remains the best solution for the interconnection of a reduced number of nodes. The deployment of sensor networks in other contexts may give better performance using a mesh routing.

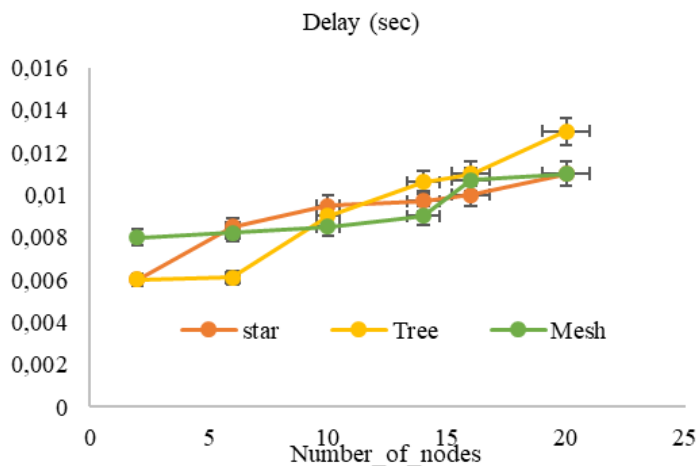


Figure 14: Delay in function of number of Nodes

In the following section, we propose to divide the domestic network into several PANs, to use a single mobile coordinator that moves between different PANs using a play car with a remote control during specific times of the day to collect information from different sensors. The next section studies the different possible mobility patterns for a coordinator, in terms of speed, pause time, and

also the movement trajectory. The results of the previous section demonstrate that the tree topology is very suitable for deploying sensors to monitor building, for this reason, the tree routing will be deployed in the next evaluation related to mobility.

In general, ZigBee networks are more designed to work in mesh. However, results presented in figure 11, 12, 13, 14 show that the tree topology offers better results for different performance metrics. We can notice that the tree topology provides a higher bandwidth. So it is revealed to be more suitable for medium density networks, like houses. The packets follow the path through the tree topology to the destination, which is revealed very practical in a HAN network, where the distribution of nodes is already defined. Also, it does not require any routing tables to send the packet to the destination. Despite this, it does not offer redundancy, like mesh topology, so in case of a node failure, the entire network is compromised. Although, in a house building and term of financial cost, the star topology remains the most appropriate, especially if the network is flat and covers a restricted area. But, in reality, and given the size of an area that can be covered by the peripheral in a HAN, the coordinator cannot access all the supervised devices. The tree topology remains practical if the building, for example, is with several floors and one or several routers is/are available to deliver data to the coordinator qualified as a "root". Although mesh network offers several advantages like: data can be transmitted simultaneously between nodes, there is a high redundancy, expansion and modification can be done without interrupting the network, but the overall cost is high and practically not suitable in HANs, giving the node density and the physical environment where they are deployed.

6.2 Results related to Motion pattern impact for different network topologies and traffic distribution without pause time

Results presented in figures [15,18], show the performance study of the network according to different mobility schemes. We consider two cases, with and without pause time. Performance metrics are given according to E2E delay, throughput, Loss Rate, network lifetime, and average connectivity duration to each PAN during recording time. Figures [15-18] bellow represent the case without pause time, while figures [19-22] resumes result for the case with pause time.

The lowest loss rate (figure 15) is obtained when the coordinator follows the predefined path, which is configured in such a way that it crosses the centers of the different PANs via the shortest paths. the different PANs contain sensor nodes, routers and also sleeping coordinators. Sleeping coordinators may represent here a state of failure. In the predefined trajectory the mobile sensor is configured to reach the different positions at different intervals, the traffic distribution is also configured to suit the crossing of the coordinator. The octagonal and circular trajectories also show a low loss rate compared to the random trajectory.

Figure 16 shows that the shortest delay of the forwarding of the useful data is also obtained successively for the predefined trajectory, circular and finally the octagonal. This is mainly due to the distance between the coordinator and the sensor/router nodes for each mobility scheme.

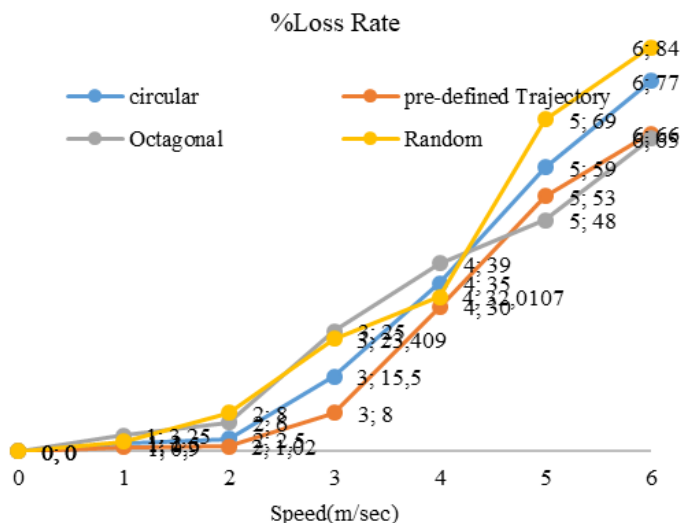


Figure 15: Percentage of Loss rate in function of increasing Coordinator speed without Pause time

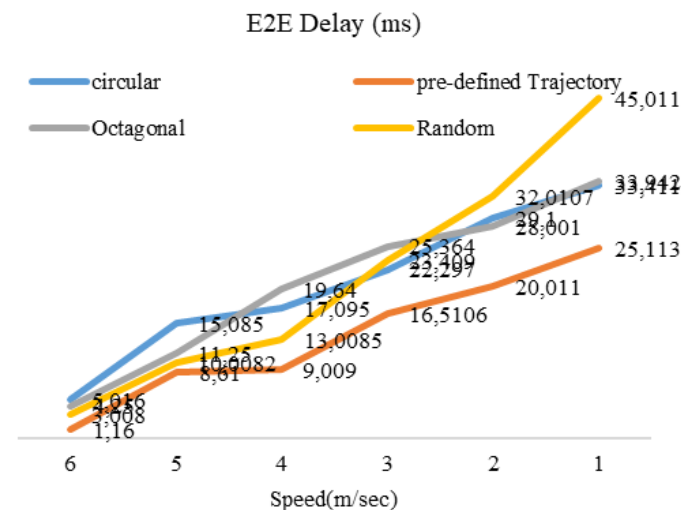


Figure 16: End to End delay(ms) in function of increasing Coordinator speed without Pause time

Regarding the lifetime of the network shown in the figure 17, presented by the average residual energy remaining at the end of the experiment, we can notice that following the predefined trajectory, it is the octagonal trajectory that offers a higher lifetime compared to the circular and the random trajectories. The octagonal trajectory allows the coordinator to reach the closest positions to the centers of the PANs being traversed, which means that at the time the data are sent, the coordinator is within the radius of coverage of a PANi, and the particularity of the trajectory is that the coordinator takes a significant amount of time to traverse the PANi, and given its speed, the coordinator reaches the next PAN before the start of data transmission. As a result, the nodes do not quickly deplete as a consequence of the increase in their transmission power in order to reach the coordinator in case of a distant coordinator.

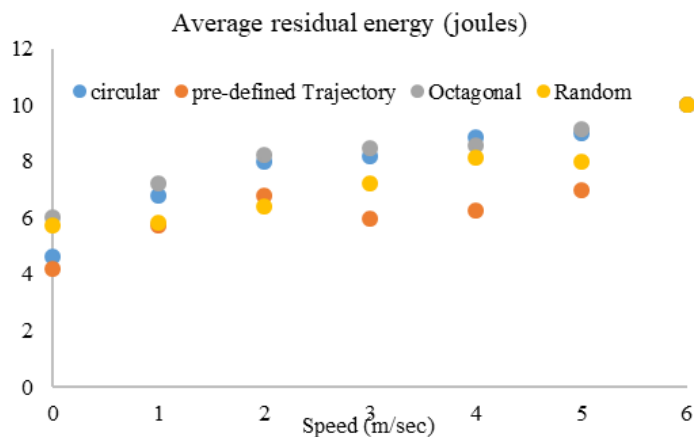


Figure 17: Average of residual energy in function of increasing Coordinator speed without Pause time

Figure 18 shows the average connectivity time to the coordinator per PAN. the highest time is obtained for the predefined trajectory and the octagonal trajectory. the same interpretation as for the case of the residual energy is valid also in this case.

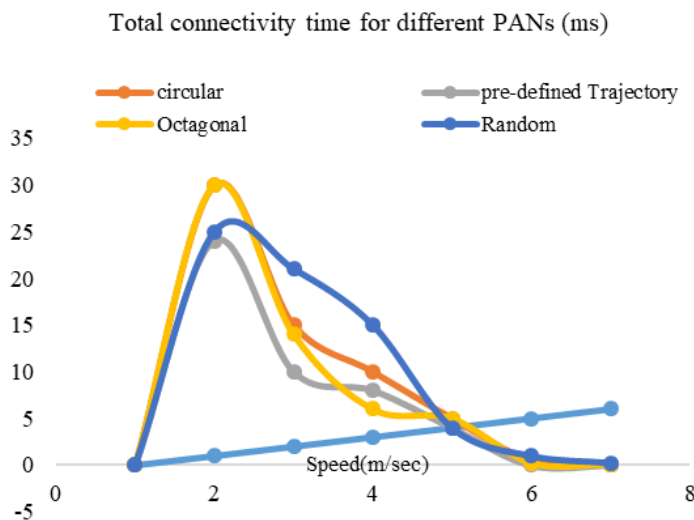


Figure 18: Total connectivity time for a coordinator to different PANs in function of increasing Coordinator speed without Pause time

6.3 Results related to Motion pattern impact for different network topologies and traffic distribution with pause time

The figure 19 shows the average connectivity time to the coordinator per PAN. the highest time is obtained for the predefined and the circular trajectories, especially for high speeds. The result was significantly improved compared to experiments without pause time. the average connectivity has increased by 20% after the introduction of the pause time. The coordinator stops for a moment before continuing his movement, which allows nodes to maintain their connectivity with the coordinator for a longer period of time.

In the meantime, as in the case of mobility experiments without pause time, the short delay (figure 20) of the useful data is obtained

successively for the predefined trajectory, circular, octagonal and finally the random. We can also add here that the coordinator stops for a period of time that is proportional to its travel speed during data collection.

life expectancy of the wireless sensor network by leveling the difference in the remaining energy between all the elements of the network. If the coordinator is far away from a given PAN at the moment of sending the information, the emitter node deploys more energy by increasing the transmission power to reach the coordinator, which depletes the network quickly.

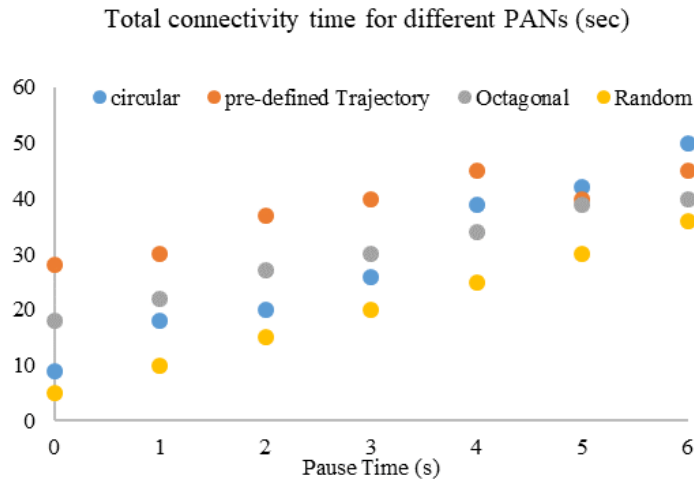


Figure 19: Total connectivity time for a coordinator to different PANs in function of increasing Coordinator speed with Pause time

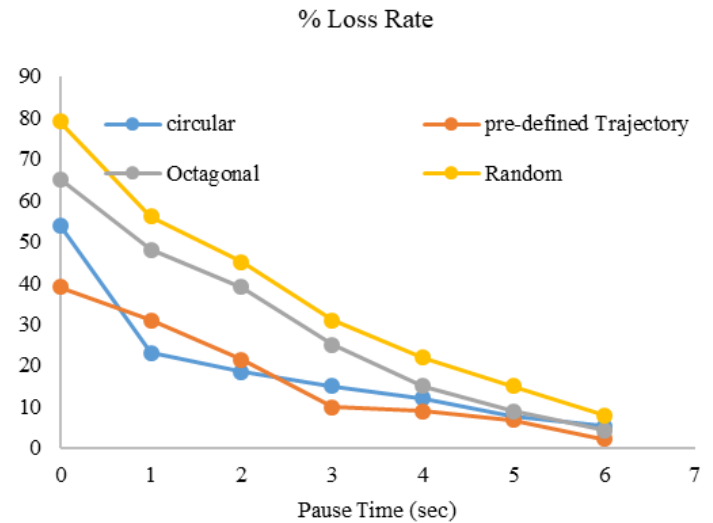


Figure 21: Percentage of Loss rate in function of increasing Coordinator speed with Pause time

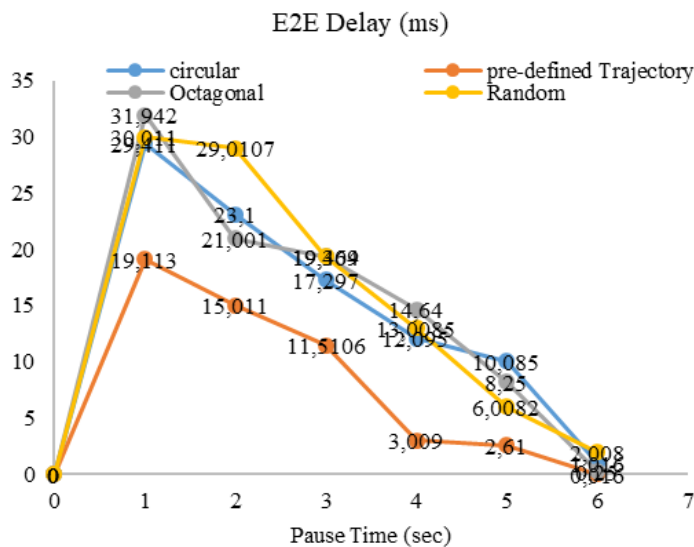


Figure 20: Total connectivity time for a coordinator to different PANs in function of increasing Coordinator speed with Pause time

Referring to figure 21, we can notice that the predefined and circular trajectory results in the best performance in term of loss rate. The results have significantly improved compared to the case without pause time. In the previous experiments, the coordinator moves without stopping, and as the speed increases, the loss rate increases, since the coordinator does not have enough time to collect the data and then transmit it to the home gateway.

By referring to figure 22, we can conclude that indeed, the methodology of managing the coordinator mobility increases the

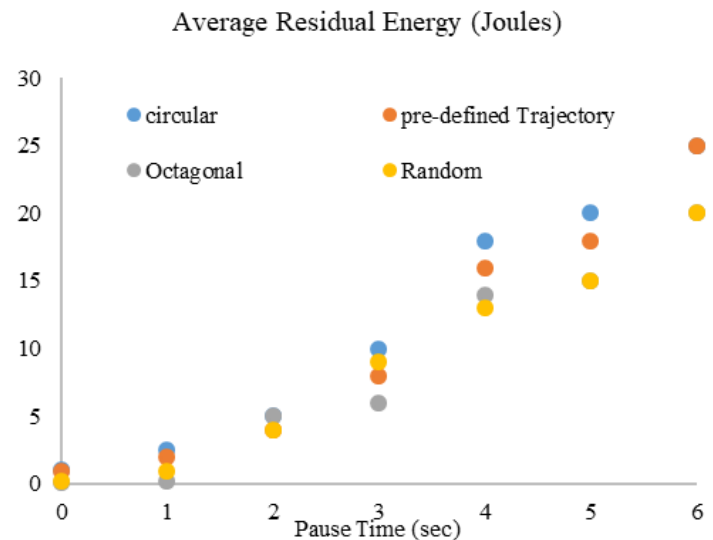


Figure 22: Average of residual energy in function of increasing Coordinator speed with Pause time

Figures [19- 22] have shown the performance results of different mobility schemes of a single coordinator for different speeds, with and without pause time. According to the obtained results, the best network performance is obtained respectively for: the predefined trajectory, circular, octagonal, and finally random path, although the latter gives the highest throughput for both the transmitter and the receiver. The circular configuration shows better performance

than the octagonal configuration, especially for high speeds. The proposed ZigBee network works very efficiently using the circular configuration mobility model. This model can even be used in a ZigBee network on a large scale. The introduction of a pause time during the coordinator movement improves significantly the network performance. The pause time was configured to allow the coordinator to remain in each PAN for a duration that increases proportionately to the mobility speed. According to the obtained results, we can conclude that, for reasonably high speed, it is imperative to have a higher pause time, to get better performances. The deployment of an adequate mobility scheme can solve several problems, related to cost nodes failure and communication centralization. However, the mobility scheme must take into account several parameters and constraints related to: traffic distribution and load, physical properties such as transmission and reception power, and also the size of the coordinator's MAC queue. We intend in our future researches to exploit, the results obtained for mobility to resolve the problem of node failure, in the case of networks where several coordinators can be deployed. A recovery algorithm will be implemented in a centralized server than manage nodes position via a monitoring system (SCADA for example), the node can be either a smart gateway, smart meter or even an SDN server. The presented study enables us to extract numerous experimental results about several configurations like: Nodes disposition, traffic distribution and motion pattern. The accuracy of the extracted results is very important since they will be used in configuring a HAN for smart grid monitoring, to ensure better performances and high efficiency, through controlling the system entries instead of setting them randomly.

7 Conclusion and Perspectives

Wireless sensor networks have greatly facilitated the different operating processes of Smart Grids. However, the diversity of characteristics of different WSN applications means that they have different transmission performance requirements. As a result, several challenges may arise depending on various application contexts, increasing the difficulty of deploying and maintaining the Smart Grid. In this paper, we have experimented with the different operating modes of HAN by considering several constraints related to the physical/logical installation of the network and also to the process of collecting and exchanging information. We started by comparing the performance of routing schemes in a ZigBee network, the obtained results attested that Mesh routing offers the best throughput, while the tree topology is more efficient for the rest of the metrics and is suitable for medium nodes density. For optimization purposes, we propose to exploit the mobility of a single coordinator to collect data from sensing nodes to be forwarded to the home gateway or directly transferred to the smart meter. We demonstrate that rapid movement during the communication process reduces transmission delay with an increase in network lifetime as long as an optimal pause time is configured according to the trajectory, mobility speed, and also traffic distribution function. In future work, we plan to implement a testbed of a HAN, that implements a centralized solution for configuring and managing nodes consumption and communication through the network for a better efficiency. The

studied solution can be implemented either on a smart meter or a separate device as an SDN server. Managing nodes failure will also be addressed.

References

- [1] Uslar, Mathias, et al, "Applying the smart grid architecture model for designing and validating system-of-systems in the power and energy domain: A European perspective" *Energies* 12.2 (2019): 258. <https://doi.org/10.3390/en12020258>
- [2] Depuru, Soma Shekara Sreenadh Reddy, et al, "Smart meters for power grid-Challenges, issues, advantages and status" in 2011 IEEE/PES Power Systems Conference and Exposition. IEEE, 2011.
- [3] M. El Brak and M. Essaïdi, "Wireless sensor network in smart grid technology: Challenges and opportunities" in 2012 6th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT). IEEE, 2012.
- [4] McBee, D. Kerry, G. Marcelo, "Utilizing a smart grid monitoring system to improve voltage quality of customers" *IEEE Transactions on Smart Grid* 3.2 (2012): 738-743.
- [5] Luan, Wenpeng, D. Sharp, S. Lancashire, "Smart grid communication network capacity planning for power utilities" *IEEE PES T and D 2010*. IEEE, 2010.
- [6] Erol-Kantarci, Melike, T. Hussein, "Suresense: sustainable wireless rechargeable sensor networks for the smart grid" *IEEE Wireless Communications* 19.3 (2012): 30-36.
- [7] Le Cam, Vincent, et al. "Applications des réseaux de capteurs intelligents et de la communication sans fil l'instrumentation des structures de génie civil." (2008).
- [8] Chen, Shu-yong, et al, "Survey on smart grid technology" *Power system technology* 33.8 (2009): 1-7.
- [9] Fan, Zhong, et al, "Smart grid communications: Overview of research challenges, solutions, and standardization activities", *IEEE Communications Surveys and Tutorials* 15.1 (2012): 21-38.
- [10] Kayastha, Nipendra, et al, "Smart grid sensor data collection, communication, and networking: a tutorial", *Wireless communications and mobile computing* 14.11 (2014): 1055-1087.
- [11] Ghasempour, Alireza, "Internet of Things in Smart Grid: Architecture, Applications, Services, Key Technologies, and Challenges" *Inventions* 4.1 (2019): 22.
- [12] Devidas, A. Remanidevi, M. V. Ramesh, "Wireless smart grid design for monitoring and optimizing electric transmission in India" in 2010 Fourth International Conference on Sensor Technologies and Applications. IEEE, 2010.
- [13] Kailas, Aravind, V. Cecchi, A. Mukherjee, "A survey of contemporary technologies for Smart Home energy management" *Handbook of green information and communication systems*. Waltham, USA: Elsevier (2012): 35-56.
- [14] Hafeez, Ayesha, et al, "Smart home area networks protocols within the smart grid context" *journal of Communications* 9.9 (2014): 665-671.
- [15] Kuzlu, Murat, M. Pipattanasomporn, S. Rahman, "Communication network requirements for major smart grid applications in HAN, NAN and WAN" *Computer Networks* 67 (2014): 74-88.
- [16] Zaballos, Agustín, A. Vallejo, J. M. Selga, "Heterogeneous communication architecture for the smart grid" *IEEE network* 25.5 (2011): 30-37.
- [17] Sharma, Priya, G. Pandove, "A Review Article on Wireless Sensor Network in Smart Grid" *International Journal of Advanced research in computer science* 8.5 (2017).
- [18] L. Li, et al, "The applications of wifi-based wireless sensor network in internet of things and smart grid" 2011 6th IEEE Conference on Industrial Electronics and Applications. IEEE, 2011.
- [19] S. Pranesh, J. Pyun, "Mobility support in IEEE 802.15. 4 based mobile sensor network" *IEICE Transactions on Communications* 97.3 (2014): 555-563.
- [20] L. Chhaya, et al, "Wireless sensor network based smart grid communications: cyber attacks, intrusion detection system and topology control" *Electronics* 6.1 (2017): 5.
- [21] A. Jurenoks, M. Boronowsky, "Dynamic Coordinator Mobility Management Methodology for Balancing Energy Consumption in the Wireless Sensor Network" *Procedia Computer Science* 77 (2015): 176-183.

- [22] A. Jurenoks, and Dejan Jokib, "Coordinator Role Mobility Method for Increasing the Life Expectancy of Wireless Sensor Networks" *Applied Computer Systems* 21.1 (2017): 46-51.
- [23] E. Eylem, Y. Gu, D. Bozdag, "Mobility-based communication in wireless sensor networks" *IEEE Communications Magazine* 44.7 (2006): 56-62.
- [24] N. Shantanu, et al, "Investigation and performance analysis of some implemented features of the ZigBee protocol and IEEE 802.15. 4 Mac specification" *International Journal of Online and Biomedical Engineering (iJOE)* 13.01 (2017): 14-32.
- [25] I. Nazrul, et al, "Mobility Issue on Octagonal Structured ZigBee Network Using Riverbed" *International Journal of Communications, Network and System Sciences* 9.3 (2016): 55-66.
- [26] H. Daki, et al, "Big Data management in smart grid: concepts, requirements and implementation", *Journal of Big Data* 4.1 (2017): 1-19.
- [27] Y. Saleem, et al, "Internet of things-aided Smart Grid: technologies, architectures, applications, prototypes, and future research directions" *IEEE Access* 7 (2019): 62962-63003.
- [28] S. Emiliano, et al, "Industrial internet of things: Challenges, opportunities, and directions" *IEEE Transactions on Industrial Informatics* 14.11 (2018): 4724-4734.
- [29] D. Amarnath, S. Sujatha, "Internet-of-Things-aided energy management in smart grid environment" *The Journal of Supercomputing* (2018): 1-13.
- [30] A. Manimuthu, R. Ramesh, "Privacy and data security for grid-connected home area network using Internet of Things" *IET Networks* 7.6 (2018): 445-452.
- [31] A. Jaffar, et al, "Wireless sensor network design for smart grids and Internet of things for ambient living using cross-layer techniques" *International Journal of Distributed Sensor Networks* 15.7 (2019): 1550147719862208.
- [32] Y. Wang, et al, "Review of smart meter data analytics: Applications, methodologies, and challenges" *IEEE Transactions on Smart Grid* 10.3 (2018): 3125-3148.
- [33] K. Gram-Hanssen, S. J. Darby, "Home is where the smart is? Evaluating smart home research and approaches against the concept of home" *Energy Research and Social Science* 37 (2018): 94-101..
- [34] B. Benaouda, "La communication sans fil dans un rseau lectrique intelligent (smart grid): mthodologie de developpement" (2013).
- [35] Luque-Ayala, Andrs, S. Marvin, "Developing a critical understanding of smart urbanism" *Handbook of Urban Geography*. Edward Elgar Publishing, 2019.
- [36] Uslar, Mathias, et al, "Applying the smart grid architecture model for designing and validating system-of-systems in the power and energy domain: A European perspective" *Energies* 12.2 (2019): 258.
- [37] Mahmood, Anzar, N. Javaid, S. Razzaq, "A review of wireless communications for smart grid" *Renewable and sustainable energy reviews* 41 (2015): 248-260.
- [38] Wang, Xiaojun, et al, "A joint routing and channel assignment in multi-radio multi-channel wireless mesh networks" *International Journal of Sensor Networks* 24.3 (2017): 173-182.
- [39] Saputro, Nico, K. Akkaya, "Investigation of smart meter data reporting strategies for optimized performance in smart grid AMI networks" *IEEE Internet of Things Journal* 4.4 (2017): 894-904.
- [40] Farahani, Shahin. *ZigBee wireless networks and transceivers*. Newnes, 2011.
- [41] etin, Grcan, "IEEE 802.15. 4/ZigBee based wireless sensor network design for monitoring server rooms" in *2017 International Conference on Computer Science and Engineering (UBMK)*. IEEE, 2017.
- [42] Zheng, Jianliang, J. Myung, "A comprehensive performance study of IEEE 802.15. 4" *Sensor network operations* 4 (2006): 218-237.
- [43] Perkins, Charles, E. Belding-Royer, Samir. Das, "RFC3561: Ad hoc on-demand distance vector (AODV) routing" (2003).
- [44] Mahapatra, Bandana, A. Nayyar, "Home energy management system (HEMS): concept, architecture, infrastructure, challenges and energy management schemes" *Energy Systems* (2019): 1-27.
- [45] Foley, F. Peter, "Home area network system and method" U.S. Patent No. 6,069,899. 30 May 2000.
- [46] D. Amarnath, S. Sujatha, " Internet-of-Things-aided energy management in smart grid environment" *The Journal of Supercomputing*, 1-13, 2018.
- [47] Deshmukh, D. Akshay,B. Shinde, "A low cost environment monitoring system using raspberry Pi and arduino with Zigbee" in *2016 International Conference on Inventive Computation Technologies (ICICT)*. Vol. 3. IEEE, 2016.

Prognosis of Failure Events Based on Labeled Temporal Petri Nets

Redouane Kanazy^{1,2,*}, Samir Chafik¹, Eric Niel²

¹Pluridisciplinary Research and Innovation Laboratory, EMSI Casablanca, 20000, Morocco

²Ampere Laboratory, CNRS, INSA Lyon, Villeurbanne, 69621, France

ARTICLE INFO

Article history:

Received: 21 February, 2020

Accepted: 30 April, 2020

Online: 18 June, 2020

Keywords:

Prognosis

Prognoser

Discrete events systems

Operating modes

Temporal labeled Petri net

ABSTRACT

To reduce the risk of accidental system shutdowns, we propose to control system developers (supervisor, SCADA) a prediction tool to determine the occurrence date of an imminent failure event. The existing approaches report the rate of occurrence of a future failure event (stochastic method), but do not provide an estimation date of its occurrence. The date estimation allows to define the system repair date before a failure occurs. Thus, provide visibility into the future evolution of the system. The approach consists in modelling the operating modes of the system (nominal, degraded, failed); the goal is to follow the evolution of the system to detect its degradation (switching from nominal to degraded mode). When degradation is reported, a prognoser is generated to identify all possible sequences and more precisely those ending with a failure event. then it checks among the sequences (with failure event) which ones are prognosable. The last step of the approach is carried out in two parts: the first part consists in calculating the execution time of the so-called prognosable sequences (by optimizing the number of possible states and resolving an inequalities system). The second part makes it possible to find the minimum execution (the earliest occurrence of a failure event).

1. Introduction

The supervision applications provided to control system developers (in manufacturing production, robotics, logistics, vehicle traffic, communication networks or IT) make it possible to report the detection of a dysfunction or accidental shutdown of the system and locate its origin. The discrete event systems (DES) community has developed diagnostic methods that focus on the logical, dynamic or temporal sequence of failure events that cause this dysfunction. However, the criticality of some systems and their complexity require a method of the failure events prognosis, to report their occurrence in advance in order to avoid any damage caused by a failure.

The challenge is therefore to prevent the future occurrence of a failure event. However, which suitable modeling tool is required for this system? And knowing that more the complexity of the system increases, more its state of space increase. So, how to overcome this problem of combinatorial explosion? And what are the prognosis limits?

Several fault prognosis methods have been developed; some have adopted a stochastic approach [1] [2] [3] while others have

chosen non-stochastic [4]-[6], one for state automaton or Petri Net. These approaches are interested in prediction of failure m-steps in advance, based on a stochastic process. However, their assessment is difficult and probabilistic information is not always realistic. Others propose a prognosis approach [7] that consists of giving occurrence rates of possible traces that end with a failure event. These approaches indicate the occurrence of a future failure event, but do not specify its occurrence date. The possible occurrence date of a failure event makes it possible to plan the intervention date to repair the system before a failure occurs and thus provides visibility into the future evolution status of the system.

The challenge of each group working on this topic is to predict perfectly the future reality. [8] introduces the notion of signature of a trace, which consists to use several formal systems devoted to the description of event signature and the recognition of behaviors, called chronic. This concept has been used in diagnostic work [9], [10] and is based on error detection, localization, evaluation, recognition and response. [11] proposed a method for calculating the execution time of a trace, but it is still diagnosis-oriented. The development of a new approach of the temporal prognosis requires a modeling tool that allows the time constraints of the system (temporal prognosis) while using labels (it involves predicting an

*Redouane Kanazy, Email: redouane.kanazy@insa-lyon.fr

www.astesj.com

<https://dx.doi.org/10.25046/aj050354>

event over time). An extension of the Petri nets offers this possibility. These Petri nets are called, the Temporal Labelled Petri net (TLPN for short).

The aim is to propose a correct control of a system subject to unforeseen failures. The existing studies use the logical order of failure events occurrences to make the prognosis. In this paper, we are not only interested in the logical order of events, but also in the date of their occurrence. We assume that the system accepts three possible operating modes (nominal, degraded, and failed one). The events occurrence allows the system to switch between these modes. The event occurrence dates allow the synchronization of state switching in the model. A delay occurrence of an event, for example, can be explained by a degradation of the system. Approach's based only of a logic events occurrence cannot detect this delay. Hence the interest of a time-based prognosis approach.

Two contributions are proposed in this paper. The first one is concerned with the formal representation and the second one with the methodology of prognosis calculation. Indeed, the model is based on a TLPN. The association of events to temporal transitions will be presented. The evolution from one mode to another one will be represented by transitions firing. The firing of each transition depends on the occurrence of an event and corresponding occurrence date. The second contribution relates to the methodology of the prognosis. A prognoser is built from the TLPN model. It is an oriented state graph, which identifies all possible sequences namely those that end in a failure event. But before predicting a failure event, it is important to make sure that it is possible to do it. That's why we introduced the prognosability property whose objective is to determine the sequences ending with a failure event. Such event is called prognosable, the goal is to predict the earliest date of failure event occurrence. To calculate the execution time of these sequences and optimize the number of possible states, the resolution of an inequalities system based on works of [11]-[13] is used. The idea is to find the set of minimum values solution of the inequalities system. These values will constitute the minimum time after which the occurrence of the failure event is sure.

The paper is organized as follows: the second section is devoted to the basic concepts of Petri Nets (PN). The third section introduces temporal PNs (according to Berthomieu [14]-[19] and Popova [11]-[13], [20]-[22]). The fourth section focuses on labelled PN. In the fifth section, we discuss time-labelled PN to verify the prognosis approach in the sixth section. Thus, in this last section, the formal approach of our proposal will be presented, with an algorithm for predicting a temporal failure event and a case study, with explanations.

2. Preliminary

2.1. Petri Nets (PN)

A PN structure is a 4-uplet $R = \langle P, T, Pre, Post \rangle$ given by:

- P is a nonempty finite set of places $\{p_1, p_2, \dots, p_n\}$.
- T is a nonempty finite set of transitions $\{t, t_2, \dots, t_m\}$ with $P \cap T = \emptyset$.
- Pre is the backward incidence function that assigns to each couple (p, t) of places and transitions a non-negative integer.

$Pre: P \times T \rightarrow \mathbb{N}$, $Pre(p, t) = \omega$ is the value of the arc weight arc from the place p to the transition t .

- $Post$ is the forward incidence function that assigns to each couple (t, p) of transitions and places a non-negative integer. $Post: T \times P \rightarrow \mathbb{N}$, $Post(p, t) = \omega$ is the value of the arc weight arc from the transition t to place p .

The initial marking m_0 is an application:

$m_0: P \rightarrow \mathbb{N}$, it is labeled as an initial global system state. A marked net system $R_m = \langle R, m_0 \rangle$ is a net R with an initial marking m_0 . When the transition t is enabled, it then would be fired. From the marking m , the firing of the t leads to the new marking m' denoted by $m[t > m']$.

- The symbol $\bullet t_j$ denotes the set of all places p_i such that $Pre(p_i, t_j) \neq 0$ and t_j^* the set of all places p_i such that $post(p_i, t_j) \neq 0$. Analogously, $\bullet p_i$ denotes the set of all transitions t_j such that $post(p_i, t_j) \neq 0$ and p_i^* the set of all transitions t_j such that $Pre(p_i, t_j) \neq 0$.

2.2. Temporal Petri Nets (TPN)

Temporal Petri Nets TPN are introduced in [5], then studied by [16], [20]-[26].

Definition 1:

A Temporal Petri Net (TPN) is a net $R_T = \langle P, T, Pre, Post, m_0, I \rangle$, in which $\langle P, T, Pre, Post, m_0 \rangle$, is a marked Petri Net R_m , and $I: T \rightarrow \mathbb{Q}^+ \times (\mathbb{Q}^+ \cup \{\infty\})$ is a static firing time interval function which assigns a firing static interval $[T_{min}, T_{max}]$ to each transition t , with $T_{min} \leq T_{max}$ (T_{max} can be infinite) and \mathbb{Q}^+ is all positive or zero rational numbers.

An enabled transition t can be fired at time τ if the time elapsed since the activation of t belongs to the interval $I(t) [T_{min}, T_{max}]$.

If $T_{min} = T_{max} = 0$ (i.e., $I(t) = [0, 0]$) the transition t is called immediate otherwise it is called timed.

Thus, we can divide the set T of transitions into two subsets T^t and T^{im} [27] where T^t is the set of timed transitions and T^{im} is the set of immediate transitions with: $T^t \cap T^{im} = \emptyset$ and $T^t \cup T^{im} = T$

The aim of this distinction is to determine the firing priorities of the transitions. Firing T^{im} transitions has a higher priority than firing T^t transitions.

2.2.1. Behavior, states and reachability relation

Definition 2:

According to [1], a state of a temporal net is a pair $E = (m, I)$ in which m is a marking and the application I associates a firing temporal interval to each transition.

The initial state consists of the initial marking m_0 and the application I_0 which associates to each enabled transition its static firing temporal interval, $E_0 = (m_0, I_0)$, such that:

$$\text{if } m_0 \geq Pre(\bullet, t) \text{ then } I_0(t) = I(t) \text{ otherwise } I_0(t) = \emptyset. \quad (1)$$

Transition t may fire iff it remains logically enabled for a time interval θ included in $[T_{min}, T_{max}]$. θ is the amount of time that has elapsed since the transition t is enabled.

A transition t is enabled in a state

$$E = (m, I) \text{ iff: } m_0 \geq Pre(\bullet, t) \wedge \theta \in I(t) \wedge \forall t_k \neq t, m \geq Pre(\bullet, t_k) \Rightarrow \theta \leq \max(I(t_k)) \quad (2)$$

From E , the result of the t firing is as usual the new state $E' = (m', I')$ such that:

- 1) $m' = m + \Delta t$ with $\Delta t = Post(\bullet, t) - Pre(\bullet, t)$
- 2) For each transition t_k :
 - If t_k is not enabled by m' , then $I'(t_k) = \emptyset$;
 - If t_k is distinct from t , enabled by m , and not in conflict with t , then $I'(t_k) = [\max(0, \min(I(t_k)) - \theta), \max(I(t_k)) - \theta]$
 - Otherwise $I'(t_k) = I(t_k)$

According to [11], a state of an TPN is a pair $E = (m, h)$ in which m is a place marking (noted $p_marking$) and h is a clock vector (of dimension equal to the number of network transitions) that corresponds to the transition markings (noted $t_marking$). Thus, the $p_marking$ describes the situation of the places and $t_marking$ that of the transitions. Such a pair ($p_marking, t_marking$) describes a TPN status.

Definition 3:

Let R_T be an TPN.

- A $p_marking$ in R_T is a function $m: P \rightarrow \mathbb{N}$. A $p_marking$ in TPN is also a marking in a untimed PN.
- A $t_marking$ in R_T is a function $h: T \rightarrow \mathbb{R} \cup \{\$\}$ \$ means that the transition is not enabled.

Definition 4:

Let $R_T = \langle P, T, Pre, Post, m_0, I \rangle$ a TPN, m is a $p_marking$ and h is a $t_marking$. The pair $E = (m, h)$ is called a state in R_T if and only if:

- 1- m is a marking accessible in R .
- 2- $\forall t (t \in T \wedge Pre(\bullet, t) \not\leq m \rightarrow h(t) = \$)$.
- 3- $\forall t (t \in T \wedge Pre(\bullet, t) \leq m \rightarrow h(t) \in \mathbb{R}_0^+ \wedge h(t) \leq \max(t))$ where $\max(t)$ is the latest firing time of t .

Definition 4 shows that each transition t has a clock. If t is not enabled by the marking m , the associated clock is not activated (sign \$), If t is enabled by m , the clock of t indicates the time elapsed since the last activation of t .

The initial state is given by $E_0 = (m_0, h_0)$ avec

$$h_0(t) = \begin{cases} 0 & \text{if } Pre(\bullet, t) \leq m_0 \\ \$ & \text{otherwise} \end{cases} \quad (3)$$

A transition t is fireable from state $E = (m; h)$ (noted $E[t>$) if and only if $Pre(\bullet, t) \leq m$ and $h(t) \geq \min(t)$; (3) where $\min(t)$ is the earliest firing time of t . The firing of t leads R_T to a new state $E' = (m', h')$ (noted $E[t>E'$)

In general, each TPN has an infinite number of states, depending formulation of time.

The construction of the reachability graph of a such PN is then generally impossible. To reduce this state space and provide a finite representation of the reachability graph, two different methods are defined. [14] Introduces the notion of state classes and [11] provides a parametric description to reduce this state space without affecting network properties. This reduced report space is used to define the reachability graph of a TPN. Such a graph will provide a basis to predict failure events of the system.

2.2.2. Parametric state and parametric sequence

Let R_T be an arbitrary TPN. Either $\sigma = t_1 \dots t_n$ a firing sequence in R_T and either $\tau = \tau_0 \tau_1 \dots \tau_n$ a time sequence with $\tau_i \in \mathbb{R}^{*+}$. Then there is at least one dated sequence $\sigma(\tau) = \tau_0 t_1 \tau_1 t_2 \dots \tau_{n-1} t_n \tau_n$ of σ in R_T called the timed sequence of σ which leads the net from the initial state E_0 to a state E (noted $E_0[\sigma(\tau)>E]$ with $E = (m, h)$).

Let us consider for example the following sequence leading the network from the initial state E_0 to a state E_n :

$E_0 [2.0>E'_0 [t_1>E_1 [2.3>E'_1[t_2> \dots t_n>E_n [1.5>E'_n$
The switch from E_0 to E_1 is made in 2 time units after the firing of t_1 .

In addition to this feasible sequence, it is obvious that there is an infinity of feasible sequences leading R_T from E_0 to E , which makes the reachability graph infinite. Instead of considering fixed numbers τ_i , a variable x_i is used to denote the time elapsed between firing the transition t_i and the transition t_{i+1} in σ . Thus instead of having an infinity of execution sequences between the states E_0 and E_n , we will study a single sequence that we will call parametric sequence $\sigma(x) = x_0 t_1 \dots x_{n-1} t_n x_n$ leading the network from the state E_0 to the state E_n with $E_0 [x_0>E'_0 [t_1>E_1 [x_1>E'_1[t_2> \dots t_n>E_n [x_n>E'_n$.

Definition 5: Parametric state and parametric sequence

Let $R_T = \langle P, T, Pre, Post, m_0, I \rangle$ an TPN, $\sigma = t_1 \dots t_n$ a firing sequence in R_T , $\sigma(x) = x_0 t_1 \dots x_{n-1} t_n x_n$ its feasible sequence and B_σ the value of x . The condition for the values $B(x_i)$ results from the time intervals associated with transitions and are united into the set B_σ (5). Then, the parametric execution $(\sigma(x), B_\sigma)$ of σ and the parametric state (E_σ, B_σ) in R_T are determined by:

* When $\sigma = \varepsilon$, i.e., $\sigma(x) = x_0$.

Then $E_\sigma = (m_\sigma, h_\sigma)$ and B_σ are given by:

1- $m_\sigma := m_0$.

$$2- h_\sigma(t) = \begin{cases} x_0 & \text{if } Pre(\bullet, t) \leq m_0 \\ \$ & \text{otherwise} \end{cases} \quad (4)$$

$$3- B_\sigma := \{0 \leq h_\sigma(t) \leq \max(t) \mid t \in T \wedge Pre(\bullet, t) \leq m_\sigma\} \quad (5)$$

* Now, it is assumed that E_σ and B_σ are already defined for the sequence $\sigma = t_1 \dots t_n$.

for $\sigma = t_1 \dots t_n t_{n+1} = \gamma t_{n+1}$, and

$\sigma(x) = x_0 t_1 \dots x_{n-1} t_n x_n t_{n+1} x_{n+1}$ we put 1. $m_\sigma := m_\gamma + \Delta t_{n+1}$,

$$2. h_\sigma(t) = \begin{cases} \$ & \text{if } Pre(\bullet, t) \not\leq m_\sigma \text{ "so } t \text{ is not enabled by } m_\sigma \text{"} \\ h_\gamma(t) + x_{n+1} & \text{if } \begin{cases} Pre(\bullet, t) \leq m_\sigma \wedge \\ Pre(\bullet, t) \leq m_\gamma \wedge \\ \bullet t_{n+1} \cap \bullet t = \emptyset \wedge \\ t \neq t_{n+1} \end{cases} \text{ "t was enabled for } m_\gamma \text{ and remains enabled for } m_\sigma \text{"} \\ x_{n+1} & \text{Otherwise "because } t \text{ is newly enabled" } \end{cases} \quad (6)$$

$$3. B_\sigma := B_Y \cup \{ \min(t_{n+1}) \leq h_Y(t_{n+1}) \} \cup \{ 0 \leq h_\sigma(t) \leq \max(t) | t \in T \wedge Pre(\bullet, t) \leq m_\sigma \}. \quad (7)$$

$h_\sigma(t)$ is a sum of variables (6) ($h_\sigma(t)$ is a parametric t-marking), it is a vector of linear functions: $h_\sigma(t) = f(x)$ with $x := (x_0, \dots, x_{|\sigma|})$ B_σ is a set of conditions (7) (a system of inequalities)

Example:
Consider the temporal Petri Net

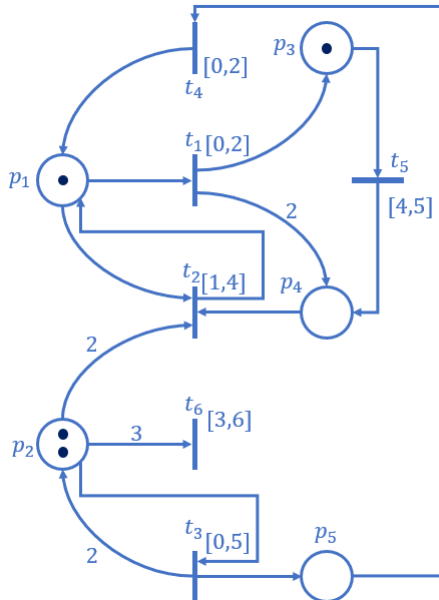


Figure 1: Model 1.

and the transition sequence $\sigma = t_1 t_3 t_4 t_2 t_3$.

$$\sigma(x) = x_0 t_1 x_1 t_3 x_2 t_4 x_3 t_2 x_4 t_3 x_5$$

$$h_\sigma = \begin{pmatrix} x_4 + x_5 \\ x_5 \\ x_5 \\ x_5 \\ x_0 + x_1 + x_2 + x_3 + x_4 + x_5 \\ \$ \end{pmatrix} \text{ and}$$

$$B_\sigma = \left\{ \begin{array}{l} 0 \leq x_0 \leq 2 \\ 0 \leq x_1 \\ 0 \leq x_2 \leq 2 \\ 0 \leq x_3 \leq 2 \\ 0 \leq x_4 \\ 0 \leq x_5 \\ x_4 + x_5 \leq 2 \\ x_0 + x_1 + x_2 + x_3 + x_4 + x_5 \leq 5 \end{array} \right\}$$

The POPOVA approach not only reduces the system's state space (considering only the essential states) [12], but also determines the time required to reach each state. By using parametric states, it is not necessary to check all possible values of the clock, and the inequation system allows to determine the minimum values of their firing times. We will take advantage of this last remark to make the prognosis as soon as possible of a failure event.

2.3. Labeled Petri net

In discrete event systems, partial observation often results in the addition of events or labels as sensor responses of the system.

Thus, a Labelled Petri Net (which we will note R_L) is a classic Petri net in which labels are associated to transitions.

Definition 6:

A Labelled Petri Net (LPN) is a net $R_L = \langle P, T, Pre, Post, m_0, \Sigma, \mathcal{L} \rangle$ in which $\langle P, T, Pre, Post, m_0 \rangle$, is a marked Petri net, Σ is the set of labels associated with transitions, $\mathcal{L}: T \rightarrow \Sigma \cup \{\varepsilon\}$ is the transition labeling function associating a label (event) $e \in \Sigma \cup \{\varepsilon\}$ to each transition $t \in T$, with ε the empty event (or silent).

Thus: $\mathcal{L}(t) = e$ means that the label of the transition t is e .

Remark: Σ can be partitioned to Σ_o and Σ_{uo} with Σ_o is the set of observable events and Σ_{uo} is the set of unobservable events

In this paper we assume that the same label $e \in \Sigma$ can be associated with several transitions, i.e., two transitions t_i and t_j with $t_i \neq t_j$ can be labelled with the same event e in a LPN.

Let Σ^* the set of all event trace Σ containing the label ε , the function of labeling transitions \mathcal{L} can be extended to sequences: $\mathcal{L}: T^* \rightarrow \Sigma^*$ such that:

$$\text{if } t_j \in T \text{ then } \mathcal{L}(t_j) = e_k \text{ for } e_k \in \Sigma; \quad (8)$$

$$\text{if } \sigma \in T^* \wedge t_j \in T \text{ then } \mathcal{L}(\sigma t_j) = \mathcal{L}(\sigma) \cdot \mathcal{L}(t_j); \quad (9)$$

Moreover, if $\mathcal{L}(\lambda) = \varepsilon$ then λ is the empty sequence.

Let σ a transition sequence and $\omega = \mathcal{L}(\sigma) \in \Sigma^*$. The labelled sequence lead to a language generated by the LPN R_L is $\mathcal{L}(R_L) = \{ \omega \in \Sigma^* \mid (\exists \sigma, m_0[\sigma >]) \mathcal{L}(\sigma) = \omega \}$. Thus, $m_1[\omega > m_2]$ means that $\exists \sigma \in T^*, \mathcal{L}(\sigma) = \omega$ where $\omega = e_1 e_2 \dots e_n$ that is, from m_1 and by firing σ , m_2 will be reached. m_2 can be noted m_σ . [27][28]. Note that the length of a sequence σ is greater than or equal to the corresponding word ω (i. e. $|\sigma| \geq |\omega|$). Indeed, if σ contains q transitions labeled by ε , then $|\sigma| = q + |\omega|$ Given the events sequence ω , the reverse labeling function $\mathcal{L}^{-1}(\omega)$ is the whole $\{ \sigma \in T^* \mid \mathcal{L}(\sigma) = \omega \}$.

Example [27]:

Let the following alphabet $\Sigma = \{e_1, e_2, e_3\}$, all transitions $T = \{t_1, t_2, t_3, t_4\}$ and the labelling function \mathcal{L} such as

$$\mathcal{L}(t_j) = \begin{cases} e_j & \text{if } j = \{1, 2\} \\ e_3 & \text{if } j = \{3, 4\} \end{cases} \quad (10)$$

Let's consider the set $\omega = \{e_1, e_3\}$, then $\mathcal{L}^{-1}(\omega) = \{ \{t_1, t_3\}, \{t_1, t_4\} \}$.

2.4. Temporal labelled Petri net

In this paper, the aim is to provide a prognosis of the occurrence date a failed event based on discrete event systems. To represent the behavior of a such system, we adopt the temporal labelled Petri net as modeling tool that represents both the events

and their occurrence dates. Let's therefore provide for each event sequence on the network a temporal signature.

The temporal labelled Petri net (TLPN) is an extension of the temporal PN [17][18] for which each transition is associated with an observable (or not) event [5] [26] [29].

Definition 7:

A TLPN is a net $R_TL = \langle P, T, Pre, Post, m_0, \Sigma, \mathcal{L}, I \rangle$ in which $\langle P, T, Pre, Post, m_0 \rangle$ is a Petri net, Σ is the set of labels associated with transitions, \mathcal{L} is the transition labelling function and I is the function associating a static time interval with each transition. A change in TLPN state can occur either on a transition firing or over an elapsed time period.

Here, the definition of state and its transition function are the same as for a TPN according to the POPOVA approach presented in section 2.2 [11] [21].

2.5. Language generated by a TLPN

Let $\sigma = t_1 \dots t_n$, a firing transition sequence in the TLPN and let $\sigma(x) = x_0 t_1 \dots x_{n-1} t_n x_n$ its feasible dated sequence [13].

We note by $\text{timed}(\sigma)$ the achievable dated sequence $\sigma(x)$: $\text{timed}(\sigma) = \sigma(x)$. conversely, we note by $\text{Logic}(\sigma(x))$ the sequence of firing in the net: $\text{Logic}(\sigma(x)) = \sigma \in T^*$

Furthermore, to avoid introducing too many different notations, the timed labelling function is introduced

Definition 8: (Timed Event Sequence: T.E.S)

Let $\sigma = t_1 \dots t_n$, a firing transition sequence in the TLPN and let $\sigma(x) = x_0 t_1 \dots x_{n-1} t_n x_n$ its feasible dated sequence

The sequence $s(\sigma) = e_1 e_2 \dots e_{n-1} e_n$ is the sequence of events associated with the transitions of the σ firing transition sequence σ .

The labelling function is extended to timed firing sequences $\sigma(x)$ as follows:

$$\mathcal{L}((t_q x_q)) = (e_q x_q), \text{ where } e_q \in \Sigma, \mathcal{L}(t_q) = e_q, t_q \in T \quad (11)$$

$$\mathcal{L}(\sigma(x))(t_q x_q) = \mathcal{L}(\sigma(x)) \mathcal{L}((t_q x_q)) = s'(x). \quad (12)$$

The sequence $s(x) = x_0 e_1 \dots x_{n-1} e_n x_n$ is a timed event sequence (T.E.S). This is the dated sequence of events, associated with the feasible dated sequence $\sigma(x)$. $s(x) = \mathcal{L}(\sigma(x))$.

Definition 9: (temporal language)

Let be TLPN noted R_TL . The temporal language generated by R_TL , noted $\mathcal{L}(R_TL)$ is defined as all the T.E.S $s(x)$ generated by R_TL since the initial marking m_0 . $\mathcal{L}(R_TL) = \{s(x) \mid m_0[\sigma(x)], \mathcal{L}(\sigma(x)) = s(x)\}$ where $\sigma(x)$ is a dated sequence available in R_TL .

3. Failure prognosis based on TLPN

The failure prognosis is intended to predict the properties of a system that are not in compliance with the specifications. The aim is to predict the occurrence of failure events in the system before their future occurrence.

The prognosis in discrete event systems has been discussed in various research papers. Most of them have developed a prognosis

approach predicting a failure event m-steps in advance, based on finite state automata [3][4][6] or Petri nets [1]-[2],[30]-[34], using stochastic and or non-stochastic ways [6][35].

Our proposed approach consists to predict a failure event n-units time in advance. The first contribution relates to a formal representation framework. The adopted modelling considers the three possible operating modes of the system, as shown in the figure 2.

- The nominal mode that contains only the set of states that represent a nominal execution of the system.
- The degraded mode groups all states in which the system operates with a tolerable degradation without influencing the behavior of the system.
- The failed mode that contains all states that represent the failed behavior of the system.

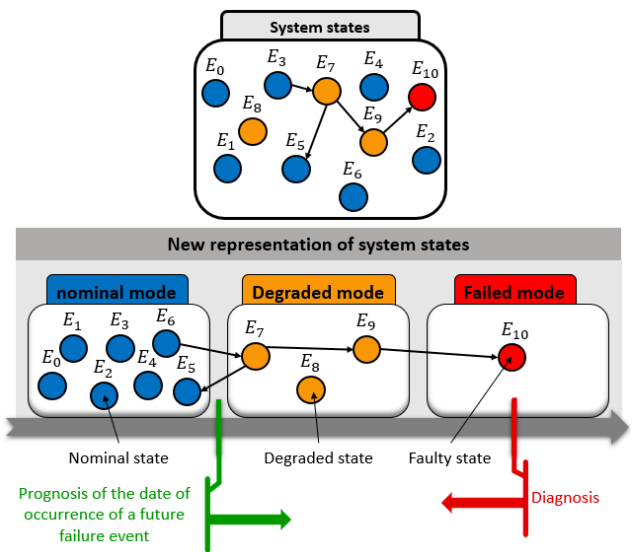


Figure 2 also shows the interest of the prognosis because it aims to explain the causality. Indeed, the diagnosis cannot prevent a failure situation, whereas the prognosis offers more visibility on the future evolution of the system and makes it possible to act before a fault occurs. Our purpose consists to determine a prognosis within an operating mode managing context.

To model such behavior, we propose an extension of the Temporal Labelled Petri nets within a context of operating modes. This extension provides an ability to represent temporal constraints and labels in the modeling process. Figure 3 shows an example of operating modes of a system based on a TLPN model. Switching state is conditioned by the firing of transitions. A transition is fired if it is enabled.

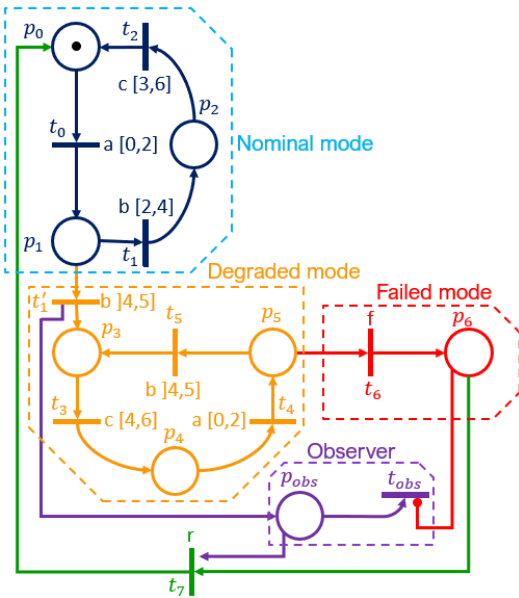
The prognosis will need an observer module constrained by a place (p_{obs}) and transition (t_{obs}). This module has no influence on the behavior of the system, it only observes the occurrence time of a failure event (figure 3).

To do this, we suppose that:

- Only one transition is fired at the same time;

- Only one mode is active at the same time;
- The PN is safe;
- we assume that the firing of transitions is immediate and there is no firing delay;
- All TLPN events are observable.

After firing the transition, the TLPN changes from $E=(m,h)$ to the state $E' = (m', h')$ (see definition 4).



Definition 10:

The extended TLPN (ETLPN) is $R_{TL} = \langle P, T, Pre, Post, m_0, \Sigma, \mathcal{L}, I \rangle$ where:

$P = P_n \cup P_{deg} \cup P_{fail} \cup P_{obs}$ with P_n is the set of nominal places, P_{deg} is the set of degraded places, P_{fail} is the set of failed places and P_{obs} is the observer place.

$T = T_n \cup T_{deg} \cup T_{fail} \cup T_{obs} \cup T_{rep}$, the set of transitions.

If $\exists p_i \in (P_n \cup P_{fail})$ and $t_j \in (T_n \cup T_{fail})$, such that: $m(p_i) \geq pre(p_i, t_j)$, then $m(p_{obs}) = 0$ and t_{obs} is not enabled, otherwise, $m(p_{obs}) = 1$ and t_{obs} is enabled.

The model of Figure 3 is an ETLPN with $\Sigma = \{a, b, c, f, r\}$, with: f is a failure event and r is a repair event.

- The transition t_6 is a failed transition such as: $t_6 \in T_{fail}$ then, $\mathcal{L}(t_6) = f$.
- The transition t_7 is a repair transition such as: $t_7 \in T_{rep}$ then, $\mathcal{L}(t_7) = r$.

By firing the t'_1 transition the system switches to a degraded mode marking thus p_{obs} , that is $m(p_{obs}) = 1$. The p_{obs} place remains marked until the system switch to a failed mode.

The introduction of p_{obs} and t_{obs} doesn't influence the behavior of the system. Their interest will be explained in the following section.

To represent sequences ending with a failure event, we use the both notions of parametric state and sequences allow to construct the reachability graph which contains only the essential states, i.e.

the time associated with each timed transition enabled of a state $E = (m, h)$ is a natural integer. However, knowing the behavior of the network in the "essential" states is sufficient to determine at any time the overall behavior of the network. (cf. [12] [22]).

The advantage of this approach is the application of linear optimization (generated by the system of inequalities in each state), which makes it possible to calculate the execution time of a sequence at the earliest and at the latest.

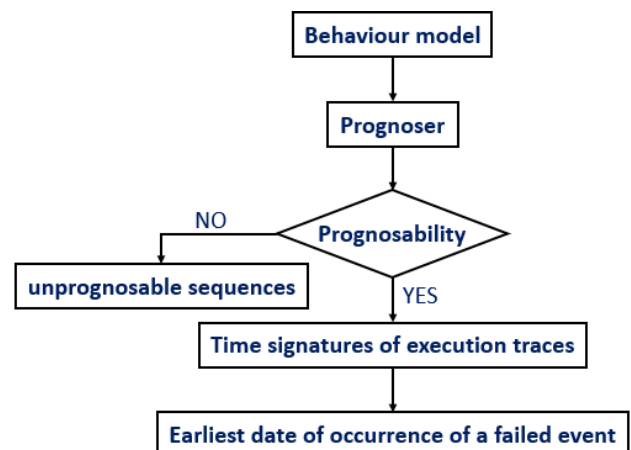
Clock times must be accumulated to progress from a state E of the net to a failed state E' . To do this, an observer model is introduced to the model in order to record the cumulative time between E and E' . This observer model has no impact on the behavior of the system, it just makes it possible to record the time required to progress from a non-defaulting, but not necessarily normal, to a state E' that is considered failed.

To calculate this execution time, we propose an extension (definition 11) of definition 5. But before discussing the proposed approach, we formulate the following assumptions:

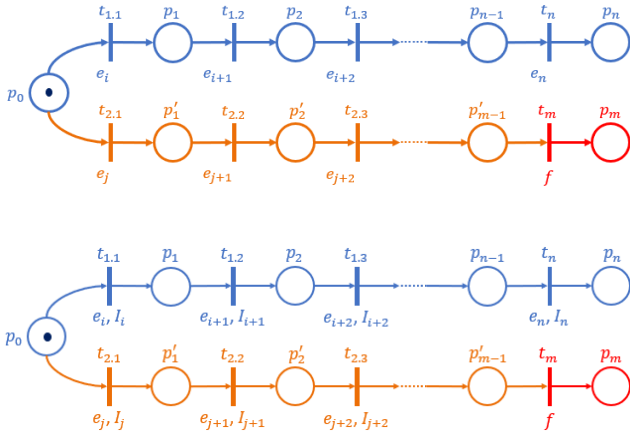
- 1- The system model is known
- 2- all events are observable. The case of prognosis under partial observation is not considered here.
- 3- The prognosis begins when the model switches from nominal mode to degraded one.

Remark: the remains the same, if the prognosis is started from any nominal state of the system.

The following framework (figure 4) describes the steps of the proposed prognosis approach. The first step, called the behavioral model, is required to describe the possible operating modes of the system (figure 3). The prognoser is an oriented state graph (figure 8), built from the system model, its role is to detect all possible traces ending with a failure event; Once the system switches from nominal to degraded mode, the prognoser must identify all the sequences of the model namely those that lead to a failure event. Such an event cannot be predicted overall in the sequences. The prognosability property is introduced to determine the sequences of failure event that can be predicted. From an inequality system, the execution time of each sequence is calculated; It called "Time signatures of execution traces". The minimal time signature will then represent the earliest date before a failure event occurs.

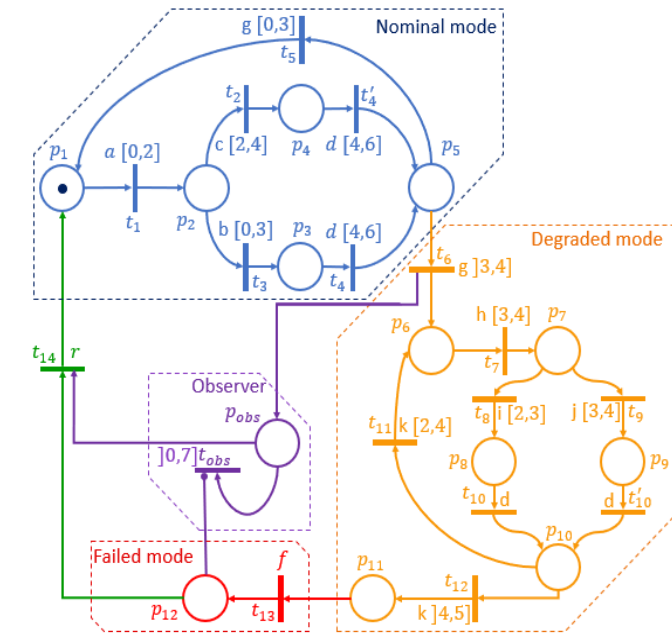


In fact, the prognosis is not possible, when two traces σ_1 and σ_2 have the same time constraints and the same sequencing of events. For example, in figure 6, $\sigma_1 = t_{1,1} \dots t_n$, $\sigma_2 = t_{2,1} \dots t_m$ and $s_1(x) = x_{1,0}e_i x_{1,1} \dots x_{1,n-1}e_n x_{1,n}$, $s_2(x) = x_{2,0}e_j x_{2,1} \dots x_{2,m-1}f x_{2,m}$ (with f a failure event) are their respective parametric event sequences.



Deciding on the future execution of σ_1 or σ_2 from a state E is conditioned by:
 $Pre(\bullet, t_i) \leq m_0 \vee Pre(\bullet, t_j) \leq m_0 \wedge t_i \neq t_j$ i.e. t_i and t_j are enabled from E .
 $e_i \in \Sigma_o \vee e_j \in \Sigma_o \wedge e_i \neq e_j$, otherwise if $e_i = e_j$ it is required that $I_{x_{1,0}} \cap I_{x_{2,0}} = 0$ ($I_{x_{1,0}}$ execution interval of $x_{1,0}$)
 In other case, we cannot prognosticate the failure event f .

If we consider in figure 5 that $e_i = e_j$, the failure event f cannot be prognosable. But if $e_i = e_j$ and the intervals of the dated sequences are different then the failure event f is prognosable (figure 6).



The resolution of the inequation system will be the last step, which calculates the time signature of execution for all the prognosable sequences. The minimum execution generated from this step represents the earliest occurrence time of a failure event.

In the example shown in figure 7, the prognosis starts from the firing transition t_6 because degraded mode will start at this place.

Indeed, if the event g occurs at earliest after 3 units time, the model switch to the degraded mode. From this state the observer place (p_{obs}) will be activated, and its corresponding transition t_{obs} becomes enabled until the event f (failure event) will occurred. Thus, the interval times associated with the transitions enabled from place p_6 , will be combined in the form of associated system of inequalities to t_{obs} while the occurrence of the failure event f of the transition t_{13} does not occur. When the event r is generated (meaning that the system is repaired), the observer place will be initialized to allow a next operating cycle.

The p_6 place is called the candidate place for the prognosis. Once this place is marked, the occurrence of the failure event can be predicted.

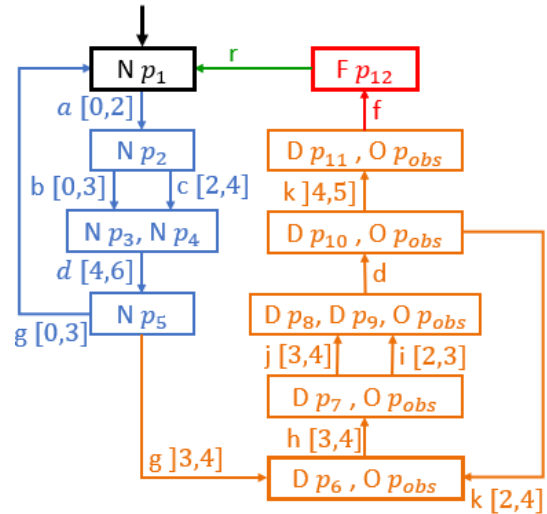


Figure 8 link to model 2 presents the prognoser model where a state is composed by the marked states of the model and their corresponding mode the nominal states are represented by N, degraded by D and failed states by F, except for Pobs place that will be associated to Observer module noted O. The prognosis of the event occurrence date is possible from any state of the prognoser. $\{N p_5\}$ for example, represent a marking of the system and its corresponding operating mode, i.e. p_5 indicates that the place p_5 is marked and N indicates that the system is in nominal mode. The occurrence of event g in interval $]3,4[$ leads the prognoser model to $\{D p_6, O p_{obs}\}$ where p_6 and p_{obs} are the marked places and D means that the system is in degraded mode and p_{obs} place marked, the prognosis process is then activated. The prognosis process is achieved by the identification of all sequences ending with an F state. According to the prognoser's model and

from $\{Dp_6, Op_{obs}\}$ the event sequences ending in a failure event are: $s(\sigma_1) = hidkf$ and $s(\sigma_2) = hjdkf$. To simplify, we don't take into consideration $hidk$ and $hjdk$ cycles.

Then, the execution time of each sequence is calculated (time signature) by applying algorithm 2. The aim is to find all the minimum solution values of the system of inequalities. These values will constitute the minimum time after which the occurrence of the failure event is certain.

Definition 11, which is an extension of definition 5, allows, from a TLPN, to recursively determine the parametric state and parametric sequence leading to a failure state, and thus generating the system of inequalities composed of the constraints obtained from the intervals associated with each enabled transition from a candidate place. But before presenting definition 11, let's first reconsider a set of enabled transitions from a m marking.

Let m be a marking of a PN. We define V_m the set of enabled transition from m as follows: $V_m = \{t_i \in T \mid m \geq \text{Pre}(\bullet, t_i)\}$,

Definition 11: (parametric state and sequence of an TLPN)
Let $R_TL = \langle P, T, \text{Pre}, \text{Post}, m_0, \Sigma, L, I \rangle$ a TLPN, m a p_marking and h a t_marking containing the time associated with each transition enabled from m . $\sigma := t_1 \dots t_n$ is a firing transition sequence in R_TL and $\sigma(x) := x_0 t_1 \dots x_{n-1} t_n x_n$ its feasible dated sequence. V_{m_σ} is the set of transitions enabled from the m_σ marking (final marking) obtained by the firing of the transition sequence σ .

Then, the parametric sequence $(\sigma(x), B_\sigma)$ of σ and the parametric state (E_σ, B_σ) in R_TL are determined by the algorithm 1.

Algorithm 1 Prognosis algorithm

Begin

* if $\sigma = \varepsilon$, i.e. $\sigma(x) = x_0$ and $s(x) = x_0$

Then $E_\sigma = (m_\sigma, h_\sigma)$ and B_σ are given by:

1- $m_\sigma := m_0$,

2- $V_{m_\sigma} = V_{m_0} = \{t_i \mid m_0 \geq \text{Pre}(\bullet, t_i)\}$

3- $h_\sigma(t) = h_0(t) = \begin{cases} x_0 & \text{if } \text{Pre}(\bullet, t) \leq m_0 \\ \$ & \text{Otherwise} \end{cases}$

4- $B_\sigma := \{0 \leq h_\sigma(t) \leq \max(t) \mid t \in T \wedge \text{Pre}(\bullet, t) \leq m_\sigma\}$

Else

repeat

We assume that E_σ and B_σ are already defined for the transition sequence $\sigma := t_1 \dots t_n$.

then $\sigma(x) := x_0 t_1 \dots x_{n-1} t_n x_n$. its corresponding T.E.S is

$s(x) = x_0 e_1 \dots x_{n-1} e_n x_n$ for $\sigma := t_1 \dots t_n t_{n+1} = \gamma t_{n+1}$, and

$\mathcal{A}(\sigma) = \mathcal{A}(\gamma)$. $\mathcal{A}(t_{n+1}) = s(\gamma)$. $\mathcal{A}(t_{n+1})$

1. $m_\sigma := m_\gamma + \Delta t_{n+1}$, with $\Delta t_{n+1} := \text{Post}(\bullet, t_{n+1}) - \text{Pre}(\bullet, t_{n+1})$

2. $V_{m_\sigma} = \{t \mid m_\sigma \geq \text{Pre}(\bullet, t)\}$

3. $h_\sigma(t) = \begin{cases} \$ & \text{if } \text{Pre}(\bullet, t) \not\leq m_\sigma \\ h_\gamma(t) + x_{n+1} & \text{if } \text{Pre}(\bullet, t) \leq m_\sigma \wedge \text{Pre}(\bullet, t) \leq m_\gamma \\ & \wedge \text{Pre}(t_{n+1}) \cap \text{Pre}(t) = \emptyset \wedge t \neq t_{n+1} \\ x_{n+1} & \text{Otherwise} \end{cases}$

4. $B_\sigma := B_\gamma \cup \{\min(t_{n+1}) \leq h_\gamma(t_{n+1})\} \cup \{0 \leq h_\sigma(t) \leq \max(t) \mid t \in T \wedge \text{Pre}(\bullet, t) \leq m_\sigma\}$

until $\mathcal{A}(t_{n+1}) = \text{cf}$

End

Let the TLPN of Figure 7 and apply definition 11

At the start $\sigma := \varepsilon$, $\sigma(x) = x_0$;

$E_\sigma = (m_\sigma, h_\sigma) = (m_0, h_0)$;

$m_0 = (p_1)^T$, means that only p_1 is marked with 1 token and 0 token in the rest of the other places;

$V_{m_0} = \{t_1\}$;

$\mathcal{A}(t_1) = a$;

$h_0(t) = (x_0, \$, \$, \$, \$, \$, \$, \$, \$, \$, \$, \$, \$, \$, \$, \$)^T$;

$B_0 = \{0 \leq x_0 \leq 2\}$.

$\sigma = t_1$;

$m_1 = (p_2)$,

$V_{m_1} = \{t_2, t_3\}$, means that only t_2 and t_3 are enabled from the marking m_1 ;

$\mathcal{A}(t_2) = b$, $\mathcal{A}(t_3) = c$;

$h_1(t) = (\$, x_1, x_1, \$, \$, \$, \$, \$, \$, \$, \$, \$, \$, \$, \$)^T$;

$B_1 = \begin{cases} 0 \leq x_0 \leq 2 \\ 0 \leq x_1 \leq 3 \end{cases}$

$\sigma = t_1 t_3$;

$m_2 = (p_3)^T$; $V_{m_2} = \{t_4\}$;

$\mathcal{A}(t_4) = d$;

$h_2(t) = (\$, \$, \$, x_2, \$, \$, \$, \$, \$, \$, \$, \$, \$, \$, \$)^T$;

$B_2 = \begin{cases} 0 \leq x_0 \leq 2 \\ 0 \leq x_1 \leq 3 \\ 0 \leq x_2 \leq 6 \end{cases}$

$\sigma = t_1 t_3 t_4$;

$m_3 = (p_5)^T$; $V_{m_3} = \{t_5, t_6\}$;

$\mathcal{A}(t_5) = \mathcal{A}(t_6) = g$;

$h_3(t) = (\$, \$, \$, \$, x_3, x_3, \$, \$, \$, \$, \$, \$, \$, \$)^T$;

$B_3 = \begin{cases} 0 \leq x_0 \leq 2 \\ 0 \leq x_1 \leq 3 \\ 4 \leq x_2 \leq 6 \\ 0 \leq x_3 \leq 3 \end{cases}$

$\sigma = t_1 t_3 t_4 t_6$;

$m_4 = (p_6 p_{obs})^T$, $V_{m_4} = \{t_7, t_{obs}\}$;

$\mathcal{A}(t_7) = h$;

$h_4(t) = (\$, \$, \$, \$, \$, \$, x_4, \$, \$, \$, \$, \$, \$, \$, \$, x_4)^T$;

$B_4 = \begin{cases} 0 \leq x_0 \leq 2 \\ 0 \leq x_1 \leq 3 \\ 4 \leq x_2 \leq 6 \\ 3 \leq x_3 \leq 3 \\ 0 \leq x_4 \leq 4 \end{cases}$

$\sigma = t_1 t_3 t_4 t_6 t_7$;

$m_5 = (p_7 p_{obs})^T$, $V_{m_5} = \{t_8, t_9, t_{obs}\}$;

$\mathcal{A}(t_8) = i$, $\mathcal{A}(t_9) = j$;

$h_5(t) = (\$, \$, \$, \$, \$, \$, \$, x_5, x_5, \$, \$, \$, \$, \$, \$, x_4 + x_5)^T$;

- [6] A. Khoumsi, "Alternative Inference-Based Decentralized Prognosis of Discrete Event Systems" In 2019 6th International Conference on Control, Decision and Information Technologies (CoDIT) (pp. 250-255). IEEE (2019, April).
- [7] Z. Ma, et al. "Marking Estimation in Labelled Petri nets by the Representative Marking Graph" IFAC-PapersOnLine 50.1 (2017): 11175-11181.
- [8] A. K. A. Toguyeni, "Surveillance et diagnostic en ligne dans les ateliers flexibles de l'industrie manufacturière" (Doctoral dissertation, Lille 1) (1992).
- [9] Y. Pencolé, "Diagnostic: étude d'un raisonnement complexe et multi-dimensionnel" (Doctoral dissertation 2018).
- [10] R. Saddem, "Diagnosticabilité modulaire appliquée au Diagnostic en ligne des Systèmes Embarqués Logiques" (Doctoral dissertation 2012).
- [11] L. Popova-Zeugmann, "Time and Petri Nets" Springer, Berlin, Heidelberg, (2013). 31-137.
- [12] Popova-Zeugmann, "Time and Petri Nets" DOI 10.1007/978-3-642-41115-1_3, © Springer-Verlag Berlin Heidelberg (2013).
- [13] L. Popova-Zeugmann, "On parametrical sequences in time Petri nets" Proceedings of the CS&P. Vol. 97. (1997).
- [14] B. Berthomieu, et al. "Problèmes d'accessibilité et espaces d'états abstraits des réseaux de Petri temporels à chronomètres." Journal européen des systèmes automatisés 39.1/3 (2005): 223.
- [15] B. Berthomieu, M. Díaz, "Modeling and verification of time dependent systems using time Petri nets" IEEE transactions on software engineering 3 (1991): 259-273.
- [16] B. Berthomieu, M. Menasche, "An enumerative approach for analyzing time Petri nets" Proceedings IFIP. (1983).
- [17] B. Berthomieu, "La méthode des classes d'états pour l'analyse des réseaux temporels" 3e congrès Modélisation des Systèmes Réactifs (MSR'2001). (2001).
- [18] B. Berthomieu, "Réseaux de Petri temporels" ETR septembre (2003): 123-153.
- [19] B. Berthomieu, F. Vernadat, S. D. Zilio, "TINA: Time Petri Net Analyzer" (2015).
- [20] L. Popova-Zeugmann, "On time Petri nets" Journal Information Processing and Cybernetics, EIK 27.4 (1991): 227-244.
- [21] L. Popova-Zeugmann, "On liveness and boundedness in time Petri nets" Proceedings of the Workshop on Concurrency, Specification and Programming (CS&P'95). (1995).
- [22] L. Popova-Zeugmann, "Essential states in time Petri nets" Humboldt-Univ. zu Berlin, (1998).
- [23] P.J Ramadge, W.M. Wonham, "the controle of discrete-event systems" Proc. IEEE, vol. 77, no 1, pp 81-98, (1989).
- [24] B. Bérard et al. "Comparison of different semantics for time Petri nets" International Symposium on Automated Technology for Verification and Analysis. Springer, Berlin, Heidelberg, (2005).
- [25] B. Berard et al. "The expressive power of time Petri nets" Theoretical Computer Science 474 (2013): 1-20.
- [26] D. Lefebvre, "Fault diagnosis and prognosis with partially observed Petri nets" IEEE Transactions on Systems, Man, and Cybernetics: Systems 44.10 (2014): 1413-1424.
- [27] F. Basile, P. Chiacchio, J. Coppola, "Identification of labeled time Petri nets" 2016 13th International Workshop on Discrete Event Systems (WODES). IEEE, (2016).
- [28] F. Basile, P. Chiacchio, J. Coppola, "Identification of time Petri net models" IEEE Transactions on Systems, Man, and Cybernetics: Systems 47.9 (2016): 2586-2600.
- [29] V. D. Aalst, W. M. P. "Timed Coloured Petri Nets and their Application to Logiges" Diss. PhD thesis, Tech. Univ. Eindhoven, (1992).
- [30] D. Lefebvre, "Fault diagnosis and prognosis with partially observed stochastic Petri nets." Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 228.4 (2014): 382-396.
- [31] D. Lefebvre, "Fault diagnosis and prognosis with partially observed stochastic Petri nets." Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 228.4 (2014): 382-396.
- [32] M. Ghazel, "Surveillance des systèmes à événements discrets à l'aide des réseaux de Petri Temporels" Diss. Ecole Centrale de Lille, (2005).
- [33] R. Ammour, et al. "Faults prognosis using partially observed stochastic Petri nets" 2016 13th International Workshop on Discrete Event Systems (WODES). IEEE, (2016).
- [34] F. Basile, P. Chiacchio, G. De Tommasi. "Fault diagnosis and prognosis in Petri Nets by using a single generalized marking estimation" IFAC Proceedings Volumes 42.8 (2009): 1396-1401.
- [35] D. Lefebvre, "Fault diagnosis and prognosis with partially observed Petri nets" IEEE Transactions on Systems, Man, and Cybernetics: Systems 44.10 (2014): 1413-1424.

Performance of Robust Confidence Intervals for Estimating Population Mean Under Both Non-Normality and in Presence of Outliers

Juthaphorn Sinsomboonthong¹, Moustafa Omar Ahmed Abu-Shawiesh^{* 2}, Bhuiyan Mohammad Golam Kibria³

¹Department of Statistics, Faculty of Science, Kasetsart University (KU), Bangkok 10900, Thailand

²Department of Mathematics, Faculty of Science, The Hashemite University (HU), Al-Zarqa, 13115, Jordan

³Department of Mathematics and Statistics, Florida International University (FIU), University Park, Miami FL 33199, USA

ARTICLE INFO

Article history:

Received: 26 March, 2020

Accepted: 29 May, 2020

Online: 18 June, 2020

Keywords:

Robust Estimators

Confidence Interval

Sample Median

Sample Trimean

Interquartile Range

Outliers

Non-Normal Distribution

Contaminated Normal Distribution

Coverage Probability

Average Width

ABSTRACT

We proposed two robust confidence interval estimators, namely, the median interquartile range confidence interval (MDIQR) and the trimean interquartile range confidence interval (TRIQR) for the population mean (μ) as an alternative to the classical confidence interval. The proposed methods are based on the asymptotic normal theorem (ANT) for the sample median (MD) and the sample trimean (TR). We compare the performance of the proposed interval estimators with the classical estimators by using a simulation study through the following criteria: (i) average width (AW) and (ii) empirical coverage probability (CP). It is evident from simulation study is that the proposed robust interval estimator performs well under both criterion and when the observations are sampled from contaminated normal distribution. However, when the observations are sampled from non-normal distributions, the classical confidence interval performs the best in the shorter width sense, but the coverage probability tends to be smaller than the two proposed robust confidence interval estimators for all sample sizes. For illustration purposes, two real life data sets are analyzed, which supported the findings of the simulation study to some extent.

1. Introduction

The classical methods in statistical inference, such as confidence intervals estimation, are widely used by researchers in many disciplines. In the usage of the confidence intervals, the assumptions such as normality and no presence of outliers must be satisfied. Unfortunately, these assumptions are rarely met when analyzing real data in many fields of research such as engineering, data science, medical, public health, biological etc. The confidence intervals provide better information that of point estimator about the population characteristic of interest. The performance of confidence intervals for the appearance of outliers and under non-normal assumption have drawn much attention among the researchers. A variety of procedures are exist in the literature to construct the confidence interval (CI) for the population mean (μ), though the classical normal confidence interval is widely used. Nevertheless, the classical normal confidence interval requires normality assumption which most of the data do not follow in

reality, particularly in presence of outliers. Thus, the robust estimators, which are less affected from non-normality assumption or outliers, are introduced in this paper in order to overcome such situations.

Student's-t confidence interval for the population mean (μ) has been used for a long period of time. It has an approximate $(1 - \alpha)$ coverage probability (CP) under the condition of positively skewed distribution or there are some outliers in the data. However, this coverage probability may be improved by developing different confidence interval methods. The bootstrap confidence interval [1] is another method to construct the confidence intervals for the population mean which many researchers are suggested. The construction of this confidence interval has concerned about resampling technique which is complicated procedures and it has a good performance in theoretical coverage probability, but it tend to be erratic in actual practice depend on the distribution of the bootstrap estimator. Further, this method hard to implement in practices because it is not easy to compute without the statistical programming [2], while

*Moustafa Omar Ahmed Abu-Shawiesh, Department of Mathematics, The Hashemite University (HU), Jordan, E-Mail: mabushawiesh@hu.edu.jo

the two robust confidence intervals that are proposed in this paper are easy to implement in practices. The two robust confidence intervals for the population mean (μ) are proposed based on robust location and scale estimators in the case of non-normal distributions and contamination of outliers in the data set. We compare the performance of proposed robust methods with that of the classical Student's confidence interval using coverage probability and average width for non-normal distributions (symmetric and skewed ones) via a Monte-Carlo simulation study. For more on robust estimators, we refer [3], Abu-Shawiesh [4, 5] among others.

The organization for the remaining of this paper are the following: Section 2 is represented the proposed confidence intervals. A Monte-Carlo simulation study has been conducted in section 3. Two real-life data are analyzed for the implementation of several methods in Section 4. Section 5 provides some concluding remarks.

2. Proposed Interval Estimators

2.1. The Classical Confidence Intervals for the Population Mean

A random sample X_1, X_2, \dots, X_n of size n is taken from the population that is normally distributed with mean (μ) and variance (σ^2). Then, the $(1 - \alpha)$ 100% classical confidence interval (CI) for the population mean (μ), for known σ is defined by (1).

$$C.I. = \bar{X} \pm Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \quad (1)$$

where $Z_{1-(\alpha/2)}$ is the $(1 - (\alpha/2))^{th}$ percentile of the standard normal distribution. However, in real life, it is unlikely that the population standard deviation (σ) is known, and then an estimate of σ is needed. To do that, we can use the sample standard deviation (S) instead of the unknown population standard deviation (σ) and apply the normal distribution to construct the $(1 - \alpha)$ 100% classical confidence interval (CI) for the population mean (μ) which is given by (2).

$$C.I. = \bar{X} \pm Z_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \quad (2)$$

Since the classical confidence interval requires the normality assumption, it is unlikely that it will give good results when data are not normal. Therefore, we suggested two robust confidence interval estimators, namely, the median interquartile range confidence interval (MDIQR-CI) and the trimean interquartile range confidence interval (TRIQR-CI) and they are discussed as follows:

2.2. The Robust Confidence Intervals

We propose two robust modifications of the classical normal interval estimator for the population mean (μ) in the case of non-normal distributions and presence of outliers. They are simple adjustments based on robust estimators for location and scale parameters. The proposed robust confidence intervals for the population mean (μ) are introduced in these subsections:

2.2.1. The Median Interquartile Range Confidence Interval

In this confidence interval (MDIQR-CI), we estimate the population mean (μ) by the sample median (MD) and the population standard deviation (σ) by interquartile range (IQR). The standard error of the sample median (MD), that

is $S.E.(MD) = \sigma_{MD} = 1.253 \sigma / \sqrt{n}$, is used in the construction of this interval estimator. Thus, the $(1 - \alpha)$ 100% MDIQR-CI confidence interval for the population mean (μ) is given by (3).

$$\begin{aligned} CI_{MDIQR} &= MD \pm Z_{1-\frac{\alpha}{2}} S.E.(MD) \\ &= MD \pm Z_{1-\frac{\alpha}{2}} \frac{1.253 \sigma}{\sqrt{n}} \\ &= MD \pm Z_{1-\frac{\alpha}{2}} \frac{1.253 IQR}{\sqrt{n}} \end{aligned} \quad (3)$$

where, the sample median (MD), is defined by (4) as follow:

$$MD = \begin{cases} X_{(\frac{n+1}{2})} & \text{if } n \text{ is odd} \\ \frac{X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}}{2} & \text{if } n \text{ is even} \end{cases} \quad (4)$$

2.2.2. The Trimean Interquartile Range Confidence Interval

In this confidence interval (TRIQR-CI), we estimate the population mean (μ) by the sample trimean (TR) and the population standard deviation (σ) by interquartile range (IQR). The standard error of the sample trimean (TR), that is $S.E.(TR) = \sigma_{TR} = 1.097 \sigma / \sqrt{n}$, is used in the construction of this confidence interval. Then, the $(1 - \alpha)$ 100% TRIQR-CI confidence interval for the population mean (μ) is given by (5).

$$\begin{aligned} CI_{TRIQR} &= TR \pm Z_{1-\frac{\alpha}{2}} S.E.(TR) \\ &= TR \pm Z_{1-\frac{\alpha}{2}} \frac{1.097 \sigma}{\sqrt{n}} \\ &= TR \pm Z_{1-\frac{\alpha}{2}} \frac{1.097 IQR}{\sqrt{n}} \end{aligned} \quad (5)$$

where,

$$TR = \frac{1}{2} \left(Q_2 + \frac{Q_1 + Q_3}{2} \right) = \frac{1}{2} \left(MD + \frac{Q_1 + Q_3}{2} \right) \quad (6)$$

is the sample trimean and Q_1, Q_2 and Q_3 are the first, second (sample median) and third quartiles, respectively [6].

3. The Simulation Study

A simulation study has been conducted to compare the performance of three interval estimators. The simulation method is one of techniques to implement for a theoretical performance comparison and the results of the study are usually very close to the ones of the exact case when using a large number of iterations. In order to make the comparisons among three confidence intervals, two performance criteria—the coverage probability (CP) and the average width (AW)—of the confidence intervals are considered. If the confidence interval that is compared among the three confidence intervals has a smaller width, it indicates this confidence interval is a better method for the same level of the coverage probability. For a higher coverage probability, the confidence interval indicates a better method when the widths are the same level. We used SAS version 9.4 programming to conduct this simulation study. We consider the widely used 95% confidence intervals for this simulation. We consider in equals to 10, 20, 30, 40, 50 and 100 were generated 100,000 times for each situation. For each data set of the samples, the common 95% confidence intervals were constructed for the three methods. The coverage probability (CP) and the average width (AW) of the confidence intervals are found by using respectively:

$$CP = \frac{\#(L \leq \theta \leq U)}{100,000} \quad \text{and} \quad AW = \frac{\sum_{i=1}^{100,000} (U_i - L_i)}{100,000} \quad (7)$$

To compare the performance of the interval estimators, the same types of distributions are used as in [7–9]; symmetric, skewed and contaminated normal ones. So, there are three cases for the simulated observations as follows:

Case (a): Skewed Distributions

In the skewed distribution cases, we will simulate observations from the gamma distribution, given by (8):

$$f(x; \alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & , \quad x > 0 ; \alpha, \beta > 0 \\ 0 & , \quad \text{otherwise} \end{cases} \quad (8)$$

where α and β are the shape and scale parameters respectively. The mean of the distribution is given by $\mu = \alpha/\beta$ and the variance of the distribution is given by $\sigma^2 = \alpha/\beta^2$. Without loss of generality, β is set to unity and if α increases then the gamma distribution will approach to the normal distribution. For this simulation study, we consider, $\alpha = 1, 2, 4, 8$ and $\beta = 1$.

Case (b): Symmetric Distributions

In the symmetric distribution cases, we will simulate observations from the student t-distribution, $t(k)$, where k is the numbers of degrees of freedom with probability density function (pdf) given by (9):

$$f(x; k) = \frac{\Gamma((k+1)/2)}{\sqrt{k\pi} \Gamma(k/2)} \frac{1}{((x^2/k) + 1)^{(k+1)/2}} , \quad -\infty < x < \infty \quad (9)$$

where mean of the distribution is zero and the variance, $\sigma^2 = k/(k-1)$. The t-distribution is one type of a symmetrical distribution and bell shaped around 0, but it has heavier tails than the normal distribution. Additionally, as the number of the degrees of freedom (k) increase, the t-distribution will approach to the normal distribution. For the simulation purposes, we will consider $k = 4, 10, 30, 50$.

Case (c): Contaminated Normal Distribution

In this case, we will simulate observations from mixture distribution that is called the contaminated normal distribution (CND) where artificial outliers are introduced in the data to assess the sensitivity of the three different interval estimators to the presence of outliers. The contaminated normal probability density function is given by (10):

$$f(x; \mu, \sigma) = (1 - \delta) N(\mu, \sigma^2) + \delta N(\mu, \lambda\sigma^2) \quad (10)$$

where $X \sim N(\mu, \sigma^2)$ denote the normal PDF, $(1 - \delta)$ and δ be the mixing probabilities, and the standard deviation of the wider component is defined as $\lambda > 1$. The main distribution of a data set is generated from the normal distribution $N(\mu, \sigma^2)$ and slightly contaminated by a wider distribution. This paper determines $\delta = 0.1, 0.2$ and 0.3 which represents 10%, 20% and 30% "contamination" respectively, and assigns $\lambda = 5^2, 10^2$ as the scale multipliers. In this section, we consider an uncontaminated standard normal distribution, $N(0, 1)$. The following six cases are constructed the PDF of the contaminated normal distribution (CND) as the linear combination of $N(0, 1)$ and $N(0, 5^2)$ densities as shown in (11) to (13), and the PDF of a contaminated normal distribution is the linear combination of $N(0, 1)$ and $N(0, 10^2)$ densities as shown in (14) to (16):

Case 1: A situation that comprises of 90% of simulated observations are sampled from $N(0, 1)$ distribution and 10% from a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 5^2$, $N(0, 5^2)$, is generated. This will give approximately 10% artificial outliers.

$$CN(0, 5^2)_{.10} = 0.9 N(0, 1) + 0.1 N(0, 5^2) \quad (11)$$

Case 2: A situation that consists of 80% of simulated observations are sampled from the standard normal distribution, $N(0, 1)$, and 20% from a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 5^2$, $N(0, 5^2)$, is generated. This will give approximately 20% artificial outliers.

$$CN(0, 5^2)_{.20} = 0.8 N(0, 1) + 0.2 N(0, 5^2) \quad (12)$$

Case 3: A situation that consists of 70% of simulated observations are sampled from the standard normal distribution, $N(0, 1)$, and 30% from a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 5^2$, $N(0, 5^2)$, is generated. This will give approximately 30% artificial outliers.

$$CN(0, 5^2)_{.30} = 0.7 N(0, 1) + 0.3 N(0, 5^2) \quad (13)$$

Case 4: A situation that consists of 90% of simulated observations are sampled from the $N(0, 1)$ distribution and 10% from a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 10^2$, $N(0, 10^2)$, is generated. This will give approximately 10% artificial outliers.

$$CN(0, 10^2)_{.10} = 0.9 N(0, 1) + 0.1 N(0, 10^2) \quad (14)$$

Case 5: A situation that consists of 80% of simulated observations are sampled from $N(0, 1)$ distribution and 20% from a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 10^2$, $N(0, 10^2)$, is generated. This will give approximately 20% artificial outliers.

$$CN(0, 10^2)_{.20} = 0.8 N(0, 1) + 0.2 N(0, 10^2) \quad (15)$$

Case 6: A situation that consists of 70% of simulated observations are sampled from $N(0, 1)$ distribution and 30% from a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 10^2$, $N(0, 10^2)$, is generated. This will give approximately 30% artificial outliers.

$$CN(0, 10^2)_{.30} = 0.7 N(0, 1) + 0.3 N(0, 10^2) \quad (16)$$

The simulation study results for all considered cases are presented in Table 1 to Table 4 and Figure 1 to Figure 4. The results in Table 1 and Figure 1 show the performances of skewed distribution cases that the observations are generated from gamma distribution with α equals 1, 2, 4, 8 and β equals 1. It is found that the coverage probabilities of the three confidence intervals tend to be lower than the nominal level (0.95) when the shape parameter equals 1, 2 and 4 for almost all sample sizes. When a shape parameter equals 8, the coverage probabilities of MDIQR and TRIQR confidence intervals are greater than the nominal level (0.95) for most of the sample sizes, whereas this of the classical confidence interval tends to be lower than the nominal level for all sample sizes. For all the shape and scale parameters of the gamma distribution, it is found that the classical interval estimator has the smallest average width of the confidence interval among the comparative confidence intervals for all sample sizes.

The simulated results in Table 2 and Figure 2 show the performances of symmetric distribution cases that the observations are generated from the Student's t-distribution with DF equals 4, 10, 30, 50. It is observed that the coverage

Table 1: Coverage probability (CP) and average width (AW) of the 95% CIs for gamma distributed data

PDF	n	Confidence Interval Methods					
		Classical CI		MDIQR-CI		TRIQR-CI	
		CP	AW	CP	AW	CP	AW
G(1, 1)	10	0.8695	1.15	0.8485	1.70	0.8360	1.49
	20	0.9045	0.84	0.8251	1.21	0.8290	1.06
	30	0.9182	0.69	0.7697	0.98	0.7889	0.86
	40	0.9239	0.61	0.7162	0.85	0.7559	0.75
	50	0.9300	0.54	0.6558	0.76	0.7123	0.67
	100	0.9390	0.39	0.3796	0.54	0.5064	0.47
G(2, 1)	10	0.8937	1.66	0.9125	2.64	0.8977	2.31
	20	0.9181	1.20	0.9143	1.89	0.9079	1.66
	30	0.9280	0.99	0.8943	1.54	0.8929	1.35
	40	0.9325	0.86	0.8752	1.34	0.8824	1.17
	50	0.9348	0.77	0.8521	1.20	0.8649	1.05
	100	0.9428	0.55	0.7152	0.85	0.7727	0.74
G(4, 1)	10	0.9034	2.38	0.9409	3.91	0.9272	3.42
	20	0.9269	1.72	0.9520	2.80	0.9448	2.45
	30	0.9332	1.41	0.9453	2.29	0.9400	2.01
	40	0.9369	1.23	0.9394	1.99	0.9387	1.74
	50	0.9393	1.10	0.9306	1.78	0.9330	1.56
	100	0.9440	0.78	0.8789	1.26	0.8981	1.10
G(8, 1)	10	0.9104	3.38	0.9526	5.65	0.9401	4.94
	20	0.9311	2.44	0.9685	4.06	0.9613	3.55
	30	0.9366	2.00	0.9671	3.32	0.9618	2.90
	40	0.9400	1.74	0.9659	2.88	0.9631	2.52
	50	0.9409	1.56	0.9624	2.58	0.9606	2.26
	100	0.9455	1.10	0.9453	1.83	0.9510	1.60

Table 2: Coverage probability (CP) and average width (AW) of the 95% CIs for t-distributed data

PDF	n	Confidence Interval Methods					
		Classical CI		MDIQR-CI		TRIQR-CI	
		CP	AW	CP	AW	CP	AW
t(4)	10	0.9259	1.63	0.9712	2.33	0.9602	2.04
	20	0.9399	1.18	0.9856	1.65	0.9792	1.44
	30	0.9433	0.98	0.9878	1.33	0.9815	1.17
	40	0.9453	0.85	0.9901	1.16	0.9856	1.01
	50	0.9474	0.77	0.9904	1.03	0.9862	0.90
	100	0.9474	0.55	0.9921	0.73	0.9886	0.64
t(10)	10	0.9207	1.34	0.9676	2.15	0.9562	1.88
	20	0.9369	0.96	0.9835	1.53	0.9779	1.34
	30	0.9418	0.79	0.9859	1.25	0.9809	1.09
	40	0.9431	0.69	0.9885	1.08	0.9848	0.95
	50	0.9447	0.62	0.9892	0.97	0.9862	0.85
	100	0.9476	0.44	0.9912	0.69	0.9895	0.60
t(30)	10	0.9182	1.25	0.9665	2.08	0.9555	1.82
	20	0.9372	0.89	0.9838	1.49	0.9786	1.30
	30	0.9413	0.73	0.9860	1.22	0.9820	1.06
	40	0.9436	0.64	0.9880	1.06	0.9855	0.92
	50	0.9452	0.57	0.9888	0.94	0.9861	0.83
	100	0.9469	0.40	0.9902	0.67	0.9888	0.59
t(50)	10	0.9176	1.23	0.9663	2.06	0.9547	1.80

20	0.9362	0.88	0.9832	1.48	0.9778	1.30
30	0.9404	0.72	0.9851	1.21	0.9809	1.06
40	0.9432	0.63	0.9878	1.05	0.9850	0.92
50	0.9442	0.56	0.9887	0.94	0.9861	0.82
100	0.9476	0.40	0.9905	0.67	0.9893	0.58

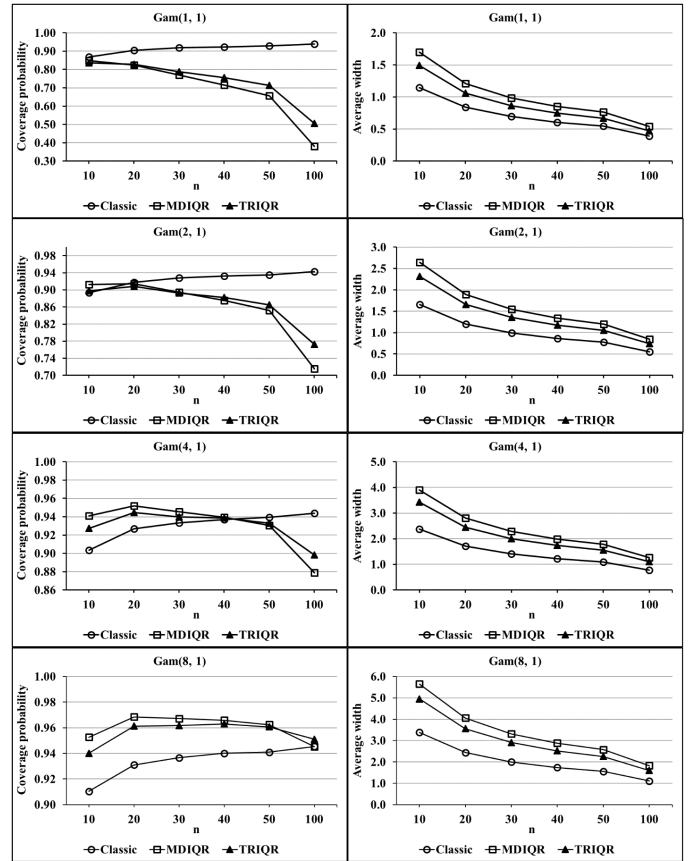


Figure 1: Coverage probabilities and average widths of the three confidence intervals for gamma distributed data

probabilities of the MDIQR and TRIQR interval estimators tend to be greater than 0.95, whereas this of the classical confidence interval tends to be lower than 0.95 for all sample sizes and all the numbers of DFs for the t-distribution. When considering the average width of interval estimators, it is found that the classical confidence interval has the smallest value among the comparative interval estimators for all sample sizes and DFs.

The simulated results in Table 3 and Figure 3 demonstrate the performances of contaminated normal distribution cases that the observations are generated from the linear combination of $N(0, 1)$ and $N(0, 5^2)$ densities with 90%, 80% and 70% of observations are sampled from the $N(0, 1)$ distribution and respectively of 10%, 20% and 30% are sampled from a $N(0, 25)$ distribution. The simulation study shown that the coverage probabilities of the MDIQR and TRIQR interval estimators tend to be greater than 0.95, whereas this of the interval estimators is about 0.95 for all sample sizes and all percentages of the artificial outliers. For the case of linear combination of $N(0, 1)$ and $N(0, 5^2)$ densities, the TRIQR confidence interval has the smallest average width among the comparative interval estimators for all sample sizes and all percentages of the artificial outliers.

That is, the 95% TRIQR interval estimator tends to have the best performance for both criteria—coverage probability and average width of the confidence interval—in this case.

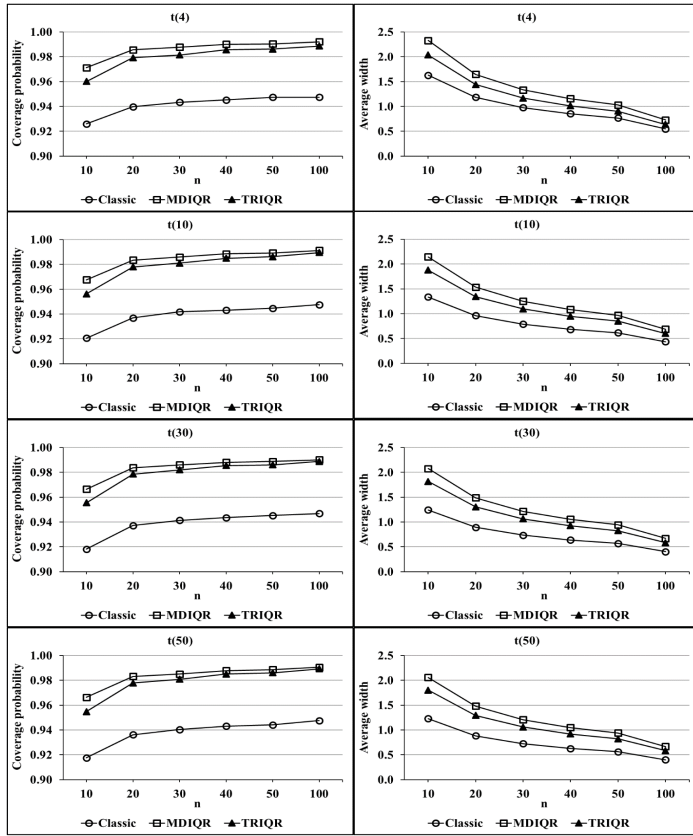


Figure 2: Coverage probabilities and average widths of the three confidence intervals for t-distributed data

Table 3: Coverage probability (CP) and average width (AW) of the 95% CIs for contaminated normal distributed data as the linear combination of $N(0, 1)$ and $N(0, 5^2)$ densities

CND	N	Confidence Interval Methods					
		Classical CI		MDIQR-CI		TRIQR-CI	
		CP	AW	CP	AW	CP	AW
$CN(0,5^2)_{10}$	10	0.9443	2.06	0.9682	2.27	0.9566	1.98
	20	0.9547	1.52	0.9845	1.62	0.9784	1.42
	30	0.9523	1.27	0.9859	1.33	0.9812	1.16
	40	0.9528	1.11	0.9888	1.15	0.9854	1.01
	50	0.9512	1.00	0.9891	1.03	0.9855	0.90
	100	0.9494	0.71	0.9912	0.73	0.9894	0.64
$CN(0,5^2)_{20}$	10	0.9509	2.75	0.9725	2.57	0.9610	2.25
	20	0.9495	2.02	0.9869	1.83	0.9806	1.60
	30	0.9484	1.67	0.9880	1.49	0.9821	1.30
	40	0.9478	1.46	0.9903	1.29	0.9856	1.13
	50	0.9483	1.31	0.9905	1.15	0.9857	1.01
	100	0.9479	0.93	0.9925	0.81	0.9892	0.71
$CN(0,5^2)_{30}$	10	0.9435	3.32	0.9780	3.11	0.9685	2.73
	20	0.9433	2.42	0.9896	2.12	0.9833	1.86
	30	0.9436	2.00	0.9907	1.70	0.9840	1.49
	40	0.9443	1.74	0.9920	1.47	0.9861	1.29
	50	0.9464	1.57	0.9927	1.31	0.9865	1.15
	100	0.9482	1.11	0.9938	0.92	0.9886	0.84

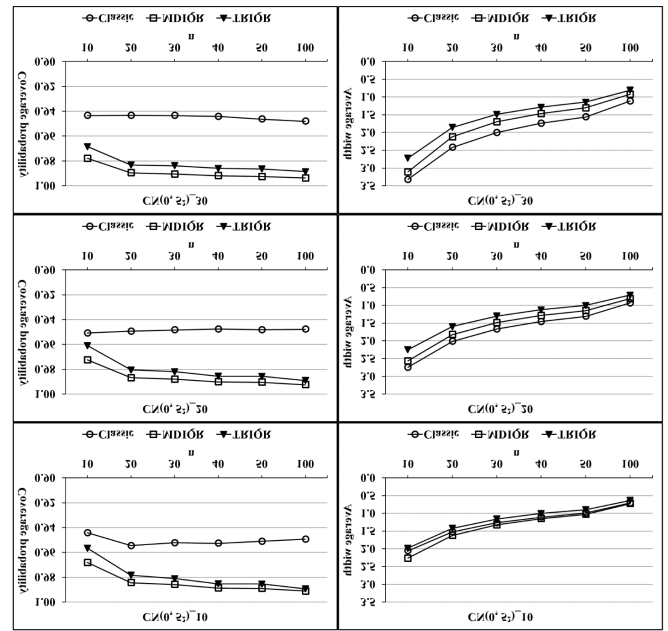


Figure 3: Coverage probabilities and average widths of the three confidence intervals for contaminated normal distributed data as the linear combination of $N(0, 1)$ and $N(0, 5^2)$ densities

The simulated results in Table 4 and Figure 4 show the performances of contaminated normal distribution cases that the observations are generated from the linear combination of $N(0, 1)$ and $N(0, 10^2)$ densities with 90%, 80% and 70% of observations are sampled from the $N(0, 1)$ distribution and respectively of 10%, 20% and 30% are sampled from a $N(0, 100)$ distribution. In this case, the TRIQR confidence interval performs the best efficiency among the three interval estimators for all sample sizes and all percentages of the artificial outliers because the coverage probability of this interval estimator is greater than 0.95 and it has the smallest average width of interval estimator. In addition, the efficiency of MDIQR confidence interval is similar to this of TRIQR interval estimator. In this case, it is found that the classical interval estimators is not robust to outliers—that is, it has the highest average width of interval estimator and the coverage probability of it is smaller than the two proposed robust methods for almost all sample size, especially for a large percentage of outliers.

Table 4: Coverage probability (CP) and average width (AW) of the 95% CIs for contaminated normal distributed data as the linear combination of $N(0, 1)$ and $N(0, 10^2)$ densities

CND	n	Confidence Interval Methods					
		Classical CI		MDIQR-CI		TMIQR-CI	
		CP	AW	CP	AW	CP	AW
$CN(0, 10^2)_{10}$	10	0.9665	3.46	0.9693	2.30	0.9576	2.02
	20	0.9766	2.63	0.9850	1.65	0.9786	1.44
	30	0.9704	2.21	0.9865	1.35	0.9814	1.18
	40	0.9650	1.95	0.9891	1.17	0.9854	1.02
	50	0.9604	1.76	0.9894	1.04	0.9856	0.91
	100	0.9520	1.27	0.9915	0.74	0.9894	0.65
$CN(0, 10^2)_{20}$	10	0.9775	5.08	0.9746	2.68	0.9629	2.34
	20	0.9609	3.78	0.9885	1.90	0.9818	1.67

	30	0.9524	3.14	0.9890	1.54	0.9828	1.35
	40	0.9497	2.75	0.9913	1.33	0.9858	1.17
	50	0.9498	2.47	0.9911	1.19	0.9859	1.04
	100	0.9493	1.77	0.9931	0.84	0.9891	0.74
$CN(0, 10^2)_{-30}$	10	0.9633	6.36	0.9832	3.85	0.9752	3.37
	20	0.9457	4.66	0.9930	2.39	0.9876	2.09
	30	0.9452	3.86	0.9927	1.84	0.9860	1.61
	40	0.9454	3.36	0.9942	1.58	0.9876	1.39
	50	0.9471	3.02	0.9941	1.40	0.9873	1.23
	100	0.9486	2.15	0.9951	0.98	0.9886	0.86

4.1. Example 1: Melting Points of Beeswax Data

The data of this example is considered from [10] (cited in [11]), p.378) and introduced by [12]. Table 5 provides data representing the melting points (°C) of beeswax obtained from 59 sources.

The statistical summary of the melting points (°C) of beeswax data was calculated and given below in Table 6.

Table 6: Statistical summary for the melting points (°C) of beeswax data

Statistics	Abbreviations	Values
Sample Mean	\bar{X}	63.589
Sample Median	MD	63.530
Sample Trimean	TR	63.564
Sample Standard Deviation	S	0.347
Inter-Quartile Range	IQR	0.475

According to [12], it is known that the population mean of the melting point of beeswax (μ) is about 63.580 °C. The histogram, density plot, Boxplot and normal Q-Q plot are displayed in Figure 5. As can be observed, a goodness-of-fit test for normality assumption by using the Kolmogorov-Smirnov (K-S) statistical test provides a p-value is greater than $\alpha = 0.05$ (KS = 0.086, p-value > 0.150), we conclude that the data are met a normal distribution assumption. The plots in Figure 5 are consistent with the above conclusion.

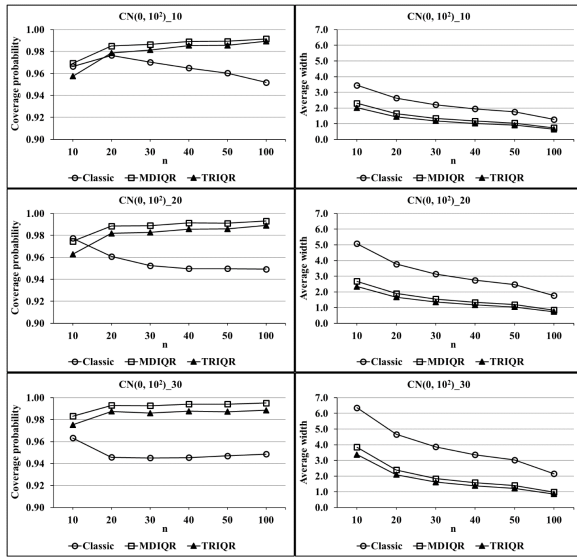


Figure 4: Coverage probabilities and average widths of the three confidence intervals for contaminated normal distributed data as the linear combination of $N(0, 1)$ and $N(0, 10^2)$ densities

Table 5: Melting points of beeswax data

No.	X	No.	X	No.	X	No.	X
1	63.78	16	63.92	31	64.42	46	64.12
2	63.83	17	63.86	32	63.50	47	63.03
3	63.88	18	63.13	33	63.84	48	63.66
4	63.78	19	63.08	34	64.21	49	63.34
5	63.50	20	63.30	35	64.40	50	63.34
6	63.41	21	63.51	36	62.85	51	63.56
7	63.45	22	63.56	37	63.27	52	63.92
8	63.63	23	63.93	38	63.36	53	63.68
9	63.36	24	63.69	39	64.27	54	63.60
10	63.92	25	63.40	40	64.24	55	63.50
11	63.30	26	63.83	41	63.61	56	63.92
12	63.60	27	63.51	42	63.31	57	63.39
13	63.58	28	63.43	43	63.10	58	63.53
14	63.27	29	63.43	44	63.86	59	63.13
15	63.36	30	63.05	45	63.50		

4. Application with Real Data

We consider two real-life examples from normal and non-normal distributions to illustrate the findings of the paper in this section.

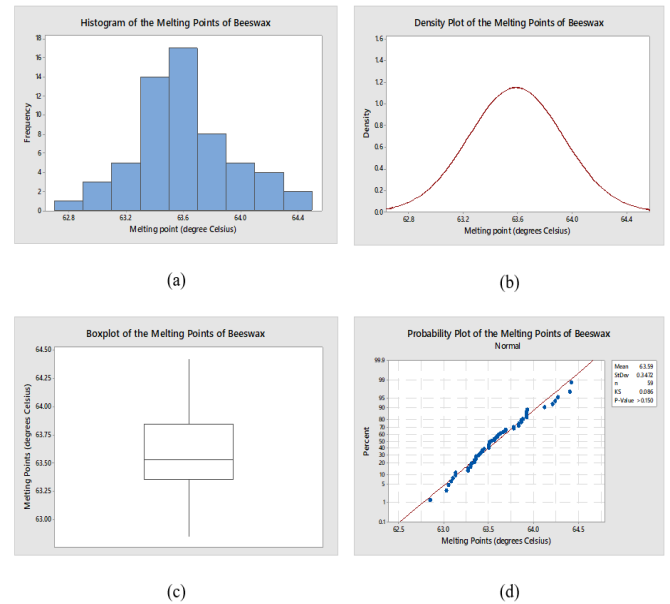


Figure 5: Plots for the melting points (°C) of beeswax data

The 95% interval estimator of μ and the corresponding widths for the proposed intervals are given below in Table 7.

Table 7: The 95% CIs for the population mean (μ) of the melting points (°C) of beeswax data

Methods	Confidence Interval Limits		Widths
	Lower Limit	Upper Limit	
Classical	63.500	63.678	0.178
MDIQR	63.378	63.682	0.304
TRIQR	63.431	63.697	0.266

It is observed from Table 7 that all the interval estimators include the true population mean ($\mu = 63.580$). The classical interval estimator has the shortest interval width followed by CI_{TRIQR} and CI_{MDIQR} , so both the classical and proposed interval estimators did well. Hence, these results are consistent with the simulation study.

4.2. Example 2: Urinary Tract Infections (UTI) Data

The data of this example represent the duration of male patient urinary tract infections (UTIs) in days and presented in Table 8. It was considered by various researchers, among them, [13–15] are notable. The summary statistics of the urinary tract infections (UTIs) data are displayed in Table 9. The histogram, density plot, Box-plot and normal Q-Q plot are given in Figure 6. As it can be observed, the Kolmogorov-Smirnov (K-S) statistical test provides a p-value less than $\alpha = 0.01$ ($KS = 0.212$, $p\text{-value} < 0.010$), which indicates that the data do not follow normal distribution. The plots in Figure 6 supported the above conclusion. It is noted from Santiago and Smith (2013) that the data are well fitted to an exponential distribution with a mean time of $\mu = 0.2100$ days.

Table 8: Urinary tract infection (UTI) data

No.	X	No.	X	No.	X
1	0.57014	19	0.12014	37	0.27083
2	0.07431	20	0.11458	38	0.04514
3	0.15278	21	0.00347	39	0.13542
4	0.14583	22	0.12014	40	0.08681
5	0.13889	23	0.04861	41	0.40347
6	0.14931	24	0.02778	42	0.12639
7	0.03333	25	0.32639	43	0.18403
8	0.08681	26	0.64931	44	0.70833
9	0.33681	27	0.14931	45	0.15625
10	0.03819	28	0.01389	46	0.24653
11	0.24653	29	0.03819	47	0.04514
12	0.29514	30	0.46806	48	0.01736
13	0.11944	31	0.22222	49	1.08889
14	0.05208	32	0.29514	50	0.05208
15	0.12500	33	0.53472	51	0.02778
16	0.25000	34	0.15139	52	0.03472
17	0.40069	35	0.52569	53	0.23611
18	0.02500	36	0.07986	54	0.35972

Table 9: Statistical summary for the urinary tract infections (UTIs) data

Statistics	Abbreviations	Values
Sample Mean	\bar{X}	0.2103
Sample Median	MD	0.1424
Sample Trimean	TR	0.1580
Sample Standard Deviation	S	0.2119
Inter-Quartile Range	IQR	0.2431

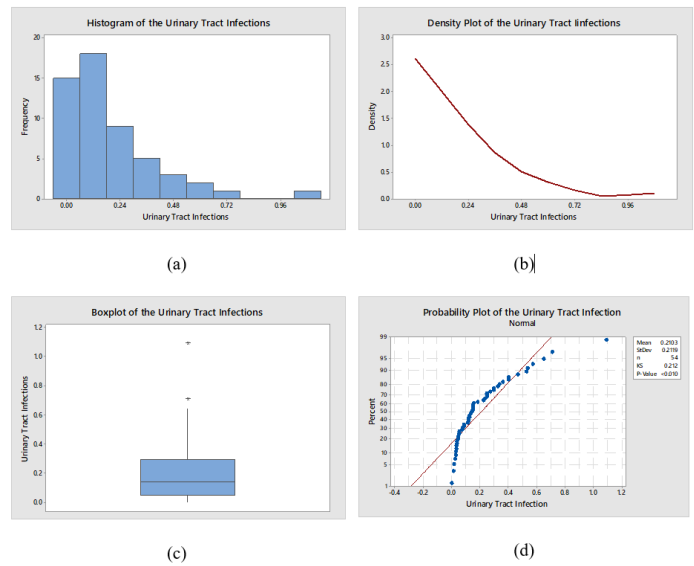


Figure 6: Plots for the urinary tract infections (UTIs) data

The 95% interval estimator of μ and the corresponding widths for all proposed interval estimators are given below in Table 10.

Table 10: The 95% CIs for the population mean (μ) of the urinary tract infections (UTIs) data

Methods	Confidence Interval Limits		Widths
	Lower Limit	Upper Limit	
Classical	0.1538	0.2668	0.1130
MDIQR	0.0612	0.2236	0.1624
TRIQR	0.0869	0.2291	0.1422

It is observed from Table 10 that all interval estimator capture the true population mean ($\mu = 0.2100$). The classical interval estimator has the shortest width followed by CI_{TRIQR} and CI_{MDIQR} , so both the classical and proposed interval estimators performed well. These results are consistent with the simulation study.

5. Some Concluding Remarks

For estimation of the population mean (μ), two robust interval estimators, namely, the median interquartile range (MDIQR-CI) and the trimean interquartile range (TRIQR-CI) are proposed in this paper. The simulation study evident that both criteria—coverage probability (CP) and average width (AW)—of the proposed robust interval estimators tend to have a good performance when observations are sampled from the contaminated normal distribution, especially for the high percentage of outliers and the main distribution is contaminated with the wider distribution. However, when observations are sampled from non-normal distributions, gamma and t-distributions, the classical confidence interval tends to have the best performance for the average width criterion, whereas the coverage probability of this tends to be smaller than those of the proposed robust interval estimators. Two data sets are analyzed to illustrate the performance of the interval estimators, which supported the simulation study. Finally, the proposed robust interval estimators are easy to compute, not computer intensive and promising, so that they can be recommended for the practitioners when these compare with the bootstrap confidence

interval that suggested by [1]. As mention in the introduction section that bootstrap confidence interval complicates to implement in practices because it is not easy to compute without the statistical programming [2].

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Authors are thankful to anonymous referees and editor for their valuable comments and suggestions, which certainly improved the quality and presentation of the paper. Author, B. M. Golam Kibria wants to dedicate this paper to his most favorite teacher and guardian, late Prof. A. B. M. Abdus Sobhan Miah, Department of Statistics, Jahangirnagar University for his wisdom, constant inspiration during student life and affection that motivated him to achieve this present position.

References

- [1] B. Efron, R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, 1993.
- [2] J. Carpenter, J. Bithell, "Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians" *Stat. Med.*, **19**(9), 1141–1164, 2000.
[https://doi.org/10.1002/\(SICI\)1097-0258\(20000515\)19:9<1141::AID-SIM479>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1097-0258(20000515)19:9<1141::AID-SIM479>3.0.CO;2-F)
- [3] M. L. Tiku, A. D. Akkaya, *Robust estimation and hypothesis testing*, new age international (P) limited, 2004.
- [4] M. O. Abu-Shawiesh, "A simple robust control chart based on MAD" *Math. Stat.*, **4**(2), 102-107, 2008.
doi:10.3844/jmssp.2008.102.107
- [5] H. W. Akyüz, H. Gamgam, A. Yalçinkaya, "Interval estimation for the difference of two independent nonnormal population variances" *GU. J. Sci.*, **30**(3), 117-129, 2017.
- [6] H. F., Weisberg, *Central Tendency and Variability*. Sage University Paper, USA: Series on Quantitative Applications in the Social Sciences, 1992.
- [7] C. M. Borrer, D. C. Montgomery, G. C. Runger, "Robustness of the EWMA control chart to non-normality" *J. Qual. Technol.*, **31**(3), 309–316, 1999.
<https://doi.org/10.1080/00224065.1999.11979929>
- [8] Z. G. Stoumbos, M. R. Jr. Reynolds, "Robustness to non normality and autocorrelation of individual control charts" *J. Stat. Comput. Sim.*, **66**(2):145–187, 2000. doi: 10.1080/00949650008812019
- [9] M. O. Abu-Shawiesh, F. M. Al-Athari, H. F. Kittani "Confidence interval for the mean of a contaminated normal distribution" *J. Appl. Sci.*, **9**(15), 2835-2840, 2009. doi: 10.3923/jas.2009.2835.2840
- [10] J. White, M. Riethof, I. Kushmir, "Estimation of microcrystalline wax in beeswax" *J. Assoc. Off. Anal. Chem.*, **43**(4), 781-790, 1960.
<https://doi.org/10.1093/jaoac/43.4.781>
- [11] J. A. Rice, *Mathematical Statistics and Data Analysis*. California, Duxbury Press, 2007.
- [12] W. Panichkitkosolkul, "Confidence interval for the coefficient of variation in a normal distribution with a known population mean after a preliminary t test". *KMITL Sci. Tech. J.*, **15**(1), 34 – 46, 2015.
- [13] E. Santiago, J. Smith, "Control charts based on the exponential distribution: adapting runs rules for the t chart" *Qual. Eng.*, **25**(2), 85-96, 2013.
<https://doi.org/10.1080/08982112.2012.740646>
- [14] M. Aslam, N. Khan, M. Azam, C. H. Jun, "Designing of a new monitoring t-chart using repetitive sampling" *Inf. Sci.*, 269, 210-216, 2014.
<https://doi.org/10.1016/j.ins.2014.01.022>
- [15] M. Azam, M. Aslam, M., C. H. Jun, "An EWMA control chart for the exponential distribution using repetitive sampling plan" *Operations Research and Decisions*, **27**(2), 5-19, 2017. doi: 10.5277/ord170201

A Solution Applying the Law on Road Traffic into A Set of Constraints to Establish A Motion Trajectory for Autonomous Vehicle

Quach Hai Tho^{1,*}, Huynh Cong Phap², Pham Anh Phuong³

¹University of Arts, Hue University, 530.000, Vietnam

²Vietnam - Korea University of Information and Communication Technologies, Da Nang University, 550.000, Vietnam

³Faculty of Information Technology, University of Education, Da Nang University, 550.000, Vietnam

ARTICLE INFO

Article history:

Received: 12 May, 2020

Accepted: 06 June, 2020

Online: 18 June, 2020

Keywords:

Autonomous Vehicle

Model Predictive Control

Path Planning

Motion Planning

Intelligent transportation systems

ABSTRACT

With a model predictive control approach and to set the motion trajectory for an autonomous vehicle in situations where emergency braking cannot be performed, in this paper, we propose a solution to apply the law on road traffic into a set of constraints and thereby build an objective function to create motion trajectory for autonomous vehicle. The newly created trajectory is created to improve performance and enhance the ability to avoid obstacle but ensure an optimal global trajectory. The performance of this solution is assessed through simulation with different scenarios, from which there are applied research orientations on the problem of autonomous vehicle in practice.

1. Introduction

In recent years, many researches on autonomous vehicle problems have been carried out based on basic components such as navigation systems, environmental perception, planning and control [1,2,3]. In these components, motion planning is an important function to determine motion process of the vehicle. It provides the coming target of the vehicle by using the information received from the environment and navigation system. Therefore, the planning components must consider not only the vehicle elements but also the legal, ethical, and environmental changes through perception data received in the system to ensure reliability and safety when participating in traffic.

One of the challenges and tasks that researchers are interested in the field of autonomous vehicles is to create the optimal motion trajectory. It includes certain criteria such as creating smooth movement, creating comfort and achieving good energy efficiency, and must meet the restrictions arising during the operation of the vehicle such as the provision about the road traffic law.

The problem of determining the optimal path in complex environmental conditions has been studied with different approaches [4] such as potential field technology, searching techniques on graphs, or model-based techniques [5], in which the

approach of interest in research is model-based. These methods are widely used for structured road and optimal trajectory that will be determined from a set of candidate trajectories. However, one difficulty faced by model-based motion planning solutions is how to effectively sample candidates in trajectory space. Since the model-based methods basically reach the lower end of the final target trajectory with optimal techniques, they require a sufficient amount of resources to obtain a large number of candidates to find the global optimal trajectory. Therefore, the optimal of model sample trajectories does not allow real-time application.

In order to solve the above-mentioned difficulties, in recent years, the planning methods of continuous air-space planning with the method based on optimal control and model predictive control (MPC) [6,7] has attracted research interest. The MPC approach usually uses nonlinear optimization technique, which will solve the problem with an iterative process for optimal control of the predictive horizon and that is the advantage of this method. From which the system built is capable of handling constraints imposed to ensure safety and facilitate movement planning in complex environmental conditions. This method can meet the requirements implementing real time.

In this paper, we propose a solution to set the motion trajectory for autonomous vehicle for a given period of time with constraints

Corresponding Author: Quach Hai Tho, Email: qhaltho@hueuni.edu.vn

www.astesj.com

<https://dx.doi.org/10.25046/aj050356>

built from the rules of road traffic law. The goal is to improve not only the calculation efficiency but also the uncertainty in the perceptual data of the proposed environment and vehicle system. The experimental simulation process in obstacle avoidance situations, overtaking other vehicles and decision to change lanes, etc. is carried out in a short distance, so that the problem is solved focusing on a prediction interval, a single MPC phase.

In the next section, we will introduce the set of constraints and the basic principles of the constraint clause, which are built from the provisions of road traffic law when participating of autonomous vehicle. Finally, we present the empirical part with the simulation process and conclude with some suggestions for further research directions for the problem of autonomous vehicle.

2. Building solutions for establishing motion trajectory

In the process of developing solutions to set the motion trajectory for autonomous vehicle, the factors need to be calculated and considered such as the dynamic system, the set of constraints in the operation of the vehicle, limitations on the operating environment such as the structure of the road system, the regulations of the law, the obstacles on the road ... The aim of the problem is to find a trajectory that is safe and feasible.

In this paper, we propose an approach to use model predictive control (MPC) to perform the motion trajectory planning, in which the expected trajectory is updated in subsequent stages of the model and build the cost function with logical constraints for the problem of generating the optimal motion trajectory.

However, in this paper, we only solve this motion planning problem in the case the autonomous vehicle is moving on a straight path and has a lane mark for overtaking. At the same time, the process of determining the trajectory is carried out in a predictive range of a predictive control phase, from which we can model roads, vehicles and other objects in the Cartesian coordinate system (x, y) overall, where the x -axis is the vertical direction and the y -axis is the horizontal direction of the road. Each vehicle and other objects are described by their position (S_x, S_y) within the boundary of the lane. The system state representations include the orientation angle θ_r , the distance from the boundary to the vehicle d_r , the reference arc length S_r and $K_r(S_r)$ are parameters of the arc curvature of the reference curve.

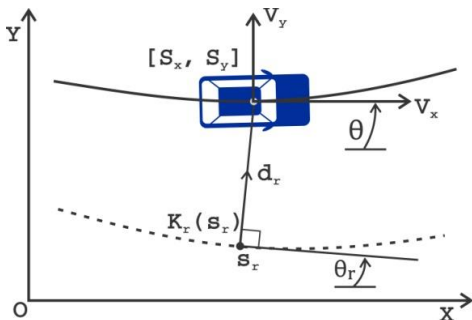


Figure 1. Vehicle model and reference curve

To facilitate the performance, in this paper we show the structure of the road as a system of roads defined by adjacent lanes of arbitrary shape and curvature. We assume and consider the i ($lane_i$) to be a path defined by the left boundary (B_{Li}) and the right boundary (B_{Ri}). Each such path is defined as a polyline and is a combination of all lanes at a given time interval ($lane(s) = \cup_i lane_i$).

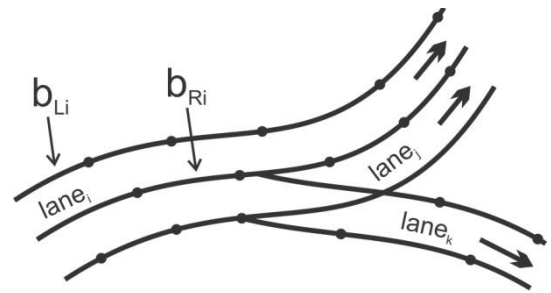


Figure 2. Road system model

The main idea when building a motion trajectory solution is to use a model predictive control (MPC) method, in which the set of constraints is constructed as logical propositions, including constraints on legal and ethical issues of traffic behavior. Based on the provisions of Vietnam traffic law, we perform logical clauses to introduce the constraint set in traffic as follows:

Rule 1: If the vehicle discovers obstacle ahead, it must slow down.

Rule 2: If it overtakes another vehicle, it must do so on the left side.

Rule 3: If it overtakes another vehicle on the right side, it must satisfy the condition that the vehicle ahead has given a signal to turn left or is turning left.

Rule 4: If it is traveling in an area where overtaking is prohibited or where the weather conditions do not ensure safe, it cannot overtake another vehicle.

Rule 5: If it changes lane or changes direction, it must give a turn signal.

Rule 6: If it changes direction, it must give way when detecting a vehicle ahead that is rudimentary vehicles and pedestrians.

Rule 7: If it moves in a residential area, the movement speed of the vehicle does not exceed 50 km/h and outside the residential area is 80 km/h.

Rule 8: If it performs changing lane and overtaking, it cannot do so continuously.

However, the issue of motion control decision depends on many situations, different states of the object types, and from which the vehicle's motion trajectory is set appropriately. Therefore, for each class the object can be mapped to a single homotopy layer of trajectory. To describe these homotopy layers, we will also build with corresponding logical propositions. For example, the logical proposition of the homotopy class is described when meeting obstacles as follow:

Rule 9: Vehicle in motion, if it overtakes other vehicles, it must overtake on the left side or on the right side.

Rule 10: Vehicle in motion, if it avoids obstacle, it will overtake to the left or to the right of the obstacle and reduce its movement speed.

As we all know, the essence of predictive control is to use an explicit model of the object to calculate optimal variables controlled through optimization methods. Specifically, to control the prediction for an object, we need to perform the following steps: Building a predictive model, defining target functions and constraint conditions, and finally solving the optimal problem. At the same time a number of predetermined conditions such as

communication between vehicles and future trajectories are determined [8, 9] during the construction of the trajectory of the motion plan.

With the idea of using the model predictive control and the first part of the proposed vehicle model with the state vector ω , control vector u as the basis for the research problem as follows:

$$\omega = [s_x, v_x, s_y, v_y]^T \quad (1)$$

$$u = [a_x, a_y]^T \quad (2)$$

where s_x, s_y represent vertical and horizontal position, v_x, v_y are velocity and a_x, a_y are vehicle acceleration along the x, y axes of the vehicle in inertial frame.

Then the dynamic model of the vehicle is represented with zero matrices of proper dimension, as follows:

$$\dot{\omega}(t) = \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \omega + \begin{bmatrix} B & \mathbf{0} \\ \mathbf{0} & B \end{bmatrix} u \quad (3)$$

where $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

In this problem, we make the assumption that the control vector u is a constant function at each time step τ . Therefore, the dynamical model of the vehicle is represented approximately with initial values including state vector $\omega(k)$ and control vector $u(k)$ in time interval $[k\tau, (k+1)\tau]$ as follows:

$$\omega(k+1) = \begin{bmatrix} A^d & \mathbf{0} \\ \mathbf{0} & A^d \end{bmatrix} \omega(k) + \begin{bmatrix} B^d & \mathbf{0} \\ \mathbf{0} & B^d \end{bmatrix} u(k) \quad (4)$$

where $A^d = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix}, B^d = \begin{bmatrix} \frac{1}{2}\tau^2 \\ \tau \end{bmatrix}$

In order to achieve computational efficiency with the limitations of the vehicle's dynamic system as well as the provisions about motion direction when overtaking obstacles on the road, it is necessary to consider the constraints on the state $\omega(\cdot)$, the input control signal $u(\cdot)$ and the motion direction of the vehicle $\theta(\cdot)$ must meet the following conditions so that when the motion trajectory is built, it is feasible as follows:

$$\omega \in [\omega_{\min}, \omega_{\max}] \quad (5)$$

$$u \in [u_{\min}, u_{\max}] \quad (6)$$

$$\theta \in [\theta_{\min}, \theta_{\max}] \quad (7)$$

In which, the upper and lower bound values are calculated as follows:

$$\omega_{\min} = [0, 0, s_{\min_y}, v_{\min_y}]^T \quad (8)$$

$$\omega_{\max} = [\infty, v_{\max_x}, s_{\max_y}, v_{\max_y}]^T \quad (9)$$

$$u_{\min} = [a_{\min_x}, a_{\min_y}]^T \quad (10)$$

$$u_{\max} = [a_{\max_x}, a_{\max_y}]^T \quad (11)$$

$$\theta = \arctan (v_y/v_x) \quad (12)$$

$$v_y \in [v_x \tan(\theta_{\min}), v_x \tan(\theta_{\max})] \quad (13)$$

In addition, in this problem we incorporate a set of constraints that are the provisions of the road traffic law. These provisions are considered as logical clauses and will perform the conversion into a set of linear inequalities with integer variables.

The transformation is done as follows: in the literals clause is represented as an indivisible statement with a linear expression on the state variables. The literals use logical operations like conjunction (\wedge), disjunction (\vee), exclusive - OR (\oplus), implication (\rightarrow), equivalence (\leftrightarrow), negation (\bar{A}) to represent. At the same time, these literals will be represented as a binary variable $\delta(\cdot)$ That only accepts 0 or 1 values, if a proposition is true then $\delta(\cdot)$, otherwise it is false then $\delta(\cdot) = 0$.

For example, to perform the constraint laws regulating about speed (Rule 1), travel speed (Rule 7) along with the structural constraints of the vehicle as follows: To ensure safety when participating in traffic then at position $x \in [200m, 1000m]$ from the starting position, will limit the vehicle movement speed to $v \leq 50km/h$ (Figure 3), there will be 3 literals defined $P_1 = [x(k) \geq 200]$, $P_2 = [x(k) \leq 1000]$ and $P_3 = [v_x(k) \leq 50]$, and the logical clauses expressed at any point in the segment are $\forall k \geq 0, (P_1 \wedge P_2) \Rightarrow P_3$.

Or we represent it after assigning a binary variable $\delta_i(k) = 1 \Rightarrow P_i$ with $i \in \{1,2,3\}, \forall k \geq 0$ as follows:

$$\delta_1(k) = 1 \Rightarrow x(k) \geq 200 \quad (14)$$

$$\delta_2(k) = 1 \Rightarrow x(k) \leq 1000 \quad (15)$$

$$\delta_3(k) = 1 \Rightarrow v_x(k) \leq 50 \quad (16)$$

$$-\delta_1(k) + \delta_3(k) \leq 0 \quad (17)$$

$$-\delta_2(k) + \delta_3(k) \leq 0 \quad (18)$$

$$\delta_1(k) + \delta_2(k) - \delta_3(k) \leq 1 \quad (19)$$

By the Big-M method [10], we can convert inequalities (14), (15) and (16) into the following:

$$x(k) \geq 200 - M(1 - \delta_1(k)) \quad (20)$$

$$x(k) \leq 1000 - M(1 - \delta_2(k)) \quad (21)$$

$$v_x(k) \geq 50 - M(1 - \delta_3(k)) \quad (22)$$

where M is a positive constant with great value. In this problem, we choose $M = 10^9$

Let $K = T/\tau$ be the number of time steps in the prediction horizon with T be the time interval, $\omega = \{\omega(0), \dots, \omega(K)\}$ and $u = \{u(0), \dots, u(K)\}$ are state vectors, control trajectory vectors

in the given time horizon, ω_r is the reference trajectory of the vehicle, which may depend on time, status, or traffic law.

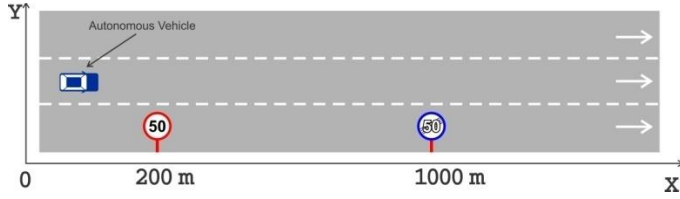


Figure 3. Speed limit ramp

To determine the objective function for this solution, we will introduce a new value vector variable $\delta(k) = \{0,1\}^m$ and δ_r , where $\delta(k)$ is a collection of all the binary variable results from the rebuilding the provisions of law on road traffic into linear inequalities and δ_r is the reference trajectory for the binary variables where we can make options on some binary states.

Thus, at time $t = 0$, the optimal problem of the model predictive control with the objective function can be written as follows:

$$\min_{u, \delta} J(\omega, u, \delta) = \sum_{k=0}^K (\|\omega(k) - \omega_r(k)\|_Q^2 + \|\delta(k) - \delta_r(k)\|_R^2 + \|u(k)\|_S^2 + \|\Delta u(k)\|_W^2) \quad (23)$$

Conditions satisfied

$$\omega(k+1) = \begin{bmatrix} A^d & 0 \\ 0 & A^d \end{bmatrix} \omega(k) + \begin{bmatrix} B^d & 0 \\ 0 & B^d \end{bmatrix} u(k) \quad (24)$$

$$\text{with } k = [0, \dots, K-1]$$

$$\omega(k) \in [\omega_{\min}, \omega_{\max}] \text{ with } k = [0, \dots, K] \quad (25)$$

$$\omega_{\min} = [0, 0, s_{\min y}, v_{\min y}]^T \quad (26)$$

$$\omega_{\max} = [\infty, v_{\max x}, s_{\max y}, v_{\max y}]^T \quad (27)$$

$$u(k) \in [u_{\min}, u_{\max}] \text{ with } k = [0, \dots, K] \quad (28)$$

$$u_{\min} = [a_{\min x}, a_{\min y}]^T \quad (29)$$

$$u_{\max} = [a_{\max x}, a_{\max y}]^T \quad (30)$$

$$\theta \in [\theta_{\min}, \theta_{\max}] \text{ with } \theta = \arctan(v_y/v_x) \quad (31)$$

$$v_y(k) \in [v_x(k) \tan(\theta_{\min}), v_x(k) \tan(\theta_{\max})] \quad (32)$$

$$\text{with } k = [0, \dots, K]$$

$$\Delta u(k) = u(k) - u(k-1) \text{ with } k = [0, \dots, K] \quad (33)$$

$$C \begin{bmatrix} \omega \\ \omega_r \\ \delta \end{bmatrix} \leq D \quad (34)$$

where the matrices C, D, Q, R, S and W are positive weight matrices of proper dimension, the final constraint (34) is the set of all provisions of road traffic law into linear inequality is represented in matrices.

The objective function (23) will be quadratic if the values of ω_r and δ_r are independent or linearly dependent on other variables. Therefore, solving this optimization problem will result in the global optimal trajectory being used so that it will then move to the controller at a lower level in tracking this trajectory.

Thus, given this optimal problem with the correlation and integration issues between vehicle dynamics and the provisions of the road traffic law, this solution can be used to effectively handle during the motion planning process with different situations during travelling. In order to achieve the optimal motion trajectory, it is necessary to process the cost function to a minimum and at the same time constraints including vehicle dynamics constraints and rules of participating in traffic after converting into logical constraints, it should be separated.

From the idea raised for this problem and in order to evaluate the effectiveness of the solution for the specific operating environment of the vehicle, the cost function is designed to optimize operational control with the initial condition that the vehicle velocity is constant and the horizontal deviation will change over time during the travel distance. Therefore, the expression $\|\delta(k) - \delta_r(k)\|_R^2$ in the optimal problem (23) will not be considered, so the cost function built in this study is given as follows:

$$J = \sum_{k=0}^K (q_1(v_x(k) - v_r)^2 + q_2(y(k) - y_r(k))^2 + q_3(v_y(k))^2 + q_4(a_x(k))^2 + q_5(a_y(k))^2 + r_1(a_x(k) - a_x(k-1))^2 + r_2(a_y(k) - a_y(k-1))^2) \quad (35)$$

3. Experimental assessment

In order to ensure the reliability, objectivity and effectiveness of the solution, we have conducted simulations with different scenarios such as assessments on the roads with speed limit signs, determining the motion trajectory of the vehicle when there is an obstacle or a motion trajectory of the vehicle when overtaking in the same direction.

In the process of checking and evaluating the given solution, we conduct empirical simulation of processes in the Matlab environment. In which the calculation (using the SI measurement system) with the sampling period $\tau = 0.2s$ and the predictive horizon $T = 10s$. During the simulation, vehicles with recorded trajectory will be considered vehicles ahead and the autonomous vehicle will be behind these vehicles. The parameters used in the experimental process are given as follows:

$$\omega_{\min} = [0, 0, -4, -2]^T, \omega_{\max} = [1500, 80, 10, 5]^T, u_{\min} = [-5, -1]^T, u_{\max} = [5, 1]^T, \theta_{\min} = -0.5, \theta_{\max} = 0.5, q_1 = 1, q_2 = 3, q_3 = 5, q_4 = 2, q_5 = 4, r_1 = 6, r_2 = 16.$$

The location of the experimental vehicle is in the right lane with the original position at the coordinates (0, 0), the original and constant speed during the experiment of the vehicle is $v_x(0) = v_r = 60\text{km/h}$. The trajectory switches to the I/O controller with a simulation time of 0.05s.

Scenario 1: In this scenario, the autonomous vehicle will move into a road with limited speed (Figure 4). At the beginning, the traveling speed of autonomous vehicles is stable $v_r = 60\text{km/h}$. When entering the prescribed speed range, with the input values and constraints for the control system will be able to make the vehicle's movement speed change. The control system will perform the braking operation to reduce vehicle speed to a safe limit with restrictions on road traffic rules. With such a driver assistance process, it shows the advantages of controlling the longitudinal force, horizontal force and acceleration the vehicle, as well as implementing the vehicle's movement plan to be complete when coming to the road systematic motion velocity regulation.

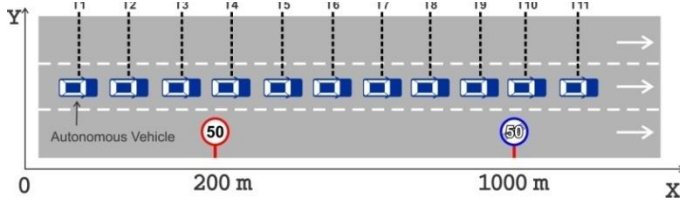


Figure 4. The vehicle moves over the speed limit segment

During the simulation process, with the parameters shown, the construction of constraints according to Rule 7 is carried out as follows:

When the vehicle moves into position $T4 \in [200m, 1000m]$, it will limit the movement speed of the vehicle is $v \leq 50\text{km/h}$, there will be 3 literals defined $P_1 = [x(k) \geq 200]$, $P_2 = [x(k) \leq 1000]$ and $P_3 = [v_x(k) \leq 50]$, and the logical propositional represented at all times in the segment is $\forall k \geq 0, (P_1 \wedge P_2) \Rightarrow P_3$. The constraints set is constructed to include:

$$x(k) \geq 200 - M(1 - \delta_1(k))$$

$$x(k) \leq 1000 - M(1 - \delta_2(k))$$

$$v_x(k) \geq 50 - M(1 - \delta_3(k))$$

$$-\delta_1(k) + \delta_3(k) \leq 0$$

$$-\delta_2(k) + \delta_3(k) \leq 0$$

$$\delta_1(k) + \delta_2(k) - \delta_3(k) \leq 1$$

where M is a positive constant with great value. In this simulation, we choose $M = 10^9$

The simulation results show that the movement speed of the vehicle from position T4 to position T10 according to the vehicle's planning of movement decreases below the speed of $v_r \leq 40\text{km/h}$, so the motion controller has been effectively applied.

Scenario 2: In this scenario, the autonomous vehicle will move in the same lane with other vehicles and the autonomous vehicle will perform the right overtaking for the vehicle ahead (Figure 5). At the beginning, the autonomous vehicle and other vehicles are in one lane with the distance from autonomous vehicle to the vehicles ahead is 30m, the initial speed of the vehicles ahead is 30km/h and of autonomous vehicles is 60km/h.

At the time of T3, the vehicle in front of the vehicle moved at a constant speed, in this situation, the vehicle would either self-drive or reduce the speed of $v_r \leq 30\text{km/h}$ or perform an

overtaking operation or an accident will occur. During the simulation, we choose to overtake and to ensure that overtaking does not violate road traffic rules, autonomous vehicle must perform overtaking in accordance with Rule 2.

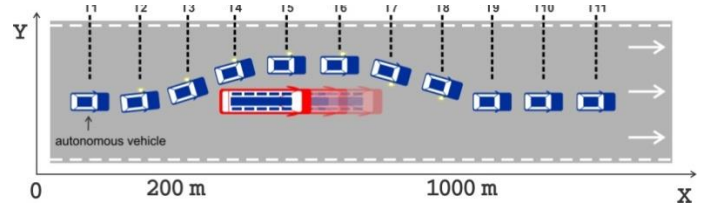


Figure 5. Autonomous vehicle passing to the right of the vehicle ahead

To set the control input values, we will describe the vehicles in front surrounded by rectangles. Each different object will have rectangles of different sizes and positions respectively. In this simulation, the position of the vehicle in front is $S_{obs} = (400,1)$ and the rectangle surrounding the vehicle in front is $L_{obs} = 12.5\text{m}$ and the width of $W_{obs} = 2.9\text{m}$. The rectangle around the vehicle in front of O_{obs} is defined in the coordinate system as $[x_o(k) - L_{obs}, x_o(k) + L_{obs}] \times [y_o(k) - W_{obs}, y_o(k) + W_{obs}]$ with $\forall k \geq 0$.

Thus, the constraint set is constructed to avoid collisions by overtaking on the right side as follows:

$$\delta_1(k) = 1 \Rightarrow x(k) \leq x_o(k) - L_{obs}$$

$$\delta_2(k) = 1 \Rightarrow x(k) \geq x_o(k) + L_{obs}$$

$$\delta_3(k) = 1 \Rightarrow x(k) \leq y_o(k) - W_{obs}$$

$$\delta_4(k) = 1 \Rightarrow x(k) \geq y_o(k) + W_{obs}$$

$$\delta_1(k) + \delta_2(k) + \delta_3(k) + \delta_4(k) = 1$$

or represented by the following inequalities:

$$x(k) \leq (x_o(k) - L_{obs}) - M(1 - \delta_1(k))$$

$$x(k) \geq (x_o(k) + L_{obs}) - M(1 - \delta_2(k))$$

$$x(k) \leq (y_o(k) - W_{obs}) - M(1 - \delta_3(k))$$

$$x(k) \geq (y_o(k) + W_{obs}) - M(1 - \delta_4(k))$$

$$\delta_1(k) + \delta_2(k) + \delta_3(k) + \delta_4(k) = 1$$

where M is a positive constant with great value. In this simulation, we choose $M = 10^9$

The simulation results show that the moving speed of the vehicle from position T1 to position T11 according to the movement plan of the vehicle is almost unchanged $v_r \approx 60\text{km/h}$, but the steering angle of the vehicle begins to change from position T2 to position T5 is $\theta(k) \approx -0.3\text{rad}$, from position T6 to position T8, it decreases to $\theta(k) \approx 0.3\text{rad}$ and from position T9 the steering angle $\theta(k) = 0\text{rad}$. The change of steering angle can be explained as follows, from position T2 the vehicle begins to move the steering angle to the right from the original movement direction to avoid collision and proceed to overtake to the right side

of the object. After overtaking the object, at the position of T6, the autonomous vehicle will move back to the left side to return e original direction of movement. From position T9, the autonomous vehicle has passed the object and adjusted the angle back to the original to return to the original motion trajectories.

Scenario 3: In this scenario, the autonomous vehicle will move through an intersection, no traffic lights, no priority roads and no priority vehicles (Figure 6). As illustrated in Figure 6, there will be the following situations: i) the autonomous vehicle crosses the intersection before other vehicles, ii) the autonomous vehicle overtakes the intersection after vehicle 01 and before vehicle 03, iii) the autonomous vehicle overtakes the intersection after vehicle 03 and before vehicle 01, iv) the autonomous vehicle crosses the intersection after all the vehicles. In these situations, because the direction of vehicle 02 does not affect the movement of the autonomous vehicle, the presence of vehicle 2 is not considered in these situations.

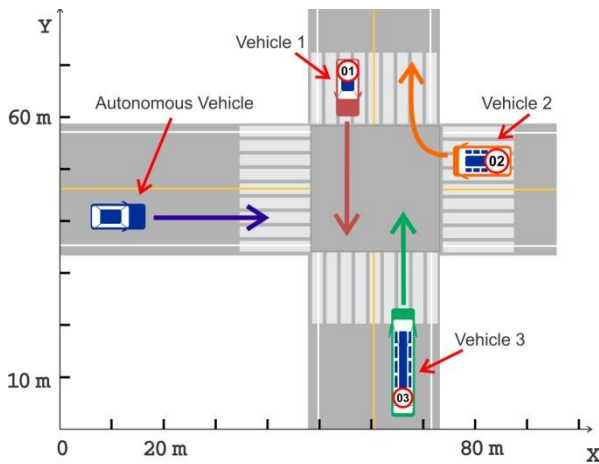


Figure 6. The vehicle moves into the intersection

To perform this simulation, we need to understand some rules when traveling through the intersection according to the provisions of Vietnam's road traffic law as follows: Is there a vehicle to the intersection?, is there a priority vehicle or not?, is there a priority path?

Vehicles are considered to enter the intersection when the front wheel has crossed the white line of the zebra crossing. In any circumstance, a vehicle that enters the intersection is given priority to go first.

For priority vehicles, the order of vehicles prioritized includes fire engines, military vehicles, police vehicles, ambulances, and other priority vehicles such as dyke protection vehicles, convoy of vehicles with police leading the way, vehicles on missions to overcome natural disasters, epidemics or vehicles on emergency duty as prescribed by law. Priority vehicles will be moved in front of other vehicles, meaning that the non-priority vehicles must give way to the priority vehicles in traffic.

For priority roads, vehicles located on non-priority roads must give way to vehicles on priority roads and should be based on the "intersection with priority roads" signs.

If there is no vehicle in the direction, the following shall be prescribed: at the intersection, priority shall be given to the vehicle that there is no vehicle to the right hand direction, at the roundabout, the priority shall belong to the vehicle that there is no vehicle to the left hand direction. Finally, for the priority turning direction, it

is regulated that the vehicle that turns right goes first, then the vehicle goes straight and finally the vehicle turns left.

In this simulation scenario, the simulated vehicles are running at a speed of 50km/h, the position of the vehicles considered in the coordinates respectively is the autonomous vehicle at position (10,45), vehicle 01 at position (45.65), vehicle 02 at position (80.55) and vehicle 03 at position (65.10), the center of the intersection is at (60.55).

The problem of this simulation is performed as follows: Let the file $[L_{av}, H_{av}]$ be the part of the autonomous vehicle and $[L_{obj}, H_{obj}]$ is the path of other vehicles in the intersection area. To avoid collisions between vehicles, only one vehicle is at any time in the intersection. Therefore, if $x \in [L_{av}, H_{av}]$, $x_{obj} \notin [L_{obj}, H_{obj}]$ with $\forall t \in (t_{min}(obj), t_{max}(obj))$, ie if the autonomous vehicle is already in the intersection, other vehicles cannot move into the intersection to avoid collisions at the intersection, where the values of t_{min} and t_{max} are symbols of the starting time and the last exit time of the vehicle while in the intersection.

When planning movement for an autonomous vehicle, we need to monitor whether other vehicles have entered the intersection area or not. First we will calculate the values of $t_{min}(obj) = \mathcal{T}(L_{obj} - x_{obj}, v_{r(obj)}, a_{r(obj)}^{max})$ with $a_{r(obj)}^{max} > 0$ is assumed to be continuous increase and $t_{max}(obj) = \mathcal{T}(H_{obj} - x_{obj}, v_{r(obj)}, a_{r(obj)}^{min})$ with $a_{r(obj)}^{min} < 0$ is assumed to be a continuous decrease of other vehicles, where $\mathcal{T}(x, v, a)$ is a time function to travel over a distance of path x , with the current velocity v and acceleration of vehicle unchanged a . In case the vehicle cannot travel the entire distance x in a finite time, this time value will be determined as $+\infty$, the values $a_{r(obj)}^{max}$ and $a_{r(obj)}^{min}$ are introduced to replace the constant acceleration value in order to determine the error limit during the evaluation of the motion of other vehicles.

The prediction horizon $k_{min}(obj)$ and $k_{max}(obj)$ of other means are calculated as follows:

$$k_{min}(obj) = \lceil t_{min}(obj) / \tau \rceil$$

$$k_{max}(obj) = \lceil t_{max}(obj) / \tau \rceil$$

where τ is the number of time steps in the prediction horizon.

Thus, with \mathcal{K} being the prediction horizon of the autonomous vehicle, the constraint set is built to avoid collisions when the autonomous vehicle moves through intersections written with the following cases:

1. If $k_{min}(obj) \leq k_{max}(obj) \leq \mathcal{K}$, it means the autonomous vehicle will have to cross the intersection before another vehicle moves into the intersection area, or the autonomous vehicle must stay outside the intersection until another vehicle leaves the intersection area, then the constraint set will be:

$$\delta_1(k) = 0 \Rightarrow x^{k_{min}(obj)} \geq H$$

$$\delta_1(k) = 1 \Rightarrow x^{k_{max}(obj)} \leq L$$

2. If $k_{\min(obj)} \leq \mathcal{K} \leq k_{\max(obj)}$, it means the autonomous vehicle will have to go through the intersection before another vehicle moves into the intersection area, or the autonomous vehicle must be outside the intersection to the end of the prediction horizon, then the constraint set will be

$$\delta_1(k) = 0 \Rightarrow x^{k_{\min(obj)}} \geq H$$

$$\delta_1(k) = 1 \Rightarrow x^{\mathcal{K}} \leq L$$

3. If $\mathcal{K} \leq k_{\min(obj)} \leq k_{\max(obj)}$, it means the autonomous vehicle will have to stay outside the intersection until other vehicles have moved through the intersection, no constraints are considered at this time.

The simulation results show that the speed of the autonomous vehicle decreases when preparing to enter the intersection area and then returns to the original speed, while moving through the intersection after all other vehicles. For a comparative basis, we tried to change the position of the autonomous vehicle at the new positions (45.45). Observing the process we realized that the autonomous vehicle moved through the intersection behind vehicle 1 and 2, but in front of vehicle 3 and the speed of vehicle 3 decreases when entering the intersection to make way for the autonomous vehicles to pass.

4. Conclusions

This paper proposes a solution to create the optimal trajectory with a model predictive control approach and the constraint set built from road traffic law. This approach is suitable for complex environmental conditions because these constraints can arise from various aspects of motion planning that must comply with traffic rules.

The motion trajectory of the vehicle is created from this solution aiming to improve performance and enhance the ability to avoid obstacles but still ensure optimal global trajectory. Simulation results with situations such as avoiding obstacles, passing other vehicles, moving through intersections ... have shown that this solution achieves the set goal of improving calculation efficiency and handling of uncertainty in the perceptual data of the environment. However, with the use of a second-order linear vehicle model to create motion trajectories, this solution is only advantageous when applied in the case of the autonomous vehicle moving on a straight path. For large curves, the current model built may not be accurate.

Evaluating of safety conditions, the proposed solution has been effective. This motion planning method has created an optimal trajectory that allows an autonomous vehicle to drive along the road and avoid obstacles safely. To improve the efficiency of calculations and consider the uncertainty of perceptual data and positioning on the general trajectory, the performance of this proposed solution depends on the probabilistic model of the system to create adaptive fields. Therefore, probabilistic analysis and the representation of a driving situation need to be more integrated into actual vehicle application.

In the future, in order to increase the reliability of this solution, the settings that have been experimented by simulation will be transferred to the real environment with experimental vehicles equipped with sensors, and when experimenting on reality, a number of factors to analyze the stability of the system will be

added so that the behavior of traffic participants is more accurately forecasted. Extensive implementation of this driver-assisted solution for semi-autonomous or fully autonomous vehicles in vehicle control systems will be able to minimize the amount of damage and create a safe movement plan for the future.

References

- [1] Bevan, Gollee and O'Reilly, Trajectory generation for road vehicle obstacle avoidance using convex optimization, *Proc IMechE Part D: J Automobile Engineering*, pp.455–473, 2010. DOI: 10.1243/09544070JAUTO1204.
- [2] Jo K, Kim J, Kim D et al, Development of autonomous car – Part I: distributed system architecture and development process, *IEEE Trans Ind Electron*, pp.7131–7140, 2014. DOI: 10.1109/TIE.2014.2321342
- [3] Zhang D, Li K and Wang J, Radar-based target identification and tracking on a curved road, *Proc IMechE Part D: J Automobile Engineering*, pp.39–47, 2012. DOI: 10.1177/0954407011414462
- [4] A. Houenou, P. Bonnifait, V. Cherfaoui et al, “Vehicle trajectory prediction based on motion model and maneuver recognition.” in *Proc of the IEEE Conference on Intelligent Robots and Systems*, pp. 4363 – 4369, 2013. DOI: 10.1109/IROS.2013.6696982
- [5] Fraichard and Howard, Iterative motion planning and safety issue, *In: Eskandarian A Handbook of intelligent vehicles. London: Springer*, pp. 1433–1458, 2012. DOI: 10.1007/978-0-85729-085-4_55
- [6] M. Werling, S. Kammel, J. Ziegler et al, “Optimal trajectories for time-critical street scenarios using discretized terminal manifolds.” *The International Journal of Robotics Research*, vol. 31, no. 3, pp. 346–359. 2012. DOI: 10.1177/0278364911423042
- [7] S. J. Anderson, S. C. Peters, T. E. Pilutti et al, “An optimal-control-based framework for trajectory planning, threat assessment, and semi-autonomous control of passenger vehicles in hazard avoidance scenarios.” *Transactions on International Journal of Vehicle Autonomous Systems*, vol. 8, no. 2/3/4, pp. 190 – 216, 2010. DOI: 10.1504/IJVAS.2010.035796
- [8] M. Althoff and J. M. Dolan, “Online verification of automated road vehicles using reachability analysis”, *IEEE Trans. on Robotics*, vol. 30, no. 4, pp. 903 – 918, 2014. DOI: 10.1109/TRO.2014.2312453
- [9] M. Althoff, D. Heß, F. Gamber, “Road occupancy prediction of traffic participants.” in *Proc. of IEEE Conference on Intelligent Transportation Systems*, pp. 99–105, 2013. DOI: 10.1109/ITSC.2013.6728217
- [10] Richard W. Cottle, Mukund N. Thapa, *Linear and Nonlinear Optimization (2nd ed.)*, International Series in Operations Research & 15Management Science, 2017, DOI: 10.1007/978-1-4939-7055-1.

Degradation Process in Pipeline and Remaining Useful Lifetime Estimation Based on Extended Kalman Filtering

Med Hedi Moulahi^{*1}, Faycal Ben Hmida²

¹Universit de Tunis, Institut Suprieur des Etudes Technologiques de Nabeul, 8000 Mrezka, Tunisia

²Universit de Tunis, Ecole Nationale Suprieure des Ingnieurs de Tunis, 2001 Tunisia

ARTICLE INFO

Article history:

Received: 20 February, 2020

Accepted: 30 May, 2020

Online: 23 June, 2020

Keywords:

Degradation modeling

Wiener process

RUL estimated by EKF

ABSTRACT

Degradation measurements are often treated and analyzed for improvement the reliability of system. Our objective in this paper is to study a class of non linear systems, whose dynamics and observations are non linear functions of the state. Obviously, we develop an Extended Kalman Filtering (EKF) algorithm for detecting the additive failures in a two tank system. However, the EKF algorithm is used to estimate the state vector of pipeline system based on all collected measures history. Such as degradation process (clogging, partial blockage) is considered and can be described by a Wiener process. For reasons of improvement reliability and security, it is necessary to predict the Remaining Useful Life (RUL) of pipeline. It follows that a major preventive maintenance actions. Furthermore, we can evaluate the RUL based on Monte Carlo simulation and compare the results.

1 Introduction

One of the most serious problems found in many chemical industry and its components degrade over time (e.g. wear, corrosion, erosion and fatigue). Among the components, we consider the pipelines, which are widely used in industrial plants networks and water distribution networks. The pipeline system are considered, one of the components more susceptible to failure in many industrial applications and that deteriorate over time. This deterioration occurs as a result of the damaging effects of fluid mixture produced from a reservoir or caused by the environment effects. Which is called crud fluid and it contains a variety of substances of different chemical structure that include hydrocarbon and non hydrocarbon components. Furthermore, the surrounding environment may cause a corrosion effect of pipe cross-section. As mentioned above the wax layer will build up in layers and can block the pipeline. Which can be reduce the pipeline reliability and safety level. This paper is an extension of work originally presented in conference name [1]. The aim of this paper is to present an approach to quantifying the reduction in reliability and safety. However, we designed a new technique for predict the remaining useful life time for clogging pipelines at point in time and a specific distance. We will focus our study mainly on cross section clogging pipelines. It may become a serious problem as a pipeline ages. Based on literature, several theories have been proposed to explain the reliability and the safety of pipelines circumferential strength. It follows that a failure level

threshold is dened for stochastic deterioration process model, which must not be exceeded for economical or security reasons. Obviously, the accumulated deposit on the pipeline wall causes the growth of a thickness of wax layer, leading to higher pressure drop and/or decreased flow rate [2]. Through periodic inspection, the growth of wax layer defect can be monitored, it follows that the reliability of pipeline system can be improvement. The loss of efficiency of the pipeline is then viewed as a change in the open loop system. Improvement reliability and safety of pipeline can be studied based on prognostic analysis. It can be defined as the prediction of future characteristic of the system such the Remaining Useful Life. Current research on diagnosis and prognostic are based on estimation the non observable system state. One of the most technic used is the stochastic filtering approaches, which gives an estimation of the system state recursively. Therefore, we can benefit from these, which give an estimation on-line with reliable performances for the parameters of a given model or a degradation path. However, Kalman Filtering (KF) and Extended Kalman Filtering (EKF) have been successfully applied to fault prognosis respectively in linear system dynamics model and nonlinear system dynamics model [3, 4]. According to the literature,[5]-[8]about degradation modeling can be mathematically described with a continuous process in terms of time, Lu[9] use the convex degradation model for the growth rate of fatigue cracks. Wang [10] studies a class of Wiener processes with random effects for degradation data.

^{*}Med Hedi Moulahi, ISET Nabeul 8000 Mrezka, Tunisia, (+216)98502162 & mohamedhedia.moulahi@isetn.rnu.tn

The remainder of this paper is organized as follows: In section 2, a two tank system modeling and it treats the stochastic deterioration in pipeline and explains why the Wiener process is the most appropriate candidate for clogging evolution. Section 3 is devoted to diagnosis and analysis the failure in pipeline system based on the EKF algorithm. Section 4 discusses two methods for estimation of the RUL, the first one is used to predict the RUL by EKF and the second method used Monte Carlo simulation to estimate the RUL. Finally, section 5 makes some concluding remarks.

2 Degradation Modeling in Pipeline System

Theoretically, in our case for modeling the movement of small particles in fluids, which can cause an additive accumulation of wax overtime. Therefore, a characteristic feature of this process in the context of structural reliability is that a pipeline's resistance alternately increases and decreases. For this reason, the Brownian motion is adequate for this deterioration modeling which is non-monotone. We can assume that degradation increment increases in time with random noise. In the literature, several theories have been proposed to describe the phenomenon of degradation by a stochastic model or a deterministic model [11, 12]. Several publications have appeared in recent years documenting the degradation evolution that can be described by Wiener process or gamma process [13]. In the following work, we consider the stochastic failure model. To solve this problem, many researchers have proposed various methods of modeling, mainly the stochastic model can be continuous or discrete [14, 15]. Wiener processes are widely applied for modeling the degradation process in engineering systems. Many physical phenomena are described by the Wiener processes, when the cumulative damage is non-monotonic. For example in our case, clogging or partial blockage in cross area pipeline is then viewed as a non-monotonic degradation process. According to the literature the Wiener process with a linear drift is frequently used to model the non-monotonic degradation process [16, 17]. Much research on the non-monotonic effect has been done about degradation process based on the nonlinear model, which was proposed in many literatures [10, 18, 19].

2.1 A Two Tank System Modeling

In our study, we consider a two tank system with pipeline. However, the section area of the first tank is noted by A_1 and the second one is noted by A_2 . The fluid is assumed as incompressible (ρ is assumed constant), which is pumped into the first tank at the top by motor pump. Through the pipeline, the outflow from the first tank fills the second tank. The overall tank system is shown in Figure 1. We assume the fluid level of tank 1 and tank 2 are measured by a level measurement sensor and controlled by adjusting the pump motor control input not shown in this case. The aim of this modeling is to control the state of the pipeline and to intervene before having the failure. In the following we assume an abrupt change at time and a given distance in this cross-sectional area of the pipeline, which is then viewed as a change in the tank fluid level. From this modeling it can be given a mathematical model in engineering system. In

this way we can simulate the evolution of the clogging in pipeline, obviously we can improve the pipeline reliability based on the estimation of the RUL.

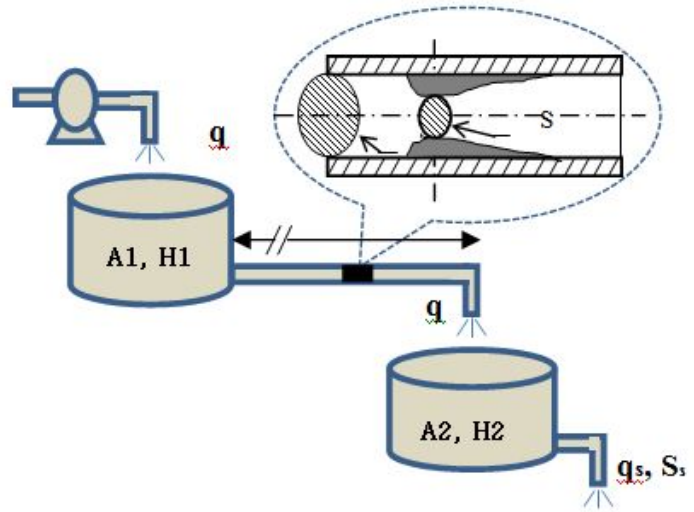


Figure 1: A double tank with pipeline

Table 1: Nomenclature

Parameter	Description
A_i	Tank cross-sectional area ($i=1,2$)
H_i	Fluid level ($i=1,2$)
p_0	Atmospheric pressure (Pa)
p_1	Pressure at the bottom (Pa)
ρ	Fluid density (Kg/m^3)
g	Acceleration of gravity (m^2/s)
Z_0	Level fluid in the upper tank (m)
Z_1	Level fluid at the bottom tank (m)
V	Fluid velocity (m/s)
h	State level at time t in tank (m)
S_s, S_i	Pipeline area (m^2)
S_c	Pipeline area clogging (m^2)
q	flow rate (m^3/s)

Let us describe the following failure mode encountered in petrochemical industrial applications these are:

1. clogging and partial blockage in pipeline cross section. Which are caused by impurities and an additive accumulation of wax overtime, see Figure 3.
2. Decreasing pump performance due to cavitations,
3. Dry running.

To control pressure losses, clogging, leaks, corrosion in the pipeline. We assume that there are two pressure sensors between the upstream and downstream tank. This is illustrated in Figure 2. We will make the following assumptions: The clogging is supposed located at distance ($d < L$) and the pressure losses are neglected along the pipeline.

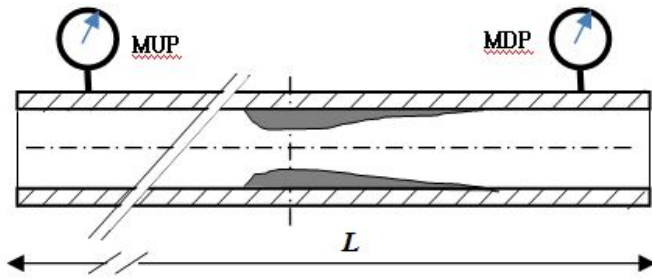


Figure 2: showing the deposit profile of wax

MUP: Measurement of upstream pressure.
MDP: Measurement of downstream pressure.

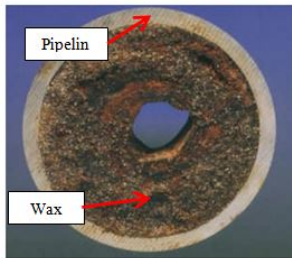


Figure 3: Clogging inside the pipeline

In order to describe the overall tank fluid level control system, we have some assumptions: The fluid is incompressible (ρ is constant), the temperature distribution of the incoming fluid is assumed constant, the flow rate is considered permanent. Under all these considerations, we use the mass balance equation, Torricelli rule and Bernoulli rule it follows that the process can be described by the following equations,

$$A \frac{dh(t)}{dt} = \sum flow \quad (1)$$

$$p_1 - p_0 = (\rho g Z_0 - \rho g Z_1) = \rho g h_1 \quad (2)$$

$$V = \sqrt{2gh} \quad (3)$$

Using the mass balance equation, we give the flow rate at each pipeline outlet and the following equations,

$$q = V.S \quad (4)$$

$$\begin{cases} \frac{Adh_1}{dt} = S_i V_i - q_c, \\ \frac{Adh_2}{dt} = q_c - q_s \end{cases} \quad (5)$$

$$\begin{cases} \frac{q_c}{S_c} = \sqrt{2gh_1}, \\ \frac{q_s}{S_s} = \sqrt{2gh_2} \end{cases} \quad (6)$$

Using(5) and(6), the mathematical model can be written as follow:

$$\frac{dh_1}{dt} = \frac{S_i}{A} V_i - \frac{S_c}{A} \sqrt{2g} \cdot \sqrt{h_1} \quad (7)$$

$$\frac{dh_2}{dt} = \frac{S_c}{A} \sqrt{2g} \cdot \sqrt{h_1} - \frac{S_0}{A} \sqrt{2g} \cdot \sqrt{h_2} \quad (8)$$

The implementing of equation (7) and equation (8), we can be rewritten the discredited model for a sampling period T_e .

$$h_1(k+1) = h_1(k) + T_e \left[\frac{S_i}{A} V_i(k) - \frac{S_c(k)}{A} \sqrt{2g} \cdot \sqrt{h_1} \right] \quad (9)$$

$$\begin{aligned} h_2(k+1) &= h_2(k) \\ + T_e \left[\frac{\sqrt{2g} S_c(k)}{A} \cdot \sqrt{h_1(k)} - \frac{\sqrt{2g} S_s}{A} \cdot \sqrt{h_2} \right] \end{aligned} \quad (10)$$

The state equation has the nonlinear terms (a square term), it follows that a nonlinear system. If we consider the process noise $v(k)$ and the measurement noise $w(k)$, the state equation $X(k+1)$ and measurement equation $z(k)$ can be written as follow:

$$X(k+1) = f_k(X_k, u_k, k) + v(k) \quad (11)$$

$$z(k) = HX(k) + w(k) \quad (12)$$

Where $X_k = \begin{bmatrix} h_1(k) \\ h_2(k) \end{bmatrix}$, $H = [1, 1]$ and $q_i(k) = k_u \cdot u_k$

k_u amplification gain, u_k Control voltage pump. From the above expression, we deduce the state equations of pipeline system.:

$$X(k+1) = X(k) + T_e D(k) \sqrt{X(k)} + T_e b(k) u(k) + v(k) \quad (13)$$

$$z(k) = HX(k) + w(k) \quad (14)$$

Where $D(k) = \begin{bmatrix} \frac{-S_c(k) \sqrt{2g}}{A} & 0 \\ \frac{S_c(k) \sqrt{2g}}{A} & \frac{-S_s \sqrt{2g}}{A} \end{bmatrix}$, $b = \begin{bmatrix} \frac{1}{A} \\ 0 \end{bmatrix}$

2.2 Wiener process for degradation modeling

For modeling the stochastic deterioration process, we can use Brownian motion with drift, as far as the author knows, the Wiener process was applied frequently in many works [14, 17]. In this work, for more control the evolution of pipeline state, the monitoring pressure differential Δp can be designed as a degradation indicator in pipeline [15]. In the following, the Wiener process is used to model the clogging damage in pipeline. Which is considered a no monotonic degradation processes. However, we assumed that the degradation increases linearly in time with random noise. In practice, the degradation processes of pipeline systems are affected by partial blockage or clogging formulated by the crud fluid and external operating environments or loads. Let's call again the properties of Wiener process to model the pipeline clogging evolution, in order to lifetime and reliability analysis. 1. $X(0) = 0$;

2. The process $\{X(t)\}$ has stationary and independent increment,

3. The process $\{X(t)\}$ is assumed normally distributed with $\mu = 0$ and variance $\sigma^2 t$ for any $t > 0$. Considering the time variables $t, u > 0$, the random variable $X(t+u) - X(u)$ and $X(t) - X(u)$ for $t > u$, have a normal density with mean 0, and variance $\sigma^2 t$ and $\sigma^2(t-u)$ respectively, In this way we define the probability density function as $f(x, t)$ of $\Pr\{X(t+u) - x(u) \leq x\} = \Pr\{X(t) \leq x\}$ is as follows

$$f(x, t) = \frac{1}{\sqrt{2\pi t} \cdot \sigma} e^{-x^2 / (2\sigma^2 t)} \quad (15)$$

and its Laplace transform is

$$\int_{-\infty}^{+\infty} f(x, t) e^{-sx} = e^{s^2 \sigma^2 / 2} \quad (16)$$

When $\sigma = 1$, $B(t) \equiv X(t) / \sigma$, is a standard Brownian motion because $V\{X(t) / \sigma\} = t$. For any $0 \leq t_0 < t_1 < \dots < t_n < t$, using the properties of independent and stationary increments, it is evident to write the probability

$$\begin{aligned} \Pr\{X(t) \leq x | X(t_0) = x_0, \dots, X(t_n) = x_n\} \\ = \Pr\{X(t) \leq x | X(t_n) = x_n\} \\ = \Pr\{X(t) - X(t_n) \leq x - x_n\} \end{aligned} \quad (17)$$

In this way, we assume that the process has a Markov property. Hence, its distribution function can be written as follows:

$$\begin{aligned} \Pr\{X(t) - X(t_n) \leq x - x_n\} \\ = \frac{1}{\sqrt{2\pi}(t - t_n)\sigma} \int_{-\infty}^{+\infty} e^{-u^2 / (2\sigma^2(t - t_n))} du \end{aligned} \quad (18)$$

Where the increment $X(t) - X(t_n)$ has a pdf with mean $\mu = 0$ and variance $\sigma^2(t - t_n)$ for any $t > t_n$, that does not depend on t_n .

$$Z(t) = \mu t + \sigma B(t) \quad (19)$$

Where $B(t)$ is a standard Brownian motion representing the stochastic dynamics of the degradation process, then $Z(t)$ is called a Wiener process with drift parameter μ and variance σ^2 . The pipelines failure is caused by the fluctuation of the fluid pressure-depression along the time P_{\min} and P_{\max} , that are repeated within an interval of time. Moreover, in [12] we deduce a minimum wax removing pressure. Obviously these pipelines are unfortunately usually designed for ultimate limits resistance. It is noted here the evolution of the internal pressure can be caused a stochastic stress in cylindrical pipeline. Moreover, in the work [5] and [19] are used a nonlinear Wiener process with random effects to model the degradation process in pipeline. A nonlinear Wiener process aims to model the heaping and movement of small particles in fluids with tiny fluctuations in pipeline. To assess the severity of the clogging in pipeline system and its impact on the residual life, it is required to analysis the degradation process characteristic. Therefore, the pipelines degradation can increase or decrease gradually under time. From now on we assume that the small increase or decrease for degradation pipeline over a small time interval be have similarly to the random walk of small particles heaping in cross sectional pipeline. Furthermore, the random effects are widely used in [10], [18] and [19] extended the degradation model in (19) to the following form:

$$Z(t) = \lambda \Lambda(t) + \sigma_B \cdot B(\Lambda(t)) \quad (20)$$

Where,

λ : is the drift parameter

$\Lambda(t)$: is a positive non decreasing function, we can use the time-scale transformation function $\Lambda(t) = t^\theta$ and $\Lambda(t) = e^{\theta t} - 1$

σ_B : is the diffusion parameter,

$B(t)$: is the standard Brownian motion.

If $\Lambda(t) = t$ is a positive function, the nonlinear model becomes a linear model given by (19).

However, a fundamental problem related to this kind of degradation models of nonlinear Wiener-process. Actually, nonlinearity and stochastically are two important factors contributing to the degradation processes of complex systems. When the degradation process $Z(t)$, $t \geq 0$ hits a failure threshold value L of the item was considered. The item's lifetime T is defined as:

$$T = \inf\{t | Z(t) \geq L\} \quad (21)$$

According to the concept of the First Hitting Time, it is evident, when the Wiener process path reaching a threshold level L , it can be obey an inverse Gaussian distribution. Obviously, when we consider the nonlinear degradation process as shown in (20) and the drift parameter λ is considered a random effect variable given by $\lambda \sim N(\mu_\lambda, \sigma_\lambda)$. Moreover, according to [18] the probability distribution function (pdf) of the life time is given by (22) and similarly, the cumulative distribution function (CDF) of the life time is given by (23).

$$\begin{aligned} f_T(t) = \frac{L}{\sqrt{2\pi \cdot \Lambda^3(t) (\sigma_\lambda^2 \Lambda(t) + \sigma_B^2)}} \times \\ \exp\left(-\frac{(L - \mu_\lambda \Lambda^2(t))}{2\Lambda(t) (\sigma_\lambda^2 \Lambda(t) + \sigma_B^2)}\right) \end{aligned} \quad (22)$$

and

$$\begin{aligned} F_T(t) = \Phi\left(\frac{\mu_\lambda \Lambda(t) - L}{(\sigma_\lambda^2 \Lambda^2(t) + \sigma_B^2 \Lambda(t))}\right) + \\ \exp\left(\frac{2\mu_\lambda L}{\sigma_B^2} + \frac{2\sigma_\lambda^2 L^2}{\sigma_B^4}\right) \times \\ \Phi\left(-\frac{2\sigma_\lambda^2 L^2 \Lambda(t) + \sigma_B^2 (L + \mu_\lambda \Lambda(t))}{\sigma_B^2 \sqrt{\sigma_B^2 \Lambda(t) + \sigma_\lambda^2 \Lambda^2(t)}}\right) \end{aligned} \quad (23)$$

when we consider the linear degradation process as shown in (19) and the drift parameter λ is considered a random effect variable given by $\lambda \sim N(\mu_\lambda, \sigma_\lambda)$. Then, the probability distribution function (pdf) and the cumulative distribution function (CDF) of the life time are given by:

$$\begin{aligned} f_T(t) = \frac{L}{\sqrt{2\pi \cdot t^3 (\sigma_\lambda^2 t + \sigma_B^2)}} \times \\ \exp\left(-\frac{(L - \mu_\lambda t)^2}{2t (\sigma_\lambda^2 + \sigma_B^2)}\right) \end{aligned} \quad (24)$$

and

$$F_T(t) = \Phi\left(\frac{\mu_\lambda t - L}{(\sigma_\lambda^2 t^2 + \sigma_B^2 t)}\right) + \exp\left(\frac{2\mu_\lambda L}{\sigma_B^2} + \frac{2\sigma_\lambda^2 L^2}{\sigma_B^4}\right) \times \Phi\left(-\frac{2\sigma_\lambda^2 L^2 t + \sigma_B^2 (L + \mu_\lambda t)}{\sigma_B^2 \sqrt{\sigma_B^2 t + \sigma_\lambda^2 t^2}}\right) \quad (25)$$

2.3 Parameters Estimation using MLE

The Maximum Likelihood Estimation (MLE) is the most widely used. It is a method of estimating the parameters of a model. This maximizes the agreement of the selected model with the observed data. We assume that $Z_{i,j}$ is a degradation indicator measurement (In this case the pressure differential see Figure 2) at the i^{th} items (We assume many pipelines are observed) at time j , where $i = 1, 2, 3, N$ and $j = 1, 2, 3, r$, r is the last observation time. The degradation paths is based on linear Wiener process with parameters (μ, σ) and that are assumed the same for all items. For each increment $\Delta Z_{i,j} = Z_{i,j+1} - Z_{i,j}$ of each item follows a normal distribution $N(\mu\Delta t_{i,j}, \sigma^2\Delta t_{i,j})$. We assume that increment $\Delta Z_{i,j}$ is independent, identically components. Similarly it is assumed normally distributed for all. Now we can derive the density function of Brownian motion process, which is given by:

$$f(\mu\Delta t_{i,j}, \sigma^2\Delta t_{i,j})(\Delta Z_{i,j}) = \frac{1}{\sqrt{2\pi\sigma^2\Delta t_{i,j}}} e^{-\frac{(\Delta Z_{i,j} - \mu\Delta t_{i,j})^2}{2\sigma^2\Delta t_{i,j}}} \quad (26)$$

Z is a Markovian process, then the maximum likelihood estimator can be used. The parameters vector $\theta = (\mu, \sigma)$ can be evaluated once we know the transition density function of Z . According to the measurements of degradation for each item i , $\Delta Z_i = (\Delta Z_{i,1}, \Delta Z_{i,2}, \Delta Z_{i,3}, \dots, \Delta Z_{i,r})$. For item i , the likelihood function for item i is :

$$L_i(\theta) = \prod_{j=1}^r \frac{1}{\sqrt{2\pi\sigma^2\Delta t_{i,j}}} e^{-\frac{(\Delta Z_{i,j} - \mu\Delta t_{i,j})^2}{2\sigma^2\Delta t_{i,j}}} \quad (27)$$

For the i^{th} item, the log likelihood can be written as follow:

$$l_i(\theta) = \ln L_i(\theta) = \ln\left(\prod_{j=1}^r \frac{1}{\sqrt{2\pi\sigma^2\Delta t_{i,j}}} e^{-\frac{(\Delta Z_{i,j} - \mu\Delta t_{i,j})^2}{2\sigma^2\Delta t_{i,j}}}\right) \quad (28)$$

The degradation measurements vector are independent, then we can write,

$$l(\theta) = \ln(\Delta Z_1, \dots, \Delta Z_N) = \sum_{i=1}^N \ln(f_i(\Delta Z_{i,1}, \dots, \Delta Z_{i,r})) \quad (29)$$

$$l(\theta) = \sum_{i=1}^N \ln\left(\prod_{j=1}^r \frac{1}{\sqrt{2\pi\sigma^2\Delta t_{i,j}}} e^{-\frac{(\Delta Z_{i,j} - \mu\Delta t_{i,j})^2}{2\sigma^2\Delta t_{i,j}}}\right) \quad (30)$$

Where $l(\theta) = \sum_{i=1}^N l_i(\theta)$, $f_i/l_i(\theta)$ the pdf is divided by log likelihood of increments, that is corresponding to each item, and $f/l(\theta)$ is the pdf divided by log likelihood of increments corresponding to all increments. In this way we obtain, we write the MLE $\hat{\theta} = (\hat{\mu}, \hat{\sigma})$ are found by maximizing $l(\theta)$, and using the partial derivative of the log likelihood function of (24), by respecting the derivation variables μ and σ , from this we deduce these equations,

$$\frac{\partial l(\theta)}{\partial \mu} = \sum_{i=1}^N \sum_{j=1}^r \frac{\Delta Z_{i,j} - \mu\Delta t_{i,j}}{\sigma^2} = 0 \quad (31)$$

$$\frac{\partial l(\theta)}{\partial \sigma} = -\frac{rN}{\sigma} \sum_{i=1}^N \sum_{j=1}^r \frac{(\Delta Z_{i,j} - \mu\Delta t_{i,j})^2}{\sigma^3\Delta t_{i,j}} = 0 \quad (32)$$

We can now apply the MLE method for $\theta = (\mu, \sigma)$ is given by:

$$\hat{\mu} = \frac{\sum_{i=1}^N \sum_{j=1}^r \Delta Z_{i,j}}{\sum_{i=1}^N \sum_{j=1}^r \Delta t_{i,j}} \quad (33)$$

$$\hat{\sigma} = \sqrt{\frac{1}{rN} \sum_{i=1}^N \sum_{j=1}^r \frac{(\Delta Z_{i,j} - \mu\Delta t_{i,j})^2}{\Delta t_{i,j}}} \quad (34)$$

2.4 First Hitting Time Concept and RUL Distribution

Based on the work presented in [20]-[22], the RUL of a system is dened as the length from the current time to the failure time. If we consider the random variable T as a stopping time for $B(t)$, $t \geq 0$, and for any t , it follows that to decide whether T has occurred or not by observing the path of $B(s)$, $0 \leq s \leq t$. For any t the sets $\{T \leq t\} \in F_t$ and given a level threshold L , the time of reaching this level could be more than once due to the random nature of Wiener process $Z(t)$. In the following, we provide the computation of a first passage time hits level threshold T_L and the RUL distribution of Wiener process paths. Suppose T_L the first passage time of Wiener process $Z(t)$ hits level L , if the maximum of $Z(t)$ at time t is greater than L , then the path of Wiener process took value L at some time before t . Let consider T_L the rst time when the degradation path $Z(t)$ reaches the threshold level, it follows that, we deduce the distribution of the maximum and the minimum of Wiener process on $[0, t]$. $M(t) = \max_{0 \leq s \leq t} Z(s)$ and $m(t) = \min_{0 \leq s \leq t} Z(s)$. identically, the distribution of the first hitting time of L , $T_L = \inf\{t > 0 : Z(t) = L\}$. From Theorem [22], and for any $L > 0$,

$$P_0(M(t) \geq L) = 2.P_0(Z(s) \geq L) = 2\left(1 - \Phi\left(\frac{L}{\sqrt{t}}\right)\right) \quad (35)$$

Proof: For the events $M(t) \geq L$ and $T_L \leq t$ are the same. When the maximum of Wiener process at time t hits a failure threshold L . If Wiener process took value L at some time before t , therefore the

maximum will be at least L . Since $\{Z(t) \geq L\} \subset \{T_L \leq t\}$, thus we have also computed the probability as follows:

$$P(Z(t) \geq L) = P(Z(t) \geq L, T_L \leq t) \quad (36)$$

As $Z(T_L) = L$

$$\begin{aligned} P(Z(t) \geq L) \\ = P(T_L \leq t, (Z(T_L + (t - T_L)) - Z(T_L)) \geq 0) \end{aligned} \quad (37)$$

Let T_L is a finite stopping time, and from the strong Markov propriety in [22], from this the random variable $\hat{Z}(s) = Z(T_L + s) - Z(T_L)$ is assumed independent of F_{T_L} and has a normal distribution, so we have

$$P(Z(t) \geq L) = P(T_L \leq t, \hat{Z}(t - T_L) \geq 0) \quad (38)$$

If we had s independent of T_L , then

$$\begin{aligned} P(T_L \leq t, \hat{Z}(t) \geq 0) &= P(T_L \leq t) P(\hat{Z}(s) \geq 0) \\ &= P(T_L \leq t) \frac{1}{2} = P(M(t) \geq L) \frac{1}{2} \end{aligned} \quad (39)$$

That is to say for any $L > 0$

$$\begin{aligned} P(M(t) \geq L) &= 2P_0(Z(t) \geq L) \\ &= 2 \left(1 - \Phi \left(\frac{L - x - \mu t}{\sigma \sqrt{t}} \right) \right) \end{aligned} \quad (40)$$

In reliability engineering, there are some researches involving the degradation based failure time T_L prediction for a component. Which is defined as the time at which the degradation path first reaches a threshold L . However, the distribution of the first passage time T_L plays an important role for predicting the remaining useful lifetime. Furthermore, RUL is often used as a decision indicator in the optimal maintenance strategies. According to [23], the inverse Gaussian distribution can be used to calculate the probability density function of the conditional first passage time, when the degradation path is modeled by a Wiener process with positive drift. According to what is said before, let T_L can be the first passage time for a fixed threshold $L > 0$ by $Z(t)$ which is the linear degradation process given in (19). Then T_L can be considered as a random variable described by the inverse Gaussian distribution as follow $T_L \sim IG\left(\frac{L}{\mu}, \left(\frac{L}{\sigma}\right)^2\right)$. We should first mention inverse Gaussian distribution (IG) is a two parameter continuous distribution given by its density function as follow:

$$f(t, \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi}} t^{-3/2} \exp\left(\frac{-\lambda}{2\mu^2 t}(t - \mu)^2\right), t > 0 \quad (41)$$

The mean $\mu > 0$ and the shape parameter $\lambda > 0$. If we consider a random variable X which is governed by the inverse Gaussian distribution and can be written as follow $X \sim IG(\mu, \lambda)$. However, the inverse Gaussian distribution describes the probability density function of the conditional first passage time, when the degradation path reaches a level L . Then $T_L \sim IG\left(\frac{L}{\mu}, \left(\frac{L}{\sigma}\right)^2\right)$ has inverse Gaussian

distribution, so put these parameters into(35). Therefore, we have probability density function given by:

$$f(t, \mu, \sigma) = \frac{L}{\sqrt{2\pi\sigma^2 t^3}} \exp\left(\frac{-(L - \mu t)^2}{2\sigma^2 t}\right) \quad (42)$$

The first passage time T_L satisfies the following function, can be written as

$$\begin{aligned} F(x, \mu, \sigma) &= P(T_L \leq t) \\ &= \int_0^t \frac{L}{\sqrt{2\pi\sigma^2 x^3}} \exp\left(\frac{-(L - \mu x)^2}{2\sigma^2 x}\right) dx \end{aligned} \quad (43)$$

Given (36), it is possible to calculate the RUL distribution. If we take the observing data from the degradation process at time t , which is at position $z(t)$, then the probability that RUL is less than a predened period h can be written as follow:

$$\begin{aligned} P(RUL_{z(t)} \leq h) \\ = \int_0^h \frac{L - z(t)}{\sqrt{2\pi\sigma^2 x^3}} \exp\left(\frac{-(L - z(t) - \mu x)^2}{2\sigma^2 t}\right) dx \end{aligned} \quad (44)$$

3 Diagnosis in pipeline using EKF

The suboptimal filter known as the Extended Kalman filter (EKF) is frequently used for non-linear state estimation problems. Since (13) and (14) that represent the model of pipeline system, that is non-linear model. Before applied the EKF algorithm, we must linearize the state equations around the actual value of state estimated for each time step. A linear approximation process is done by using a Taylor series approximation. According to (11) and (12), we carried out the non-linear model of the pipeline system. From now on we apply the Taylor series approximation, in order to linearize the functions that given by equation (13) and equation (14)

$$\begin{aligned} f(X_k, u_k) \\ \approx f(\hat{X}_k, u_k) + dF_k(X_k - \hat{X}_k) + dG_k(u_k - \bar{u}_k) \end{aligned} \quad (45)$$

$$\text{Such as: } dF_k = \frac{\partial f(X_k, \bar{u}_k)}{\partial X} \Big|_{X_k = \hat{X}_k}, \quad dG_k = \frac{\partial f(\hat{X}_k, u_k)}{\partial u} \Big|_{u_k = \bar{u}_k},$$

$C_k = f(\hat{X}_k, \bar{u}) - dF_k \hat{X}_k$. The following derivation terms are found:

$$dF_k = \begin{bmatrix} 1 - \frac{TeSc}{2A} \frac{\sqrt{2g}}{\sqrt{h_1}} & 0 \\ \frac{TeSc}{2A} \frac{\sqrt{2g}}{\sqrt{h_1}} & 1 - \frac{TeS_2}{2A} \frac{\sqrt{2g}}{\sqrt{h_2}} \end{bmatrix}, \quad dG_k = [Te \frac{S_i}{A} 0].$$

In the work [24] and in related references it was defined the EKF algorithm.

Obviously, the prediction step is defined by

$$\hat{X}_{k|k-1} = f_k(\hat{X}_{k|k-1}, u_k) + v_k \quad (46)$$

$$P_{k|k-1} = Q_{k|k-1} + dF_k P_{k-1|k-1} dF_k^T \quad (47)$$

Moreover, the updating equation can be implemented recursively as follow

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k [z_k - (H_k X_k + w_k)] \quad (48)$$

$$P_{k|k} = P_{k|k-1} - K_k H_k P_{k|k-1} \tag{49}$$

$$K_k = P_{k|k-1} H_k^T [H_k P_{k|k-1} H_k^T + R_k]^{-1} \tag{50}$$

It is interesting to consider the error in the prediction of z_k from its past z_{k-1} . This error is known as the prediction residual on the innovation. This latter term comes from the fact that we can write:

$$r_1(k) = ep_1 = h_1(k) - h_{1e}(k) \tag{51}$$

$$r_2(k) = ep_2 = h_2(k) - h_{2e}(k) \tag{52}$$

The simulation data are taken from [25, 26] are given in Table 2.

Table 2: Nomenclature

Physical parameters	Description
$A_1 = A_2 = A = 16(m^2)$	Identical sections for the two tanks
$S_0 = 1/32(m^2)$	Tank outlet section N2
$S_i = 1/4(m^2)$	Tank outlet section N2
$g = 9,81(m/s^2)$	Acceleration of gravity
$T = 5(s)$	Sampling period
$S_{max} = 1/4(m^2)$	Section of the pipe linking the two tanks
$N = 100$	Time (Year)
$H_1(1) = 0.5(m)$	Level of the fluid in the first reservoir
$H_2(1) = 0.01(m)$	Level of the fluid in the second reservoir
$q_i = 5(m^3/s)$	Flow rate

The numerical value of state noise for simulation are $v = Q.rand(1, n)$ and $Q = diag(0.002, 0.05, 0.0001)$ The numerical value of measurement noise for simulation are $w = R.rand(2, n)$ and $R = diag(0.01, 0.01)$.

4 System without Fault (No clogging)

According to the EKF algorithm given by (46) and (52), we illustrate and show the simulation results through a two tank system without degradation process. These simulation results are given by Figure 4 and Figure 5. According to the algorithm of EKF given by(46)-(52), which contains the fluid level and flow rate we can generate the statistical residual. According to the estimated value of level tank, it is possible to quantify the influence of wax deposition action. The simulation results as shown in Figure 4 and Figure 5. However, the pressure and flow rate in pipeline system are commonly measured in order to monitoring the characteristic pipeline system. These equations in algorithm of EKF incorporate a measurement value into a priori estimation to obtain an improved a posteriori estimation. The simulations results were carried out in the absence of clogging in pipeline system to show the fluid level evolution in two tank in Figure 4.

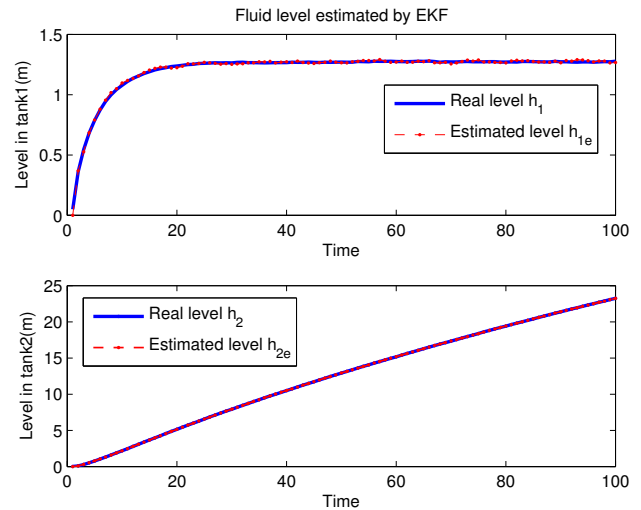


Figure 4: Fluid level in Two Tank system

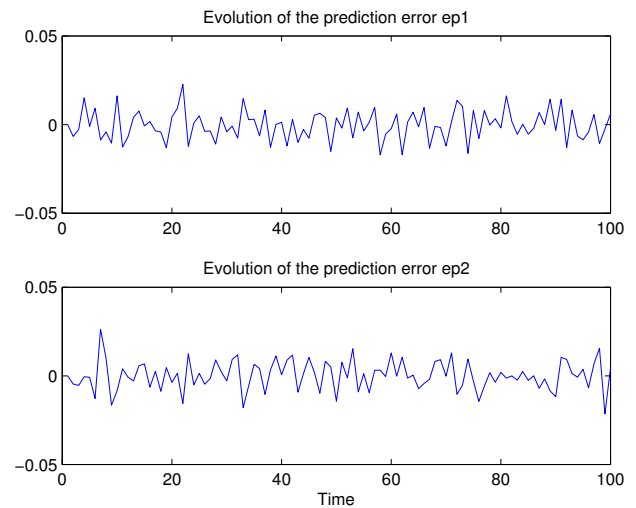


Figure 5: Residuals without faults(No clogging)

Consider Figure 5, which plots the residue observed from EKF for diagnosis, that indicates the system is no faulty. Moreover, it shows the filter performance measures in terms of the innovation sequence. We know that if the filter is working correctly, then the innovations have a zero mean Gaussian white noise with a covariance (S_r). However, we can verify that the filter is consistent by applying the following two procedures. Check that the innovation is consistent with its covariance by verifying that the magnitude of the innovation is bounded by $\pm 2 \sqrt{S_r}$. Verify that the innovation is unbiased and white.

The statistical characteristics of the residual without faults are given in Table 3:

Table 3: The statistical characteristics of the residual

Residual	Mean	Standard deviation
r_1	$\mu_1 = -0.000242$	$\sigma_1 = 0.0087$
r_2	$\mu_2 = 0.0000321$	$\sigma_2 = 0.0102$

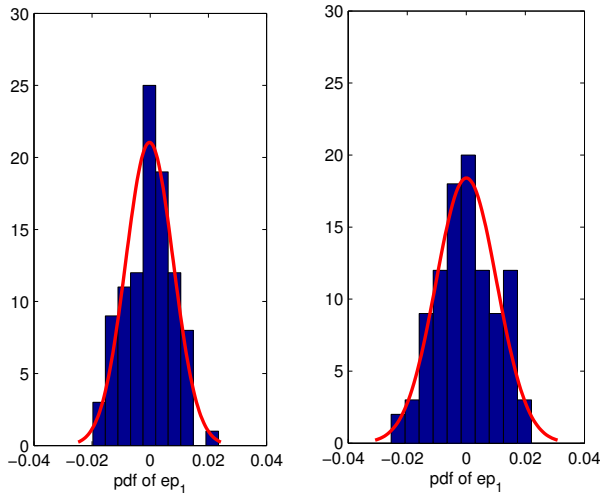


Figure 6: Residuals statistical characteristics

4.1 System including clogging by Steps

In order to validate the EKF algorithm for dynamic failure estimation. Consider Figure 2, which shows the deposit profile of wax. We will be interested in clogging pipeline, that is assumed at distance d projected on the longitudinal axis ($d < L$), where the clogging can form a critical region of stress concentration and L is the length of the pipeline. In practice, this type of defect is unobservable and no measurable. For this reason, we consider two pressure sensors to control the increase or decrease the pressure downstream and upstream of the obstruction. Before use the degradation model by Wiener process that can be assumed and described by (n) steps to model the cumulative damage in pipeline. The model of degradation by steps can be described by (53), this results in a decreasing of cross sectional area and it model a clogging evolution in pipeline.

$$S_c(t) = \begin{cases} \lambda_1 \cdot S_{c_{max}}, 0 < t < t_1 \\ \lambda_2 \cdot S_{c_{max}}, t_1 < t < t_2 \\ \vdots \\ \lambda_i \cdot S_{c_{max}}, t_{i-1} < t_i < t_{i+1} \\ \vdots \\ \lambda_n \cdot S_{c_{max}}, t_n > t_{n-1} \end{cases} \quad (53)$$

where $0 < \lambda_i \leq 1$, is a ratio between $S_c(t)/S_{c_{max}}$, see Figure 2. Thus, t_{i-1} and t_i denote the initiative and terminal time under cumulative wax step respectively. However, in the following we have $n = 3$ steps, then $\lambda_1 = 0.3$, $\lambda_2 = 0.5$ and $\lambda_3 = 0.7$. Which, the first clogging level is assumed to occur at time $t = 30$ unit of time.

$$S_c(t) = \begin{cases} 0.3 \cdot S_{c_{max}}, 0 < t < t_1 \\ 0.5 \cdot S_{c_{max}}, t_1 < t < t_2 \\ 0.7 \cdot S_{c_{max}}, t > t_2 \end{cases} \quad (54)$$

For a given form of evolution clogging level in cross-sectional area in pipeline Figure 9 and according to the EKF algorithm, it follows that the fluid level in tank1 and tank2 are affected by this clogging. Furthermore, the quality of the state estimation of the deposit profile of wax is accepted and well. The EKF algorithm

is used to estimate recursively the pipeline system state, when we consider a deteriorating system with noisy measurement of h_1 and h_2 . Figure 7 shows the fluid level evolution in two tank system. Moreover, Figure 8 depicts the prediction errors, which contains a statistical information for diagnosis.

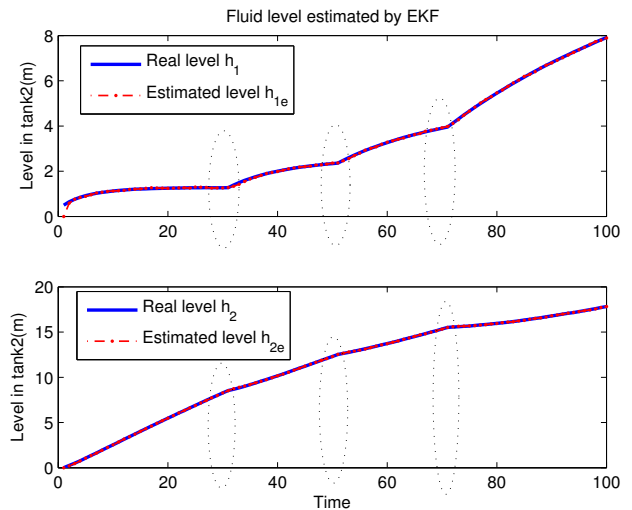


Figure 7: Fluid level in two tank

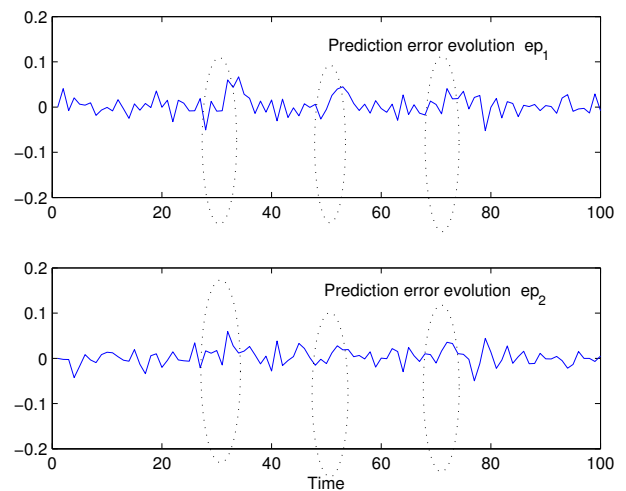


Figure 8: Residuals with three steps of faults

Consider Figure 9, which plots section degradation prole estimated by EKF. If we compare these results with Figure 4 and Figure 5, we can show the effect of clogging pipeline in the characteristics of pipeline system. One can notice that the performance of our proposed methodology of estimation and failure diagnosis are perfectly adopted. For a better lecture of Figure 9, we can conclude that the EKF algorithm is applied very well in order to estimate the state of the cross-sectional area in pipeline.

The problem consists to model the evolution of the wax in pipeline and to determine the minimum cross-section area which corresponds to a maximum pressure. That can cause a crack or an explosion of the pipeline system. For more reality and improve the clogging damage model, it is more appropriate to describe the type of failure with a continuous process.

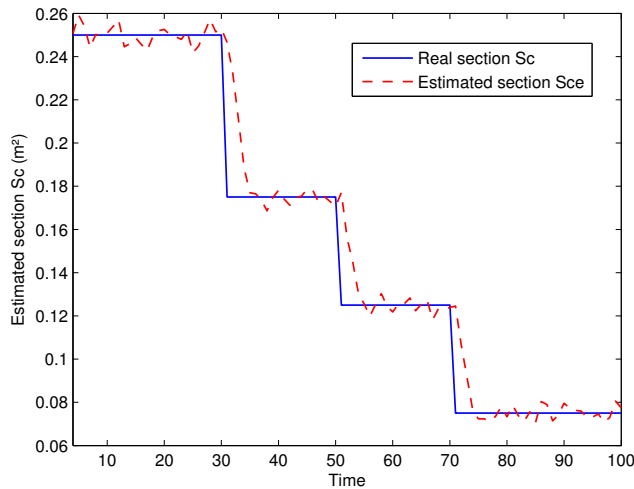


Figure 9: Section profile estimated by EKF

From this Figure 10, it can be seen that random evolutions of the internal section $Sc(t)$ of the pipeline for a given distance ($d < L$). The time at which the prognosis is made it can be noted by t_{pro} . However, the remaining residual life of the pipeline system at t_{pro} will be noted $RUL_{t_{pro}}$. It equals to the time elapsed between $RUL_{t_{pro}}$ and the moment when the pipeline system is failed. Noted that T be the date of failure, then the RUL prediction at the time of prognosis is assumed as a random variable which can be defined for $t_{pro} < T$ (see Figure 10).

$$RUL_{t_{pro}} = T - t_{pro} \quad (55)$$

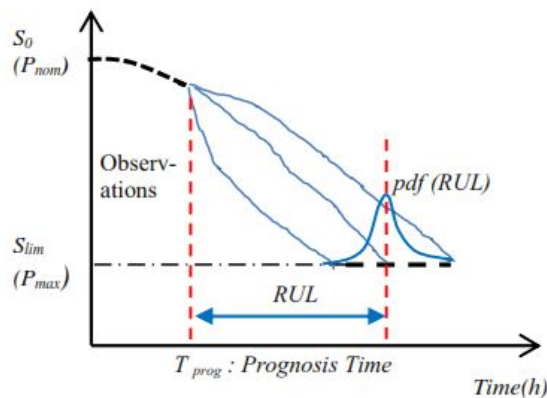


Figure 10: Remaining Useful Lifetime prediction

S_{lim} : Minimum value of Sc , when a lower value, there will be an explosion in pipeline, which can dene the threshold level. According to the literature [2, 27], who presented models for evaluation of the wax appearance at given point. From the authors work [28, 29], we can define the specific criterion for pipeline resistance as follow.

$$P_{max} = Sy + 68.95 (MPa) \quad (56)$$

where

Sy : is the yield strength of the pipe material (MPa).

P_{max} : flow strength of the pipe material (MPa)

5 RUL Predicted by EKF and Monte Carlo simulation

The main objective in this section is to present an approach to evaluate the reliability. However, the RUL for deteriorating pressurized pipelines at any distance d and time t require to modeling this degradation process and predicted its evolution on line. We will focus mainly on cross-sectional area of pipeline, where the deposit profile of wax is shown in Figure 2. Based on paragraph (2.2) the degradation process models can be applied easily. From (19),the evolution of cross-sectional area of pipeline can be described by new form of Wiener process.

$$Sc(t) = Sc(t - 1) + \mu.t + \sigma B(t) \quad (57)$$

In order to assess clogging pipeline evolution, there are many techniques for estimation, KF and EKF is the most widely used for linear and non linear systems. Theoretically, we can estimate the RUL of a pipeline containing wax defects using two techniques EKF and Monte Carlo simulation.

5.1 RUL Predicted by EKF

In order to predicted the RUL by EKF that is computed using two steps the first one requires to estimate the clogging level evolution on cross-sectional area $Sc(t)$ of pipeline, trough using the EKF algorithm for non linear system. The second require to use the estimated path of the clogging level and using (44) it may be evaluated the RUL cumulative distribution function. However, the prognostic of degradation path estimated by EKF and illustrated in Figure 12 is used. For a predefined period h and given the current degradation status ($Time, Sc(t)$), we can calculate the RUL distribution. For example from Figure 12, we are taken three points of current degradation path $M_1(20, 0.230)$, $M_2(30, 0.222)$ and $M_3(40, 0.215)$. Figure 13 depicts the cumulative RUL distribution in each point M_i . However, the red curve M_1 is the RUL distribution when monitoring time is 20 unit of time and 0.225 current level degradation. As we can see, the later the observing time, the higher possibility that the cumulative wax in cross-sectional area Sc of pipeline would within the predened monitoring time period.

For evaluation of RUL distribution, a philosophy of RUL estimation based on first passage time distribution, which is given by (44) in paragraph(2.4). In order to analysis the impact of Wiener parameters in degradation path, we demonstrate that with Matlab simulation. However, when μ is small in comparison with σ , we conclude that the drift has a greater impact on the Winer process. Conversely, if σ is small in comparison with μ , then noise dominates in the behavior of the Winer process. For more analysis, we generate ten realizations of Sc paths with μ and σ are estimated before. From (27) and (28), we can estimate the parameters corresponding to the data set. Using the MLE method the estimated parameters are given by: $\hat{\mu} = 0.1525$ and $\hat{\sigma} = 0.052$. In the first time, we apply the EKF algorithm in Section 3. In order to predict the systems future state using (40-44), that enable to use the system information, through the measurement value, measurement error and system noise, it is possible to obtain a trend optimal estimation of the pipeline system state. Therefore, EKF is a recursive algorithm

used for estimation, which needs to save the system state value and covariance matrix at the last time every step of estimation. For level value of degradation process path at unit of time, it is possible to predict the RUL of the pipeline system. The failure threshold value is set to respect the criterion of reliability and safety given in literature [28, 29].

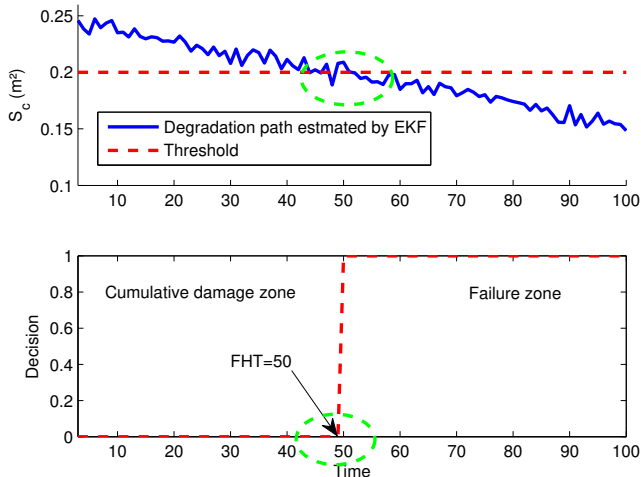


Figure 11: Degradation paths of S_c

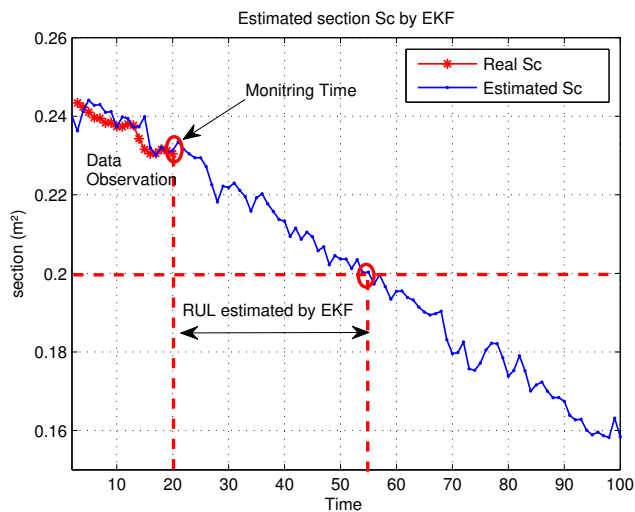


Figure 12: RUL prediction by EKF

Use the criterion in (56) to prove the pipeline system reliability. For example, for $S_{c_{lim}} = 0.2(m^2)$ obviously, the simulation results are shown in Figure(11) and Figure(12). That can give more information about prognostic. Consider Figure (11) which plots the degradation path of S_c , which is predicted by EKF algorithm. Given a threshold level $S_{c_{lim}}$, we can make decision between two hypothesis at time $t = 50$ unit of time, working zone and failure zone. Figure (12) illustrates the RUL prediction by EKF algorithm. When the path of S_c exceed the failure threshold at $t = 55$ unit of time, then we can compute the RUL from time of prognostic $t_{pro} = 20$ unit of time until failure time. However, from (55) the $RUL = 55$ (unit of time).

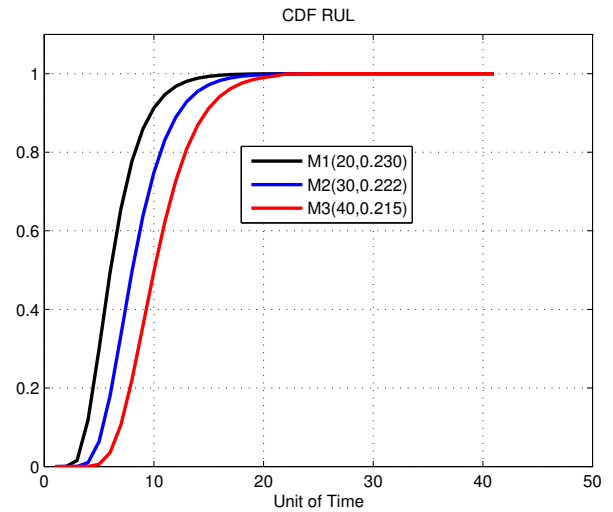


Figure 13: CDF of RUL

5.2 RUL Estimated by MC

The second method requires to generate many paths by Monte Carlo simulation and according to paragraph (2.4) we can evaluate the first hitting time and plot the RUL pdf. Using the Wiener model to generate some paths describing the true degradation process with parameters estimated before $\hat{\mu} = 0.1525$ and $\hat{\sigma} = 0.052$. After that, we apply Monte Carlo simulation for a given threshold level to depict the RUL distribution. For more analysis, we generate the degradation paths of S_c using the presented approach based on Wiener model, we carried out several numerical simulations including the procedures of initial parameters estimation. Moreover, assume that S_c is simulated randomly for $N = 10$ realizations independent and identically tested and that are based on Wiener process with positif drift. In order to analysis the degradation process evolution, we simulate the testing data set using Matlab see Figure 14 and Figure 15. For simplicity and reasons of understanding the evolution of the degradation process, we simulate the first path with mean and variance parameters above and initial condition of $S_c = 0.25m^2$.

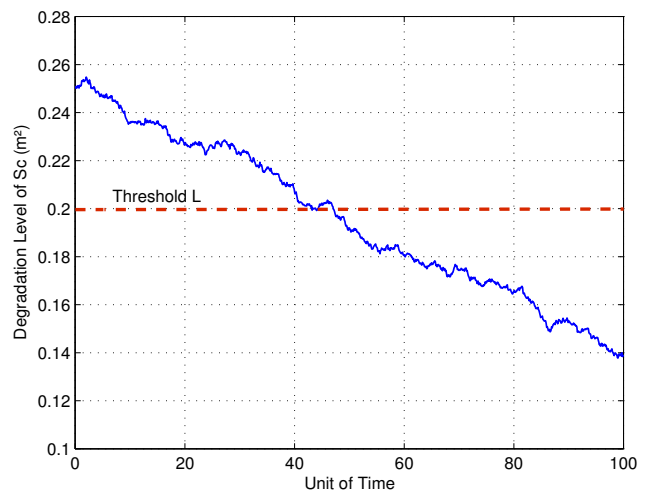


Figure 14: Degradation path of S_c

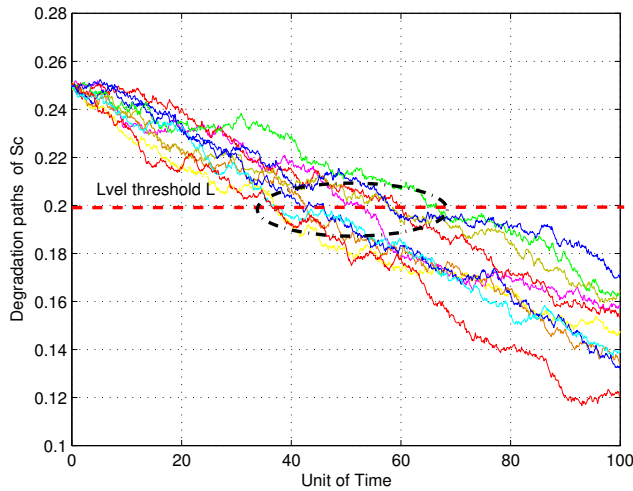


Figure 15: Degradation paths of Sc by MC (N=10)

In the following, we are taken $N = 50$ and $N = 100$ random samples are generated and we want to figure out the time when the degradation paths hits the failure threshold $L = 0.2(m^2)$. It follows that the distributions of first hitting time computed by Monte Carlo simulations, which are compared with analytical function of FHT in (42) as can be seen from Figure 16 and Figure 17. We can say that the analytical function of FHT based on Wiener process approximately pick up the true value in degradation process. However, it is not very reasonable to make such deterministic conclusion only according to figures results. An advanced evaluation need to be carry out to judge the accuracy of the proposed method and the model of degradation process. Nevertheless, there are some relevant problems to be addressed.

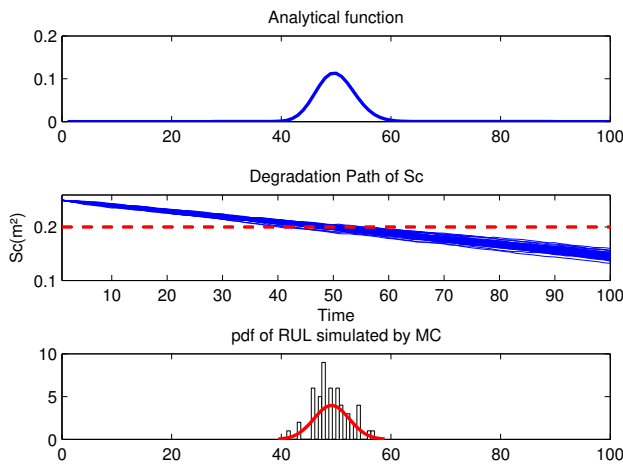


Figure 16: First Hitting Time and pdf RUL (N=50)

However, the choice of the initial parameters μ and σ of the degradation process is important for evolution of the Wiener dynamic drift. For this reason, a simulation with Matlab demonstrate that when μ is small in comparison with σ , then we conclude that drift has a greater impact on the dynamic of Wiener process. Further-

more, if σ is small in comparison with μ , then the noise dominates in the behavior of the Wiener process. But it is very obvious from these figures that the Wiener processes are more spread out when its parameters are fluctuated. Moreover, it appears that the paths has regions where motions looks like they have decreased trend with random fluctuations. It seems to me in this practical case, the degradation does not follow a linear drift.

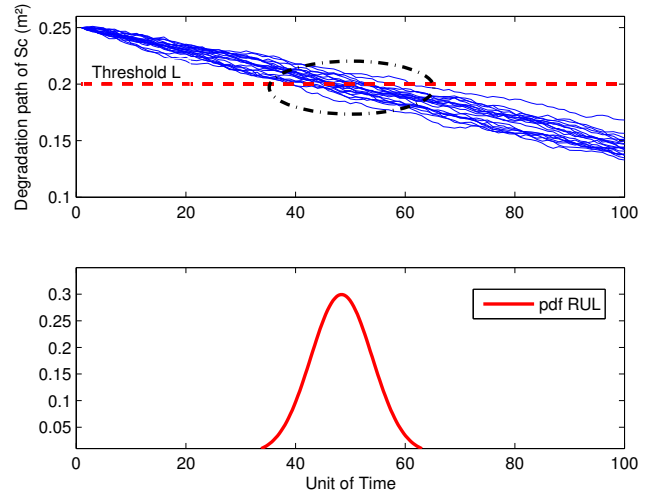


Figure 17: Monte-Carlo Simulation and pdf RUL (N=100)

Figure 18 shows the actual pdfs RUL at different observation. The first curve in the figure is the pdf RUL when monitoring time is $t = 0$ unit of time and the last curve shows the pdf RUL when monitoring time is $t = 5$ unit of time. It is clear that, the latter observing time, the rather that pipeline would have clogging higher possibility. Similarly, we can estimate the mean of the RUL as an useful input for a preventive maintenance activity.

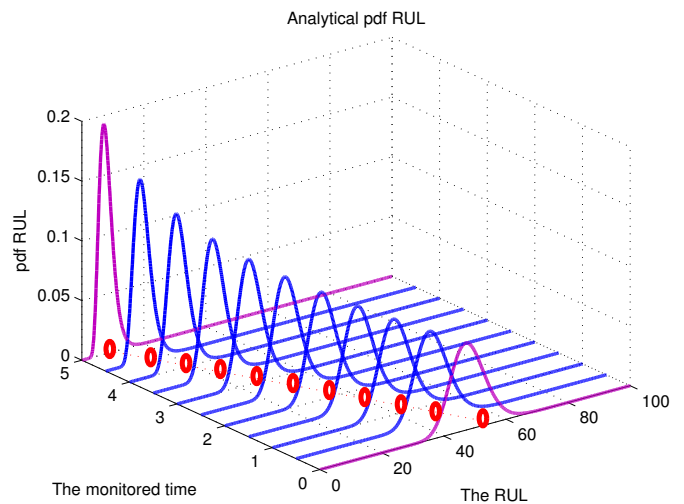


Figure 18: Analytical pdf of the RUL prediction

6 Conclusion

In this paper, we have proposed a new idea that allows to model the degradation process in two tank system. We have focused our mod-

eling particular in pipeline clogging component. We only consider a Wiener process without random effects has a linear drift. But in practice, a nonlinear model may be more appropriate for complex system. Based on the Wiener process model, the unknown parameters in this model are estimated using the MLE approach. However, at the beginning of modeling we have used a simple model to describe the cumulative damage which causes by the small particles in crud fluid in pipeline system. In this context, we have assumed the clogging of cross sectional area of pipeline by walk steps. By using the Extended Kalman Filter algorithm, we have effectively estimated the state parameters system. There fore, the EKF algorithm has been successfully applied to predict recursively the Remaining Useful Life Time of the pipeline containing wax defect. Based on the concept of the First Hitting Time, the Monte Carlo simulation method is used to estimate the probability density function. The effectiveness of these methods are valid though numerical simulations. Moreover, the results show that the work is promising and opens many perspectives for future research.

Acknowledgment The authors would like to thank the editor and anonymous reviewers for their valuable and constructive suggestions that lead to considerable improvements of this paper. The research reported here is partially supported by Laboratory LISIER.

References

- [1] M. M.Hedi, B. H.Fayal , C. Abdelkader. Using Wiener Model Damage in Two Tank System and Prediction of the Remaining Useful Lifetime.15th International Multi-Conference on Systems, Signals and Devices (SSD),2018. IEEE.<https://doi.org/10.1109/SSD.2018.8570671>.
- [2] P.R.S. Mendes,S. L.Braga. Obstruction of pipelines during the flow of waxy crude oils. Journal of Fluids Engineering, **118** (4), 722-728,1996. <https://doi.org/10.1115/1.2835501>.
- [3] P.Lall,R.Lowe, K.Goebel. Extended Kalman filter models and resistance spectroscopy for prognostication and health monitoring of leadfree electronics under vibration. IEEE Transactions on Reliability, **61** (4), 858-871,2012. <https://doi.org/10.1109/ICPHM.2011.6024324>
- [4] R. K.Singleton,E. G.Strangas, S.Aviyente. Extended Kalman filtering for remaining-useful-life estimation of bearings. IEEE Transactions on Industrial Electronics, **62** (3), 1781-1790,2015. <https://doi.org/10.1109/TIE.2014.2336616>
- [5] K. Doksum, A. Hoyland, Models for variable stress accelerated life testing experiments based on wiener process and the inverse gaussian distribution. Technometrics, **34**,74-82,1992. <https://doi.org/10.2307/1269554>
- [6] G.A. Whitmore, Estimation degradation by a Wiener diffusion process subject to measurement error, Lifetime Data Analysis **1**, 307-319, 1995. <https://doi.org/10.1007/BF00985762>.
- [7] N.D. Singpurwalla,"Survival in dynamic environments", Statistical Science **1** 86-103,1995. <https://doi.org/10.1214/ss/1177010132>.
- [8] G.A. Whitmore, F. Schenkelberg, Modelling accelerated degradation data using Wiener diffusion with a time scale transformation, Lifetime Data Analysis,**3**, 27-43, 1997. <https://doi.org/10.1023/A:1009664101413>.
- [9] C.J.Lu, W.O.Meecker,"Using Degradation Measures to Estimate a Time-to-Failure Distribution" Technometrics, **35**(2),161-174, 1993. <https://doi.org/10.1080/00401706.1993.10485038>.
- [10] X. Wang, "Wiener processes with random effects for degradation data", Multivariate Analysis,**101**,340-351, 2010. <https://doi.org/10.1016/j.jmva.2008.12.007>.
- [11] P. Paris,F. Erdogan. "A critical analysis of crack propagation laws". Journal of basic engineering, **85**(4), 528-533, 1963. <https://doi.org/10.1115/1.3656900>.
- [12] P. S. Mendes, A. M. B.Braga, L. F. A. Azevedo, K. S. Correa. "Resistive force of wax deposits during pigging operations". Journal of Energy Resources Technology, **121**(3), 167-171,1999. <https://doi.org/10.1115/1.2795977>.
- [13] W. Kahle,S. Mercier,C. Paroissin. Degradation processes in reliability. John Wiley and Sons,2016.
- [14] T. Nakagawa. Stochastic processes: With applications to reliability theory. Springer Science and Business Media,2011.
- [15] N. Matta, Y. Vandenboomgaerde, J. Arlat. Supervision, surveillance et sret de fonctionnement des grands systemes. Lavoisier, 2012.
- [16] S.Tang,X.Guo,Z.Zhou."Mis-specification analysis of linear Wiener process-based degradation models for the remaining useful life estimation". Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, **228**(5), 478-487,2014. <https://doi.org/10.1177/1748006X14533784>.
- [17] X.S.Si,W. Wang, C. H.Hu, M. Y.Chen,D. H.Zhou. A Wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation. Mechanical Systems and Signal Processing, **35**(1-2), 219-237,2013. <https://doi.org/10.1016/j.ymsp.2012.08.016>.
- [18] S. J.Tang, X. S.Guo, C. Q.Yu, Z. J.Zhou, Z. F.Zhou, B. C.Zhang. "Real time remaining useful life prediction based on nonlinear Wiener based degradation processes with measurement errors". Journal of Central South University, **21**(12), 4509-4517,2014. <https://doi.org/10.1177/1748006X14533784>.
- [19] Z. S. Ye,Y. Wang, K. L. Tsui, M.Pecht. Degradation data analysis using Wiener processes with measurement errors. IEEE Transactions on Reliability, **62**(4), 772-780, 2013.<https://doi.org/10.1109/TR.2013.2284733>.
- [20] F. Jędrzejewski. Modeles aleatoires et physique probabiliste. Springer Science and Business Media,2009.
- [21] G. F. Lawler. Introduction to stochastic processes. Chapman and Hall/CRC, 2006.
- [22] F. C. Klebaner. Introduction to stochastic calculus with applications. World Scientific Publishing Company,2012.
- [23] M. Lovric. International Encyclopedia of Statistical Science.687688, Springer, 2011.
- [24] M. S. Grewal, A. P. Andrews. Kalman ltering: Theory and practice using matlab, wiley, 2001.
- [25] GOMM, J. Barry. Fault Detection in a Multivariable Chemical Process by Monitoring Process Dynamics. IFAC Proceedings Volumes, **27** (5),171-176, 1994. [https://doi.org/10.1016/S1474-6670\(17\)48023-2](https://doi.org/10.1016/S1474-6670(17)48023-2)
- [26] M. Blanke, M. Kinnaert, J. Lunze, M. Staroswiecki, J. Schroder. Diagnosis and fault-tolerant control,**2**, Springer, 2006.
- [27] J. A. Svendsen. Mathematical modeling of wax deposition in oil pipeline systems. AIChE Journal, **39** (8),1377-1388, 1993. <https://doi.org/10.1002/aic.690390815>.
- [28] M. Ahammed, R. Melchers. Reliability estimation of pressurised pipelines subject to localised corrosion defects. International Journal of Pressure Vessels and Piping, **69** (3),267272, 1996. [https://doi.org/10.1016/0308-0161\(96\)00009-9](https://doi.org/10.1016/0308-0161(96)00009-9).
- [29] M. Ahammed. Probabilistic estimation of remaining life of a pipeline in the presence of active corrosion defects. International Journal of Pressure Vessels and Piping, **75** (4),321329,(1998). [https://doi.org/10.1016/S0308-0161\(98\)00006-4](https://doi.org/10.1016/S0308-0161(98)00006-4).

Spot Toyota: Design and Development of a Mobile Application for Toyota's Promotion Actions to the Young Audience

Nuno Martins^{1,*}, Joel Enes²

¹Polytechnic Institute of Cavado and Ave / ID+, 4750-810 Barcelos, Portugal

²Polytechnic Institute of Cavado and Ave, 4750-810 Barcelos, Portugal

ARTICLE INFO

Article history:

Received: 21 April, 2020

Accepted: 29 May, 2020

Online: 20 June, 2020

Keywords:

Mobile App

UI & UX Design

Digital Design

Communication Design

Toyota

ABSTRACT

This project aims to demonstrate the importance that digital media can have to the development of loyalty programs, namely in creating empathy and proximity relationships between brands and their target audience: young people. This study consisted of the creation of a digital platform for Toyota Portugal, named Spot Toyota, to communicate actions promoted by the car brand, especially during music festivals. With the support of advertising agency Caetsu, this mobile application was developed to bring the brand closer to a younger audience – festival fans – with potential interest in two Toyota fleet car models: Aygo and C-HR. Through strategies typical of loyalty programs, such as the awarding of vouchers or coupons, the accumulation of points or winning prizes, a system was produced with the main focus on attracting users to the platform in a continuity perspective. The working process of this investigation resulted in the design of a smartphone application, based not only on the analysis of other examples present in the market but also on the understanding of crucial subjects such as loyalty programs, UX and UI design, application of personas models, creation of wireframes and workflows, and development of usability tests.

1. Introduction

This project arose from a proposal, from the Portuguese advertising agency Caetsu, to develop a digital application (app) for the Toyota car brand.

This app, entitled "Spot Toyota", was originally created in July 2018, by an internal department of Toyota Portugal. The main purpose of the app was to create connections of proximity and empathy between the Toyota brand and the younger audience (often at music festivals). Thus, it was intended that this app worked as a loyalty product and not as a direct way of promoting or selling cars.

The study initially focused on searching mobile applications in the automotive sector. However, since no considerable supply was found in this sector, other platforms such as Cartão Continente, MB Way, McDonald's and Galp EvoDrive, designed under the same assumptions as Spot Toyota, were taken into account. In addition to the consumer-perceived value, other key

concepts for research were also analysed, such as loyalty programs, usability, UX and UI design.

After understanding the problem, the work moved on to the platform design phase. In this sense, essential methods and guidelines were put into practice in the process of designing the mobile application. Realizing the value that this solution could provide to the market was one of the first steps. Then, the project advanced to the creation of *persona* models in order to understand the pattern of the platform users.

Subsequently, the phase of developing information architecture began, which would be fundamental to the creation of the app *wireframes* and *workflows*. In this step, the screen sketches were conceived, distributing the graphic elements according to the importance and hierarchy of the selected information. To the design, were observed the user actions in the intended tasks accomplishment.

The platform creation evolved to the interface design phase.

Following the brand guidelines, the design of the screens began with a more considered and rigorous focus on visual issues, the

*Corresponding Author: Nuno Martins, nunomartins.com@gmail.com

graphic elements distribution, the interface communication, the color selection, shape, and information. This process aimed to create a credible product and offer the best user experience. Finally, the phase of usability tests was initiated, the main purpose of which was to detect problems in the interface and develop alternatives to improve its performance.

2. General and specific objectives

The main objective of this project arose from a real need on the part of Toyota Portugal to create a digital platform that would help the brand get closer to one of its target segments, the young audience. In this sense, this platform was created for an audience potentially interested in two models of Toyota's fleet in the Portuguese market: Toyota Aygo and Toyota C-HR.

Consideration was given to developing a mobile application (app) as a privileged channel to communicate for the brand's actions with the defined target, in order to capture a maximum of leads; to support lead conversions in effective customers; and to retain both customers (buyers) and fans of the brand (not necessarily buyers, but individuals capable of influencing their close circle).

As a result of the introduction and implementation of RGPD (General Data Protection Regulations), Toyota was forced to efface the contact data the company used to communicate brand information and actions. The app to be developed, thereby, in addition to trying to regain lost contacts, would mainly serve to establish and consolidate a regular empathy with the public.

The goal was not only to develop a conceptual study of a mobile app through good design practices, but also to build a tangible product, with added value, that could be implemented and be a reference in this automotive market [1].

To achieve the general objective (sustainable development of an app), a set of specific objectives was addressed:

- Identifying, understanding and systematizing the goals proposed by Toyota for the app: Through the meetings held with Caetsu, the company responsible for monitoring the project, a detailed analysis of the work context was prepared. For the app development, it became necessary to understand and systematize all the raw information, transforming it into raw material to start the project.
- Listing the added-value offers that Toyota can make through the app (discounts, points, coupons, among others): As the mobile app had well-defined objectives, it was essential to understand the type of functionalities and specific offers that it would have to contemplate. In order for the brand to be successful in approaching its target audience, the app would have to be perceived as an offering of added value to its users. From this point of view, the focus shifted to examining loyalty program techniques, such as discounts, points and coupons, and designing attractive means to encourage the use of the app.
- Designing an app that contains relevant and useful information and functionalities for the user, with an appealing and intuitive interface, culminating in a high degree of usability satisfaction.

3. Methodologies

In an initial phase, an in-depth inquiry of the problem was carried out through relevant data collection and analysis.

A first stage of information gathering was carried out in a meeting with the Caetsu team, the agency responsible for monitoring the Toyota project, in which the main objectives, needs and expectations concerning the app were discussed. Correspondingly, schemes and mental maps, which defined the essential factors to conceive the platform's contents, were produced. It became crucial to understand the typology of mobile application offerings in the automotive sector, but above all, other applications built on the same assumptions as Spot Toyota. Analysing these examples, we were able to understand the range of features they offer as a reference for the work to be developed in the app under study. Having said that, it was essential to understand what is a loyalty program, how it works, and how the app could use these strategies to bring the user closer to the platform.

In a second phase, and after this analysis of the state of the art, the work was developed with special focus on interface design, experience and usability heuristics.

In the first stage of this second phase, the *personas* study began, which focused on the construction of profiles based on real users. Mirroring the target, the *personas* represented the various types of users that can use the developed app. Then, the project proceeded to the information architecture phase. The content present in the app was mapped, as well as the paths and links between the various stages of interaction within the platform. This work would lead to the construction of the app *wireframes* and *workflows*. The project sketches were drawn using basic representations to facilitate the perception and thinking of the interface division. In this follow-up, the *workflows* were fundamental to organize the *wireframes*, and to perceive step by step (in an illustrated way) the user's navigation intentions.

Later, the project evolved to the graphical construction of the interface. Respecting the Toyota brand identity standards, the screens design began. The work focused on information hierarchy, readability, color and contrast. On this step, there was a rigorous focus on visual issues in order to create a credible product and offer the best possible user experience.

Finally, based on user feedback, the usability testing phase was carried out. Through methods to test and observe user behavior, it was possible to detect interface problems and develop enhanced alternatives. These improvements meant a better user's perception, and easier and faster navigation within the contents.

After this process of app development, data collection, review and feedback from users, started the implementation of the programming language, a task that was assigned to the Caetsu team. Once finished, the app became available on the marketplaces to be installed and used according to the defined purposes.

4. State of the Art

4.1. Loyalty programs

Loyalty programs, in general, are sales techniques that guarantee consumers benefits from their purchasing behavior, while allowing sellers and service providers to build a long-term relationship with the customer [2]. These consumer benefits can take the form of discounts, cash, bonuses or even access to special

services such as in the case of magazines and newspapers that offer exclusive content for loyal readers.

With the development and investment of the technology industry in the smartphone and mobile applications market, a favorable environment was created for the growth and evolution of loyalty programs. Mobile applications provide a highly segmented channel for two-way communication, becoming powerful drivers of customer loyalty.

“Mobile applications are the main channel for most, and it is clear that they present themselves to businesses as an unprecedented opportunity to interact directly with customers. They are easily accessible through an icon on the main screen; they do not require URL's and currently provide the best experience on mobile devices.” [3]

Spot Toyota, despite not presenting itself as a service aimed at converting sales or with the objective of becoming exclusively a loyalty program, sought important guidelines and tools for creating empathy with its users in these marketing techniques. It was in this context that it became essential to leverage all the advantages of a mobile application (app) to boost and develop direct communication channels with the defined target audience. For instance, with a mobile application, the process of sending individual and segmented messages, known as "push notifications"¹ is simpler. With a great distance between the company and the user, this conversation mode can make a difference in terms of gaining customer loyalty.

With people spending much of their time on smartphones and applications, companies have the opportunity to increase their customer loyalty by offering a personalized application experience and making the customer feel valued [3].

4.2. Apps in the automotive sector

After an analysis focused on the offer of mobile applications in the automotive industry, especially in the Portuguese market and in Google Play (Android) and App Store (iOS), it was understood that mobile applications can be divided into three major typologies: Lifestyle; Remote Control; and Utilities.

Some brands that can be considered in the Premium segment, such as Volvo, BMW and Mercedes, make a digital magazine available to their customers in a mobile application format. With a refined and high-quality graphic presentation, they approach a well-defined target, giving access to exclusive news about the brand, and offering a wide range of information about services and accessories, inspirational articles, films, photography and links to explore their vehicles universe.

In the second typology, it is possible to include an entire range of applications aimed at controlling the car digital system itself. In these cases, after pairing the smartphone with the vehicle, certain levels of control access are given to the system: it is possible to access and monitor the GPS and select options on the music player, among some functionalities. An example in this regard is the Volkswagen Media Control app.

At the last typology, we can inscribe the automotive industry applications with a great presence in stores, distributed by countries or markets. This is the class of applications for car daily control, in which it is possible to: monitor the vehicle condition; manage duration, distance and consumption of trips; and record the entire maintenance plan, miles and guarantees.

Although applications with the purpose of increasing loyalty and offering value to attract new customers are recurrent in the digital market, we may declare can be in that the automotive sector has not yet adopted this strategy, addressing all its offer to existing vehicle owners. Given the potentially disruptive advantages granted by modern-day technology, such approach can be considered a misuse.

4.3. Reference Apps in the area of loyalty

In the universe of mobile applications, there is a wide range of examples to observe. However, this part of the research has not focused on the automotive sector because, as we have seen before, there has not been much supply in this area so far. Therefore, the approach focused on other business sectors in Portugal with models that could serve as guidance examples for Spot Toyota.

The "Cartão Continente" app, belonging to one of the major Portuguese hypermarkets' chains, has more than 500K installations in Google Play (Google Applications Store) and it is a reference app in consumer loyalty. In a clear graphic approach and adequately mirroring the brand identity, it offers a set of tools that capture the user's interest and promote its use. It makes available numerous discount coupons and offers when buying products and services. Similarly, the "Galp EvoDrive" app, from a fuel supplier, is identified by the constant providing of discount coupons to users in order to attract them to the shops. A relevant feature of this application is the possibility of sharing discount codes between users. This way, the brand benefits through P2P (person-to-person) advertising and invites further usage of the app.

With a graphic interface based on a colorful chromatic palette, which can be assumed to be dedicated to the young audience, the McDonald's app stands out. It also makes use of loyalty tools, such as discount coupons on certain products and the accumulation of points for purchases made.

In the services context, "Via Verde", from a company that provides an electronic toll collection system, presents another type of benefit to consumers. In this case, the attribution of points and discounts is not made through what the brand sells, but through advantages in its entire network of travel and hotel partners.

As another example within this particular selection of reference mobile applications, is "MB Way", an app that offers a product with interesting tools for attracting users. Although it is designed mainly to facilitate money transfers, it uses playful interfaces (gamification), creating an environment for the users to engage with recreational features and game dynamics, both individually and in connection with other users. According to Cook [4], when customers are driven to participate, or to act in some way, they are more likely to feel that they will be rewarded for their action,

¹ "Push Notification" is a notification that the user receives through a smartphone application, tablet or browser without requesting it.

thereby reinforcing the interaction. The "MB Way" is comprehensive when it comes to loyalty techniques, however, certain usability issues are evident in terms of the clarity of perception regarding all possible interactions.

Regarding the interfaces of the applications presented, it is possible to observe a careful and contained use of the most expressive and complex graphic elements, creating a simplified and intuitive language. Mostly designed on clear backgrounds, the interfaces show concern towards clarity in terms of comprehending all features and interactions. Due to the large volume of information and the fact that the success of these applications depends on guaranteeing a good user experience, it can be considered that all these platforms are designed with special focus on usability and accessibility issues. As Norman [5] states, "Good design requires good communication, especially from the machine to the person, indicating what are the possible actions, what is happening and what is about to happen".

With these examples, we understood how the strategies and functionalities that these applications use to enhance the loyalty of their users are put into practice. We can consider that the component of discount coupons and accumulation of points became a common and transversal practice in almost all applications analysed, highlighting the importance of rewarding the user for using the app. In accordance, Spot Toyota used these proven models in the market as benchmarks to design its own features set.

5. App development

This project was developed from an already existing product that was still basic, with a strong need for development.

After discussing with Caetsu and Toyota Portugal managers what was intended for the app, a first prototype was developed. The prototype contemplated all the functionalities and needs discussed in the first meeting between the research team and the Caetsu and Toyota Portugal managers.

Although it was developed on a poorly sustained basis, with a lack of definition of *workflows* and critical paths, the first prototype was fundamental for the second meeting with Caetsu and Toyota. It was the basis for discussion towards defining the design and functionalities of the final project.

The prototype presentation contributed to a more objective discussion about the app's problems and to the suggestion of new improvement ideas, namely: the design of a graphic interface with lesser red tone; the change of the typographic font to the official Toyota font (the Toyota Text Font); a greater presence of the Toyota Spot logo on the various screens; the redesign of some tab bar icons as well as some buttons.

After this second meeting, the entire app structure design was started. After collecting all the first prototype inputs, the final *workflows* were designed, the *personas* created, the use cases studied and the functionalities to contain were listed. Then, the redesign phase of the *wireframes* has started, based on the corrections discussed.

To improve and understand the interaction with the platform, usability tests using real users were developed. With the collection of their feedback, it was possible to apply the necessary

adjustments and changes to evaluate and verify the user experience and interface design, thus solving most usability problems encountered [6-10].

It is important to note that all changes and decisions made during this whole process were constantly communicated to those responsible for monitoring the project, with the final validation, as well as the definition of other necessary corrections or changes [11,12].

5.1. Personas

When creating a digital product such as an app or a website, it is essential to know who will use it. In this sense, and to anticipate the problems that may arise from its use, it is necessary to understand the user and his needs [13].

The *Personas* are fictional characters, created based on a market analysis. They represent the various types of users who can have a similar comportment when using the service or product developed.

This process of idealization leads us to understand the needs, experiences, behaviors and goals of users [14]. *Personas* simplify the design process, in the perspective that all processes are developed on the basis of ensuring the best user experience for the identified target audience [15].

All this survey of the user's characteristics, needs and objectives is translated into relevant information, which helps us to calculate the route the user takes on the platform.

Through the development of the *personas*, we try to create empathy with the target users and correspond to their needs, which are tested in prototypes. The creation of these profiles is based on real users, i.e. all the conversion of needs into effective functionalities of the application derives from the real needs of users [16].

The development of the *personas* for this project (Table 1) has gone through study and contact with different people who fit the target audience of Spot Toyota. The insights provided by Caetsu and Toyota Portugal, who had already gathered detailed information about the defined target market, along with a series of informal conversations with recurring festival attendees took by the researchers, made it possible to embody the *personas* portraits.

Profiles were constructed based on the same interests and assumptions: a young, active audience, who had already participated in music festivals or had shown interest in them [13].

It was added some personal information that summarized what matters in the *persona* who relates to the product.

To add a demographic profile describing the personal and professional background, the user environment, and the psychographics, like interests, motivations, and pain points, was also important.

The following step was to attribute the end goal, the motivation factor that inspires the action. This determines what the *persona* wants or needs to fulfil.

To conclude, it was created a scenario to describe how the *persona* interacts with the product in a particular context to achieve his end goals.

Table 1: Analysis grid of different *personas*.

RAQUEL 18 years old	PEDRO 25 years old	JOANA 32 years old	TIAGO 20 years old	ABEL 26 years old
Fan of indie rock music; about to enter university education; loves cats.	Addicted to live concerts; intern in a marketing studio; fan of promotions.	Loves summer festivals; loves photography ; works in an architecture studio.	Camping enthusiast; studying computer engineering; motorsports fan.	Enjoys a good concert; Guitar teacher; Environmentalist.
Use case: Raquel, for financial reasons, participates in all competitions and activities in which she can win tickets to concerts and festivals. Uses apps and groups in social networks to be up to date with all the news.	Use case: João adheres to all services that earn him points and gets rewarded for it. If there are no associated costs, he installs all the applications that may bring him advantages in the short or medium term.	Use case: Joana is addicted to new technology and all that is practical and time efficient. She uses digital versions of tickets for concerts, trips, and offers.	Use case: Tiago, as a fan of the automotive world, ardently follows brands and their activities, to have the possibility to see new car models first hand, and to be able to go for test-drives.	Use case: Ricardo, not a big fan of digital media, uses some applications that his friends impose upon him to validate codes that they forward. Since he is indifferent to them, he ends up helping them.
Objectives: - Follow news and campaign releases; - to receive notifications to the second of all news.	Objectives: - Earn points to be able to convert into Toyota prizes. - Be rewarded for daily use of the app.	Objectives: - Use festival coupons to enjoy Toyota rides; - Participate in brand actions just for the pleasure of the challenge.	Objectives: - Be informed of Toyota brand actions; - Know what actions will be undertaken where, and which campaigns will be associated with automobiles.	Objectives: - Validate codes sent by friends; - Send codes for friends to validate.
Frustrations : - Being afraid of losing some important information or launching a hobby. - Having to look for information on scattered platforms.	Frustrations : - Having difficulty converting points into advantages; - Knowing that if would require a long time for any reward to be effectuated.	Frustrations : - No quick and intuitive access to coupons; - Not being able to access and utilize through smartphone.	Frustrations : - Access outdated information; - Unclear and disorganized information.	Frustrations : - Time-consuming and complicated process; - Need to go through complex registration processes in order to access the app.

5.2. Information architecture

After studying the competitors and analysing the *personas*, we moved on to the project information architecture phase. This process became decisive for the final product development, since it was the main stage of content organization for the app.

Based on the first conversations with Caetsu, and taking into account the existing product (Spot Toyota in browser version), the design and structuring of the flowchart began.

In order to meet the brand's purposes, it was fundamental to set up reward methods and techniques. The intention was to encourage the use of the app and consequently connect the user to the environment that the brand had imagined for the project. This system was developed based on the good references in the market and the cases analysed in the state of the art section of this article.

Thus, based on the research, it was defined that the points would be awarded as per the following:

- **For the daily use of the app:** once a day, and after the user enters the app, points are automatically generated and added to the balance of the user presented in his/her personal area.
- **For the use of vouchers or coupons:** upon the brand's interest in the actions where it made itself present, vouchers or coupons can be awarded to the user, which, after use and validation, turn into more points.
- **By invitation:** in the personal area of the app, there is a section called "Invite friends". Here, and by sharing a code with other contacts, a user can add points after at least one of those contacts has validated the code sent.
- **By inserting a code:** in precisely the same place mentioned in the previous point, there is another section entitled "Insert promotional code". If the user has received a sharing code from a contact or by some promotional action of the brand, he/she gains points after the insertion and validation by the app of that same code.

In addition to these point attribution methods, another specific system was developed for the project: through the interaction with the app, the user progresses on a scale of five levels. On reaching the last stage, he/she is rewarded and goes back to the initial step. With this process, the user is able to add points always with a concrete objective.

We concluded that flowchart had translated into an added value for the perception of the whole project hierarchy and organization, as well as a fundamental work base for the construction of the corresponding *wireframes*.

5.3. Wireframes

The next phase of the app construction was the wireframe design. We can define this concept as a low fidelity representation in the initial design phase of a certain visual product. It is the outline of the structure of a digital project, using basic representations, such as lines or circles, in order to facilitate the ideation of the interface division. Normally, *wireframes* are used to arrange the content and functionality of a page, reflecting the user needs and paths of interaction. Because it is a sketch, this type of work allows the solutions presented to be subjected to constant changes, until a suitable result of the app is obtained in the final interface design [17,18].

For this project, although the product would eventually be made available cross-platform (including Android), we chose to develop the app in iOS, the mobile operating system for iPhones

and iPads from Apple Inc. The choice was made in line with the convenience and experience of the researcher responsible, because it was the system with which he was best familiar. The knowledge of its usability standards, through the daily use of the platform, thus made the project construction process more conscious and reliable.

5.4. Workflows

Workflows are sequences of steps a user goes through to accomplish a certain task. In this type of project, workflows are fundamental for the wireframe's organization, in order to understand, step by step, the user's navigation intentions. This is the flow that is established for the tasks' realization. For the workflows development we use the Sketch software for aiding in the rapid prototyping of the various screens, and testing all the interaction between wireframes and workflows.

In this sense, we highlighted some of the critical concerns of the personas created, in order to perceive the applicability of the designed flows and how these minimized the expected constraints. Based on the cases of Joana, Pedro and Ricardo (Table 1), we developed the workflows related to the tasks that could be most relevant to their concerns.

For example, in Joana's case, one of her frustrations was the fear of not having access to the coupons quickly and intuitively. Since the beginning of the project, and being one of the main features of this app, the access to vouchers and coupons would have to be as upfront and evident as possible, creating a high degree of user satisfaction.

To understand how this workflow works, we present the following example: "Toyota Rides" is one of the activities practiced by Toyota at summer festivals. Through the fleet of cars available at festivals, hitchhikers are given rides between points of interest and the concert venue. This app also appears to speed up all registration and validation processes at time of enjoying the ride. In this case, after getting into the Toyota car, Joana needs to access the mobile phone and show the travel coupon in order to get the ride.

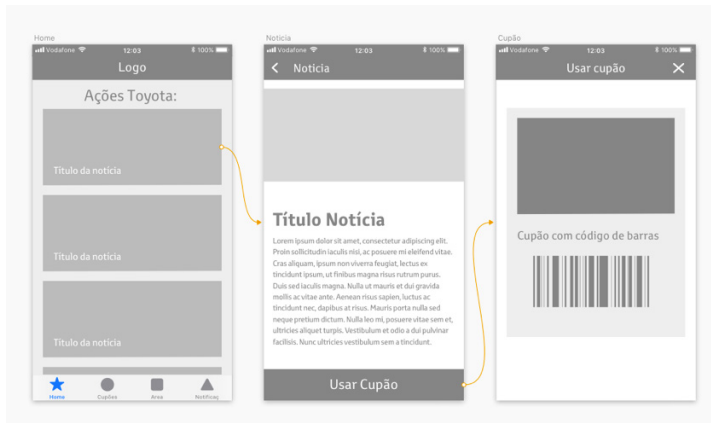


Figure 1: Workflow vouchers/coupons 01

To increase the assertiveness in carrying out this task, we have created, for most of the cases, two possibilities to access vouchers and coupons. During a summer festival, such as the Vodafone Paredes de Coura, there are brand activities that matter to the target market, such as the "Toyota rides". As illustrated in Fig 1, the user

can access the details in the news concerning the hitchhiking, and from there is redirected to the coupon associated with the action. However, and as a faster and more direct way of accessing, the user always has a direct button to the vouchers and coupons section available in the lower tab bar (Fig. 2). From here, all available coupons are displayed, and the corresponding details can be accessed just by clicking on "Use Coupon".

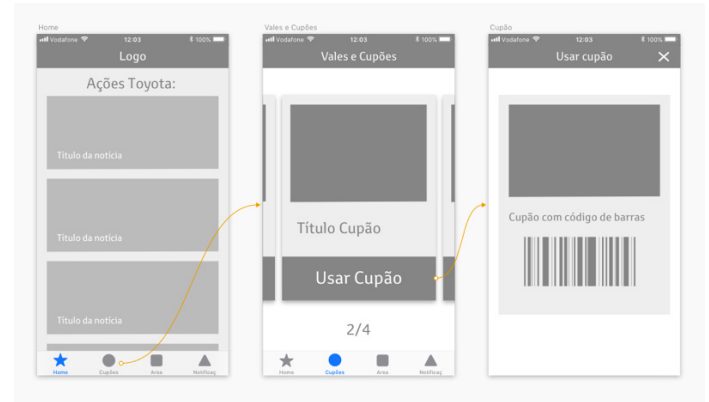


Figure 2: Workflow vouchers/coupons 02

This is an app aimed at gaining and maintaining user loyalty. The associated system of validation and acquisition of points, management of vouchers, and the awarding of rewards provides the app with a strong playful aspect, and is intended to contribute further towards connecting present and potential future users with the Toyota brand.

5.5. Interface design

After the wireframes and workflows design, we moved on to the rigorous design phase of the interface. The screens were designed with a set of graphic and usability concerns, with the goal of creating a credible product that reflected the values of the Toyota brand, and which could provide a best user experience. According to Tidwell [8], users do not trust amateur platforms where there are no efforts taken to create a pleasant appearance.

Since there was already a Spot Toyota version (but only a browser version), a number of aspects had to be taken into account for the restructuring of the entire platform, including the "Spot Toyota" logo (Fig. 3), regarding which it was decided to be kept unchanged.



Figure 3: Spot Toyota logo.

The graphical approach was developed based on the standards and guidelines defined by Toyota. Although this is an independent product from the car sales business, we have endeavored from the outset to maintain visual consistency with Toyota's identity in

Portugal. To achieve this goal, we worked with the guidelines provided in the new Toyota brand visual identity manual [19,20].

For the color system, the work was based on the reference colors defined by Toyota's identity manual: "Primary Color" (grey "PMS Cool Grey 5") and "Highlight Colors" (white, red "PMS 186", and dark grey "PMS CG 11"). In addition to the color, the icons (Fig. 4) adapted for the platform were also conceptualized. We selected those that we considered as having a greater relevance with the app environment. The icons were thus kept as simple line drawings, that were easy to read and carried a minimalist and youthful visual language.



Figure 4: Application icons.

For the Toyota Display and Toyota Text typefaces (original typefaces designed exclusively for Toyota), specific indications in the Toyota brand's visual identity manual were considered.

The layouts design was shaped based on these visual elements, such as icons, color and text. It was a constantly evolving process, however, at each step, Caetsu's approvals were taken into account, alongside with conclusions from user tests; and strategies that were considered appropriate for the target audience [21,22].

The first designed layouts had a higher incidence of the color red. With the work development, red was replaced with other colors of the pre-defined chromatic palette. With this change, it was possible to create greater harmony and clarity on the screens, particularly in those containing images, simplifying the interface graphic environment and minimizing the use of unnecessary contrasts.

In this follow-up, and through constant articulation with Caetsu's perspectives, changes were implemented that would end up significantly altering the graphic aspect. Although the structure has been closed since the *wireframe* phase, it was necessary to optimize certain graphic aspects in order to achieve the best result.

As we can see in figures 5 and 6, the tab bar was one of the elements in which the color was changed from red to light grey.

As it was felt that the Spot Toyota logo needed further prominence, a bar in the same tone was also implemented, increasing the contrast considerably. This change ended up highlighting the menu titles in the nav bar in the same way.

Regarding the homepage, it was concluded that the news/action cards needed to highlight more their button functions. With these elements redesign, an attempt was made to remove any doubt that these areas were clickable, pushing this action.

Other changes focused mainly on the emphasis given to certain titles or text fields, the fonts weight on certain buttons, chromatic coherence in the "call to action" style, as well as certain improvements in content.

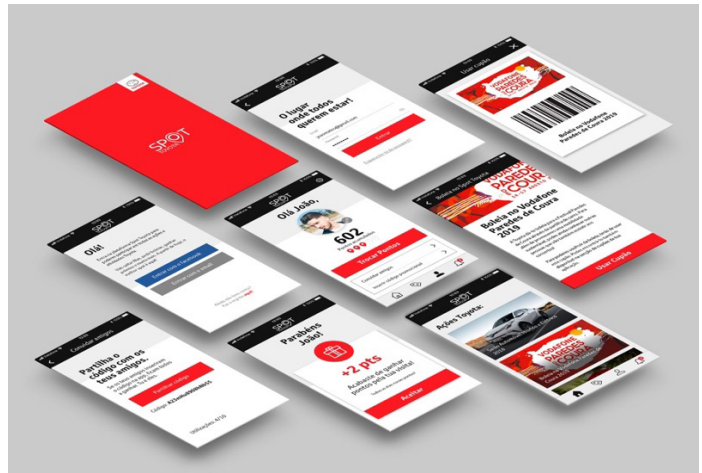


Figure 5: Final version screens.

In figure 6, we present through the example of the "home page" screen, the result of the changes made in the interface graphics.

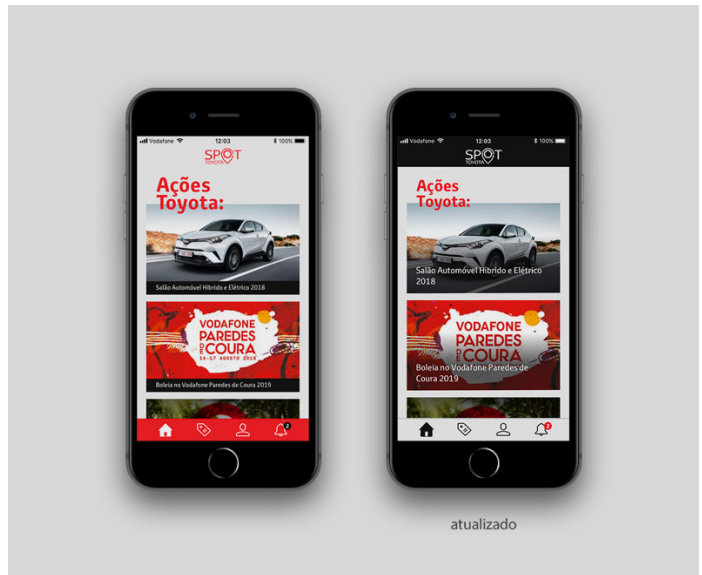


Figure 6: The first version (left) and the final version (right) of the "home page" screen.

5.6. Prototyping

The entire prototyping process proved essential for the communication between the entities involved in the project, Caetsu and Toyota. Since the first meeting, the prototype was an indispensable tool for the evaluation of the work developed and of what would be the real functioning of the mobile application (app).

For the realization of the prototype, we used the Marvel web application. By mapping the clickable zones per screen, we assign the function, the type of interaction and the animation we want for the target screen. Once this process was concluded, based on the *workflows* developed, we had a prototype capable of transmitting what was intended for the final app, and therefore a fundamental tool for testing with users [23,12].

Besides this type of prototyping, some simple animations were created to illustrate the app dynamics. Using these dynamic

images, it became easier to present to the programming team what we wanted (Fig. 7).

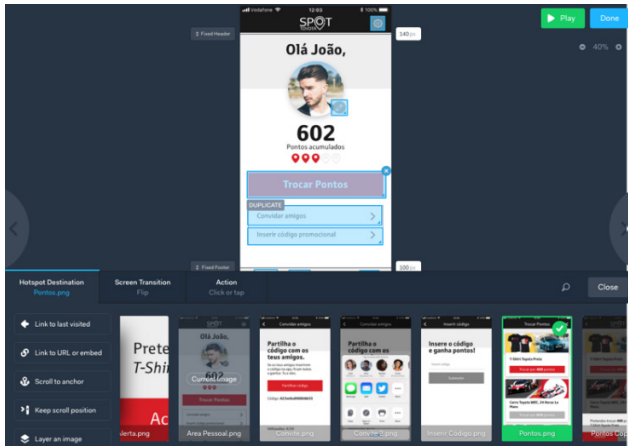


Figure 7: Marvel Platform. Connections between screens.

5.7. Usability tests

Usability tests were developed beginning from the first phase of interface designs. In conjunction with Caetsu, and as the project evolved, decisions were made based on constant user feedback. This process allowed the app to reach a final phase with considerable optimization of the user experience and interface, as well as a reduced number of errors and inconsistencies to be detected at a later stage.

For concluding usability tests with greater accuracy in the analysis method, ten users with an approximate profile to the project's target audience were selected: young users, between the ages of 17 and 32, with some connection degree to the music festival environment.

In this process, user reactions were observed throughout their interaction with the app. For this procedure, we used the last developed online prototype.

For the tests, a set of different tasks to be performed by the test users was drawn up.

Task list:

1. Log in to the app until you find Toyota shares/news;
2. Access the coupon "Hitchhiking on Vodafone Paredes de Coura 2019";
3. Exchange 400 points for a black Toyota T-shirt;
4. Enter a promotional code;
5. Access the news "Hitchhiking on Vodafone Paredes de Coura 2019";
6. Edit your data profile.

After the presentation of these tasks, we asked users to indicate the difficulties encountered, suggestions for improvements, questions, and finally a personal evaluation of the app through a short inquiry form.

According to the observation made to the participants, we concluded that everyone was able to accomplish the proposed tasks. It was also found that the time taken to perform the tasks was

www.astesj.com

identical, with minor differences justifiable by the different levels of users' digital literacy. As expected, we depict these results as the outcome of constant proximity to users during the project development.

In order to assess the level of these same users' satisfaction, we also used, as mentioned above, a short inquiry form. This was provided after completing the usability tests and enquired about the interaction quality and the information appearance, alongside the usefulness and the global appreciation.

The usability tests were extremely important as they allowed us to analyse and evaluate the project with regard to navigation issues, understanding and interface involvement. It proved to be a fundamental process to improve the app usability.

Based on the collected feedback, a high degree of overall satisfaction was obtained, both in terms of the app usage and perception of its added value. These assessments validated and helped to optimize content structuring, information organization, aesthetic and graphical approach, as well as the functionalities developed to capture the users' interest.

6. Conclusions

The main objective of this research was to design a mobile application (app) capable of stimulating a closer relationship between the Toyota brand and its target audience, young people. Thus, through this app, it was intended to explore the potential of technology and digital communication for the loyalty programs development.

Based on the identification and analysis of the state of the art, it was possible to establish objectives that led to the conception of a totally differentiating and pioneering product in the automotive sector. Developed on a real need, and already tested in a previous version, this app aims to create value for Toyota, with effective implementation capacity in the market.

It became essential to understand all the needs, objectives and goals that Caetsu and Toyota Portugal had for the project. Only through this multidisciplinary work, reflected upon the objective articulation of ideas and communication between the entities involved, was it possible to successfully carry out this research project.

The methodology, once properly systematized, made it possible to explore and understand the context of use and the user, as well as Toyota's intentions and expectations. We used several methods throughout the platform development process, including the understanding of UX and UI design, the application of the *personas* method, the design of information architecture and the creation of *wireframes* and *workflows*.

The implementation of Spot Toyota is expected to contribute effectively to a targeted and closer communication between the brand and its audience. From a selective list of features, users will be given access to exclusive brand information; rewards will be given for the use of the app; unique features will be made available for participation in brand actions, and finally, access will be given to a set of tools for awarding points that can be exchanged for products or services.

For Toyota, it will provide the opportunity to regain contacts lost by the entry into force of RGPD; increase its contact database;

develop lead capture strategies; get to know better its target audience in terms of habits, behaviors and needs; and finally establish strong and lasting relationships with its target group.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

This research was supported by the Polytechnic Institute of Cavado and Ave [IPCA] and the Research Institute for Design, Media and Culture [ID+].

References

- [1] N. Martins, D. Brandão, D. Raposo (Eds.), *Perspectives on Design and Digital Communication: Research, Innovations and Best Practices*, Cham: Springer, 2021 (in press). DOI: 10.1007/978-3-030-49647-0. ISBN: 978-3-030-49646-3.
- [2] M. Erbschloe, "Customer Loyalty Programs". *Salem Press Encyclopedia*. [Online]. Available: <http://search.ebscohost.com/login.aspx?direct=true&db=ers&AN=89163635&site=eds-live>. [Accessed: 23-Jul-2019]
- [3] J. Oliver, "Forward Thinking Accounting Firms Turn to Apps to Boost Client Loyalty." *Credit Control*, 39(1/2), 30–32. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=heh&AN=128028330&site=eds-live> [Accessed: 5-Nov-2019]
- [4] W. Cook, "Five Reasons Why You Can't Ignore Gamification". *Incentive*, 187(1), 22–23, 2013. <https://doi.org/10.1248/bpb.23.762>
- [5] E. Jun, H. Liao, A. Savoy, L. Zeng, and G. Salvendy, "The design of future things", by D. A. Norman, basic books, New York, NY, USA. *Hum. Factors Man.*, 18: 480-481, 2008. <https://doi.org/10.1002/hfm.20127>
- [6] C. Vargas-Irwin, J.P. Donoghue "Automated spike sorting using density grid contour clustering and subtractive waveform decomposition", *Journal of Neuroscience Methods*, vol. 164, 1-18, 2007. <https://doi.org/10.1016/j.jneumeth.2007.03.025>
- [7] S. Krug, Don't Make Me Think! In K. Whitehouse & L. Brazieal (Eds.), *Don't Make Me Think! A Common Sense Approach to Web Usability* (Second). Berkeley: New Riders, 2006.
- [8] A. Dickinson, J. Arnott, S. Prior, and A. Dickinson, 'Methods for human-computer interaction research with older people', *Behav. Inf. Technol. is*, vol. 26, no. 4, 343–352, 2007.
- [9] P. Morville & L. Rosenfeld, "Information Architecture for the World Wide Web". In *IEEE Transactions on Professional Communication* (Third, Vol. 43), 2006. <https://doi.org/10.1109/tpc.2000.826425>
- [10] R. A. Virzi, "Refining the test phase of usability evaluation: How many subjects is enough?" *Special Issue: Measurement in human factors. Human Factors*, 34(4), 457–468, 1992.
- [11] E. Stevens, "What Is User Experience (UX) Design? Here's What You Need To Know", 2018 [Online]. Available: <https://careerfoundry.com/en/blog/ux-design/what-is-user-experience-ux-design-everything-you-need-to-know-to-get-started>. [Accessed: 8-May-2019]
- [12] J. Nielsen, "10 Usability Heuristics for User Interface Design", *Nielsen Norman Group*, 1995. [Online]. Available: <https://www.nngroup.com/articles/ten-usability-heuristics>. [Accessed: 14-Oct-2019].
- [13] Raven L. Veal. "How to define a User Persona" [Online]. Available: <https://careerfoundry.com/en/blog/ux-design/how-to-define-a-user-persona>. [Accessed: 29-Mar-2019]
- [14] R. Dam & T. Siang, "Personas: A Simple Introduction". *Interaction Design Foundation*. [Online]. Available: <https://www.interaction-design.org/literature/article/personas-why-and-how-you-should-use-them>. [Accessed: 28-Mar-2019]
- [15] N. Martins & T. Araújo, "The contribution of design in supporting the pregnancy process: The study of a mobile application towards a more informed relationship between the pregnant woman and the healthcare professional" in *ARTECH 2019 - Proceedings of the 9th International Conference on Digital and Interactive Arts, Braga, 2019*. Ed: ACM. pp. 191-197. <https://doi.org/10.1145/3359852.3359857>.
- [16] Interaction Design Foundation, "What are User Personas?", 2018. [Online]. Available: <https://www.interaction-design.org/literature/topics/user-personas>. [Accessed: 11-Jul-2019]
- [17] Guilizzoni, P. "What Are Wireframes?" *Balsamiq Wireframing Academy*, 2019. [Online]. Available: <https://balsamiq.com/learn/resources/articles/what-are-wireframes>. [Accessed: 22-Jul-2019]
- [18] J. Bernal, G. Flores, R. Vargas, "Diseñas: The web app for learning design terms for deaf students" in *DIGICOM 2019 - 3rd International Conference on Design and Digital Communication: Proceedings, Barcelos, 2019*. Ed: IPCA. pp. 315-324. <https://digicom.ipca.pt/docs/DIGICOM2019-Proceedings.pdf>. ISBN: 978-989-54489-5-1.
- [19] J. Tidwell, *Designing Interfaces*. In M. Treseler & R. Monaghan (Eds.), *Animal Genetics* (Second Edi, Vol. 39). Sebastopol: O'Reilly Media, 2011.
- [20] Toyota. *Novo Manual de Identidade Visual da Marca Toyota*. Portugal, 2012.
- [21] J. K. Liker, D. Meier, *The Toyota Way Fieldbook: A Practical Guide for Implementing Toyota's 4Ps*. New York, London: McGraw-Hill, 2006.
- [22] S. Kujala, R. Mugge, T. Miron-Shatz, "The role of expectations in service evaluation: A longitudinal study of a proximity mobile payment service" *International Journal of Human-Computer Studies*, Vol. 98, 51-61, 2017 <https://doi.org/10.1016/j.ijhcs.2016.09.011>.
- [23] C.P. Jansen, S.F. Donker, D.P. Brumby, A.L. Kunc "History and future of human-automation interaction", *International Journal of Human-Computer Studies*, vol. 131, 99-107, 2019. <https://doi.org/10.1016/j.ijhcs.2019.05.006>

Monte Carlo Estimation of Dose in Heterogeneous Phantom Around 6MV Medical Linear Accelerator

Zakaria Aitelcadi^{1,*}, Mohamed Reda Mesradi¹, Redouane El Baydaoui¹, Ahmed Bannan¹, Abdennacer Ait Ayoub², Kamal Saidi¹, Saad Elmadani¹

¹Hassan First University of Settat, High Institute of Health Sciences, Laboratory of Sciences and Health Technologies, 26000, Settat, Morocco

²Clinique spécialisée Menara, 40000, Marrakech, Morocco

ARTICLE INFO

Article history:

Received: 12 May, 2020

Accepted: 18 June, 2020

Online: 20 June, 2020

Keywords:

Monte Carlo

MLC

GATE

Artefacts

PTO

Heterogeneous

ECLIPSE

AAA

ABSTRACT

In this work, we completed a validation of the Varian Clinac IX equipped with the High Definition Multi-Leaf Collimator (HD 120 MLC) instead of the removable jaws, using GATE Monte Carlo Platform version 8.2. We validated the multileaf collimator (MLC) geometry by simulating two dosimetric functions (Percentage Depth Dose (PDD) and Dose Profile (DP)), for 6MV photon beam energy and different field sizes (3x3, 4x4, 6x6, 8x8, 10x10, 12x12, 15x15, and 20x20 cm²). We then compared the results with measurements realized with two detectors, namely the cylindrical ionization chamber and the micro-diode PTW silicon. By applying the Relative Dose Difference method (RDD), we noted a less than 2% and 1% agreement for the field sizes (10x10, 12x12, 15x15, 20x20 cm²) and (3x3, 4x4, 6x6, 8x8 cm²) respectively. Moreover, to evaluate the relevance of Monte Carlo method in a heterogeneous media, particularly in small field sizes (1x1, 2x2, 3x3 cm²), we have simulated three clinical studies based on the Physical Test Objects (PTOs) that are the equivalent slabs of lung and bone included in a water phantom. We noticed that the simulated PDDs exhibit two significant irregularities in the interface between water and lung. To eliminate these phenomena, we have used the "setMaxStepSizeInRegion" parameter implemented in GATE. We also noticed an important difference of 5% corresponding to the small field sizes, between homogeneous and heterogeneous simulated PDDs. We used the RDD method in this case as well. Moreover, we observed a difference between 1-4% between the simulated PDDs and the calculated ones by ECLIPSE Treatment Planning System (TPS). These results indicate that GATE (8.2) is useful in dosimetry with heterogeneous situations as well such as bone and lung.

1. Introduction

In clinical radiotherapy, most TPS are calibrated in a homogeneous media with densities equal to 1. However, some organs have strong heterogeneities such as bone and lung. Hence a better precision requires a corrective dose in conventional TPS. In this context, Monte Carlo simulations present a real alternative allowing enhanced precision related to the transport of high energy photons, particularly in heterogeneous media. However, complex MC simulations require a great amount of computing resources and are time-consuming. Consequently, the optimization of the computation time is necessary. In our study,

we used lung and bone equivalent slabs included in a water phantom as Physical Test Objects (PTOs).

Moreover, modern radiotherapy also uses complex beam shapes. For this purpose, we modeled a Varian Clinac IX 6MV photon beam energy with the High Definition Multi-Leaf Collimator (HD 120 MLC). This instrument can hold up to 120 pairs of leaves that move independently to allow the output of a complex beam shape. In practice, there are three types of MLCs, namely type A (e.g. Scanditronix and Siemens) [1], type B (e.g. Elekta) [2] and type C (e.g. Varian) [3]. The three are distinguished by their leaf's size, speed of movement, and the

*Corresponding Author: Zakaria Aitelcadi, z.aitelcadi@gmail.com

transmission factor related to their arrangement and their geometry.

In this study, we have used the PTOs to evaluate the relevance of the MC method in the case of small radiation fields used in the context of small tumors. We realized this objective in two parts. In the first part, we have modeled a Varian Clinac IX 6MV photon beam energy to take into account the MLC based on our previous work [4] and using the geometric data provided by the manufacturer [3]. Thus, we compared the simulated dosimetric functions (PDD and DP) for different field sizes (3x3, 4x4, 6x6, 8x8, 10x10, 12x12, 15x15 and 20x20 cm²) to the measured ones using the RDD method [5]. In the second part, we have simulated three PTOs geometries (water+bone, water+lung, and water+bone+lung). Then we compared the PDDs of the heterogeneous media with the homogeneous ones. We also compared the simulated PDDs with the ones calculated by ECLIPSE TPS, based on the Anisotropic Analytic Algorithm (AAA). Finally, we've been interested in the optimization of the artefact phenomenon at the interfaces.

2. Material and Methods

2.1. Measurements including the HD 120 MLC

In this study, the Varian Clinac IX 6MV photon beam we used is equipped with the High-Definition Multi-Leaf Collimator (HD 120 MLC). PDD measurements were made at 100 cm of the Source Surface Distance (SSD), with a pitch of 0.1 mm in a depth between 0 and 5 cm in water and 1 mm for depths greater than 5 cm. In the case of DP, we chose 0.1 mm as a pitch in the penumbra region. In measurements with fields greater than or equal to 4x4cm² a cylindrical ionization chamber (Exradin type A28), with a volume of 0.125 cm³ was used, while for fields less than 4x4cm², measurements were carried out using a micro-diode silicon detector PTW with a volume of 0.03 cm³ placed in a Doseview standard 3D solid water phantom of Standard Imaging.

2.2. HD 120 MLC modelling

The geometry and components materials of the HD 120 MLC have been implemented in GATE (V8.2) code using the manufacturer's data [3]. It is formed by two blocks (A and B) that can hold 60 independent leaves oriented according to the Y-axis. Each block holds 28 external leaves «half leaves» (0.5 cm width) and 32 internal «quarter leaves» (0.25 cm width). They both are placed at 100 cm from the source [6,7]. Furthermore, the ends of the leaves are rounded with a 16 cm radius, their thickness is 6.9 cm and they are spaced from each other with a distance of 0.0047 cm.

The 32 internal leaves are positioned according to an alternative pattern (Figure 2): a drop with its fine end oriented towards the source (or "Target leaf") then its neighbor whose fine end is oriented this time towards the isocenter of the accelerator (or «Isocenter leaf»), these two types of leaves differ by the distance tongue and groove which is worth 0.1 and 0.01 cm respectively to create a vertical play between the tongue of a leaf and the groove of the adjacent leaf [8].

Besides, in GATE (V8.2), we introduced four types of leaves (quarter isocenter, quarter target, half isocenter, and half target) (Figure 2). Which we located them at the origin of the marker placed in the entrance of the photon target. Then the leaves were repeated with the possibility of rotation and translation around their center and Y axis, and X, Y, and Z respectively. Figure 3 shows the GATE model of the Varian Clinac IX, including the HD 120 MLC [8].

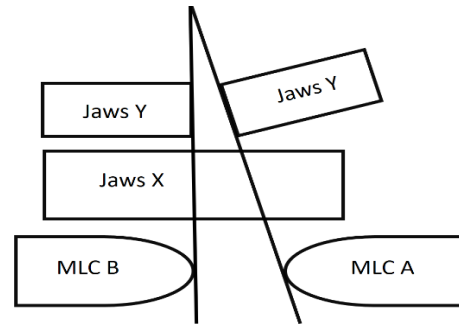


Figure 1: Varian Clinac MLC illustration

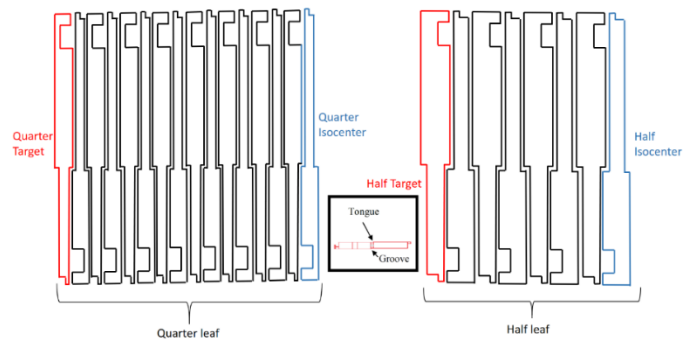


Figure 2: Schematic presentation of the HD 120 MLC, with Target and Isocenter leaf for each type

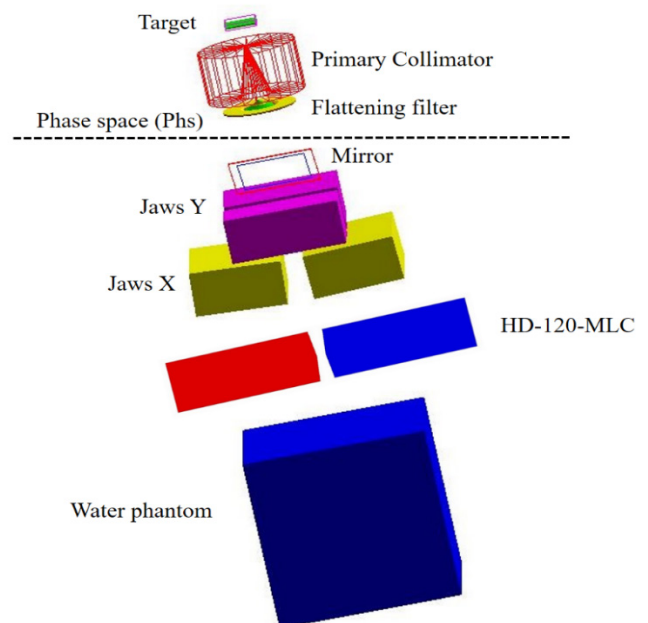


Figure 3: Varian Clinac IX accelerator head GATE modeled including the HD 120 MLC

On the other hand, in GATE (V8.2), the dosimetric functions (PDD and DP) corresponding to the different field sizes (3x3, 4x4, 6x6, 8x8, 10x10, 12x12, 15x15 and 20x20 cm²) were realized using the "DOSE ACTOR". PDD was normalized at depth (D_{max}) where the deposited dose is maximal. We compared the simulation results with those measured using the RDD method [6]. The latter consists of evaluating the relative dose difference between an experimental value and a theoretical reference value using equation 1. The dose difference should be less than 3% in the build-up region and less than 1% for most depths ranging from maximum dose depth (D_{max}) to 30 cm. In equation 1, D_c is the calculated absorbed dose, and D_m is the measured dose (reference). Indeed, to obtain good statistics 9.10^9 particles were generated from a phase space (Phs) previously used as a source [4] and directed into a water phantom of the same size used in measurements.

$$RDD (\%) = 100 * \frac{D_c - D_m}{D_m} \quad (1)$$

Table 1: Material characteristics for lung and bone equivalent slabs

Materials	Density (g/cm ³)	Width (cm)	Composition (%)
Lung slab	0.31	10	H (8.31), C (60.08), N (2.71), O (23.04), Mg(4.8), Cl(1.02)
Bone slab	1.91	5	H (3.30), C (25.37), N (0.91), O (35.28), Mg(3.36), P (8.82), Cl(0.03), Ca(22.91)

2.3. Heterogeneity Study

In this work, we conducted a study of heterogeneity using the two PTOs (lung and bone equivalent slabs). The three heterogeneous geometries studied are illustrated in Figure 4. The Phantom 1 includes a lung equivalent slab with a size of (30 x 30 x 10 cm³) placed at 5 cm from the entrance. The Phantom 2 contains a slab bone with a size of (30 x 30 x 5 cm³) placed at 5 cm. The Phantom 3 includes two slabs: bone and lung located at 5 cm and 10 cm respectively.

To evaluate the ability of GATE to predict the dose distribution in a heterogeneous media, we compared the simulated PDDs with the three phantoms with a homogenous one. This concern six different fields sizes (1x1, 2x2, 3x3, 10x10, 15x15 and 20x20 cm²). In GATE, the composition and the density of lung and bone equivalent slabs were given by the manufacturer (Table 1). We compared GATE results with the ones obtained with the ECLIPSE TPS by applying the RDD method. We note that we used the same GATE geometry in ECLIPSE and that calculations were performed using the AAA algorithm [9,10].

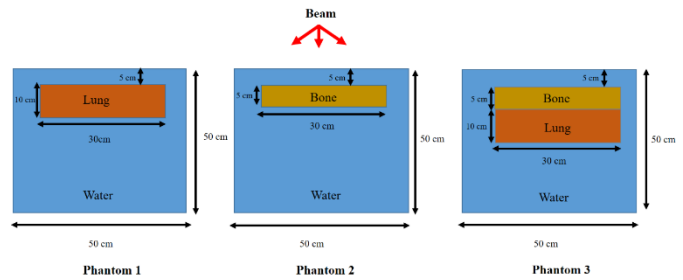


Figure 4: Two PTOs geometries using for three different studies

3. Results and Discussion

3.1. Static validation of the HD 120 MLC

Figures 5 and 6 show the simulated PDDs and DPs compared to those measured for the field sizes (3x3, 4x4, 6x6, 8x8, 10x10, 12x12, 15x15, and 20x20 cm²). Table 2 illustrates the results of this comparison. We note that the PDD results exhibit an agreement of less than 2% for the most points, while with the DP results the differences are around 1% in the build-up region and 2% outside. These simulations indicate that the accuracy with MLC is better than with the removable jaws used in our previous work [4].

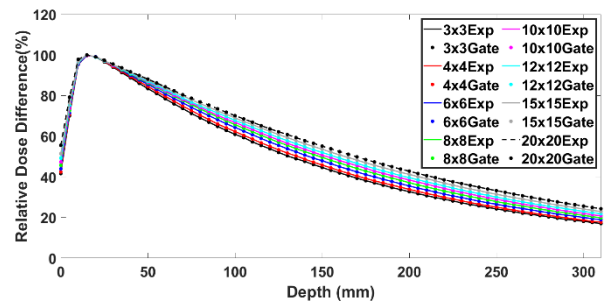


Figure 5: PDDs defined by the 120 HD MLC for different field sizes

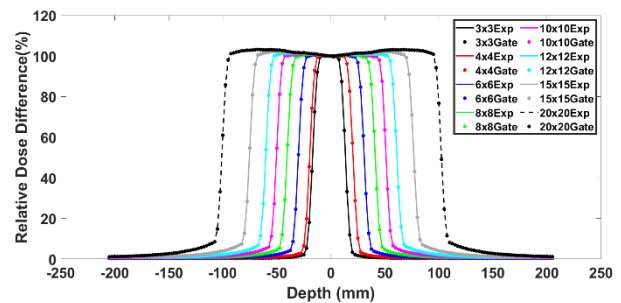


Figure 6: DPs defined by the 120 HD MLC for different field sizes

3.2. Heterogeneity study

3.2.1. Artefact phenomenon

In the literature, few studies were interested in the artefact phenomenon observed in MC simulations at the boundary between two biological matters [11], for example between water and lung. For this purpose, we simulated the PDD using phantom 1. The results indicated two significant irregularities in the

interface between water and lung (Figure 7). The "setMaxStepSizeInRegion" is the key factor for this phenomenon provided in GATE. It is defined as the maximal step size of charged particles. Thus, we performed simulations by varying this parameter within the range of 10-50 μm . Table 3 shows the calculation time for each value corresponding to a step of 10 μm . Adjusting the "setMaxStepSizeInRegion" with the recommended cutoff value [4], the two artefacts are gone.

Table 2: Average RDD between (PDD & DP) calculated and measured ones for different field sizes

Field sizes (cm ²)	RDD (%)	
	PDD	DP
3x3	0.1666	0.1952
4x4	0.1836	0.1936
6x6	0.258	0.4256
8x8	0.3839	0.8934
10x10	0.5289	0.9612
12x12	0.7145	1.0142
15x15	0.9236	1.0958
20x20	1.0147	1.3541

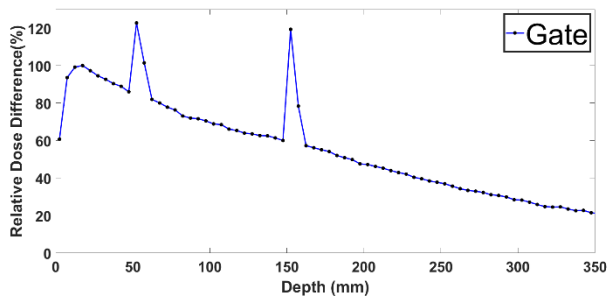


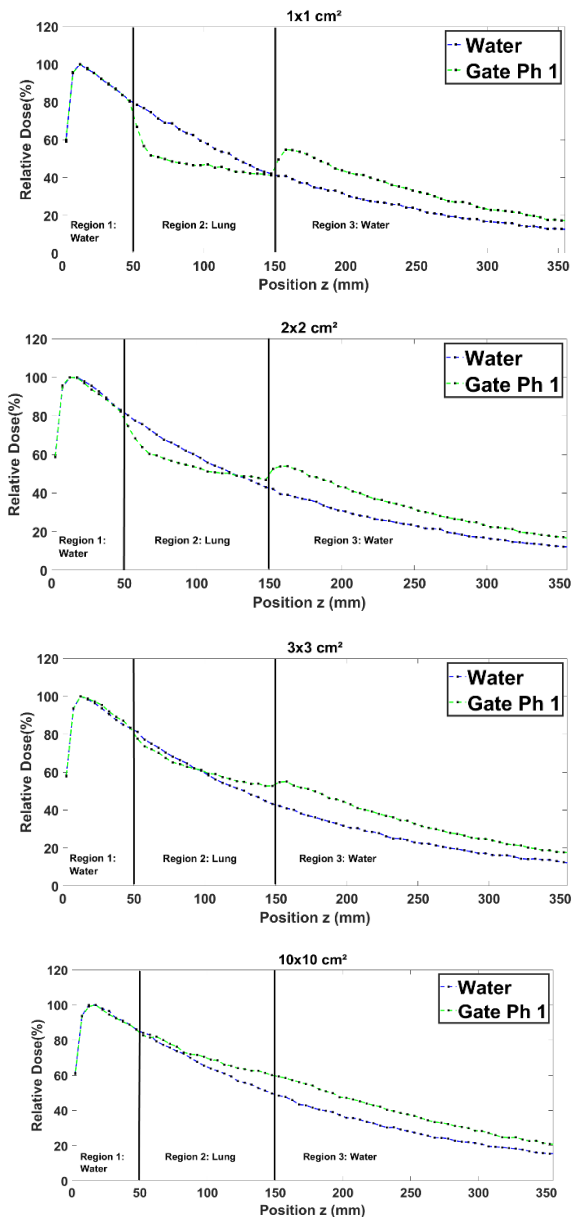
Figure 7: Artefact phenomenon in the boundaries between Water and Lung equivalent slab for 10x10 cm² field size

Table 3: CPU timing for "setMaxStepSizeInRegion" values

"setMaxStepSizeInRegion" value (μm)	CPU timing (h)
10	30
20	20
30	48
40	53
50	59

3.2.2. Heterogeneity study compared to homogenous one

In Figure 8, the fact that the lung has a weaker density, this doesn't lead to any change in the PDD in region 1 therefore, the two PDD curves (water and water-lung) are almost identical for all field sizes studied in this region. In region 2, the photon's attenuation is weaker, and the fact that in small fields (1x1, 2x2, and 3x3 cm²) there is almost no lateral electron equilibrium, lung PDDs are lower than the one in water. This electronic disequilibrium is due to the Compton effect [12]. Indeed, when the electron range produced by the Compton Effect is half of the field size, the electrons produced will transfer their energies outside the radiation field from where the electronic balance is lost. However, for a large field sizes (10x10, 15x15 and 20x20 cm²) the photons' attenuation and the lateral electron equilibrium becomes significant. This is due to the field sizes increase, consequently, lung PDDs become relatively higher than the ones in the water. Indeed, in region 3, the fact that for all field sizes, the PDDs in the lung are relatively higher than ones in the water, is mainly due to the lower density of lung in region 2.



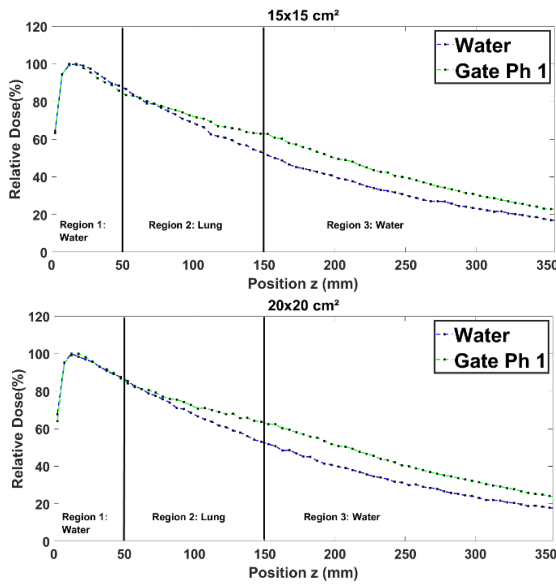


Figure 8: Comparison between the PDD in a homogenous phantom with phantom 1 for 1x1, 2x2, 3x3, 10x10, 15x15 and 20x20 cm² field sizes.

Figure 9 shows that in region 1, for all field sizes the presence of bone seems to not affect the PDDs curves. In region 2, although bone has a higher attenuation coefficient, the related PDDs seem to be relatively weaker than the ones in the water. This is observed for all field sizes, owing to the Compton scattering effect [12]. Indeed, in region 3, the higher attenuation coefficient of bone in region 2, makes that PDDs in bone are still relatively weaker than the ones in water for all field sizes.

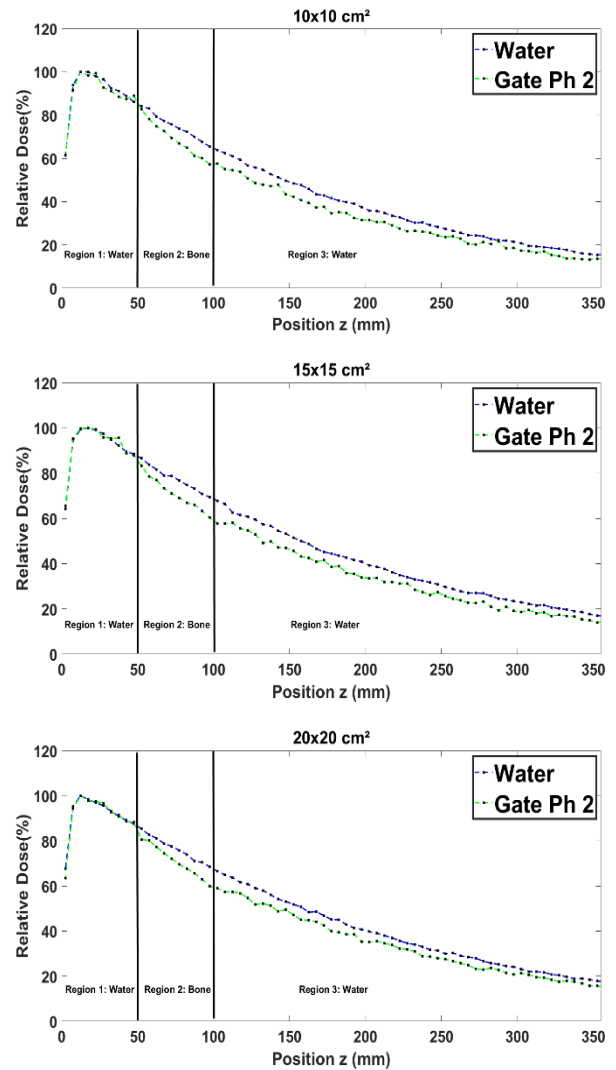
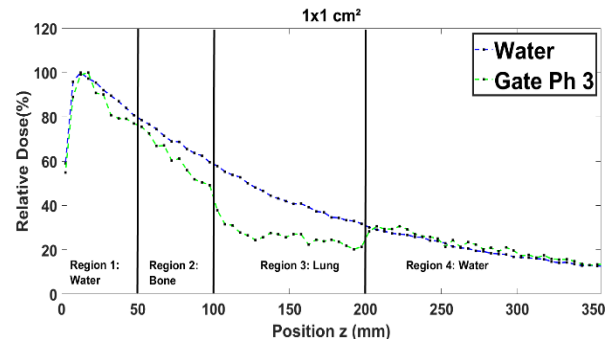
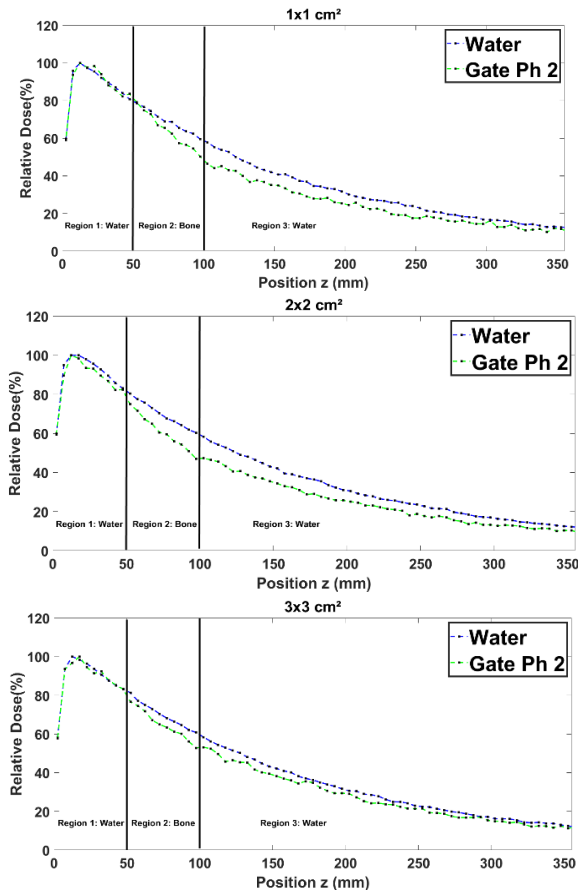


Figure 9: Comparison between the PDD in a homogenous phantom with phantom 2 for 1x1, 2x2, 3x3, 10x10, 15x15 and 20x20 cm² field sizes.

Figure 10 shows that in region 1 the two PDDs curves are almost identical, despite the presence of the two bone and lung slabs. In regions 2 and 3, the related PDDs seem to be relatively weaker than the ones in the water, owing to the Compton scattering effect and the low density of lung. This is observed for all field sizes. In region (4) the fact that for all field sizes, the PDDs in the presence of lung and bone are relatively higher than the ones in water alone, results primarily to the low density of lung in region 3.



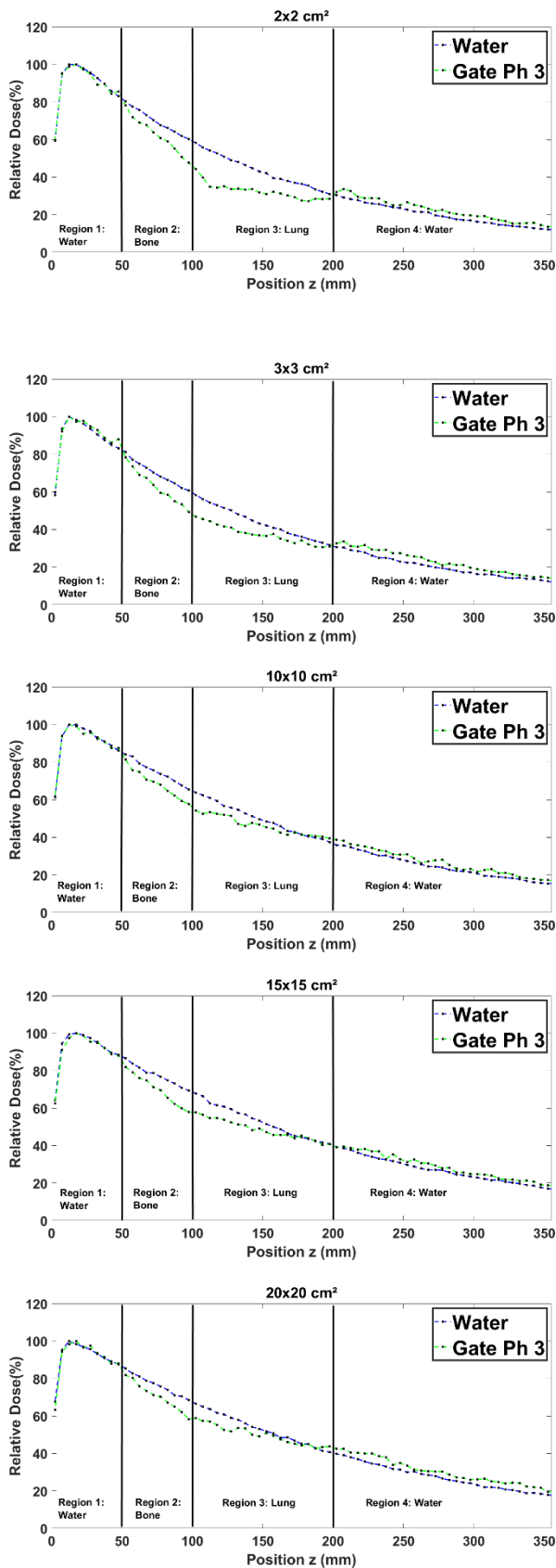


Figure 10: Comparison between the PDD in a homogenous phantom with phantom 3 for 1x1, 2x2, 3x3, 10x10, 15x15 and 20x20 cm² field sizes.

3.2.3. Heterogeneity comparison study between GATE and ECLIPSE (AAA)

Figure 11 shows that the PDDs obtained by ECLIPSE and GATE from Phantom 1 are in general closely similar with a difference less than 1%, except for the 1x1 cm² field where ECLIPSE PDD exceeds GATE by 4.02%. This can be explained by the fact that ECLIPSE does not take into account the lateral electronic equilibrium in lung slab for very small fields. Table 4 presents the average RDD of the PDDs calculated by ECLIPSE and simulated by GATE for Phantom 1.

Table 4: Average RDD between GATE and ECLIPSE (AAA) PDDs for different field sizes in phantom 1

Field sizes (cm ²)	RDD (%)
1x1	4.02
2x2	0.83
3x3	1.27
10x10	1.68
15x15	0.53
20x20	0.63

Figure 12 shows that in region 1, the PDDs obtained from Phantom 2, are in general closely similar for all field sizes. In region 2, we note a difference of 1 to 5% explained by the fact that ECLIPSE overestimates the energy deposited by secondary electrons in bone slab [13]. Table 5 presents the average RDD of the PDDs for Phantom 2.

Table 5: Average RDD between GATE and ECLIPSE (AAA) PDDs for different field sizes in phantom 2

Field sizes (cm ²)	RDD (%)
1x1	0.26
2x2	1.31
3x3	2.31
10x10	4.08
15x15	4.03
20x20	3.75

Figure 13 shows that in region 1, the PDDs obtained from Phantom 2, are in general closely similar for all field sizes. In region 2, for small field sizes, ECLIPSE PDDs exceed the GATE ones by more than 4%. This is because ECLIPSE overestimates the energy deposited by secondary electrons in bone slab [13]. On the other hand, for field sizes greater than or equal to (10x10 cm²) the difference is 2%, owing to the presence also of lung in region 3. Table 6 presents the average difference for PDDs calculated and measured by applying the RDD method.

Table 6: Average RDD between GATE and ECLIPSE (AAA) PDDs for different field sizes in phantom 3

Field sizes (cm ²)	RDD (%)
1x1	4.23
2x2	4.59
3x3	3.61
10x10	2.00
15x15	1.91
20x20	1.74

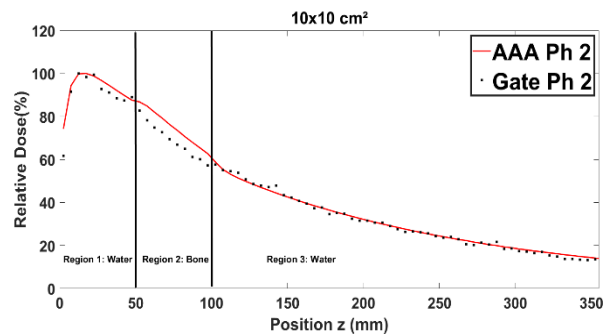
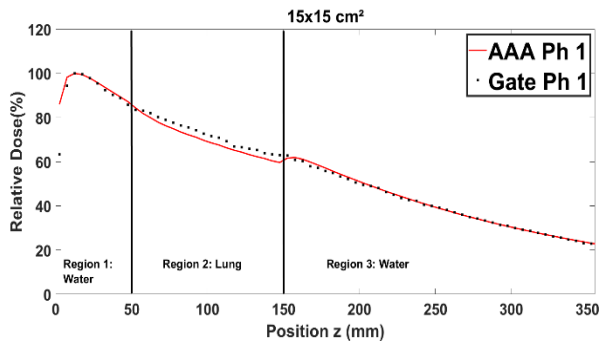
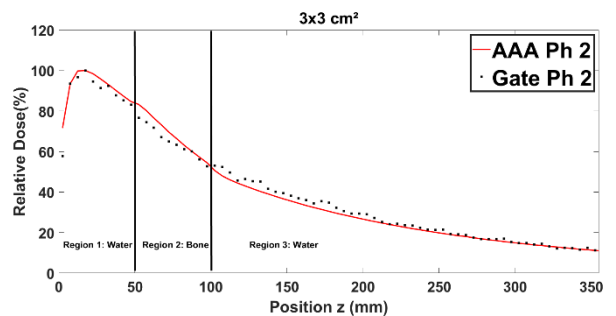
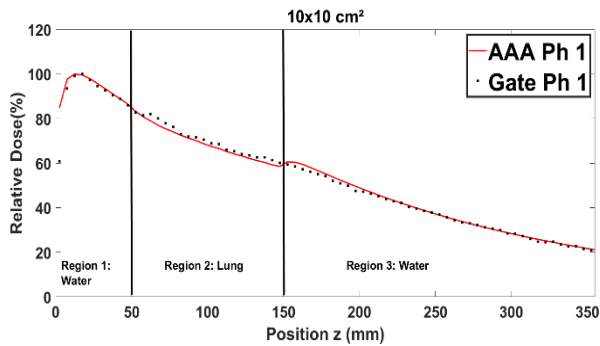
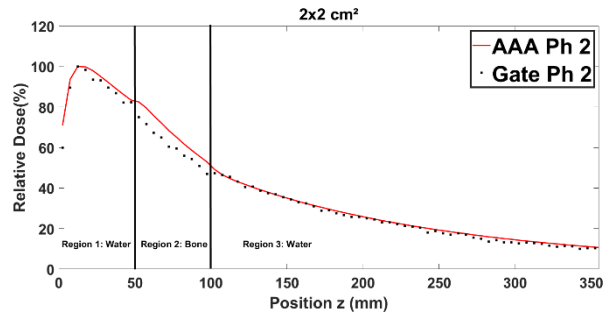
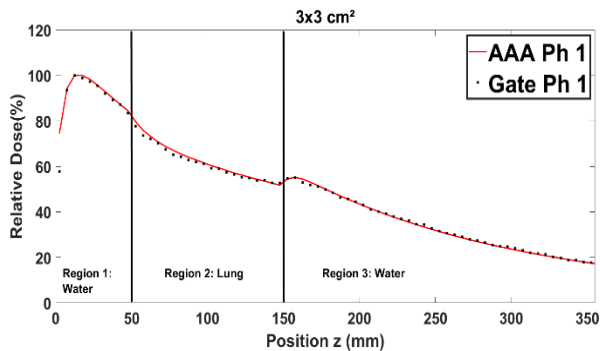
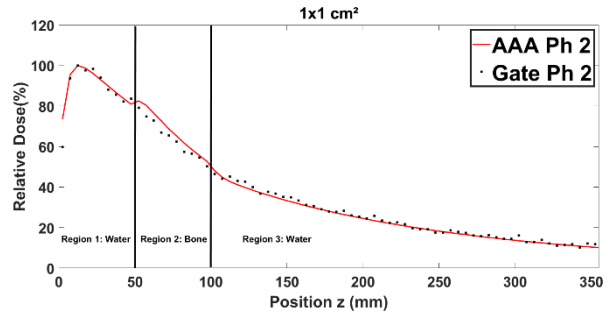
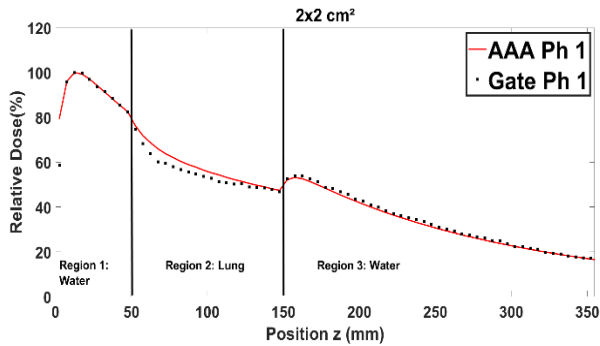
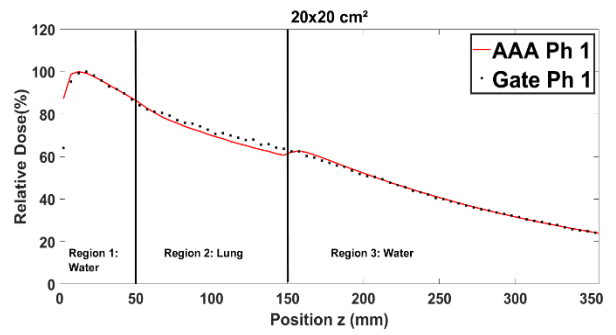
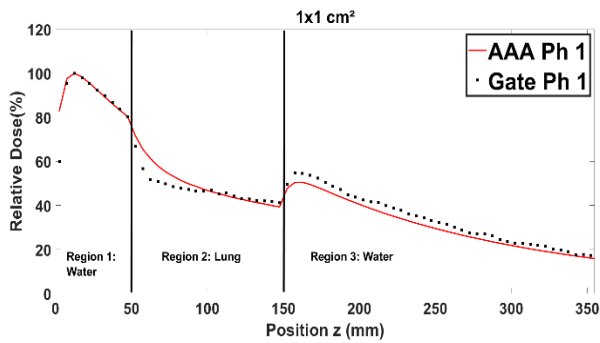


Figure 11: Comparison between GATE and ECLIPSE (AAA) PDDs in phantom 1 for 1x1, 2x2, 3x3, 12x12, 15x15 and 20x20 cm² field sizes.

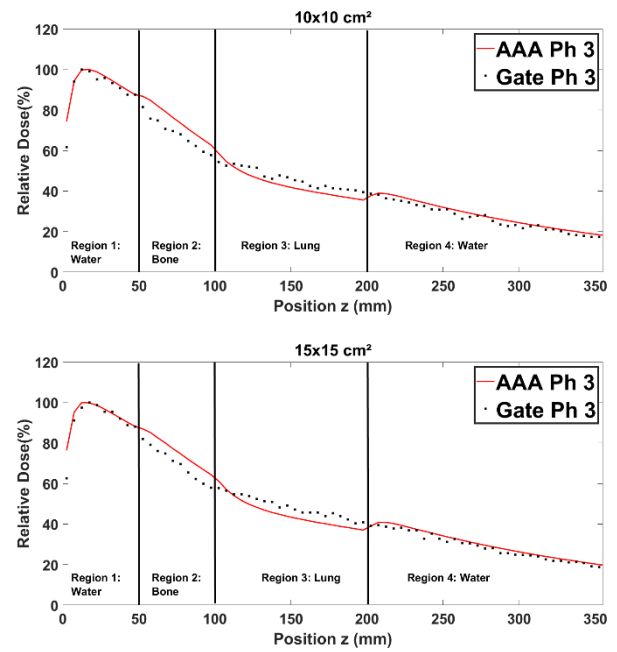
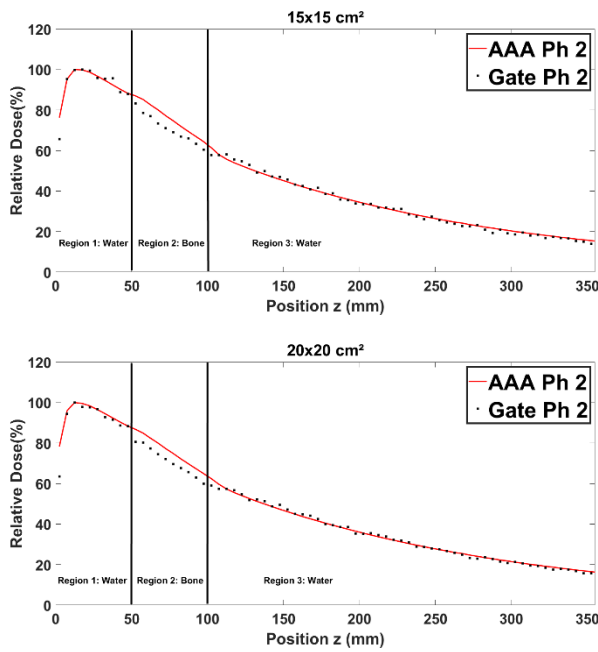


Figure 12: Comparison between GATE and ECLIPSE (AAA) PDDs in phantom 2 for 1x1, 2x2, 3x3, 12x12, 15x15 and 20x20 cm² field sizes.

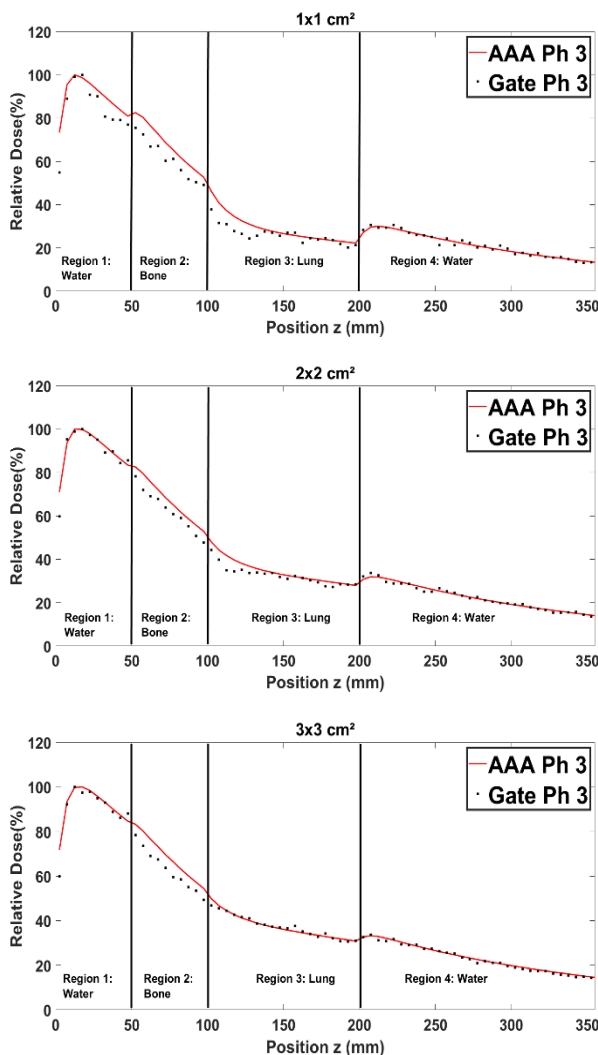


Figure 13: Comparison between GATE and ECLIPSE (AAA) PDDs in phantom 3 for 1x1, 2x2, 3x3, 12x12, 15x15 and 20x20 cm² field sizes.

4. Conclusion

In addition to our previous work, in this study, we successfully used the up-to-date version of GATE 8.2 to simulate the High Definition Multi-Leaf Collimator (HD 120 MLC) using the manufacturers'. We performed the MLC validation by comparing the dosimetric functions (PDD and DP) measured in a water phantom with those simulated for different field sizes using the relative dose difference method (RDD). Results show an agreement of less than 2% between simulated and measured functions. On the other hand, we demonstrated by using three studies based on the Physical Test Object (PTOs), the capacity of GATE 8.2 code to reproduce the dosimetric function in heterogeneous media such as lung and bone. Moreover, we showed that GATE exceeds ECLIPSE in the assessment of dose in heterogeneous media, since the latter does not take into account the lateral electronic disequilibrium. Indeed, the optimization of the "setMaxStepSizeInRegion" parameter in GATE led to eliminate the phenomenon of artefact in the interface between lung and bone. In conclusion, the static validation of our MLC 120 HD model, as well as the results of the dose distribution in heterogeneous media will lead us in future research for a dynamic

validation to test the feasibility of the clinical application of the Monte Carlo simulations.

Conflict of Interest

We declare that we have no conflict of interest.

References

- [1] AAPM. Basic applications of multileaf collimators. report of the AAPM radiation therapy committee Task Group No 50 AAPM Report No 72 (Madison, WI: American Institute of Physics by Medical Physics Publishing). 2001.
- [2] Heath, E, Seuntjens, J. Development and validation of a BEAMnrc component module for accurate Monte Carlo modelling of the Varian dynamic Millennium multileaf collimator. *Phys Med Biol* 2003;48(24):4045-4063. <https://doi.org/10.1088/0031-9155/48/24/004>
- [3] Michael K. Fix, Werner Volken, Daniel Frei, Daniel Frauchiger, Ernst J. Born and Peter Manser, Monte Carlo implementation, validation, and characterization of a 120 leaf MLC. *Med. Phys.*, 38(10) : 5311-5320, 2011. <https://doi.org/10.1118/1.3626485>
- [4] Z. Aitelcadi, A. Bannan, R. El baydaoui, MR. Mesradi, A. Halimi, S. Elmadani, Feasibility of external radiotherapy dose estimation in homogenous phantom using monte carlo modeling. *JATIT*, 98(8) :1151-1162, 2020.
- [5] Cho S.H., Vassiliev O.N., Lee S., Liu H.H., Ibbott G.S., and Mohan R. Reference photon dosimetry data and reference phase space data for the 6 mv photon beam from varian clinac 2100 serie linear accelerators. *Med. Phys.*, 32(1) :137-148, 2005. <https://doi.org/10.1118/1.1829172>
- [6] Huq, MS, Das, IJ, Steinberg, T, Galvin, JM. A dosimetric comparison of various multileaf collimators. *Phys Med Biol* 2002;47(12):N159-170. <https://doi.org/10.1088/0031-9155/47/12/401>
- [7] J. V. Siebers, P. J. Keall, J. O. Kim, and R. Mohan, "A method for photon beam Monte Carlo multileaf collimator particle transport," *Phys. Med. Biol.* 47(17), 3225 (2002). <https://doi.org/10.1088/0031-9155/47/17/312>
- [8] C. Borges, M. Zarza-Moreno, E. Heath, N. Teixeira and P. Vaz, Monte Carlo modeling and simulations of the High Definition (HD120) micro MLC and validation against measurements for a 6MV beam. *Med.Phys.*, 39(1), 2012. <https://doi.org/10.1118/1.3671935>
- [9] Fogliata, A, Vanetti, E, Albers, D, et al. On the dosimetric behaviour of photon dose calculation algorithms in the presence of simple geometric heterogeneities: comparison with Monte Carlo calculations. *Phys Med Biol* 2007;52(5):1363-1385. <https://doi.org/10.1088/0031-9155/52/5/011>
- [10] Panettieri, V, Barsoum, P, Westermark, M, Brualla, L, Lax, I. AAA and PBC calculation accuracy in the surface build-up region in tangential beam treatments. Phantom and breast case study with the Monte Carlo code PENELOPE. *Radiother Oncol* 2009;93(1):94-101. <https://doi.org/10.1016/j.radonc.2009.05.010>
- [11] Poon, E., Verhaegen, F., 2005. Accuracy of the photon and electron physics in GEANT4 for radiotherapy applications. *J. Med. Phys.*, 2005, 32(6), 1696-1711. <https://doi.org/10.1118/1.1895796>
- [12] L.A.R. da Rosa, S.C. Cardoso, L.T. Campos, V.G.L. Alves, D.V.S. Batista, A.Facure, Percentage depth dose evaluation in heterogeneous media using thermoluminescent dosimetry. *JOURNAL OF APPLIED CLINICAL MEDICAL PHYSICS*, 11(1), 2010, 117-127. <https://doi.org/10.1120/jacmp.v11i1.2947>
- [13] Cardoso SC, Alves VGL, da Rosa LAR, Campos LT, Batista DVS, Facure A, Monte Carlo Simulation of Bony Heterogeneity Effects on Dose Profile for Small Irradiation Field in Radiotherapy. *PLoS ONE*. 2010; 5(5): e10466. <https://doi.org/10.1371/journal.pone.0010466>

5G mm-wave Band pHEMT VCO with Ultralow PN

Abdelhafid Es-Saqy^{*1}, Maryam Abata¹, Mohammed Fattah², Said Mazer¹, Mahmoud Mehdi³, Moulhime El Bekkali¹, Catherine Algani⁴

¹AIDSESL, Sidi Mohamed Ben Abdellah University, 30050, Morocco

²EST, My Ismail University, 50050, Morocco

³ML, Lebanese University, 6573/14, Lebanon

⁴ESYCOM, Gustave Eiffel University, CNAM, 75003, France

ARTICLE INFO

Article history:

Received: 21 May, 2020

Accepted: 07 June, 2020

Online: 20 June, 2020

Keywords:

pHEMT VCO

Local Oscillator

Low Phase Noise

MMIC VCO

5G VCO

ABSTRACT

Oscillator phase noise (PN) has a strong impact on the spectral purity of the RF signal in wireless systems and is, therefore, a main challenge when designing a local oscillator. In this paper, we propose a new approach for designing a low PN oscillator based on the Time-Invariant Linear Model of phase noise. It leads on voltage-controlled oscillator (VCO) simulated good performances: a low phase noise (PN) near -123.2 dBc/Hz@1MHz offset from the carrier, an output power of 3.26 dBm, and an output signal frequency ranging from 27.98 GHz to 29.67 GHz. Low power-consumption (51mW) and small size (0.237 mm²) benefit from MMIC UMS foundry (United Monolithic Semiconductors) and 0.15 μ m-pHEMT GaAs technology.

1. Introduction

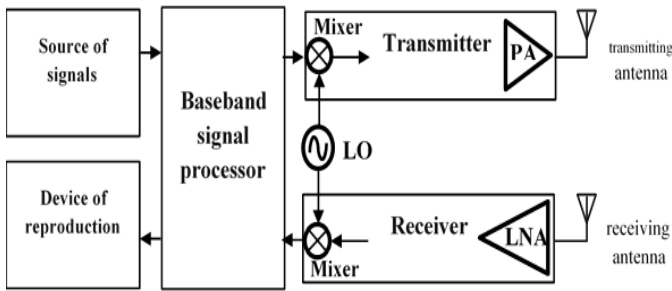
The development of fifth-generation mobile networks (5G) is currently of great interest for telecommunication firms. This new generation will be the gateway to self-driving cars, virtual and augmented reality, Internet of Things, and other future technologies [1], therefore, it requires large contiguous blocks of spectrum.

The microwave frequency spectrum has become fully occupied with time. Therefore, in order to fulfill the explosive increase in broadband transmission requirements, new frequency allocation is needed for 5G applications. This new generation of mobile communications should allow download speeds up to 10Gbit/s, with a latency less than 1ms [2]. Fortunately, there are large frequency bands in the mmWave range that are not devoted to any other application [3]. Delegates at the World Radiocommunication Conference 2019 (WRC-19) identified five additional mm-Wave frequency bands for the deployment of 5G networks: [24.25-27.5 GHz], [37-43.5 GHz], [45.5-47 GHz], [47.2-48.2] and [66-71 GHz] [4].

In order to transpose the IF (Intermediate Frequency) signal to RF (Radio Frequency), or vice versa, wireless communication systems (Figure 1) require one or more local oscillators (LO). Thus, the local oscillator has strong impact on the spectral purity of the transmitted or received signal. Therefore, the design of LO with high purity is a major challenge. Many researchers have addressed the issue of phase noise in oscillators and have suggested several solutions to overcome this limitation [5-7], however, the impact of the proposed methods remains very limited, in term of phase noise performance, and does not meet the requirements of the new generation 5G.

In reference [5], the authors proposed an innovating architecture using two coupled VCOs and exhibiting -121.4 dBc/Hz @1MHz frequency offset Phase Noise for a chip area of 1.55 mm². In reference [6] a phase noise varying from -100 to -96 dBc/Hz@100kHz has been achieved, while the VCO core equals to 1 mm². And finally, the authors of reference [7] presented an architecture composed of four Colpitts VCOs and three millimeter-wave selectors, the phase noise varies from -100.7 dBc/Hz to -95.3 @ 1 MHz offset, while the chip size exceeds 3.7 mm².

*Corresponding Author: Abdelhafid Es-Saqy, abdelhafid.essaqy@usmba.ac.ma

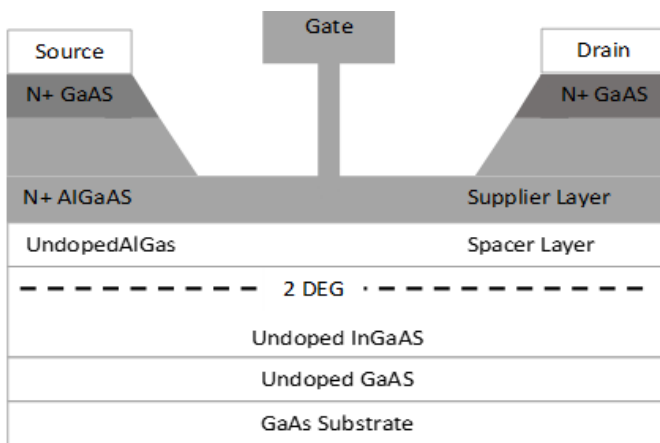


In this paper, we propose a new approach for designing a low PN oscillator presented previously in reference [8]. It is based on the Time-Invariant Linear Model (TILM) of PN [6, 8], and applied to design a MM-wave band VCO. Simulations show that the proposed architecture exhibits an extremely low phase noise level associated to a small size, compared to the ones published recently in the literature.

This paper is organized as follows. In section 2, we introduce the device technology. In section 3, a phase noise analysis is presented in order to extract an electrical model. The fourth section is dedicated to the proposed VCO architecture, while the fifth section presents the simulation results. Then we conclude.

2. GaAs pHEMT technology: PH15 UMS process

In last years, the development of III-V semiconductor materials for microwave devices has been sustained. Among these III-V semiconductors, gallium arsenide (GaAs) is the precursor, with better electronic and physical properties than silicon such as a higher electron mobility. GaAs Pseudomorphic High Electron Mobility Transistors (pHEMT) (Figure 2) are currently base of MMIC (Monolithic microwave integrated circuit) circuits thanks to mature technologies.



Due to the development of these transistors, local oscillators [9], power amplifiers [10, 11], mixers [12], and frequency multipliers [13] have shown improved performances. The 5G mm-wave VCO circuit, presented in this paper, is designed using the commercial UMS foundry (PH15 process). PH15 technology is based on a classical pseudomorphic AlGaAs/InGaAs/GaAs HEMT structure shown in Figure 2. This 5G mm-wave process features typically 110 GHz- f_T cut-off frequency, 640mS/mm peak transconductance, 220 mA/mm maximum drain current, -0.7 V

pinch-off voltage and beyond 4.5 V gate-drain breakdown voltage. As the gate length is 0.15 μm , we take 30 μm -gate width. The PH15 process includes two metal interconnect layers, TaN resistors of 30 Ω/square , SiN MIM capacitors of 330 pF/mm², airbridge and via holes. Table 1 shows typical datas of the pHEMT transistor [14].

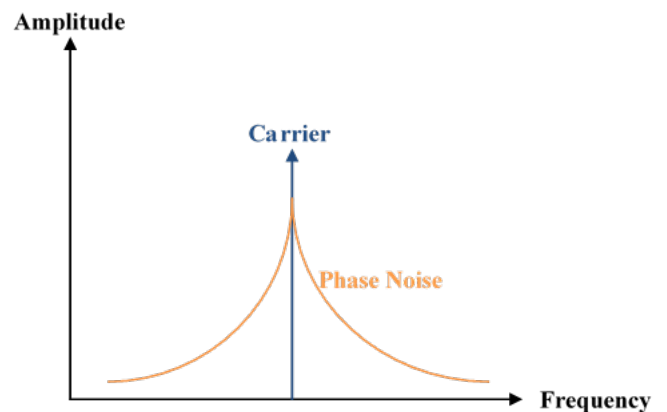
Table 1: Typical data of the pHEMT Transistor

Parameter	Value
Power Density	300 mW/mm
Gate Length	0.15 μm
Ids (gm max)	220 mA/mm
Ids sat	550 mA/mm
V _{BDS}	> 4.5 V
Cut off frequency	110 GHz
V _{pinch}	-0.7 V
Gm max	640 mS/mm
Noise	0.5 dB @ 10 GHz 1.9 dB @ 60 GHz
Gain	14 dB @ 10 GHz 6 dB @ 60 GHz

3. Voltage Controlled Oscillator Phase Noise

In wireless communication systems, a pure sine wave at a single frequency is an ideal case, i.e., a Dirac delta function at a single frequency. However, additive noise from propagation environment and circuits modulates the oscillator, introducing frequency fluctuations. These fluctuations spread the signal power at adjacent frequencies of the carrier frequency resulting in noise sidebands (Figure 3), generally named phase noise since it can be, in the time domain, represented as a random variation of the phase.

Recent wireless communication systems require radio frequency carriers of very high spectral purity. The quality of the VCO becomes a determining factor in the quality of the entire system. It would be very difficult to transmit a signal at very high frequencies and with complex modulations without a very low PN VCO.



The up-converted amplitude noise is a critical source to phase noise, but there are other contributors of phase noise in a LO [15], we cite for example: flicker noise, 1/f noise, thermal noise and shot noise. In order to design a LO that combines both low PN

and small size, it is mandatory to find the relationship between PN and the circuit parameters. This relationship is given by the equation (1) [6]:

$$L(f_0, \Delta f) = \frac{2kT}{c} \cdot \frac{f_0}{Q} \cdot \frac{1+F}{A_0^2} \cdot \frac{1}{\Delta f^2} \quad (1)$$

where $L(f_0, \Delta f)$ is the single sideband phase noise (SSPN) at Δf offset frequency from the carrier f_0 , k is Boltzmann constant, T is the absolute temperature, C is the resonator capacitance value, Q is the resonator quality factor, F is noise factor and A_0 is the output signal amplitude.

From equation 1, PN of the LO is reduced when increasing the quality factor of the resonator circuit, in order to improve the amplitude of the output signal [6] or increasing the value of the capacitance of the circuit. In this work, we focused on a capacitance increase. Therefore, a capacity $C1$ has been added, in order to enhance the capacitance value of the resonator circuit and thus improve the PN. However, as the oscillation frequency depends on the circuit capacitance, the inductance value $L1$ must be reduced to keep the oscillation frequency value (figure 5).

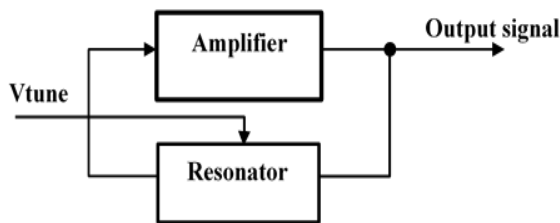
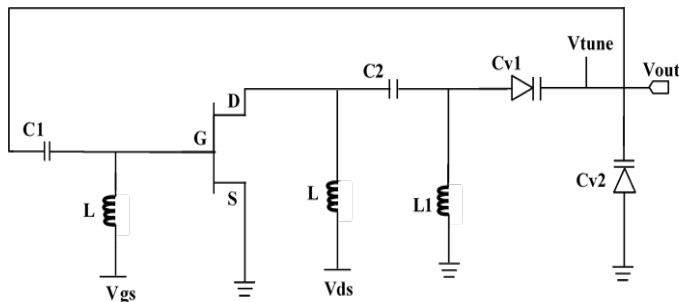


Figure 4: Synoptic diagram of a VCO



4. Voltage Controlled Oscillator Design

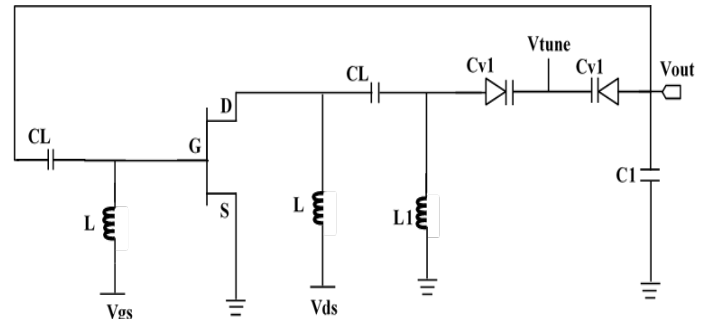
4.1. Colpitts Architecture

An oscillator (Figure 4) can be modeled by an amplifier to compensate for the energy losses, and a resonator to select the oscillation frequency [13]. The tank circuit of the Colpitts Oscillator contains a capacitive divider (two capacitors in series) and an inductance $L1$, while the amplifier circuit is composed of transistors with their bias elements. In [16], two varactors $Cv1$ and $Cv2$ replace fixed capacitors for purpose of tune the oscillation frequency with $Vtune$ (figure 5). The simulation results of this VCO Colpitts show a varying oscillation frequency between 26.6 and 28.85 GHz, a PN near -96.07 and -113.07dBc/Hz@100kHz and @1MHz offset frequency respectively, and a fundamental output power of 9 dBm @ $Vtune= 6$ V. Although the VCO has

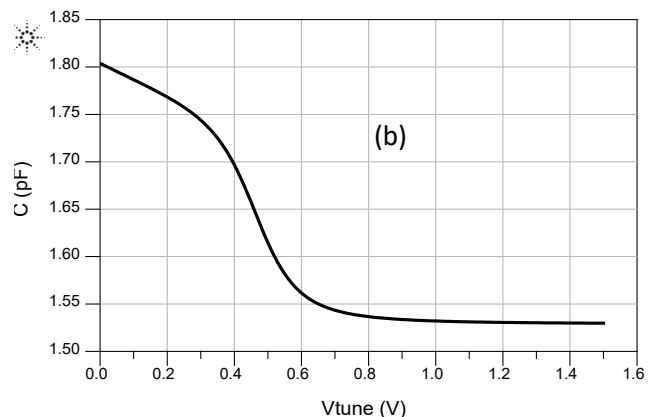
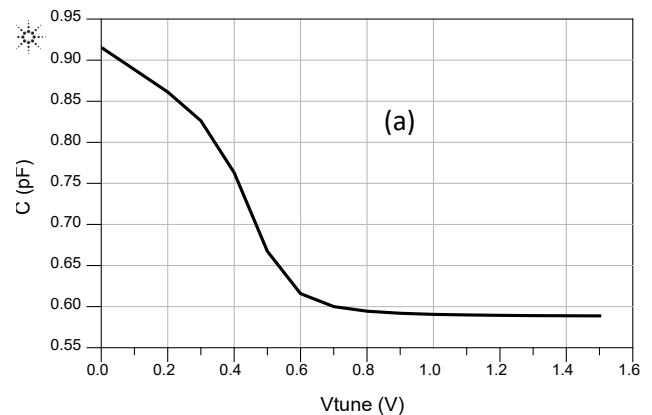
an acceptable FoM of -172.82dBc/Hz, the PN must be improved to fulfill the 5G requirements.

4.2. Low PN VCO Architecture

In order to design a best performing VCO, an innovative architecture is proposed, studied and designed in this work. This circuit is based on the back-to-back structure of the varactors presented in [5] and on the Colpitts structure studied in [13] and [16]. Therefore, we maintained the same active part used in [13] and [16] and a capacity $C1$ has been added to the resonator circuit in order to enhance the capacitance value of the resonator circuit (Figure 6).



Graphs in Figure 7 show the variation of the capacitance of the resonator versus the control voltage $Vtune$ with and without $C1$. From the two graphs, we can observe that the capacitance value of the circuit is clearly increased.



4.3. pHEMT VCO Layout

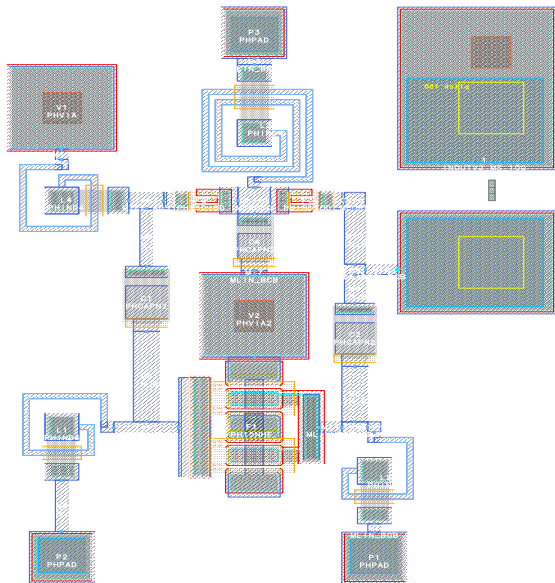
At mm-wave frequencies, parasitic capacitances and inductances can have a strong impact on the final Voltage Controlled Oscillator circuit performance. For this reason, a number of optimization steps were carried out before the final VCO layout presented in Figure 8.

The GaAs technology is multilayered. The chip dimensions are 0.56 x 0.423 mm², for a surface area of 0.237 mm². We can note that it is an extremely compact circuit compared to the architectures recently published in the literature [7, 13, 17-19].

5. Voltage Controlled Oscillator Simulation

The first step when designing an oscillator is to check the convergence of the circuit. Figure 9 shows that at 28 GHz the loop gain is 1.122 and the phase is 0.048 so both oscillation conditions in amplitude and phase are satisfied.

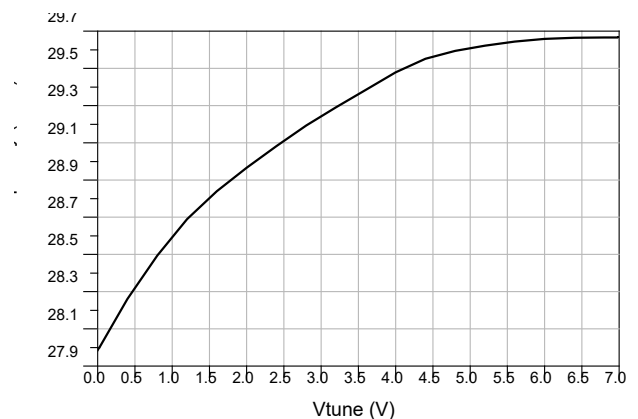
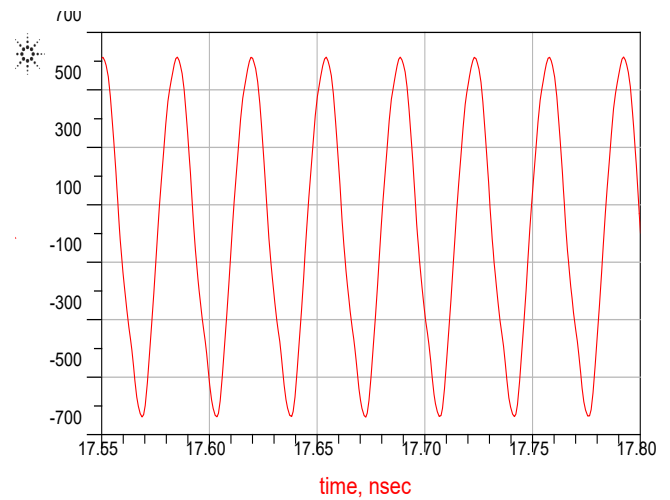
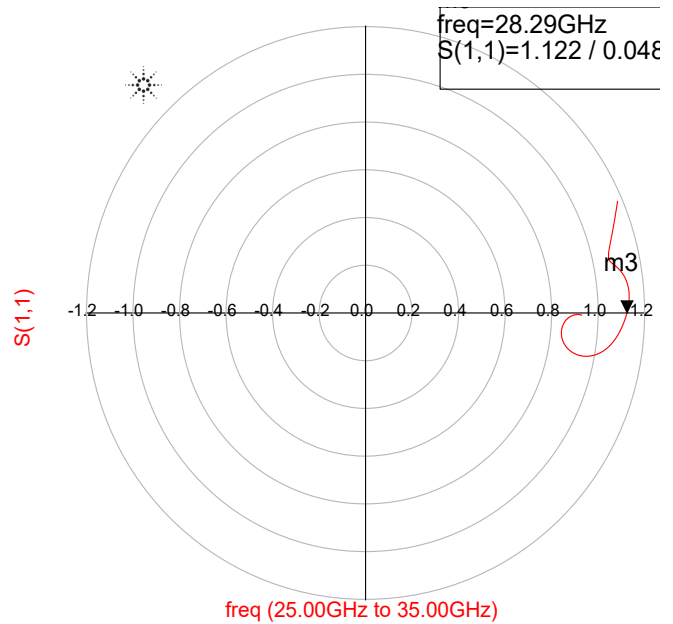
Simulation results on an electrical simulator software, show that the delivered signal is a sine wave (Figure 10) varying between 27.98 GHz and 29.67 GHz (Figure 11). Therefore, the tuning range (TR) is 1.69 GHz (about 6% of central frequency).



At the frequency 28.97 GHz ($V_{tune} = 2.00$ V), phase noise equals to -113.073, -133.078 and -153.07 dBc/Hz @ 100kHz, 1MHz and 10MHz offset frequency respectively (Figure 12). We can deduce that is an ultra-low PN level. Nevertheless, this architecture has an ultra-low PN only in a small tuning range part (around 28.95 GHz) (Figure 13), outside, the PN level is increased by 15 dB, i.e. the PN increases until it reaches the value -133.3 dBc/Hz @10MHz offset. Anyway, the PN of this architecture remains low compared to the other architectures recently published.

The fundamental output power is 3.26 dBm, on the other hand, at the first and second harmonics, it is equals to -18.8 dBm -14.2 dBm, respectively. Thus, the first and second harmonic rejections

are 22.05 dB and 17.45 dB respectively (Figure 14). The DC power consumption is about 51 mW.



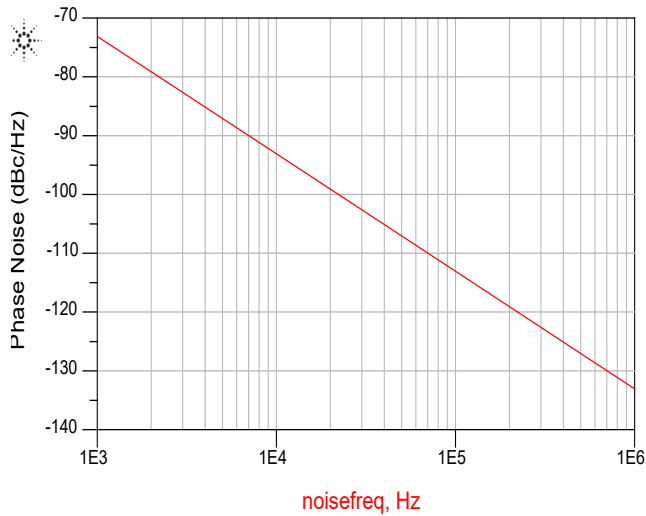


Figure 12: Single Sideband Phase Noise for Vtune=2 V

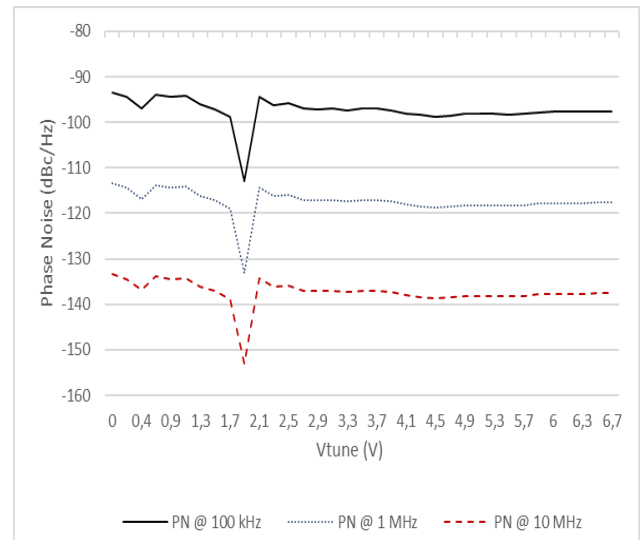


Figure 13: Single Sideband Phase Noise versus Vtune

6. Analysis of the simulation results

Table 2 presents the performance of the designed LO architecture compared to other architectures recently published in the literature. The PN of our architecture is 21.2 dB lower than the architecture proposed in [17] and beyond 22.5 dB, 22.1 dB and 14 dB lower than the architecture designed in [7], [13] and [18] respectively. Our VCO also has a relatively a wide tuning frequency range of about 1.67 GHz and a high fundamental output power of 3.26 dBm. The circuit consumption is quite low, less than half the DC power consumed by the VCO circuits studied in [7] and [19]. Finally, our LO has a good FoM of -195.3 dBc/Hz and is very small chip size compared to other architectures [7, 17-19].

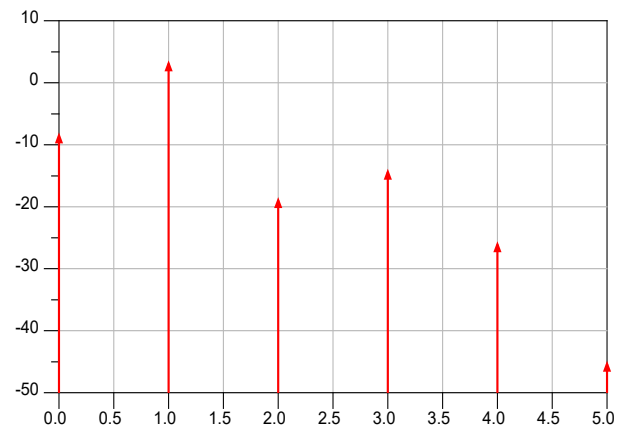


Table 2: Performance of different VCOs recently published in the literature

Ref.	Oscillation frequency (GHz)	DC power consumption (mW)	Fundamental output power (dBm)	PN (dBc/Hz) @1MHz	FoM** (dBc/Hz)	Chip area (mm ²)	Technologie
[17]	28.3	-	11.8	-102	-	0.5	0.15 μm pHEMT
[7]	29.4	124.6	4.5	-100.7	-169.11	3.75	0.13 μm SiGeBiCMOS
[13]	28.02	96	5	-100.9	-180.37	0.25	0.15 μm GaAs pHEMT
[19]	27.7	122	9.75	-113.16	-181.06	0.515	0.15 μm GaAs pHEMT
[18]	28.1	20	-	-109.2	-184.2	0.24	22 nm CMOS
This work	28.8	51	3.26	-123.2*	-195.3	0.237	0.15 μm GaAs pHEMT

(*) The average value of PN

$$(**) \text{ FoM} = L(f_0, \Delta f) + 10 \log(P_{disp}) - 20 \log\left(\frac{f_0}{\Delta f}\right)$$

Where $L(f_0, \Delta f)$ is the PN @ Δf offset from the f_0 and P_{disp} is the DC power consumption of the circuit in mW.

7. Conclusion

A new design approach for a low PN VCO is proposed, studied, and validated in this paper. This approach has enabled the design of a LO for the 5G mm-Wave band of low PN of -123.2 dBc/Hz @1MHz offset frequency from carrier, with a TR of 1.69 GHz, a fundamental output power of about 3.26 dBm associated to low DC power consumption.

References

- [1] B. Dzogovic, B. Santos, T. Van Do, B. Feng, T. Van Do, and N. Jacot, "Bringing 5G Into User's Smart Home" in 2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress, Fukuoka, Japan, 2019. <https://doi.org/10.1109/DASC/PiCom/CBDCCom/CyberSciTech.2019.00145>.
- [2] D. Milovanovic and Z. Bojkovic, "5G Mobile Networks: What is Next?" *International Journal of Communications*, 4, 1-5, 2019.
- [3] G. Kang, S. Jiang, and Z. Zhao, "Research on millimeter wave spectrum planning for 5G" *J. Phys. Conf. Ser.*, 1345, 1-5, 2019. <https://doi.org/10.1088/1742-6596/1345/3/032040>.
- [4] International Telecommunication Union, "Key outcomes of the World Radiocommunication Conference 2019" *ITU News Magazine*, 6, 1-56, 2019.
- [5] J. Hyvert, D. Cordeau, and J.-M. Paillot, "Analysis and design of a very low phase-noise Ku-band coupled VCO using a modified cascode architecture in 0.25 μm SiGe:C BiCMOS technology" *Microelectron. J.*, 75, 137-146, 2018. <https://doi.org/10.1016/j.mejo.2018.04.002>.
- [6] L. Romanò, V. Minerva, C. Samori, and M. Pagani, "Low Phase Noise, Very Wide Band SiGe Fully Integrated VCO" in 12th Gallium Arsenide applications symposium, Amsterdam, 2004.
- [7] G. Cheng, Z. Li, Z. Li, T. Han, and M. Tian, "A 22-to-36.8 GHz low phase noise Colpitts VCO array in 0.13- μm SiGe BiCMOS technology" *Microelectron. J.*, 88, 79-87, 2019. <https://doi.org/10.1016/j.mejo.2019.04.004>.
- [8] A. Es-Saqy, M. Abata, S. Mazer, M. Fattah, M. Mehdi, M. El Bekkali, 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>
- [9] A. Es-saqy, M. Abata, M. Fattah, S. Mazer, M. Mehdi, M. El bekkali, "Study and Design of a MMIC Voltage Controlled Oscillator for 5G mm-wave band Applications" *Int. J. Adv. Trends Comput. Sci. Eng.*, 9(2), 2124-2129, 2020. <https://doi.org/10.30534/ijatcse/2020/186922020>.
- [10] H. Lee, W. Lee, T. Kim, M. Helaoui, F. M. Ghannouchi, and Y. Yang, "6-18 GHz GaAs pHEMT Broadband Power Amplifier Based on Dual-Frequency Selective Impedance Matching Technique" *IEEE Access*, 7, 66275-66280, 2019. <https://doi.org/10.1109/ACCESS.2019.2917699>.
- [11] Y. Chen, C.-N. Chen, C.-C. Chiong, and H. Wang, "A Compact 40-GHz Doherty Power Amplifier With 21% PAE at 6-dB Power Back Off in 0.1 μm GaAs pHEMT Process" *IEEE Microw. Wirel. Compon. Lett.*, 29(8), 545-547, 2019. <https://doi.org/10.1109/LMWC.2019.2925716>.
- [12] C. Wang, D. Hou, J. Chen, and W. Hong, "A Dual-Band Switchable MMIC Star Mixer" *IEEE Microw. Wirel. Compon. Lett.*, 29(11), 737-740, 2019. <https://doi.org/10.1109/LMWC.2019.2940637>.
- [13] M. Kim, K. Choi, and J. Kim, "W-Band Backward Distributed Frequency Doubler Using GaAs 0.15 μm pHEMT Process," *IEEE Microw. Wirel. Compon. Lett.*, 29(6), 400-402, 2019. <https://doi.org/10.1109/LMWC.2019.2913307>.
- [14] PH15-UMS MMIC Foundry, 2015-2016 Foundry Brochures.
- [15] A. Hajimiri, "A General Theory of Phase Noise in Electrical Oscillators" *IEEE JOURNAL OF SOLID-STATE CIRCUITS*, 33(2), 179-194, 1998.
- [16] A. Es-Saqy, M. Abata, S. Mazer, I. Halkhams, M. Mehdi, and C. Algani, "Comparative Study Between Hartley and Colpitts VCO for 5G MM-Wave Band Applications" in 2019 7th Mediterranean Congress of Telecommunications (CMT), Fès, Morocco, 2019. <https://doi.org/10.1109/CMT.2019.8931343>.
- [17] B. Piernas, K. Nishikawa, T. Nakagawa, and K. Araki, "A compact and low-phase-noise Ka-band pHEMT-based VCO" *IEEE Trans. Microw. Theory Tech.*, 51(3), 778-783, 2003. <https://doi.org/10.1109/TMTT.2003.808584>.
- [18] M. A. Shehata, M. Keaveney, and R. B. Staszewski, "A 184.6-dBc/Hz FoM 100-kHz Flicker Phase Noise Corner 30-GHz Rotary Traveling-Wave Oscillator Using Distributed Stubs in 22-nm FD-SOI" *IEEE Solid-State Circuits Lett.*, 2(9), 103-106, 2019. <https://doi.org/10.1109/LSSC.2019.2929326>.
- [19] A. Es-saqy, M. Abata, M. Mehdi, S. Mazer, M. Fattah, M. El bekkali, C. Algani, "28 GHz Balanced pHEMT VCO with Low Phase Noise and High Output Power Performance for 5G mm-Wave Systems" *Int. J. Electr. Comput. Eng.*, 10(5), 2020. <http://doi.org/10.11591/ijece.v10i5.pp%25p>.

Smart Transmission Line Maintenance and Inspection using Mobile Robots

Thongchai Disyadej^{1,*}, Surat Kwanmuang², Paisarn Muneesawang³, Jatuporn Promjan¹, Kanyuta Poochinapan^{4,5}

¹Transmission Line Maintenance Department, Electricity Generating Authority of Thailand, 50300, Thailand

²Department of Mechanical Engineering, Chulalongkorn University, 10330, Thailand

³Department of Electrical and Computer Engineering, Naresuan University, 65000, Thailand

⁴Centre of Excellence in Mathematics, CHE, Si Ayutthaya Rd., Bangkok, 10400, Thailand

⁵Advanced Research Center for Computational Simulation, Chiang Mai University, 50200, Thailand

ARTICLE INFO

Article history:

Received: 21 January, 2020

Accepted: 07 June, 2020

Online: 21 June, 2020

Keywords:

Transmission line

Maintenance

Robot

ABSTRACT

The paper presents sharing several of experiences and practices on smart robotic application for overhead transmission line maintenance and inspection. First, the pilot-line pulling robot is an invention used to pull a pilot line which is an important step for additional high voltage conductor installation. The puller robot can traverse the overhead ground wire, OHGW, and pulls a lead line via a set of cradle blocks at intervals, carrying the line above the ground. The robotic puller passes over barriers below the power line, such as the road with traffic, power distribution lines, river, or vegetation making tasks achieved conveniently, safely, and rapidly without impact on nearby communities. The robot was further utilized to pull a lead line/conductor crossing over the electrical substation without interrupting energy and pull a lead line for the improvement of transmission line ground clearances. The developed pilot-line pulling robot has been accredited as the corporate best practice; the standards for innovation, operation, and maintenance are archived for works at all EGAT transmission line operation & maintenance units nationally. Moreover, EGAT was now jointly investigating with universities on a new robotic device for aerial transmission line inspection. The target of the research is to create a mobile robot prototype for inspection of overhead power lines. The inspection robot shall crawl along the ground wire and transpose autonomously across installed equipment on the ground wire, such as vibration dampers, suspension clamps, compression dead ends, etc. In addition, the inspection robot is able to take photos and videos during a transmission line inspection in both offline and online features. Using the robot, transmission line inspection's labor cost can be reduced, and the new method helps improve patrol and inspection efficiency, comparing to the conventional manpower method. Trough utilization of the new maintenance and inspection robots, utilities can minimize transmission line operation & maintenance budget.

1. Introduction

Electricity Generating Authority of Thailand (EGAT) is Thailand's leading state enterprise about electric energy under the supervision of the department of energy that has its main business in generation, transmission, and energy sales to other power

distribution utilities. In order to promote the development of the national economic and industrial field in conjunction with the "Power for Thai Happiness" corporate agenda, the electricity utility's mission is to preserve the efficiency of the transmission network for the entire country. In the view of electrical providers, the primary purpose of high voltage transmission systems is the provision of services. EGAT also has a long tradition of

*Thongchai Disyadej, EGAT, Tel. +66 53 220048, Thongchai.d@egat.co.th

developments in the service and management of the energy supply sector utilizing the robotic technology, as demonstrated in Figure 1. Northern Region Operation Division developed an Optical Ground Wire (OPGW) inspection with the robot for examining a damage spot on OPGW's strands in 2008 [1]. The OPGW inspection robot was upgraded to a better version, 2 years later, such that phase conductors can be inspected also. In 2014, Electrical Maintenance Division and a domestic university co-invented a robot prototype used to inspect boiler water-wall tubes [2]. Utilizing magnetic adhesion, the NDT inspection robot can examine the thickness and the external properties of boiler water-wall tubes. Furthermore, Electrical Maintenance Division funded the national technology center on the same year a project on a tiny robot prototype with 2-cm size or less for internal generator examination [3]. Similar to other robots on the same sector, owing to their smaller form, the robot may analyze inside the generator a distance of less than 3.5 cm. This saves generator testing period and expensive oversea testing robots. In addition, EGAT funded the national metal and materials technology center an R&D work in 2013 about a robotic system development for welding [4]. Main findings include a welding device with applications for welding process monitoring and torch motion creation for overlay welding of parts used in power plants. Afterwards, in 2017, another development for robotic welding system to be used in the repair phase for high-value power plant components was performed [5]. Recently developed [6], new transmission line maintenance and inspection robots are proposed in this paper.

220-500 kV transmission lines are developed by Chinese academy of sciences and many Chinese universities [19-22]. Three types of high voltage transmission line robots are developed by Wuhan University for single and bundled conductors [23-30].

According to presented relevant works, many robots were designed with some limitations for actual field working, e.g. their weight and dimension, specific configuration of conductors, limitation for some tower junction, difficulty for actual field works, and etc. Therefore, two transmission line maintenance and inspection robots with light weight, small size, and ease of use are designed and developed in order to be a new realistic solution for power utilities expecting more efficient O&M system.

The paper sections are arranged as the following. We present in section 2 a developed transmission line maintenance robot. We firstly summarizes problems and constraints of the existing pilot-line pulling method for additional stringing of conductor projects. Then, the development of the new maintenance robot and actual success implementations in transmission lines are presented. The next part, section 3, deals with a novel inspection robot for transmission lines. The background, design, Lab test, and field test of the inspection robot are presented. Working results of both developed robots are discussed in the next section. Evaluation and future improvement are finally concluded in section 5.

2. Transmission Line Maintenance Robot

According to the corporate social responsibility policy, transmission system division desires to improve a better approach for stringing additional power lines while communities suffer no public impacts. Thus, the team of inventors decides to evolve a research project on an advanced approach of pilot-line pulling for the power line installation that conventionally requires manpower and land/water vehicles to pull out a pilot line dragged over the field beneath the power line above, shown as Figure 2-3. The consequence is that it damages farmlands beneath the power line and causes an effect on communities, Figure 3.

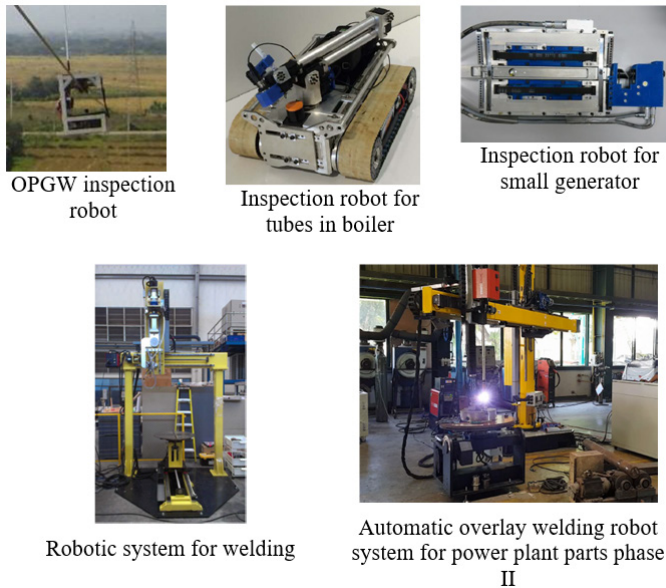


Figure 1: Examples of EGAT innovation on robots [1-5]

A summarized literature review is performed on transmission line maintenance and inspection robotics developed by many researchers. LineScout is a mobile robot developed by the Hydro-Quebec Research Institute to perform live line inspection and maintenance [7-15]. Expliner is an inspection robot developed by the HiBot Corporation, Kansai Electric Power Corporation, and Tokyo Institute of Technology for commercialization [16-17]. Also, TI is an inspection robot developed, evaluated, and demonstrated by Electric Power Research Institute under EHV condition [18]. Moreover, various robot prototypes focusing on

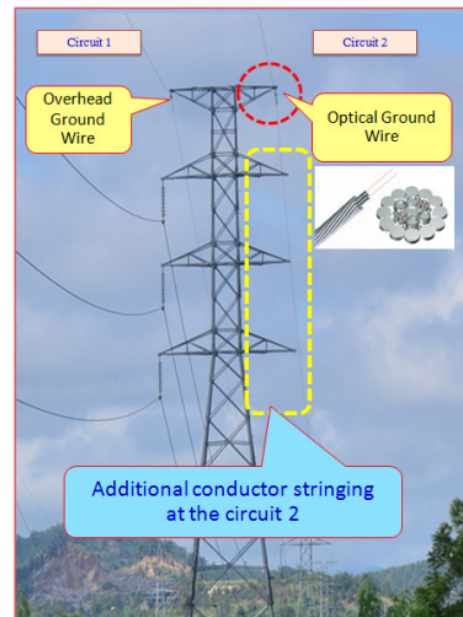


Figure 2: Conductor stringing for the circuit no. 2 to be executed on the blank crossarms of towers

Step 1: Setup a conductor puller and scaffolds; install stringing blocks on towers.



Step 2: Pull a pilot rope sliding over the ground surface along the line route.



Step 3: Pull a wire rope through stringing blocks on scaffolds/towers



Figure 3: Conventional steps for pilot-line pulling

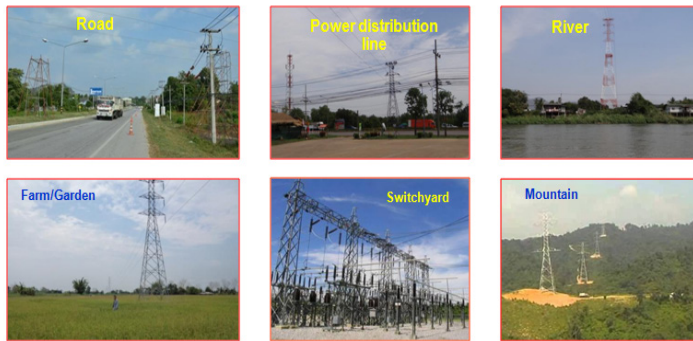


Figure 4: Constraints and challenges on pilot-line pulling

Constraints and challenges for pulling the pilot line with the existing working approach are concluded as the following:

- Interrupted traffic and vegetation damages to landowner (damage compensation to be paid)
- A great amount of workers needed co-working on different places
- Risks of injury by incidents during line pulling operations over the roads, railways, communication lines, power distribution lines, and etc.
- Scaffolds needed above the street and electric distribution line
- Electricity de-energization required when crossing over high voltage switchyards

Therefore, the researchers utilize the robotic technology to diagnose and evolve the new creative approach for better pilot-line pulling operation above all obstructions with the robot developed as a tool to pull out a pilot line through a number of cradle blocks. The complete description of the specially designed puller robot together with a series of cradle blocks is shown in Figure 5. Whilst pulling, the running line is carried on the OPGW at structures and lifts at the height more than 13 m. over the ground, in Figure 6. Thereby, communities suffer no more problems from our jobs, and damage compensation for properties does not need to be paid anymore. Moreover, construction workers can work on stringing a lead line above the highway, river, and live power line with safety, efficiency, and reliability. Additionally, the robot has been utilized further for a lead-line pulling over the live substation in the absence of service outage and also a lead-line pulling for the ground-clearance upgrading of an overhead power line.

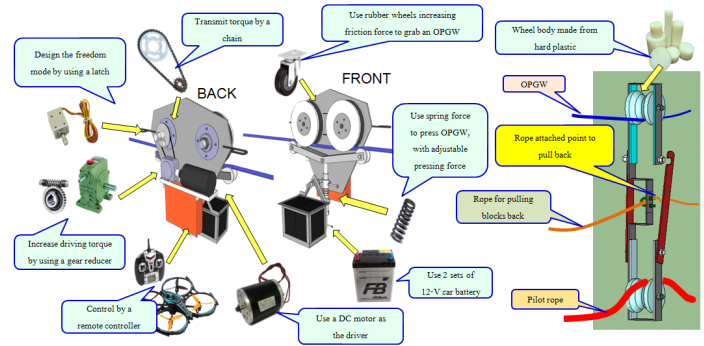


Figure 5: The composition of robot for line pulling work

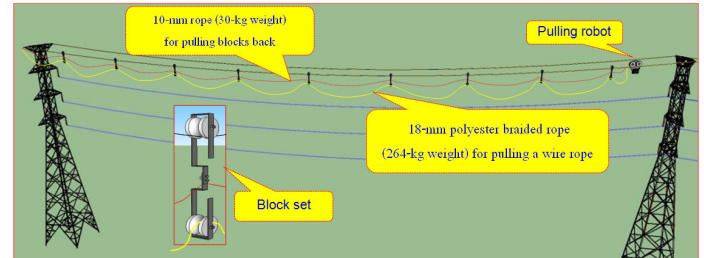


Figure 6: The complete description of the new pilot-line pulling method



Figure 7: Actual implementations of the transmission line maintenance robot

With success implementation, EGAT has innovated a special robot for pilot-line pulling over all obstructions, and it passed the pulling-force test with 60-kilogram load and 35-degree inclination. The consequences have proved the accomplishment.

3. Transmission Line Inspection Robot

3.1. Background

Throughout the world, inspection of transmission lines usually consumes workforces to achieve the routine tasks by ground patrolling along the transmission line paths. Minimization of man-hour is necessary in order to improve capability while the precision of data is still needed to be conserved. For this reason, partially/completely utilizing robotic technology into the standard line- inspection working method is of great essential. Existing means of routine power line inspection in the past have been dealing with both helicopters and UAVs to minimize cost of manpower. However, the former demands costly capital and operation budget whilst the latter suffers with many constraints, such as short battery lifetime, relevant laws, dependability in operation, and so forth. Therefore, application of an innovative inspection robot rolling along an OHGW can play a significant role in smart transmission line inspection improvement for electric utilities.

The maintenance robot in the earlier section that travels along an OHGW can be practically upgraded to an inspection robot. Nevertheless, the service robot is unable to cross through sequences of obstacles, e.g. vibration dampers, suspension clamps, compression dead ends, and other hardware. That is, the robot has ability to travel just one single span between towers only. To significantly improve inspection efficiency, a new autonomous robot negotiating and crossing successive obstacles automatically must be developed.

Crossing through each tower top is the key challenge for the robotic transmission line inspection. When the robot encounters a tower, personnel needs to climb to each tower and then manually transpose the robot to another side of span. It will be very much more efficient to use if the inspection robot can automatically navigate through tower tops. Nevertheless, autonomous crossing of obstacles repetitively requires huge efforts on advanced technology because equipment installed on the transmission structures have a lot of variety. Regarding mentioned constraints, the research team has designed an approach by developing a new mobile robot which can travel through a series of obstacles and earth wire crossarms with automatic navigation as well as feature of gathering high-quality visual information, e.g., damaged equipment and right-of-way vegetation data [31].

The aim of this research project is to create a novel transmission line section robot that can roll along an OHGW with autonomous crossing of obstacles. This working approach improves more efficiency than the existing approaches, such as ground patrols or helicopters. Photo and VDO via both offline and online (real time) modes can be recorded by the robot; it is a new mean to improve more accurate inspection than the conventional mean. In addition, for an outage event, the robot can be sent out quickly to search for an outage cause in mountainous regions, minimizing the customer's service interruption.

3.2. Conceptual design of transmission line inspection robot

We designed an overhead power line inspection robot to work with autonomous feature and use a communication network for teleoperation. Figure 8 presents the conceptual design of the inspection robot. One can observe that the control portion is divided into 3 key portions systematically connected inside: moving control, communication control, and remote control parts. Those three units link altogether operations for the entire system and link with operators through an App on a handheld equipment in the field that will track data of the robot transmitted via the cloud network system, which is illustrated in Figure 8.

- **Robot motion-control part**
Robot movement is controlled by a microprocessor inside, which is the core duty for moving towards a desired location precisely. The inner subsystem includes a system of battery power supply which will energize power to all parts and a feedback-control part that can work with a communication-control part closely for sensing the state of the robot such that the microprocessor can make a judgement to receiving information to command motion of each movement.
- **Communication control part**
This part includes an embedded microprocessor which has a role about communicating to each equipment, e.g. sensors which transmit feedback information to the movement part that identify the movement state or a camera that record images during the robot roll along each path.
- **Tele-monitoring control part**
The robot has a communication unit with a network for transmitting states back to the cloud system via a mobile cellular system for a tele-monitoring control part. M2M (machine to machine), which is a protocol commonly for the IoT world, is used as a format of Data streaming. It can work with reliability although the restriction of system bandwidth, i.e. sending data in the field to the cloud system via the App for observing.

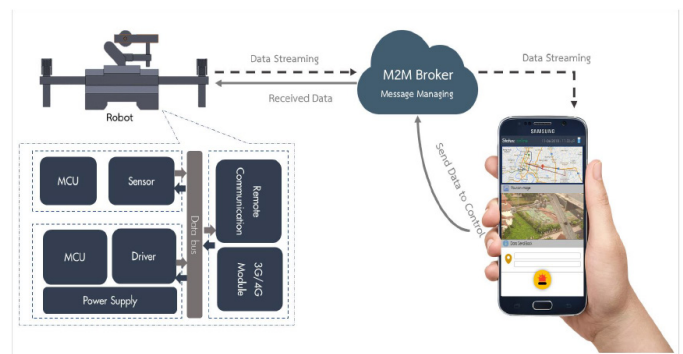


Figure 8: The conceptual design of the transmission line inspection robot

3.3. Purposed design

Figure 9 illustrates a design concept of an efficient robotic prototype for transmission line inspection with governing mechanism to transpose through obstacles on transmission lines. The model is designed and simulated using a special computer software. The robot consists of 6 linear actuators with the

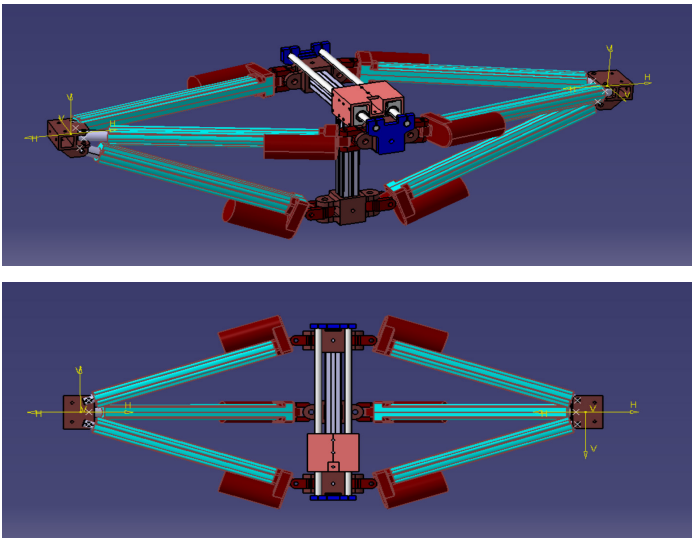


Figure 9: Design concept of the inspection robot

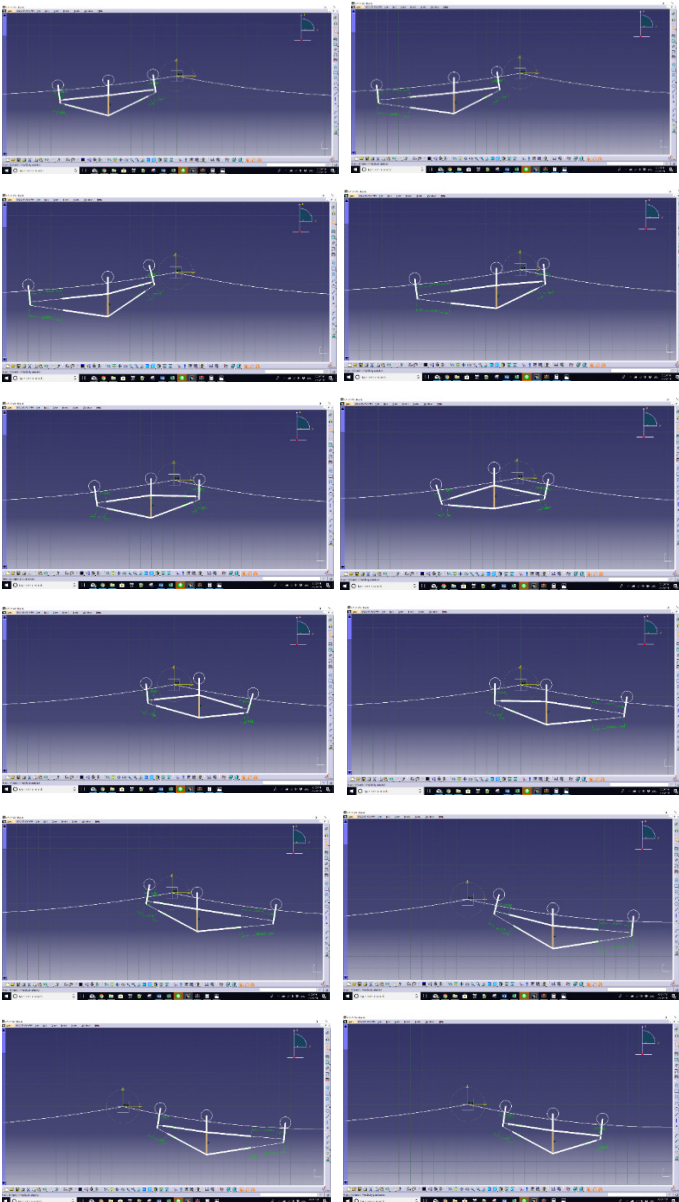


Figure 10: Transposition steps of the robot

maximum extension of 20 cm. They are used to control the front or back wheel towards the desired position. By connect 3 linear actuators together to form a parallel mechanism, the resulted actuator has very high stiffness and compact while the end-effector is still able to reach in required envelop. Moreover, there are 3 wheels for the driving of the robot. Each of wheel is installed with a driving motor which is a 200W BLDC hub motor installed inside the driving wheel set.

3.4. Motion of the robot and obstacles mitigation

The robot can shift its center of gravity to the front and rear by extending its arm to the desired direction. This enables the robot to lift the other arm out of the OHGW without losing stability.

The motion process is that the robot usually run along the OHGW and can transpore through obstacles, e.g. equipment and hardware on the structure with steps as the following:

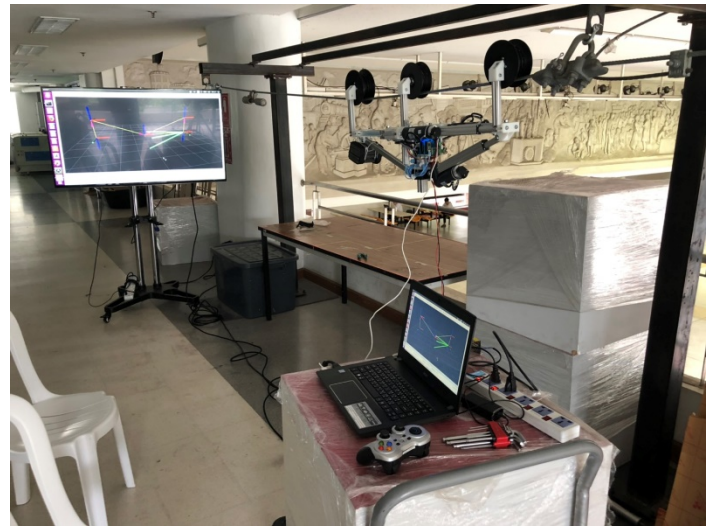


Figure 11: Laboratory test for the prototype with the line model

3.5. Kinematic model

The robot front and back arms are symmetric and identical. Each arm has three linear actuators arranged in the form of a tripod. The motion of the end effector can be described using coordinate system in the Figure 12.

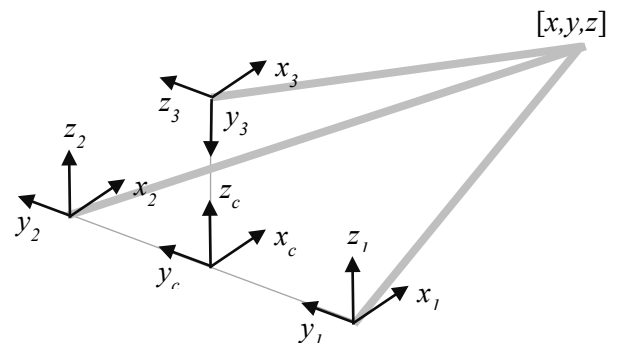


Figure 12: Coordinate system for each actuator and arm's center and also the end-effector position.

The inverse kinematic model of each actuator is an actuator length (l) as a function of the end-effector position in arm's center frame (x, y, z). There is a restriction in designing the end joint of the actuator which results in offsets between y and z rotation axis (l_1). Since all actuators are identical, we can derive the model for all actuator in each actuator's coordinate frame in the following:

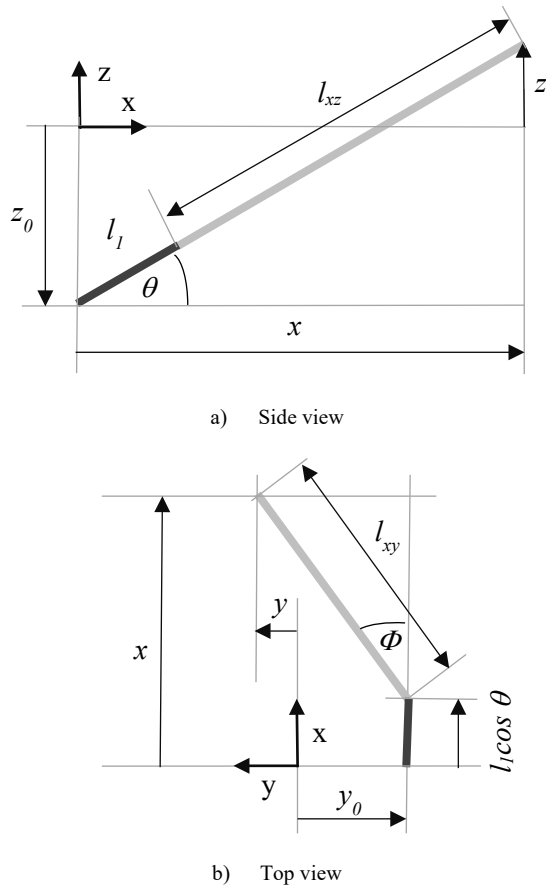


Figure 13: a) Side view and b) Top view of the mechanical model of a single linear actuator assembly

$$\theta = \tan^{-1} \left(\frac{z - z_0}{x} \right) \quad (1)$$

$$l_{xy}^2 = (y - y_0)^2 + (x - l_1 \cos \theta)^2 \quad (2)$$

$$l_{xz}^2 = (z - z_0 - l_1 \sin \theta)^2 + (x - l_1 \cos \theta)^2 \quad (3)$$

$$l^2 = l_{xy}^2 + l_{xz}^2 \quad (4)$$

To verify the derived model, we commanded the end-effector to move in the form of geometric paths. We then recorded the movement using long-exposure snapshot which enables us to see the actual path taken. As a result, we found that the maximum available speed of the end-effector is 2 cm/sec. The speed beyond this number would result in linear actuator saturation and distortion in end-effector path.

3.6. Electronics

The control system for operation of robotic arms is for the front arm and the back arm. Each side consists of 3 linear actuators with DC motors and optical quadrature encoder to measure the position of the extended core. For the control system of motion, motors are

grouped with 2 motors of each set. They are controlled via a Teensy 3.2 microcontroller which is a STM32 microcontroller with 96 MHz speed and with 2 quadrature encoder counters.

Operation and control board for the control system consists of Teensy 3.2 microcontroller and 2 motor driver boards VNH5019 assembled onto the circuit board, as shown in the Figure 15.

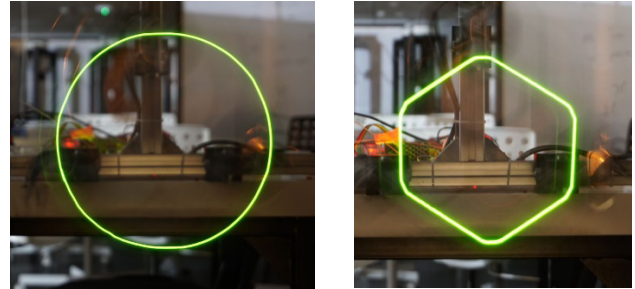


Figure 14: Actual path of the end-effector when tracing a 20mm radius circle at the speed of 2cm/s (left) and 4cm/s (right)

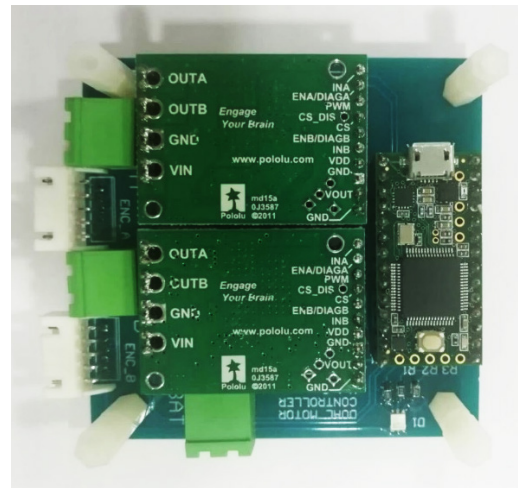


Figure 15: Operation and control board for the system

The control is commanded through the USB port to command the motor position control program with PID control that is processed by microcontroller. The diagram for the operation of the motor control set is presented in the Figure 16.

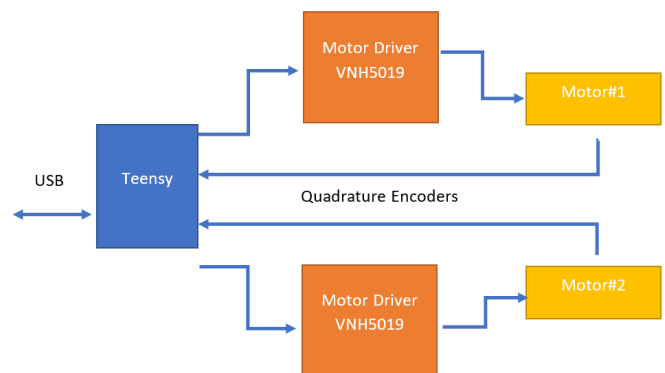


Figure 16: The diagram for the operation of the motor control set

All 3 motor control set are assembled altogether to control all 6 motors, and USB signal lines are connected through 4-port USB hub.

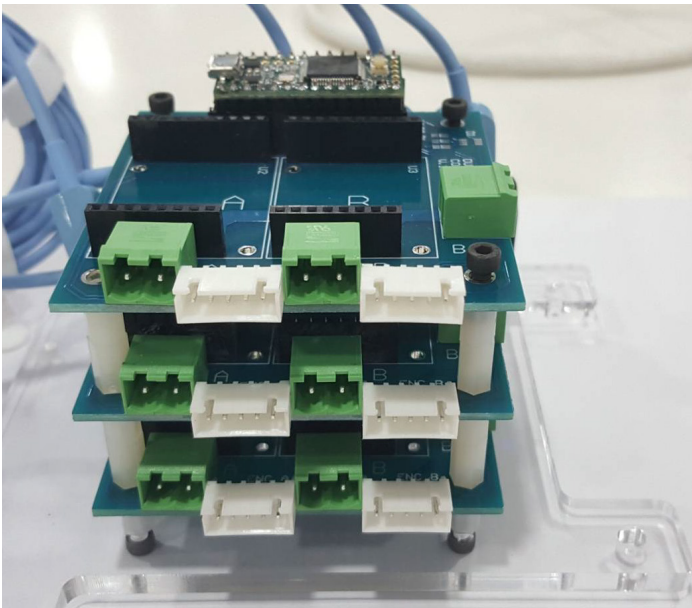


Figure 17: All motor control set are assembled altogether

Moreover, there is the fourth controller board. This board is used to control 3 BLDC motors. It is also equipped with inertial measurement module and load cells analog to digital converter to measure weight on each wheel for center of gravity calculation.

3.7. Software and control system

The main processor for this robot is an Industrial PC with Intel Core I5 processor running Ubuntu 18.04 as the operating system.

Operation and control program for the motion of the robot is constructed on the basis of Robotics Operating System (ROS) with the control structure in the Figure 18 below.

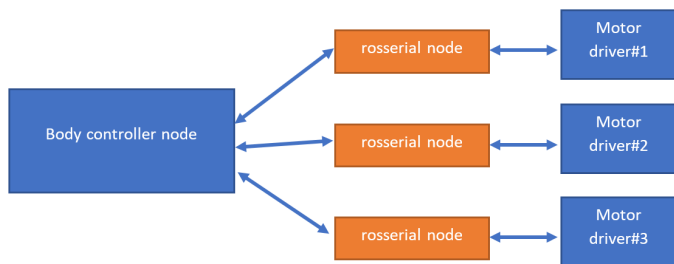


Figure 18: Operation and control program for the robot motion control



Figure 19: The developed inspection robot during set-up before the field test



Figure 20: Actual field test on a 115 kV transmission line

4. Results

The invented pilot-line pulling robot was utilized to line pulling operations at many 115-kV transmission lines for different of maintenance jobs during 2016-2019. The results showed its capability on pulling the lead line/conductor along the OHGW effectively and was accredited for real workings with no more harm to landowner and can operate with safety and reliability. Examined results after achieving the tasks have proved about the robot's performance on a pilot line pulling over the field with success. The robot can cut the personnel cost up to 34 less people and save the payment for damage's properties. For a direct output, the development of the robotic pilot-line pulling on the power line above the field can be achieved with no public concerns. For an indirect output, the electrical transmission network is safer and more dependable. Figure 7 demonstrates real implementation for inventive robot utilizations that pass the street, river, farmland, mountains, live switchyard, and electric power line. Also, the novel pilot line pulling robot was accredited as best practice by EGAT. The standards for invention, working, and maintenance were archived as references for transmission line operation and maintenance units nationally.

The inspection robot for transmission lines was specially designed for inspecting overhead power lines, running via the OHGW. Autonomous feature is a key challenge for developing the robot. Movement test was performed well on the simulated tower top for suspension type in the laboratory. After that, an actual field test was performed on July 2019 on suspension type and tension type towers of the real 115-kV power line. The observed results showed that the prototype of robot can run and transpose through actual obstacles like tested in the laboratory.

5. Conclusions

There is already evidence of different types of advanced robots designed and built by EGAT for the maintenance and inspection of high voltage transmission lines. The smart approach of work presently proposes the essential implementation of deploying innovative robots for approaches towards the smarter energy utility industry. The developed robots were successfully utilized to the actual transmission line in Thailand, causing tasks achieved with safety and reliability and without customer's service interruption. In addition, it helps improve effectiveness by saving manpower and be an effective approach rectifying inspection precision while comparing to the convention one, such as ground patrolling. By utilization with the smart robots, power companies will save on routine line maintenance and inspection

expenditure. Future development with highly advanced technology could be performed on the next generation of robot that can automatically climb the tower by itself.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The researches were supported by the Electricity Generating Authority of Thailand and the Center of Excellence in Mathematics, the Commission on Higher Education, Thailand.

References

- [1] EGAT Northern Region Operation Division, "The optical ground wire inspector robot: phra-siva vehicle," Conference of Electric Power Supply Industry, 2008.
- [2] FIBO, "A prototype of inspection robot for water wall tubes in boiler," EGAT Final Research Report, 2014.
- [3] NECTEC, "Development of a robot with thickness less than 2 centimeters for generator inspection," EGAT Final Research Report, 2014.
- [4] MTEC, "Development of robotic system for welding," EGAT Final Research Report, 2013.
- [5] MTEC, "Automatic overlay welding robot system for power plant parts phase II," EGAT Final Research Report, 2017.
- [6] T. Disyadej, J. Promjan, K. Poochinapan, T. Mouktonglang, S. Grzybowski and P. Muneesawang, "High Voltage Power Line Maintenance & Inspection by Using Smart Robotics," 2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 2019, pp. 1-4, doi: 10.1109/ISGT.2019.8791584.
- [7] S. Montambault and N. Pouliot, "LineScout Technology: Development of an Inspection Robot Capable of Clearing Obstacles While Operating on a Live Line," ESMO 2006 - 2006 IEEE 11th International Conference on Transmission & Distribution Construction, Operation and Live-Line Maintenance, Albuquerque, NM, 2006, pp. , doi: 10.1109/TDCLLM.2006.340744.
- [8] N. Pouliot and S. Montambault, "LineScout Technology: From inspection to robotic maintenance on live transmission power lines," 2009 IEEE International Conference on Robotics and Automation, Kobe, 2009, pp. 1034-1040, doi: 10.1109/ROBOT.2009.5152291.
- [9] N. Pouliot, P. Latulippe and S. Montambault, "Reliable and intuitive teleoperation of LineScout: a mobile robot for live transmission line maintenance," 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO, 2009, pp. 1703-1710, doi: 10.1109/IROS.2009.5354819.
- [10] S. Montambault, N. Pouliot, J. Toth and B. Spalteholz, "Reporting on a large ocean inlet crossing live transmission line inspection performed by linescout technology," 2010 IEEE International Conference on Robotics and Automation, Anchorage, AK, 2010, pp. 1102-1103, doi: 10.1109/ROBOT.2010.5509190.
- [11] S. Montambault and N. Pouliot, "Field experience with LineScout Technology for live-line robotic inspection and maintenance of overhead transmission networks," 2010 1st International Conference on Applied Robotics for the Power Industry, Montreal, QC, 2010, pp. 1-2, doi: 10.1109/CARPI.2010.5624454.
- [12] J. Toth, N. Pouliot and S. Montambault, "Field experiences using LineScout Technology on large BC transmission crossings," 2010 1st International Conference on Applied Robotics for the Power Industry, Montreal, QC, 2010, pp. 1-6, doi: 10.1109/CARPI.2010.5624413.
- [13] N. Pouliot, D. Mussard and S. Montambault, "Localization and archiving of inspection data collected on power lines using LineScout Technology," 2012 2nd International Conference on Applied Robotics for the Power Industry (CARPI), Zurich, 2012, pp. 197-202, doi: 10.1109/CARPI.2012.6473341.
- [14] S. Montambault, N. Pouliot and M. Lepage, "On the latest field deployments of LineScout Technology on live transmission networks," 2012 2nd International Conference on Applied Robotics for the Power Industry (CARPI), Zurich, 2012, pp. 126-127, doi: 10.1109/CARPI.2012.6473342.
- [15] S. Montambault and N. Pouliot, "About the future of power line robotics," 2010 1st International Conference on Applied Robotics for the Power Industry, Montreal, QC, 2010, pp. 1-6, doi: 10.1109/CARPI.2010.5624466.
- [16] P. Debenest et al., "Expliner - Robot for inspection of transmission lines," 2008 IEEE International Conference on Robotics and Automation, Pasadena, CA, 2008, pp. 3978-3984, doi: 10.1109/ROBOT.2008.4543822.
- [17] P. Debenest and M. Guameri, "Expliner — From prototype towards a practical robot for inspection of high-voltage lines," 2010 1st International Conference on Applied Robotics for the Power Industry, Montreal, QC, 2010, pp. 1-6, doi: 10.1109/CARPI.2010.5624434.
- [18] A. Phillips, E. Engdahl, D. McGuire, M. Major and G. Bartlett, "Autonomous overhead transmission line inspection robot (TI) development and demonstration," 2012 2nd International Conference on Applied Robotics for the Power Industry (CARPI), Zurich, 2012, pp. 94-95, doi: 10.1109/CARPI.2012.6473343.
- [19] Tang Li, Fang Lijin and Wang Hongguang, "Obstacle-navigation control for a mobile robot suspended on overhead ground wires," ICARCV 2004 8th Control, Automation, Robotics and Vision Conference, 2004., Kunming, China, 2004, pp. 2082-2087 Vol. 3, doi: 10.1109/ICARCV.2004.1469485.
- [20] Tang Li, Fang Lijin and Wang Hongguang, "Development of an inspection robot control system for 500KV extra-high voltage power transmission lines," SICE 2004 Annual Conference, Sapporo, 2004, pp. 1819-1824 vol. 2.
- [21] Tang Li, Fu Shuangfei, Fang Lijin and Wang Hongguang, "Obstacle-navigation strategy of a wire-suspend robot for power transmission lines," 2004 IEEE International Conference on Robotics and Biomimetics, Shenyang, 2004, pp. 82-87, doi: 10.1109/ROBIO.2004.1521756.
- [22] W. Hongguang, J. Yong, L. Aihua, F. Lijin and L. Lie, "Research of power transmission line maintenance robots in SIACAS," 2010 1st International Conference on Applied Robotics for the Power Industry, Montreal, QC, 2010, pp. 1-7, doi: 10.1109/CARPI.2010.5624428.
- [23] X. Xiao, G. Wu and S. Li, "Dynamic Coupling Simulation of a Power Transmission Line Inspection Robot with its Flexible Moving Path when Overcoming Obstacles," 2007 IEEE International Conference on Automation Science and Engineering, Scottsdale, AZ, 2007, pp. 326-331, doi: 10.1109/COASE.2007.4341691.
- [24] G. Wu, X. Xiao and Y. Lai, "A Wheel-Claw Hybrid Manipulator and its Grasping Stability for the Mobile Robot Rolling/Crawling Along Flexible Cable," 2007 International Workshop on Robotic and Sensors Environments, Ottawa, Ont., 2007, pp. 1-6, doi: 10.1109/ROSE.2007.4373964.
- [25] J. Guo et al., "Database Environment of an Inspection Robot for Power Transmission Lines," 2009 Asia-Pacific Power and Energy Engineering Conference, Wuhan, 2009, pp. 1-4, doi: 10.1109/APPEEC.2009.4918249.
- [26] G. Wu et al., "Design and Application of Inspection System in a Self-Governing Mobile Robot System for High Voltage Transmission Line Inspection," 2009 Asia-Pacific Power and Energy Engineering Conference, Wuhan, 2009, pp. 1-4, doi: 10.1109/APPEEC.2009.4918256.
- [27] Gongping Wu, Tuo Zheng, Hua Xiao and Cheng Li, "Navigation, location and non-collision obstacles overcoming for high-voltage power transmission-line inspection robot," 2009 International Conference on Mechatronics and Automation, Changchun, 2009, pp. 2014-2020, doi: 10.1109/ICMA.2009.5245998.
- [28] C. Hu, G. Wu, H. Cao and X. Xiao, "Obstacle Recognition and Localization Based on the Monocular Vision for Double Split Transmission Lines Inspection Robot," 2009 2nd International Congress on Image and Signal Processing, Tianjin, 2009, pp. 1-5, doi: 10.1109/CISP.2009.5303695.
- [29] Gongping Wu, Tuo Zheng, Zhenglie Huang, Huan Liu and Cheng Li, "Navigation strategy for local autonomous obstacles-overcoming based on magnetic density detection for inspection robot of single split high voltage transmission line," 2010 8th World Congress on Intelligent Control and Automation, Jinan, 2010, pp. 6555-6561, doi: 10.1109/WCICA.2010.5554409.
- [30] X. Xu, G. Wu, Y. He, X. Xiao, M. Liu and Q. Liu, "Navigation and control for a novel ground line inspection robot," 2010 IEEE International Conference on Mechatronics and Automation, Xi'an, 2010, pp. 1228-1232, doi: 10.1109/ICMA.2010.5589954.
- [31] Naresuan University, "Development of mobile robot for inspection of high voltage transmission line," EGAT Research Proposal, 2018.

Measurement of Employee Awareness Levels for Information Security at the Center of Analysis and Information Services Judicial Commission Republic of Indonesia

Mainar Swari Mahardika^{*,1}, Achmad Nizar Hidayanto¹, Putu Agya Paramartha¹, Louis Dwysevrey Ompusunggu¹, Rahmatul Mahdalina¹, Farid Affan²

¹Faculty of Computer Science, Universitas Indonesia, Jakarta, 10430, Indonesia

²Faculty of Economics and Business, Universitas Indonesia, Jakarta, 10430, Indonesia

ARTICLE INFO

Article history:

Received: 20 March, 2020

Accepted: 14 June, 2020

Online: 21 June, 2020

Keywords:

Information Security

Measuring IS Awareness

HAIS-Q

AHP

Palinfo

Judicial Commission

ABSTRACT

The Center for Analysis and Information Services (Palinfo) at the Judicial Commission closely related to the management of information systems which are used to process organizational internal data and information systems on public services. Data processing and network management have an information system security risk. The Judicial Commission seeks to reduce risk and improve the quality of information security. This study aims to measure employee awareness of information security at the Center of Analysis and Information Services at the Judicial Commission, which also includes the Data/IT department. The study was conducted through an arranged interview with three experts and the dissemination of information security awareness questionnaires to all Palinfo employees, amounting to 25 persons. The results of the questionnaire were evaluated using The Human Aspects of Information Security Questionnaire (HAIS-Q) and the Analytic Hierarchy Process (AHP) method. The results showed that the level of information security awareness in Palinfo and the Data/IT section was at the "average" level. There is one focus area that shows a "good" level. While in the Data/IT department, several sections that show a "good" level. Based on these results, we recommend being used in maintaining information security, namely seven policies, ten information technology approaches, and socialization/training conducted in various ways.

1. Introduction

Information is a valuable asset for an organization because information is a strategic resource in increasing business value. Therefore, the protection of information security is an absolute matter that must be taken seriously by all highest ranks of leaders to employee concerned. With the overall safety of the environment where the information is located, the integrity, availability, and confidentiality of information in the company will be guaranteed. To maintain the continuity of an organization's business, the organization needs the availability of data and information as one of the influential factors [1].

Information system security threats are actions taken both from within the system and from outside systems that can consider the balance of the information system. Threats to information security

arise from individuals, organizations, connections, and events that can cause damage to information sources. Security threats to information systems not only related from outside the company such as business opponents or other individuals and groups but can also be used from within the company [2].

According to data reports on information security incidents based on reports in 2017 showed that at the Judicial Commission there was a hacker attack that crippled several application systems and ransomware virus attacks that attacked several computers connected to the Internet network. The report shows that the role of human error is a contributing factor to information security incidents. Human error involved in information security can be in the form of opening insecure websites, opening attachments/links carelessly, downloading files without scanning, using passwords easy to guess, sharing passwords with others, losing devices or losing access to mobile devices, often connecting devices to public networks [3]. The occurrence of the security incident shows that

*Mainar Swari Mahardika, Universitas Indonesia, Jakarta, Indonesia.

Email: mainar.swari@ui.ac.id

employees are not expected to have an awareness of information security. Therefore, research needed to measure the level of employee awareness of information security.

According to the January-December 2018 Annual Report ID-SIRTII/CC found that in 2018 there were 16,939 website incidents/defacement and the .go.id domain ranked first with 30, 75% more often affected by defacement. Based on the monitoring results, there are 4,499 phishing links, of which 1,654 Indonesian domain websites have been affected or indicated for phishing. Data leak monitoring in 2018 obtained data leakage of 785,967 from domains and records. The number comprises 785,906 records / lines from 61 various .id domains. One of the domains obtained from data leakage is the domain go.id [4].

The Judicial Commission of the Republic of Indonesia is vested with two constitutional authorities, namely to conduct a selection of candidates for Supreme Court Justices and other authorities to maintain and uphold the dignity and behavior of judges [5]. With these two authorities, the Judicial Commission must be able to utilize the use of Information Technology (IT). Utilization of IT aims to make public services easily and cheaply accessible to the public. With the increasing use of ITs in carrying out their authority functions, making information security issues an important aspect.

The Center of Analysis and Information Services (Palinfo) is a center with three functions, namely the Analysis section, the Information Services section, and the Data/IT section. The Analysis section manages the analysis of decisions. Information Services section implements management and control of information relating to the internal use of the government and the general public. The Data/IT section manages and controls the information and communication technology sector. The Center of Analysis and Information Services closely related to the management of information systems that are used to process organizational data internally and information systems relating to public services. For this reason, information security awareness is very important to be carried out within the Center for Information Services and Analysis.

The background of this research stems from information security issues in the Judicial Commission that were not as expected. We divide the problem into 3 aspects, namely organization, inadequacy, and people. From the organizational aspect, the problem that occurs is that not yet implemented a comprehensive information security management system policy and not yet implemented ISO 27001 regarding information security in all sections. From the aspect of inadequacy, the problems that occur are lack of training on information security, lack of security of access to information in each room, and lack of knowledge regarding the importance of information security. And from the aspect of people, the problem that occurs is that there has not been much socialization to improve employee information security understanding, and Measuring the level of employee information security awareness has never been carried out. From the background of this problem, the thing that most concerns the researcher is the problem in the aspect of people, namely the measurement of employee awareness of information security has never been carried out. We need measurement of information

security awareness level to be carried out to determine the level of awareness of Judicial Commission employees, especially Palinfo, which level they are at. We can see the background of the problem in the fishbone diagram in Figure 1.

Therefore, the research needed to measure the level of information security awareness to identify the focus area of information security which still needs to be improved to develop a strategy for information security awareness methods. Many frameworks are used to measure information security awareness. We finally chose The Knowledge Attitude Behavior (KAB) theory developed by Kruger and Kearney (2006) and AHP (Analytic Hierarchy Process). KAB theory has often been used as a model for measuring information security [3]. We chose AHP in this research because of its superiority in terms of decision making and accommodation over attributes both qualitative and quantitative. Besides, AHP decision making able to provide more consistent results, easy to understand and use [6].

The purpose of this research is to measure the level of information security awareness among employees at the Center of Analysis and Information Services (Palinfo) of the Judicial Commission Republic of Indonesia. The author would like to measure the level of information security awareness of employees and recommend increasing information security awareness in the Center of Information and Analysis Services (Palinfo) of the Judicial Commission Republic of Indonesia.

The systematic writing of this paper consists of Introduction that contains background topic selection in the paper, Literature Review that contains theories related to selected topics, Research Methodology which contains the methodology used and the results, recommendations and conclusions of the research.

2. Related Works

Various studies related to the measurement of information security awareness have been carried out by several researchers, especially in Indonesia. Sari et al. (2014) conduct an information security awareness study for smartphone users. In this study, they developed the KAB framework. The KAB model that they use only takes on the dimensions of knowledge and behavior. Then the data they have obtained from the dimensions analyzed using the CFA model [7].

In the following year, Sari et al. (2015) conducted a similar study of smartphone users. However, there are differences with previous research. They use the KAB framework with dimensions of knowledge, attitude, and behavior. Then they do the analysis using AHP calculations [8].

Sari et al. conducted research using the same method as the researchers, the KAB and AHP methods. The difference with researchers, Sari et al. studies smartphone users while researchers study government employees in Indonesia.

Other research has been conducted by Kusumawati (2018) who researched government agency employees in Indonesia. This research uses the KAB model and MCDA calculation method. This study uses 5 focus areas [9]. The difference in research conducted by Kusumawati (2018) with researchers is that researchers used 7 focus areas and AHP calculation methods.

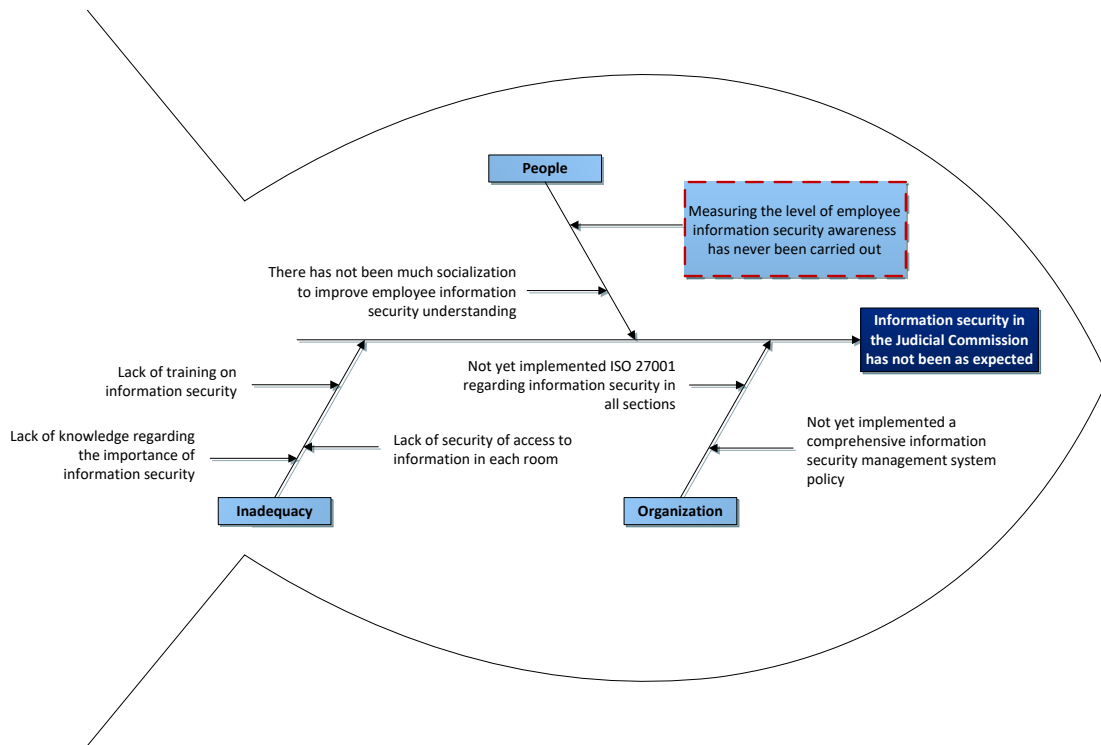


Figure 1: Fishbone Diagram Analysis

Subsequent studies have been conducted by Puspitaningrum et al. (2018) which used as the main reference for researchers. Puspitaningrum et al. (2018) conducted a study of SDPPI employees under the Ministry of Communications and Information of the Republic of Indonesia. They use the HAIS-Q framework and AHP calculations [3]. The difference from the research conducted by the researchers is that the researchers do not use the KAMI Index framework and the researchers research employees within the Judicial Commission of the Republic of Indonesia. The researcher also made a comparison among information security awareness between non-Data/IT employees and Data/IT employees.

For the framework used in this study, researchers used research written by Lund (2018) for the use of the HAIS-Q Questionnaire which contained 63 questions divided between knowledge, attitude, and behavior, and 7 focus areas [10]. Examples of questionnaires can be seen in Table 3.

3. Literature Review

3.1. Information Security

Information security is the protection of data, information, and equipment from unauthorized parties so that the information resources remain safe from all types of threats and risks. Information is an important resource in an organization, used as a material for decision making. Because of this, information must be quality. The quality of information is determined by three factors namely relevance, timeliness, and accuracy [11].

It may also be interpreted that Information is a description, statement, concept, and sign that contain values, meanings, and messages, whether data, facts or explanations that can be read, heard and seen in various forms in according to the times [12].

Information security means protecting data or information systems from prohibited use or access, and also focuses on maintaining the integrity, confidentiality, and availability of various information related to where information is stored on electronic media, paper, or other forms [12].

3.2. Information Security Awareness

According to NIST (2011) Information Security Awareness is a condition where the concern focused on information security problems. It can also be interpreted as using Information Security Awareness as a bulwark of a company in the face of current information security threats [13].

Information Security Awareness also defined as a situation in which people have a responsibility to use information derived from knowledge about information security that has been obtained. The person must also be aware of the importance of information security goals, threats, and risks. [14].

Information Security Awareness can be measured using the Human Aspect of Information Security (HAIS-Q) instrument. HAIS-Q can measure information security behavior and its validity has been recognized by many studies [15].

3.3. HAIS-Q (Human Aspects of Information Security Questionnaire)

HAIS-Q (Parsons et al., 2013) is a tool that could be used to measure employee knowledge, attitude and behavior, namely KAB Component. KAB is a benchmark for organizations that can solve various problems. For example, the use of KAB to determine the condition of an organization's information security and the use of KAB for making an organization's information technology strategy. HAIS-Q has seven focus areas including Password Management (PM), Email Use (EU), Internet Use (IU),

Social Media Use (SMU), Mobile Devices (MD), Information Handling (IH), and Incident Reporting (IR). These focus areas have their sub-focus areas [16] as can be seen in Figure 2.

3.4. AHP (Analytic Hierarchy Process)

AHP is a model that uses human subjects who are experts in their fields to make decisions. The human subject is the only input in the AHP model. Expert criteria refer to people who understand the problem posed correctly. Because it uses qualitative inputs (human perception), this model can process qualitative things besides quantitative things. Make AHP as a comprehensive decision-making model, taking into account quantitative and qualitative matters immediately [17].

Based on Thomas L. Saaty (1990), AHP is a framework for making effective decisions on complex issues. AHP helps simplify issues and speed up the decision-making process [18]. AHP is a global framework that arranges variables into hierarchies, provides relationships and values for these variables so that decision-makers can consider them and provide alternative solutions [19].

Based on Taylor (2004), AHP is used globally in a variety of problem conditions in the private and government fields. AHP is a method used to facilitate the selection of criteria and provide ratings so it can facilitate decision making [20].

4. Research Methodology

To achieve the objectives of this study, we first conduct a literature review on theories related to the topic of this research. We then compare the various measurement models to find suitable

models for measuring information security awareness. Next, we finally selected the model that will be used in this study based on previous studies is the HAIS-Q model by Parsons et al. for a table of questions. HAIS-Q model has a detailed focus compare to the others. HAIS-Q measures 7 focus areas related to measuring of employee awareness levels for information security in the organization. HAIS-Q provides a questionnaire to identify the level of information security awareness [16]. The flowchart showing the research process can be seen in Figure 3.

4.1. Questionnaire Method

The questionnaire methodology contains 3 lists of issues. The first set of questions tests the knowledge factors, the second about the attitude factors, and the third about the behavior factors. These 3 factors questions were developed by Parsons et al. and compared to 7 focus areas in the HAIS-Q model. Research questions are answered in sequential order, with a clear declaration for each question in the questionnaire using a Likert scale, from 1 shows strongly disagree until 5 shows strongly agree.

4.2. Data Collection Method

Data collection was conducted from October 2019 to December 2019 at the Center for Analysis and Information Services of the Judicial Commission of the Republic of Indonesia. In data collection activities, researchers will conduct research on information security reporting data at the Center for Information Analysis and Services by providing questionnaires to 25 companies related to their security awareness.

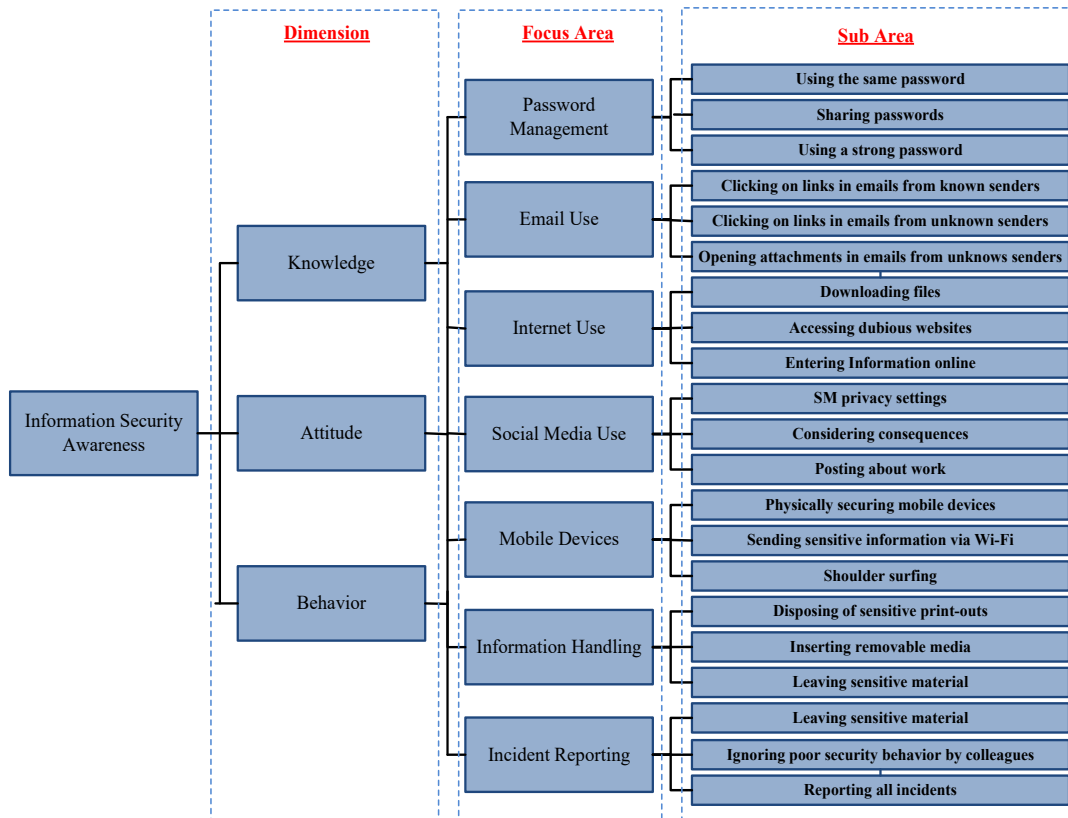


Figure 2: HAIS-Q Focus Area (Parson et al)

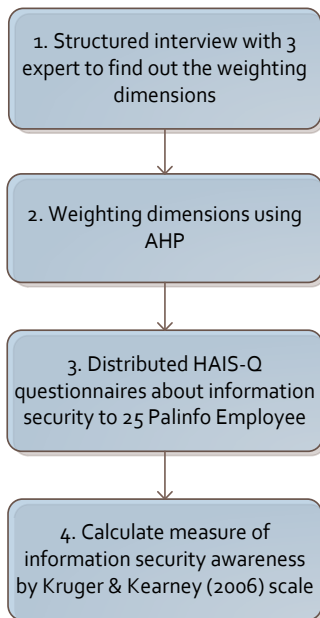


Figure 3: Research Process Flowchart

4.3. Measurement of Weight

At the first event, we asked people (experts) with have knowledge in the information security sector to fill the paired focus area matrix. In selecting most matrices, experts compare the important certain focus areas with other people. The level scale using scale 1 indicates the lowest level important, for 3 shows moderate important, for 5 shows strong important, for 7 shows very strong or demonstrated important and for scale 9 indicates the highest level important. The AHP process is used to gain information security awareness about the weight of each focus [19]. Experts fill the paired comparison focus area. The weight will then be ranked to find which focus areas have the highest information security awareness.

Next, at the second event, we calculated the scale of information security awareness after collecting questionnaires from employees. We determined the priority scale of 7 factors in HAIS-Q. While the preference scale used in each question in the questionnaire is a scale 5 which indicates the highest level (very aware) to scale 1 which indicates the lowest level (not aware) for each question in 7 HAIS-Q factors. Then we calculate the scale of 7 factors with percent of knowledge, attitude, and behavior factors. The scale obtained will be matched with a scale by Kruger & Kenney (2006) which divided into 3 levels: poor, average, and good [21] as can be seen in Figure 5.

5. Result, Discussion, and Recommendation

5.1. Result of Weighting Focus Area Dimensions

The research first, we create an AHP Hierarchy to determine the criteria used. AHP hierarchy can be seen in Figure 4. After determining the criteria, we conducted the study in an arranged interview with three experts to discover out the weighting results from seven focus area dimensions. The format of the pairwise criteria can be seen in Table 1. We then calculate the focus area that has been weighted by the expert using the AHP weighting with a comparison matrix formula. The results of the study show

that the focus area “Incident Reporting” was at first place with the highest weighting of 0,233278921, the focus area “Social Media Use” was ranked next with 0.229004904, the focus area “Information Handling” was in third place weighing 0,15646, the focus area was “Internet Use” was in fourth place weighing 0,131023552, the focus area “Email Use” was in fifth place weighing 0,115031643, the focus area “Password Management” was in sixth place weighing 0,0876223, and the focus area “Mobile Devices” was ranked the last with a total weight of 0,047578679 can be seen in Table 2. The focus areas for Incident Reporting, Social Media Use, and Information Handling are the highest. This is because the Center for Analysis and Information Services is closely linked to the management of information systems, which are used to process organizational data internally and information systems relating to public services, so that the three focus areas must be well managed so that all-important data are maintained.

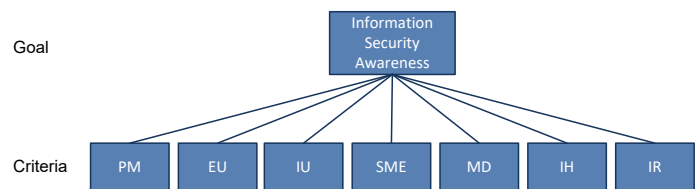


Figure 4: AHP Hierarchy

Table 1: Example AHP Pairwise Criteria

	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	
E U													X					I U

5.2. Result of Measuring Information Security Awareness

Questionnaires on information security were distributed after expert weighting of focus area dimensions. The research questionnaire was distributed to all 25 employees of the Center for Analysis and Information Services. Example questionnaire can be found in Table 3. The sample questionnaire was then collected for analysis of the data obtained. Respondent data show that the respondent’s work units are divided into sections on analysis, Information Services and Data/IT, each consisting of 8 persons. While the Administration Section consisted of only 1 person. More than half of the respondents held non-functional or general functional positions. The complete demographic of respondents can be seen in Table 4.

Table 2: Focus Area Weight Ranking

Focus Area	Weight	Ranking
Incident Reporting	0,233278921	1
Social Media Use	0,229004904	2
Information Handling	0,15646	3
Internet Use	0,131023552	4
Email Use	0,115031643	5
Password Management	0,0876223	6
Mobile Devices	0,047578679	7

Table 3: Example HAIS-Q Questionnaire

Internet Use							
Attitude	Knowledge	Behavior	SD	D	N	A	SA
While I am at work, I shouldn't access certain websites	Just because I can access a website at work, doesn't mean it's safe	When accessing the internet at work, I visit any website that I want to					

Table 4: Respondent Demography

Variable	List	Total	Percent
Work Unit	Analysis	8	32%
	Information Service	8	32%
	Data/IT	8	32%
	Administration	1	4%
Gender	Male	15	60%
	Female	10	40%
Age	21 – 30 years	6	24%
	31 – 40 years	16	64%
	41 – 50 years	3	12%
	51 – 60 years	0	0%
Position	Structural	2	8%
	Functional	6	24%
	Non-Functional	17	68%
Work Period	≤ 5 years	6	24%
	6 – 10 years	11	44%
	11 – 15 years	7	28%
	16 – 20 years	0	0%
	21 – 25 years	1	4%
Education	≥ 26 years	0	0%
	≤ SLTA/Equivalent	0	0%
	D-I – D-III	4	16%
	D-IV / S-1	16	64%
	S-2 / S-3	5	20%

To calculate the final measurements, weights and scales are used in Table 5. As explained by Kruger & Kearney (2006), the percentage of 30%, 20%, and 50% determined the weight and scale of information security awareness in this research for each dimension of knowledge, attitudes, and behavior [21].

Table 5: Weight and Awareness Scale (Kruger & Kearney, 2006)

Dimensions	Weightings
Knowledge	30%
Attitude	20%
Behavior	50%

The color map by Kruger & Kearney (2006) in Figure 5 is used to show in detail the level of awareness of information security in each focus area. The red color represents the level of "Unsatisfactory", the yellow color represents the level of "Monitor" which has potential needs to be repaired. Green represents the level of "Satisfaction".

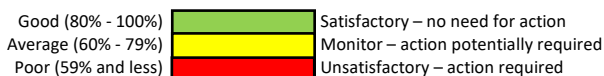


Figure 5: Scale of Information Security Awareness Colour (Kruger & Kearney, 2006)

The results of measuring the level of information security awareness in the Center of Analysis and Information Services are amount to 78.10 and included in the "average" level, which can be seen in Table 6. These findings indicate that the information security awareness of employees at the Center of Analysis and Information Services needs to be monitored regularly and action taken if needed.

Table 6: Level of Information Security Awareness of The Center of Analysis and Information Services

Focus Area	Knowledge (30%)	Attitude (20%)	Behavior (50%)	Total (%)
Password Management	82,08	79,04	79,04	79,95
Email Use	77,01	76,25	76,25	76,48
Internet Use	80,81	72,71	68,91	73,24
Social Media Use	78,03	78,28	78,79	78,46
Mobile devices	81,83	77,01	78,53	79,22
Information Handling	82,84	80,31	80,81	81,32
Incident Reporting	80,05	76,76	77,27	78,00
Total	80,38	77,19	77,09	78,10

The Center of Analysis and Information Services, as can be seen in Table 6, mostly indicates the level of "average" in terms of information security awareness. But there is an area that shows a "good" level of information security awareness, namely the "information handling" area. The area of "internet use" has the lowest weight, so it needs to be monitored more intensely. Therefore, this area requires attention monitoring to increase employee awareness. Internet use gets a low value on the behavioral dimension. Because maybe employees have the idea to open a website at working hours can become entertainment for them without considering work computers. They can contaminate with viruses through access to certain websites. What they don't know is that certain websites can carry viruses/malware that can turn off their work computers. For this reason, socialization is necessary where each employee must know the importance of maintaining information security. The Center for Information Services and Analysis also needs to develop a policy on Information Security. Not only made, the policy must be implemented effectively and must be understood by all employees. Policies must be easily accessible or available to employees to ensure that they will not ignore the policy. It should also be clear to all employees what their actual roles and responsibilities with regards to information security.

A study was also conducted to compare information security awareness among Data/IT employees. The results of measuring the level of employee awareness of information security of Data/IT are equal to 83,51 or categorized as a "good" level as can be seen in Table 7. For employees in the Data/IT section, out of a total of 8 people, 6 areas indicate the level of "good" information security namely "password management", "e-mail use", "social media use", "mobile devices", "information handling", and "incident reporting". Whereas there is only one area shows the "average" level of information security, namely "internet use". This result shows the level of information awareness among Data/IT employees is higher than that of all employees in the Center of Analysis and Information Services, the graph can be

seen in Figure 6. A better level of information awareness among Data/IT employees is possible because starting last year the Data/IT sector is implementing ISO 27001:2013 concerning information security.

Table 7: Level of Information Security Awareness of Data/IT Unit

Focus Area	Knowledge (30%)	Attitude (20%)	Behavior (50%)	Total (%)
Password Management	87,88	86,29	87,88	87,56
Email Use	82,33	81,54	83,13	82,57
Internet Use	84,71	79,17	76,00	79,25
Social Media Use	79,96	81,54	81,54	81,07
Mobile devices	88,67	83,92	85,50	86,13
Information Handling	86,29	84,71	84,71	85,18
Incident Reporting	86,29	80,75	81,54	82,81
Total	85,16	82,56	82,90	83,51

5.3. Discussion

5.3.1 Mapping Level of Security Awareness

Based on the results of the study, Data/IT employees received higher scores than employees of the Center for Analysis and Information Services (Palinfo) of the Judicial Commission of the Republic of Indonesia. These results can be compared to previous research conducted by Puspitaningrum et al. (2018) of SDPPI employees under the Ministry of Communications and Information of the Republic of Indonesia who receive an awareness value of 78,33. From the two research results it can be seen that Palinfo employees have lower information security than SDPPI employees. But the awareness of Data/IT employees are more aware than SDPPI employees. These results can help to map the level of information security awareness among government employees in Indonesia.

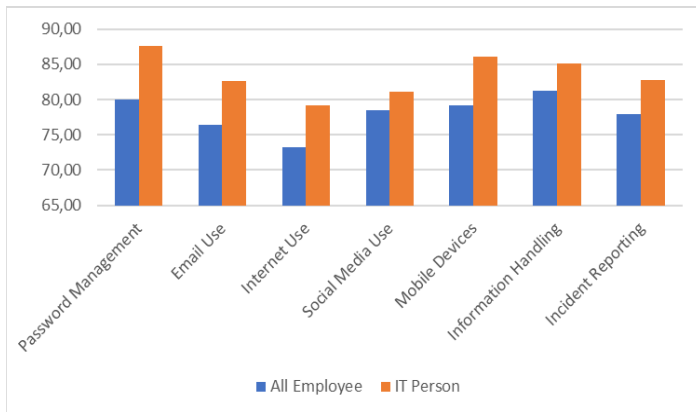


Figure 6: Level Comparison between IT Person and All Employee

5.3.2 Lesson Learned

Lesson learned is knowledge or understanding gained from experience that can be both success and failure. A lesson learned must be significant (or important, a dominant factor, the main cause) and have a real impact or be concluded that it is worthy of learning from an activity. The learning must be valid, factual,

technically correct and can be applied in the design, process, subsequent decisions to reduce or eliminate the potential causes of failure, problems whether predicted or not, setbacks, difficulties, bad luck and reinforcing results positive for example in terms of efficiency and effectiveness going forward.

In this research, lesson learned can be taken in the form of successful implementation of information security awareness. Lesson learned can be drawn from the results of information security awareness of employees in the unit of Data/IT that have shown a “good” level. Employees in the unit of Data/IT get a good result, certainly due to several factors. For this reason, researchers conducted additional interviews with the head of the Data/IT unit and Data/IT staff to find out the factors that led to the success of information security awareness in the Data/IT unit. Factors that led to the success of information security awareness in the unit of Data/IT can be seen in Table 8. These factors can be lesson learned for the Center of Analysis and Information Services (Palinfo) who still shows an “average” level awareness or lesson learned for other sections of the Judicial Commission that will implement information security awareness of employees and other organizations in order to successfully implement information security awareness as well.

5.4. Recommendations

The recommendation to increase information security awareness for employees at the Center of Analysis and Information Services is to create policies that can be applied to all focus areas, including:

- Policies about governing password security that include procedures that require employees to apply a password. Passwords must be at least 8 characters in length and a password must consist of numbers, symbols, capital letters, and lower-case letters. Employees are also required to keep their passwords confidential to anyone;
- Policies about governing the use of e-mail, including procedures requiring employees to be aware that not all emails they receive are safe;
- Policies about governing the use of the internet which include procedures for not providing access to employees to be able to download files freely. Also, policies governing employee access rights to certain websites and sanctions that must be applied if employees carelessly enter information about work on certain websites;
- Policies governing the use of mobile devices, including procedures that prevent the use of public networks for work purposes;
- Policies about governing the use of social media, including procedures for employees who cannot freely open social media accounts using office networks and there are sanctions that must be applied if employees carelessly enter information about work on their social media;
- Policies governing the handling of information, including procedures requiring employees to protect all forms of confidential work documents;
- Policies about governing incident reporting which include procedures requiring employees to report all forms of information security incidents that occurring at the

workplace and sanctions that must be applied if employees do things that jeopardize information security.

Table 8: Success Factors Data/IT Unit in Implementing Information Security Awareness

Dimension	Success Factors Data/IT Unit	Source
Knowledge	<ul style="list-style-type: none"> Data/IT employees have gained knowledge about information security based on ISO 27001 Data/IT employees already have knowledge of the rules for sharing passwords and the rules for using quality passwords Data/IT employees already have knowledge of the user's responsibility regarding email Data/IT employees already have knowledge of websites that should not be accessed and the consequences of using these prohibited websites. Data/IT employees already have knowledge of risks when using public networks Data/IT employees already have knowledge of USB that can store viruses/malware 	Interviews with Heads of Data/IT and Data/IT staff
Attitude	<ul style="list-style-type: none"> Data/IT employees already have responsibilities regarding the use of quality passwords Data/IT employees already have responsibilities regarding email security in the organization Data/IT employees already have a policy regarding the use of licensed software Data/IT employees already have responsibilities regarding the risks of using public networks Data/IT employees already have responsibilities towards outsiders visiting the office for interests in the Data/IT unit 	Interviews with Heads of Data/IT and Data/IT staff
Behavior	<ul style="list-style-type: none"> Data/IT employees have implemented information security procedures based on ISO 27001 Data/IT employees are already using passwords for personal use and using quality passwords Data/IT employees can distinguish safe and non-secure e-mail, and not open any link in the e-mail Data/IT employees are already using licensed software Data/IT employees are already using a VPN to work remotely Data/IT employees accustomed to doing regular backups of important data 	Interviews with Heads of Data/IT and Data/IT staff

Meanwhile, in terms of the information technology approach, we recommend raising awareness in focus areas that are still in the "average" area, especially in Palinfo. Our recommendations are:

- Encrypt sensitive documents/data, emails, and passwords. The recommendation is to increase the level of focus area level on e-mail use and password management;
- Routinely updating software, operating systems, applications, anti-virus, and firewalls. The recommendation is to increase the level of focus areas on internet use, e-mail use, mobile devices, and social media use;
- Use of VPN if the employee wants to access work e-mail from an outside place. This recommendation is to increase the level of focus areas on mobile devices and email use;
- Develop software that can assist employees in reporting information security incidents that occur. This recommendation is to increase the level of focus areas on incident reporting;
- Use spam filters on emails so that spam emails can be blocked. The recommendation is to increase the level of focus areas on e-mail use;
- Perform regular backups of sensitive documents/data using the correct backup procedures. The recommendation is to increase the level of focus areas on information handling;
- Access control over the use of the internet so that employees can only open websites that relate to work needs. The recommendation is to increase the level of focus areas on internet use;
- Creating a multi-layered room security using RFID technology. This recommendation is to increase the level of focus areas on information handling;
- Provides knowledge about downloading files and installing programs. The recommendation is to increase the level of focus areas on internet use and information handling; and
- Provides knowledge about information security standards that refer to ISO 27001. The recommendation is to increase the level of the entire focus area.

Strengthening information security awareness also requires socialization and training of employees about information security awareness, which is very important in organizations. Socialization can be done by various means, such as

- Socialization by sending e-mails to all employees;
- Socialization using media brochures distributed to all employees;
- Socialization by using banner media placed in strategic places which can be seen by all employees;
- Socialization by holding an open seminar attended by all employees;
- Socialization by placing advertisements on the Judicial Commission website so that employees are always reminded to continue to maintain information security. Training on information security also needs to be done, so that information security knowledge among employees increases and can be directly applied in the organization.

Several businesses, such as implementing policies, information technology, socialization and training, do need to be done. But apart from that, many other things need to be done. But apart from that, much more needs to be done so that the relevant preventive and corrective actions can be effectively applied. Learning and reflecting from the experience of organizations that have successfully developed the habit of obtaining information, the following examples are a variety of approaches that can be taken as preventive and corrective action: (1) Implement a system of rewards with a penalty (reward-punishment) for all staff and employees; (2) Top-down approach, where each leader will give instructions to his subordinates periodically to care for and implement information security procedures [22].

6. Conclusion

The results of calculating the level of information security awareness in the Center for Analysis and Information Services are at the “average/monitoring” level. This means that there are still many employees at the Center for Analysis and Information Services who do not understand the importance of information security. While the results of calculating the level of information awareness in the Data/IT section are at the level of “good/satisfactory”. Information security awareness in the field of Data/IT is better because employees in the Data / IT section have been certified ISO 27001: 2013 on information security. So they understand the importance of maintaining information security. We suggest several solutions for the Center of Analysis and Information Services to increase the level of employee awareness of information security, namely by making 7 policies, by using 7 technology approaches, by conducting socialization using 5 means of approach and by conducting training related to information security for employees. In addition, 2 approaches are also needed which can be done so that preventive and corrective actions can be applied effectively.

For future research, it is a necessary to organize research to measure information awareness among all employees at the Judicial Commission of the Republic of Indonesia, considering that information security is important not only for the Center of Analysis and Information Services (Palinfo) but also important for all employees at the Judicial Commission of the Republic of Indonesia.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this journal.

Acknowledgment

This study was supported by PUTI Prosiding 2020 grant with title “Evaluasi Master Data Management pada Organisasi, Universitas Indonesia”.

References

- [1] Information Security Awareness, ISO 27001:2013 Standard, 2013.
- [2] T.H. Purwanto, “Makalah KSI: Pentingnya Keamanan Sistem Informasi”, Fakultas Teknik, Universitas Muria Kudus, 2014.
- [3] E.A. Puspitaningrum, F.T. Devani, V.Q. Putri, A.N. Hidayanto, “Measurement of Employee Information Security Awareness: Case Study At The Directorate General of Resources Management and Postal and

- Information Technology Equipment Ministry of Communications and Information Technology” in 2018 Third International Conference on Informatics and Computing (ICIC), Palembang, Indonesia, 2018. <https://doi.org/10.1109/IAC.2018.8780571>
- [4] BSSN, Indonesia Security Incident Response Team on Internet Infrastructure/Coordination Center, Pusat Operasi Keamanan Siber Nasional, Jakarta, 2018.
- [5] Undang-Undang Dasar Negara Republik Indonesia Tahun 1945, UUD 1945, 2002.
- [6] R. Daneshvar, E. Turan, “Selecting The Best Supplier Using Analytic Hierarchy Process (AHP) Method”, African Journal of Business Management, 6(4), 1455-1462, 2012. <https://doi.org/10.5897/AJBM11.2009>
- [7] P.K. Sari, Candiwan, N. Trianasari, “Information Security Awareness Measurement with Confirmatory Factor Analysis” in 2014 International Symposium on Technology Management and Emerging Technologies, Bandung, Indonesia, 2014. <https://doi.org/10.1109/ISTMET.2014.6936509>
- [8] P.K. Sari, Candiwan, “Measuring Information Security Awareness of Indonesian Smartphone User” *Telkomnika*, 12(2), 493-500, 2014. <https://doi.org/10.12928/TELKOMNIKA.v12i2.2015>
- [9] Kusumawati, “Information Security Awareness: Study on a Government Agency” in 2018 International Conference on Sustainable Information Engineering and Technology, Malang, Indonesia, 2018. <https://doi.org/10.1109/SIET.2018.8693168>
- [10] P. Lund, “Information Security Awareness Amongst Students”, System Sciences, Luleå University of Technology, 2018
- [11] A. Kadir, Terra Ch. Triwahyuni, Pengantar Teknologi Informasi Edisi Revisi, Penerbit Andi, Yogyakarta, 2013.
- [12] R. Primartha, Security Jaringan Komputer Berbasis CEH, Penerbit Informatika, Bandung, 2018.
- [13] R. Kissel, NIST IR 7298 Revision 1, Glossary of key information security terms, National Institute of Standards and Technology, US Department of Commerce, 18, 2011.
- [14] F.J. Haecussinger, “Information Security Awareness: Its Antecedents And Mediating Effects On Security Compliant Behavior”, Georg-August-University Goettingen, 2014.
- [15] K. Parsons, D. Calic, M. Pattinson, M. Butavicius, A.M. Cormac, T. Zwaans, “The Human Aspects of Information Security Questionnaire (HAIS-Q): Two Further Validation Studies”, *Journal Computers & Security*, Volume 66, 40-51, 2017. <https://doi.org/10.1016/j.cose.2017.01.004>
- [16] K. Parsons, Mc Cormac, A. Butavicius, M. Pattinson, M. Jerram. “Determining employee awareness using the Human Aspects of Information Security Questionnaire (HAISQ)”, *Journal Computers & Security*, Volume 42, 165-176, 2014. <https://doi.org/10.1016/j.cose.2013.12.003>
- [17] B. Permadi, AHP, Pusat Antar Universitas, Universitas Indonesia, Jakarta, 1992.
- [18] T.L. Saaty, “How to Make a Decision: The Analytic Hierarchy Process”, *European Journal of Operational Research*, 48, 9-26, 1990. [https://doi.org/10.1016/0377-2217\(90\)90057-I](https://doi.org/10.1016/0377-2217(90)90057-I)
- [19] T.L. Saaty, “Decision Making with the analytic hierarchy process *International Journal of Services Sciences (IJSSCI)*, Vol. 1, No. 1, 83-95, 2008. <https://doi.org/10.1504/IJSSCI.2008.017590>
- [20] B.W. Taylor, “Introduction to Management Science”, Pearson Education Inc., New Jersey, 2004.
- [21] H.A. Kruger, W.D. Kearney, “A prototype for assessing information security awareness”, *Journal Computers & Security*, Volume 25, Issue 4, 289-296, 2006. <https://doi.org/10.1016/j.cose.2006.02.008>
- [22] R.E. Indrajit. *Keamanan Informasi dan Internet Edisi Kedua*, Preinexus, Yogyakarta, 2016.

The Application of Mobile Learning Technologies at Malaysian Universities Through Mind Mapping Apps for Augmenting Writing Performance

Rafidah Abd Karim^{*1}, Airil Haimi Mohd Adnan², Mohd Haniff Mohd Tahir², Mohd Hafiz Mat Adam¹, Noorzaina Idris³, Izwah Ismail⁴

¹Universiti Teknologi MARA, Perak Branch, Tapah Campus, Tapah Road, 35400, Malaysia

²Universiti Teknologi MARA, Perak Branch, Seri Iskandar Campus, 32610, Malaysia

³Universiti Teknologi MARA, Puncak Alam Branch, 42300, Malaysia

⁴Politeknik Ungku Omar, Ipoh, 31400, Malaysia

ARTICLE INFO

Article history:

Received: 09 April, 2020

Accepted: 07 June, 2020

Online: 25 June, 2020

Keywords:

Mind mapping applications

Mobile learning technologies

Mobile-assisted Mind Mapping

Technique Model (MMTM)

Writing performance

ABSTRACT

21st century learning focuses on the flow of information, media, and technology. In Malaysia, many university students face problems in English writing. Thus, students should be exposed to the technology training in innovative ways to produce students with a dynamic in this ever-changing world. Recently, the transformation and the evolution of mobile have created a huge impact on mobile users, as it is the current trend. Due to this matter, university students are now experiencing innovative learning development through mobile application and this can certainly improve their learning performance. The purpose of the study is to examine the application of mobile learning technologies through mind mapping applications for augmenting writing performance at Malaysian universities. The study was based on three different research theories -Flower and Hayes Writing Process Model, Radiant Thinking Theory, and Unified Theory of Acceptance and Use of Technology (UTAUT). The results of the study show that the students had positive responses towards English writing skills background, mobile technologies application background and mind mapping applications background. The proposed conceptual framework, Mobile-assisted Mind Mapping Technique Model (MMTM) supports the need for Malaysian university students to augment their writing performance. It is hoped that this study will benefit the policymakers, tertiary educators and university students in teaching and learning specifically in writing courses.

1. Introduction

The Industry 4.0 era offers another impetus for improvements in instruction and learning programs with momentous impacts, currently being a number one in ICT-related specialized training. The noteworthy impact of ICT-related advances on instruction and learning programs has gained the interests of teaching and learning experts. For them, these interesting developments will have a strong influence on the Education 4.0 initiative and learners should be readied to face opportunities and challenges that come with these developments. Though specialization, Education 4.0 needs to create future-oriented students with higher abilities in thinking

and reasoning. When the students complete their formal education, they need to be driven and ready to adapt to any kinds of disruptions in an authentic manner. Some of them would co-work with human as well as with machines. Considerably more basic than at any other time is the requirement for improved cooperation and coordinated efforts. To be relevant in a period of rapid disruptions, students and graduates must also develop aptitude that are related to self-directed improvements. that, students also need to master several soft skills. The skills, such as critical thinking, creativity, context problem solving, and cognitive flexibility are needed to thrive in the Industry 4.0. Mobile technologies have also become one of the smarter talents that ride through the wave of this era. Moreover, the thriving of this era has produced the accessible

^{*} Rafidah Abd Karim, Email: feida16@uitm.edu.my

various devices, applications, and apps which assisted the educators for instructional practices.

Thus, effective teaching and learning methods are very important in enhancing student's achievement in each course. As the technology advances today, various teaching and learning techniques and methods are introduced and applied in the field of education. The mobile convention is gaining more attention and is often applied in instructional methods especially in the higher education institutions. This paper presents the study on the mobile learning technologies through mind mapping applications for augmenting writing performance. Hence, it is critical to determine the relationship between the Mobile-assisted Mind Mapping Technique (MAMMAT) and students' writing performance. In this study, the MAMMAT is defined as a designed technique with Mobile-Assisted Language Learning (MALL) and mind mapping applications to employ students in learning and increase their motivation and performance in writing. Whilst, writing performance is a variable measured using the writing test and the scores are evaluated by the marking guidelines rubric.

1.1. Purpose

This purpose of the study was to examine the application of mobile learning technologies through mind mapping applications for augmenting writing performance among Malaysian universities. Specifically, the objectives of the study were:

- To investigate university students' English writing skills background.
- To investigate university students' mobile learning technologies application background.
- To investigate university students' mind mapping applications background.
- To propose a new conceptual framework for augmenting writing performance among Malaysian university students.

This paper presents the concept of mobile technologies application with the integration of mind map applications for developing the soft skills specifically in critical thinking, creativity, problem solving in addition to augmenting writing performance among university students.

2. Related Research Work

This section presents an overview of the related works within this study. Writing is a multiple task and process [1]. Brainstorming is a method to generate ideas. For example, students can brainstorm the ideas for thinking about a topic, to understand a topic and to decide the solutions. The brainstorming method offers numerous advantages for teaching and learning writing skills. This method can be used for assisting students to solve problems in learning. As related to the present research, students may use this method for solving problems in their essays. Before the actual writing stage is a pre-stage activity which often use the mind mapping application to brainstorm ideas. This stage refers to an action that urges students to compose as it animates their contemplations for beginning and moves them away from confronting a clear page for creating thoughts and gathering data for composing [2].

Brainstorming and mind mapping activities are common among students to help them in the planning stages of writing [3]. 128 students were investigated using a mixed method approach

based on the outcome of students' writing performance through mind mapping strategy [4]. He found that the students' performance in writing had improved due to the association of mind mapping and organizational pattern. This strategy was found to help students in developing positive attitudes for essay writing. In addition, another benefit of using mind maps for writing. They found that visual learners had a 40% higher memory rate compared to the verbal learners in writing tasks [5]. Evidently, this strategy provides several benefits for students and teachers in writing lessons.

Mobile and gadgets have the prospective to be featured as a medium of instruction in higher learning institutions. Therefore, various initiatives have been taken to diversify mobile assisted teaching approaches. A study revealed that 85 students that used the CASE tools improved their achievements in the classroom (Performance Expectancy). The tools also affected their behavioural intentions (Social Influence) to employ CASE tools in the place of effort expectancy [6]. This shows that students' performance can be enhanced by using technological tools in learning. The impact of technological advancement on education is enormous which advances the knowledge and the development of the teaching and learning environment. The effects of the teaching method using mind map sharing with digital archive data showed that this method improved the students' knowledge [7]. In addition, the utilization of mobile technology with mind mapping applications have improved the undergraduates' writing assignments [8]. In the case of this study, the paper focuses on the application of mobile technologies through mind mapping applications specifically. In this sense, some theoretical studies on writing and mobile technology were also discussed in the next section.

3. Theoretical Background

3.1. Flower and Hayes' Writing Process Model

In general, students find writing to be extremely difficult. This process involves several stages for instance, thinking, planning, drafting, and editing ideas [9]. The activities include cognitive and metacognitive dimensions of language learners. Additionally, many educators feel positive towards the process approach and think that students will greatly benefit from this approach [10]. The model by Flower and Hayes is really a reasoning-based procedural model of writing, which designates the complex nature of producing high quality written text. This model is built upon the basic premise of planning, writing, and revising, with a hierarchical progression, and goals developed and modified throughout the process.

Flower and Hayes then proposed another important cognitive model of the creative phase and [11] who reviewed and explained that there are three main features in this model of the process. The elements proposed are task environment, long-term memory, and writing developments.

Consequently, they state that the task environment can be described as writing assignments. Then, the assignments will be assessed, and they will contribute to the writer's achievement in writing. The second element, the long-term memory, refers to the background information which applies in the writing process such as topic, audience, and several ideas. Finally, the writing processes elements include organizing, decoding, and revising evaluated by

a monitor. The monitor will control the process of creating ideas, planning information, and setting objectives.

3.2. Radiant Thinking Theory

Mind maps have been considered as an all-inclusive yet widely spread field of study. Understanding the nature of the mind map concept is significant for this thesis' aim to contextualize the writing of undergraduates. Developed from Buzan's theory [12], the mind map is recognized as a learning instrument that can help learners to use their radiant thinking. The concept of mind map covers more brain potential. This delimitation appears to derive from the comprehensiveness defining the usage of mind maps such as the category of radiant thinking, colours, numbers images, exploiting lines and pictures or keywords. Thus, this theory emphasizes how to exploit radiant thinking in using the mind map for learning as in Figure 1.



Figure 1: The Mind Map (Buzan & Buzan, 1996)

Radiant Thinking was introduced as the underlying theory and philosophy behind mind mapping [13]. This theory is called Branching Association Machine (BAM), which describes the mind as always radiating from a central image. In addition, every image has a never-ending chain of branching patterns or images away from this central image. Buzan also highlighted some of purposes of mind maps in his study such as, to form an idea on a broader topic, to help learners to plan and decide ideas, and to help learners to collect and organize data. By so doing, learners will develop their efficiency in writing. They will be able to engage in reflective reflection and stimulate their thinking and memory in reflective reflection and stimulate their thinking and memory. Mind map as literally a map of the mind [14].

3.3. Unified Theory of Acceptance and Use of Technology (UTAUT)

Several researchers have addressed users' applications of new advances. They presented three constructs to the Unified Theory of Acceptance and Use of Technology (UTAUT) [15]. His team of researchers listed the constructs as hedonic drive, price value and habit which are suitable to be used for technology users. This study proposes to add context-specific constructs to this theory. Similarly, another study by [16] who added another construct for this theory which is an Information Technology (IT) as a

component to distinguish work, especially for the organizational work setting. This theory has a potential to be explored further in future research. In other words, there are several alternatives for additional constructs to be explored and added to this theory. Other researchers then explored further on this theory in academic settings. The investigations are to determine the different results from this theory when they focus on students in different academic settings. These research efforts suggest that other detailed constructs should be added to the model.

From that point onwards, seven new constructs should be the direct determiners for this model based on the review of the models. Conversely, the researchers found that the other three constructs did not contribute directly to the use of technology. However, they accepted another four constructs, and these were retained in the model.

3.4. Conceptual Framework

For this study, a conceptual model framework called the 'Mobile-assisted Mind Mapping Technique Model (MMMTM)' was developed and proposed for augmenting Malaysian university students' writing performance. This designed and proposed framework was based on three different research theories; (1) the Flower and Hayes Writing Process Model (1981), (2) Unified Theory of Acceptance and Use of Technology (UTAUT) model (2003) and (3) the Radiant Thinking Theory (1996). The conceptual framework of this study investigates the students' English writing skills background, the mobile technologies application background and mind mapping applications background. Within this context, it shows the learners' background knowledge and their attitudes influence their writing performance using the mobile technologies application and mind mapping application.

In Figure 2, the framework describes the association of three main variables. Within this context, it shows the MAMMAT as an independent variable, the writing performance as a dependent variable, and the learners' background knowledge and learners' attitude as moderator variables.

Within this framework, learners' attitudes comprise of the students' attitudes towards the use of mobile-assisted mind mapping technique that can influence their writing performance. Whilst, learners' background knowledge comprises of the knowledge of writing skills, the knowledge of mobile learning application and the knowledge of mind mapping applications. Venkatesh proposed four core and fundamental paradigms: Performance expectancy, facilitating conditions, effort expectancy together with social effect. These constructs be the direct determinants of behavioural intentions and ultimately behaviour [17].

Besides, learners' background knowledge which is the knowledge of writing, knowledge of mobile learning and knowledge of mind mapping can also give positive effect to the learners' writing performance. Research by [18] suggest that long term memory is a part that stores author's information. The information rotates on the subject, composing process, target group, and general objectives and plans for carrying out the composing job which needed to be done. To augment students' writing performance, these learners must be given access to application of the proposed framework. Hopefully, this framework will be able to augment their writing performance.

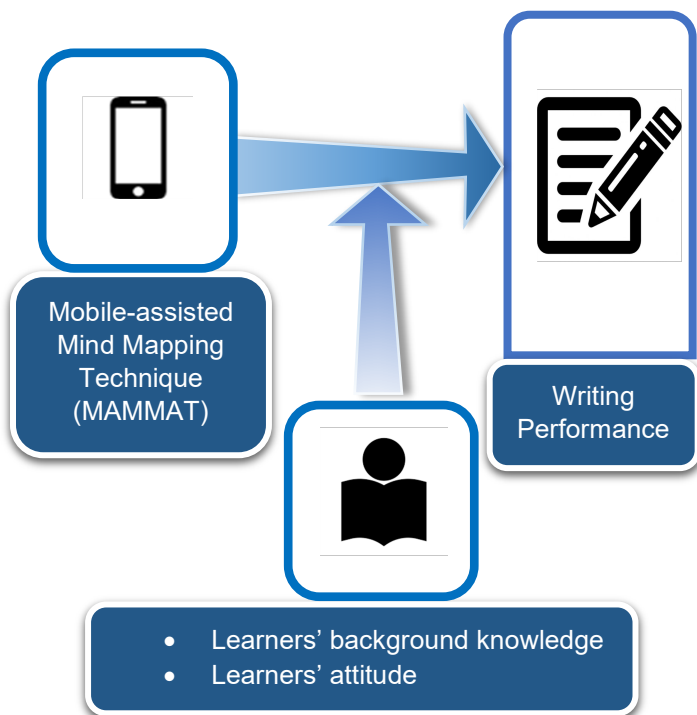


Figure 2: Conceptual Framework for Augmenting Writing Performance

4. Research Methods Adopted

This study utilizes a quantitative method which examine Malaysian university students' application of mobile learning technologies through mind mapping applications for augmenting writing performance. A questionnaire was designed to investigate the students' English writing skills background, mobile learning technologies application background and mind mapping applications background. They were asked to answer 10 items of the questionnaire. The questionnaire was divided into three sections. The first section includes four items on English writing skills background to find about the students' background on English writing skills. The second section contains three items to find about the students' application of mobile technologies for learning. The final section includes three items to find about the students' application of mind mapping for learning. The questionnaire items for all sections were developed using Likert scale, multiple choice questions and rating scale. The questionnaire was piloted, and the reliability was at 0.78.

The questionnaire was disseminated to forty-four (N=44) university students in Malaysia. 7 of university students are male and the other are 37 females. The data were collected and analysed for about one semester. Based on the study, the following research question (RQ) was formulated for the purpose of the study. After the data from the questionnaire was gathered, they were analysed by determining the descriptive statistics and percentages for each of the questionnaire's items. The data was analysed using Statistical Package for the Social Science (SPSS) software version 23. At the end, the results were presented in the form of table and charts.

The following research question (RQ) was formulated for the purpose of the study.

RQ 1: What are the university students' English writing skills skills background?

RQ 2: What are the university students' mobile learning technologies application background?

RQ 3: What are the university students mind mapping applications background?

RQ 4: What is the proposed framework for Malaysian university students to augment their writing performance?

5. Results and Discussion

Tables and charts represent the results of the questionnaire's items. The frequency and percentage analysis demonstrated about the university students of English writing skills background, mobile learning technologies application background and mind mapping applications background.

5.1. English Writing Skills Background

In the first section, the questionnaire measured the respondents' English writing skills background. The respondents were asked to answer four items in this section; (1) enjoy writing in English, (2) level of English writing proficiency, (3) writing stage difficulty and (4) reasons from stop writing. The data then were analysed for investigating the university students' English writing skills background as shown in Table 1. The responses were given based on a four-point scale.

Based on Figure 3, most of the respondents had answered 'Somewhat agree' to Item 1. The data gathered from the respondents showed that 24 respondents or 54.5% had a positive attitude in English writing. There were also 12 respondents or 27.3% who had responded that they enjoyed very much writing in English. Meanwhile, only 7 respondents or 15.9% and 1 respondent or 2.3% had answered that they somewhat did not enjoy and that they did not enjoy at all writing in English. This result shows that the students mostly enjoyed English writing because English is an international language which is important for them to acquire the language.

Table 1: English Writing Skills Background

Item	Frequency	Percentage
1. Do you enjoy writing in English?		
Strongly disagree	1	2.3
Somewhat disagree	7	15.9
Somewhat agree	24	54.5
Strongly agree	12	27.3
2. How good do you think you are at writing in English?		
Average	17	38.6
Good	23	52.3
Excellent	4	9.1
Not good at all	0	0

3. Which activity of the writing process do you find it most difficult to carry out?		
Pre-writing (Brainstorming/Getting ideas)	26	59.5
Writing (When writing/Drafting)	15	34.1
Revising (After writing)	1	2.3
Editing (When revising)	2	4.5
4. What usually stops you when you are writing an essay?		
I have no ideas/ points for my essay	26	59.1
I cannot find the right word/ expression in English	14	31.8
I do not know how to spell the word in English	0	0
I cannot find the right linking words and sentence connectors to connect my points/ideas	4	9.1

writing. None of students said that they are excellent in writing English. 4 respondents (9.1%) answered that they were ‘Excellent’ in writing English. None of students said they are not good at all in writing English, showing that there was a need for instructors to find other effective strategies in augmenting English writing abilities.

As shown in Figure 5, most of the respondents had answered ‘Pre-writing’ for Item 3. Pre-writing is the stage for brainstorming or getting ideas before we start writing. The data showed that 26 respondents or 59.5% said that ‘Pre-writing’ was the most difficult activity of the writing process to carry out. While for actual writing stage, a writer needs to write or draft their essay. The chart shows that about 15 respondents or 34.1% had responded that actual ‘Writing’ was the most difficult activity for the writing process. The next stage of the writing process is called revising stage or after writing stage. For this stage, the writer needs to revise their essay after the writing stage. For this stage, only 1 respondent or 2.3% had answered ‘Revising’ as the most challenging activity to carry out. Another activity that the writer will experience after their writing stage is the editing stage or when revising stage. 2 respondents or 4.5% had responded that the ‘Editing’ process was the most difficult activity of the writing process. This finding highlighted that most of the respondents (59.5%) said that ‘Pre-writing’ was the most challenging activity within within the writing process.

Thus, it aligns with the study which is to discover the feasibility of the framework which incorporate the use of mobile technologies and mind map applications as a brainstorming strategy in pre-writing for augmenting university students’ writing skills.

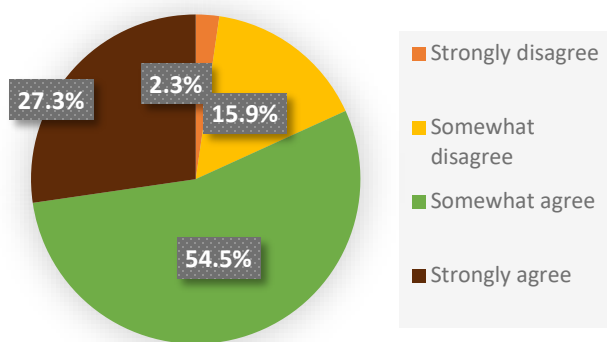


Figure 3: Enjoy writing in English

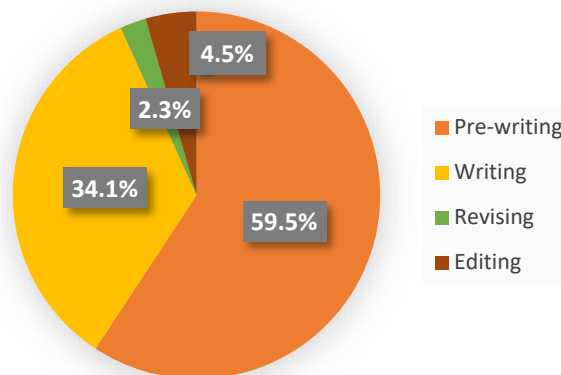


Figure 5: Writing stage difficulty

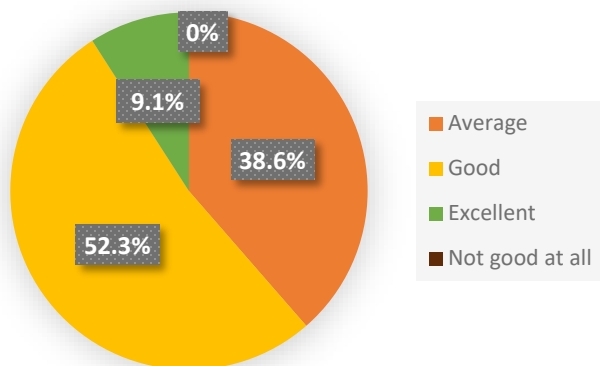


Figure 4: Level of English writing proficiency

Figure 6 below shows that 26 of the respondents (59.1%) answered that ‘I have no ideas/ points for my essay’. There were also 14 respondents or 31.8% who had responded that they cannot find the right word or correct expression in English that usually stops them when they are writing an essay. Only 4 respondents or 9.1 % said that ‘I cannot find the right linking words and sentence connectors to connect my points/ideas’. None of students said they do not know how to spell the right word. From this finding, it can be inferred that most respondents need more creative teaching techniques to help them to brainstorm ideas to start their essays. Therefore, this proposed framework is in the right path to provide students with a new technique that might be able to assist them in getting ideas to start their English language essays.

In Figure 4, the majority of the respondents answered ‘Good’ to Item 2. The data gathered showed that 23 respondents or 52.3% thought that they were good at writing in English. Whilst, 17 respondents or 38.6% felt that they were average in English

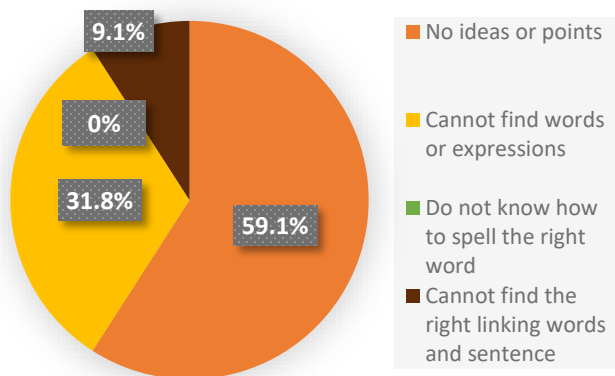


Figure 6: Reasons for stop writing

applications or ‘apps’ in the market that provide a lot of learning activities and opportunities.

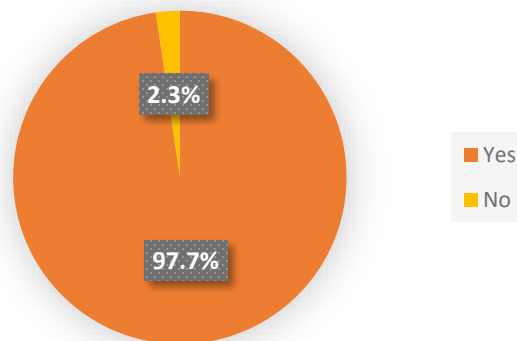


Figure 7: Use mobile devices for learning

5.2. Mobile Learning Technologies Application Background

In this section, the information collected was about the respondents’ mobile learning technologies application background as shown in Table 2. The respondents were asked to answer three items from this section; (1) use mobile devices for learning, (2) categories of mobile learning device, (3) hours per day usage. The data then were analysed for investigating the university students’ mobile technologies application background.

Table 2: Mobile Learning Technologies Application Background

Item	Frequency	Percentage
5. Do you use your mobile devices for learning?		
Yes	43	97.7
No	1	2.3
6. What mobile device do you usually use for learning?		
Mobile phone	9	20.5
Smart phone	35	79.5
Tablet	0	0
Personal Digital Assistant (PDA)	0	0
Other	0	0
7. On average, how many hours per day do you spend using the mobile devices for learning?		
None at all	1	2.3
An hour or lesser	8	18.2
One to three hours	23	52.3
Three to seven hours	12	27.3

Figure 8 below describes the categories of mobile learning devices that the respondents used frequently. Based on the chart, the majority of the respondents (35 respondents) or 79.5% used ‘smart phones’ as the mobile device for their learning. 9 respondents or 20.5% said that they relied on ‘mobile phones. No respondents state for ‘tablet’, ‘personal digital assistant (PDA)’ and ‘other’ type. Thus, the finding showed that the smart phone was chosen as the most common device amongst the respondents. The reason behind this current trend is possibly due to the function of smart phones not only as a telephone but it provides some functions like a miniature computer, internet access and GPS, and it generally uses touchscreen technology.

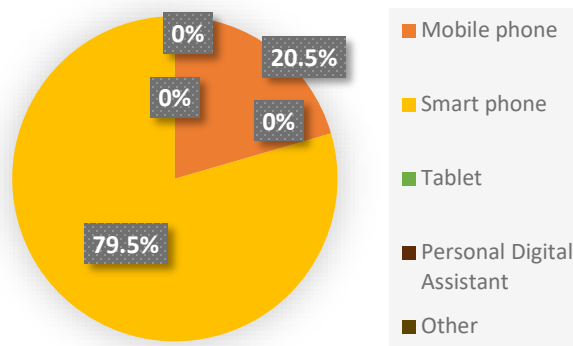


Figure 8: Categories of mobile learning device

In Figure 7, 43 respondents or 97.7% said that they experienced mobile learning. Only 1 respondent or 2.3% answered ‘No’ for Item 5. This remarkably showed that mobile devices were indeed very popular among the respondents and they commonly used these gadgets for learning purposes. This trend is becoming widespread most likely due to an increase in the number of mobile

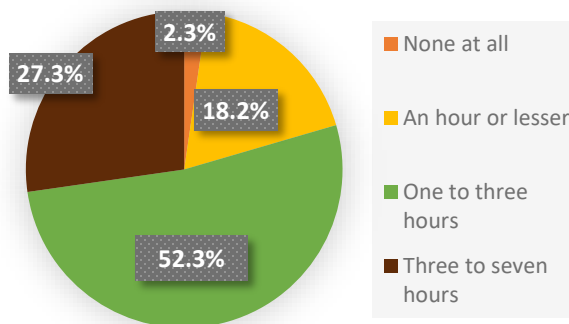


Figure 9: Hours per day of usage

From Figure 9, most of the respondents had responded ‘Three to seven hours to Item 7. The data gathered showed that 23 respondents or 52.3% said that they spent ‘Three to seven hours’ on average per day in using mobile devices for learning. Whilst, the data showed that 12 respondents or 27.3% used mobile devices for ‘One to three hours’ on average per day. Only 8 respondents or 18.2% utilised their mobile gadgets for ‘An hour or lesser’ on average per day for learning. Only 1 respondent or 2.3% of students answered, ‘None at all’.

Based on this finding, it can be concluded that most of respondents spent their time with their mobile device for learning or studies. The reason for this was likely to be that the respondents found that it was much more convenient and portable to exploit mobile devices for knowledge compared to their laptops or notebook computers.

5.3. Mind Mapping Applications Background

For this section, the respondents were required to answer a total of three items from this section. The information collected from this section; (1) like to use mind mapping applications, (2) mind mapping for learning, (3) level of mind mapping skills as shown in Table 3. The data then were analysed for investigating the university students’ mind mapping applications background.

Table 3: Mind Mapping Application Background

Item	Frequency	Percentage
8. Do you like to use mind mapping?		
Strongly disagree	0	0
Somewhat disagree	5	11.4
Somewhat agree	20	45.5
Strongly agree	19	43.2
9. Do you use mind mapping for learning?		
Yes	33	75.0
No	11	25.0
10. How good do you think you are at using mind map?		
Average	17	38.6
Good	23	52.3
Excellent	4	9.1
Not good at all	0	0

As Figure 10 illustrates, 20 respondents or 45.5% indicated that they ‘Somewhat agree’ to use mind mapping. The remaining 19 respondents or 43.2% claimed that they ‘Strongly agree’ to use mind mapping, whereas only 5 respondents or 11.4% said that they did not like to use mind mapping. Nobody answered ‘Strongly disagree’ for this item. This meant that many respondents liked to use mind mapping but perhaps they were not being encouraged much to use this technique for their learning, especially in essay writing.

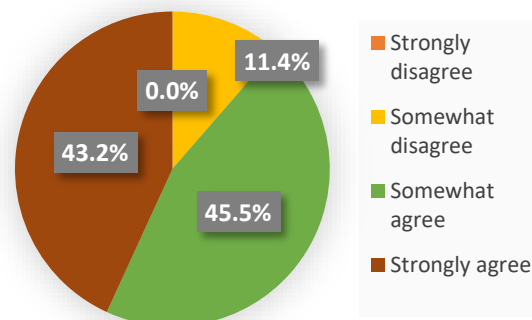


Figure 10: Like to use mind mapping applications

The results in Figure 11 clearly indicated that the majority of the respondents (33 respondents) said ‘Yes’ for Item 8. Only 11 respondents or 25% responded that they did not use mind maps for learning. This result indicated that the majority of respondents (75%) used mind mapping for learning, and this was essentially a common technique among the respondents.

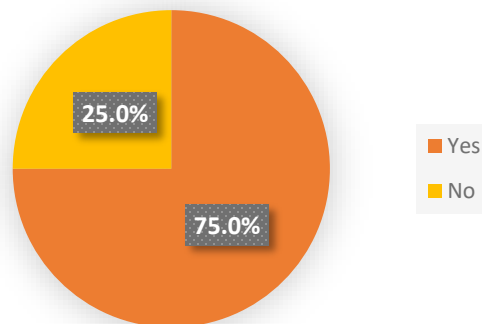


Figure 11: Mind mapping for learning

As shown in Figure 12, most of the respondents (23 respondents) or 52.3% believed that they were good at using mind mapping. About 38.6% of the respondents responded ‘Average’ to Item 9. Only 4 respondents or 9.1% had answered ‘Excellent’ at using mind mapping. This finding displayed that not many students were experts in using mind map for their learning. Therefore, most respondents still need some assistance from their instructors in guiding them to create mind maps.

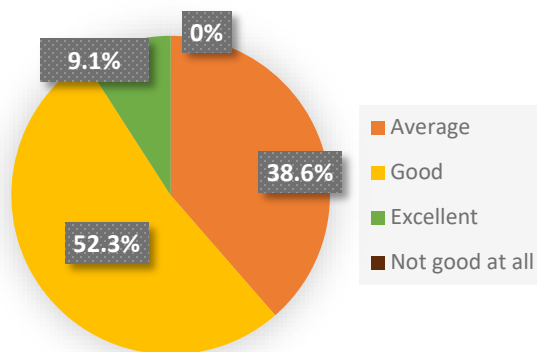


Figure 12: Level of mind mapping skills

6. Conclusion

Changing curriculum policies and solutions in accordance to the current trends to meet the growth of expertise in the sector is needed in moving towards Industry 4.0. This study shows that the utilization of mobile devices is best apt together with the use of mind maps because the mobile usage is very easy and flexible, accessible everywhere and students can learn effectively [19], [20]. Based on the empirical data, new mobile apps to augment writing skills could be developed by focusing on positive responses from the students' English writing skills background, mobile technologies applications background and mind mapping application background. Moreover, using gadgets could cultivate their skills and develop their writing performance. Therefore, the framework designed is optimised to serve a guideline for improving achievement in writing as Education 4.0 progresses.

Authorities and stakeholders within the education system together with educators on the ground should support and provide platforms and facilities to assist in the implementation of this framework such as advanced ICT and smart technology in institutions of higher learning. This framework has the potential to go a long way in helping students to become graduates who are well-versed in the era of Industry 4.0. To upgrade the appropriation of 21st century learning aptitudes by means of versatile innovation, analysts and educators must attempt to utilize portable learning innovation to improve students composing abilities [21], [22]. The utilization of mobile innovations ought to be supported for advanced Education 4.0 as it brings real possibility for the advancement of our instruction delivery systems [23]. Last but not the very least, further research ought to be done to churn out more knowledge with respect to the propagation of the current investigation to quickly propel university students' learning behaviours towards Industrial Revolution 4.0 preparedness.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] H. A. Alsamadani, "The relationship between Saudi EFL students' writing competence, L1 writing proficiency, and self-regulation." *European Journal of Social Sciences*, 16 (1), 53-63, 2010.
- [2] A. Seo, "The writing process and process writing." In J. Richards, W. Renandya (Eds.), *Methodology in Language Teaching: An Anthology of Current Practice* (pp. 315-320). Cambridge UK, Cambridge University Press, 2002.
- [3] J. Williams, *Preparing to Teach Writing: Research, Theory and Practice*, 3rd ed. New York NY, Lawrence Erlbaum Associates, 2003.
- [4] S. Hadeel, H. Al-Omari, "The Effectiveness of a Proposed Program Based on a Mind Mapping Strategy in Developing the Writing Achievement of Eleventh Grade 5 EFL Students in Jordan and their Attitudes towards Writing." *Journal of Education and Practice*, 5 (13), 88-110, 2014.
- [5] Adam, A., & Mowers, H. (2007). Get inside their heads with mind mapping. *School Library Journal*, 53(3), 24.
- [6] S. Dasgupta, M. Haddad, P. Weiss, E. Bermudez, "User acceptance of case tools in systems analysis and design: An empirical study." *Journal of Informatics Education Research*, 9 (1), 51-78, 2007.
- [7] J. Chang, P. Chiu, Y. Huang, "A Sharing Mind Map-oriented Approach to Enhance Collaborative Mobile Learning with Digital Archiving Systems." *International Review of Research in Open and Distributed Learning*, 19 (1), 1-24, 2018. DOI: <https://doi.org/10.19173/irrodl.v19i1.3168>
- [8] R. A. Karim, A. G. Abu, F. N. Mohd Khaja, "Theoretical perspectives and practices of mobile-assisted language learning and mind mapping in the teaching of writing in ESL classrooms." *Journal of English Teaching Adi*

- Buana, 2 (1), 1-12, 2017.
- [9] L. S. Flower, J. R. Hayes, "A cognitive process theory of writing." *College Composition and Communication*, 32, 365-387, 1981.
- [10] R. V. White, V. Arndt, *Process Writing*. London, Longman, 1991.
- [11] C. Bereiter, M. Scardamalia, *The Psychology of Written Composition*. Hillsdale NJ, Erlbaum Press, 1987.
- [12] Buzan, T. & Buzan, B. (1996). *The Mind Map Book: How to Use Radiant Thinking to Maximize Your Brain's Untapped Potential*. Dutton: New York. P. 59,166, 224, 225, 229.
- [13] Buzan, T. (1993). *The Mind Map Book*. London, BBC Books.
- [14] Buzan, T. & Buzan, B. (2003). *The Mind Map Book*. Revised Edition. London. BBC Press.
- [15] V. Venkatesh, J. Y. L. Thong, X. Xin, "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology." *MIS Quarterly*, 36 (1), 157-178, 2012.
- [16] Y. Sun, A. Bhattacharjee, Q. Ma, "Extending technology usage to work settings: The role of perceived work compatibility in ERP implementation." *Information & Management*, 46 (4), 351-356, 2009.
- [17] V. Venkatesh, M. G. Morris, G. B. Davis, F. D. Davis, "User acceptance of information technology: Towards a unified view." *MIS Quarterly*, 27 (3), 425-478, 2003.
- [18] L. S. Flower, J. R. Hayes, "Writing research and the writer." *American Psychologist*, 41, 1106-1113, 1986.
- [19] R. A. Karim, "Technology-Assisted Mind Mapping Technique in Writing Classrooms: An Innovative Approach." *International Journal of Academic Research in Business and Social Sciences*, 8 (4), 1092-1103, 2018. DOI: <http://dx.doi.org/10.6007/IJARBS/v8-i4/4146>
- [20] R. A. Karim, A. G. Abu, "Using Mobile-Assisted Mind Mapping Technique to Improve Writing Skills of ESL Students". *Journal of Social Science and Humanities*, 1 (2), 01-06, 2018.
- [21] R. A. Karim, A. H. M. Adnan, M. H. M. Adam, A. Zaidi, "Mobile Technology use in writing classrooms for higher education." In *MNNF Network* (Ed.), *Proceedings of the Int'l Invention, Innovative & Creative Conference, Series 1/2019* (pp. 197-202). Senawang, MNNF Network, 2019.
- [22] R. A. Karim, A. G. Abu, A. H. M. Adnan, A. D. J. Suhandoko, "The Use of Mobile Technology in Promoting Education 4.0 for Higher Education." *Advanced Journal of Technical and Vocational Education*, 2 (3), 34-39, 2018. DOI: <https://dx.doi.org/10.26666/rmp.ajtve.2018.3.6>
- [23] A. H. M. Adnan, R. A. Karim, M. H. M. Tahir, N. N. Mustafa Kamal, A. M. Yusof, "Education 4.0 Technologies, Industry 4.0 Skills and the Teaching of English in Malaysian Tertiary Education." *Arab World English Journal*, 10 (4), 330-343, 2019. DOI: <https://dx.doi.org/10.24093/awej/vol10no4.24>

A Survey and an IoT Cybersecurity Recommendation for Public and Private Hospitals in Ecuador

Maximo Giovanni Tanzado Espinoza*, Joseline Roxana Neira Melendrez, Luis Antonio Neira Clemente

Department of Computer Science, Universidad Politécnica Salesiana (UPS), Guayaquil, 010102, Ecuador

ARTICLE INFO

Article history:

Received: 12 April, 2020

Accepted: 06 June, 2020

Online: 25 June, 2020

Keywords:

IoT Cybersecurity

Health

Public Hospital and Private

IoT Architecture

Information Assets

ABSTRACT

It was analyzed the reference information on Cybersecurity architectures, models, standards, evaluations, mechanisms, and procedures applied to IoT domains, and public and private health area. The problem is the lack of proposals for IoT Cybersecurity in public and private hospitals to minimize random failures, ensure the privacy of personal data of patients, avoid the paralysis of the IoT medical network and minimize attacks on information assets. The objective is to perform a survey and an IoT Cybersecurity recommendation for public and private hospitals in Ecuador. The exploratory research was used to review references and specific analytical reasoning to end in a known scoop with a trusted solution. A survey of cybersecurity vs. competitiveness of hospitals in Ecuador resulted, a Model conceptual prototype of IoT Cybersecurity for a public or private hospital, an Architecture prototype of IoT Cybersecurity for a public or private hospital, and an Algorithm prototype of cybersecurity for IoT architecture. It was concluded that the cybersecurity standards applied to the design of IoT for a public or private hospital generates trust on information assets, preserves the confidentiality, integrity and availability of the information at the operational, tactical and strategic levels; the architecture prototype is between 59.38% and 99.71% of acceptable workload. This proposal is scalable and applicable to a public or private hospital regardless of the dimensions of areas, devices, floors, workers or other characteristics; the architecture only considers the hospital's own IoT devices and information; the devices of doctors or patients are not considered.

1. Introduction

Internet of Things (IoT) has significant popularity and growth in many areas and organizations, it is estimated that by 2030 there will be 125 billion connected devices [1].

On IoT network devices generate, deliver, monitor and process data; this data is sent and stored on private or public clouds; IoT is used in various areas such as sports, education, commerce, infrastructure, transportation and health [2]; other areas are factory, buildings, city, electrical networks, infrastructure and home [3]

In the health area, the integrity and availability of information for the care of patients have priority [2], while confidentiality guarantees the protection of information [4]; data encryption from device data delivery is required in this area..

Other research advises adopting standards, frameworks and best procedures to increase information security [5]; according to [6] to apply cybersecurity the following standards are used:

* Maximo Giovanni Tandazo Espinoza, Email: mtandazo@ups.edu.ec

International Organization for Standardization (ISO) 27001, National Institute of Standards and Technology (NIST), International Organization for Standardization 27032, International Information System Security Certification Consortium (ISC), Control Objectives for Information and Related Technologies (COBIT), Payment Card Industry Data Security Standard (PCI DSS), Health Insurance Portability and Accountability Act (HIPAA), Health Information Trust Alliance (HITRUST), North American Electric Reliability Corporation (NERC); this standards made it easier to apply an audit.

Cybersecurity includes the security of cyberspace and applies Confidentiality, Integrity and Availability to information assets, in addition to guaranteeing the privacy of the participants [4]; it is to protect the information assets to minimize the threats of the information processed, collected and transported by interconnected applications; is an element in information security [6]; cybersecurity is established on convergence of computing, engineering, information systems, networks, human and political elements [7].

IoT Cybersecurity is a strategic mechanism for improvement and changes in IoT and increases the environment security; it also ensures that an infrastructure has connected devices in a safe environment and with appropriate use by users [8].

eHealth is an area coming from the intersection between medical informatics, public health and companies, here health and information services are collected, processed, delivered and improved through the Internet [4]; with the use of ICT, medical care is improved at the local, regional and global level [9].

Basic characteristics of IoT are: comprehensive information collection, reliable transmission of information, information processing and data transformation of medical system [8].

According to [10] medical knowledge doubles every 73 days, this makes health data valuable; in addition, there are more efforts to ensure the integrity and access to patient records; among the main attacks on IoT infrastructure are: Denial of Services (DoS), remote brute force attacks, man-in-the-middle, password tracking, trojans and data manipulation.

The problem is the lack of proposals for IoT Cybersecurity in public and private hospitals to minimize random failures, ensure the privacy of personal data of patients, avoid the paralysis of the IoT medical network and minimize attacks on information assets.

The health sector must maintain the historical information of the data generators; this data must be stored, processed and visualized through the infrastructure with efficiency and security for the hospital and service providers.

Why is an IoT cybersecurity analysis and recommendation necessary for public and private hospitals in Ecuador?

To determine the appropriate models or standards that provide security to an IoT environment, it is necessary to understand the security requirements in the design of IoT on health area.

The objective is to perform a survey and an IoT Cybersecurity recommendation for public and private hospitals in Ecuador.

References about IoT cybersecurity, IoT domains and health are: Anomaly detection of IoT cyberattacks [1], Cybersecurity of healthcare IoT-based systems [2], IoT Security Mechanisms for e-Health [3], Cybersecurity education and training in hospitals [4], A Novel Model for Cybersecurity Economics and Analysis [5], A comprehensive cybersecurity audit model [6], Identifying Core Concepts of Cybersecurity [7], IoT cybersecurity research [8], Evaluating EHR and health care in Jordan [9], Blockchain Secured Electronic Health Records [10], Campus IoT collaboration and governance [11], CyberSecurity: A Review of IoT [12], IoT solutions for health monitoring [13], Security and Privacy-Preserving Challenges of e-Health Solutions [14], Self-Service Cybersecurity Monitoring [15], Standardising a moving target [16], Cybersecurity Vulnerabilities, Attacks and Solutions in the Medical Domain [17], The governance of safety and security risks in connected healthcare [18], Framework for improving critical infrastructure cybersecurity [19].

The exploratory research is used for reference review and specific analytical reasoning to end in a known scoop with trusted solution.

The results are: Cybersecurity and competitiveness survey of hospitals in Ecuador, Model conceptual prototype of IoT

Cybersecurity for a public or private hospital, Architecture prototype of IoT Cybersecurity for a public or private hospital, and Algorithm prototype of cybersecurity for IoT architecture.

It is concluded that the cybersecurity standards applied to the design of IoT for a public or private hospital generates trust on information assets, preserves the confidentiality, integrity and availability of the information at the operational, tactical and strategic levels; the architecture prototype is between 59.38% and 99.71% of acceptable workload.

2. Materials and Methods

2.1. Materials

The researchers designed a system to detect and alert unusual attacks or activities on IoT network or distributed network over smart city by using two types of intrusion; the first one is installed on software to detect abnormal activities or behavior, the second one is installed on network to detect attacks, through monitoring it reduces the probability of assaults on the network [1]. The researchers proposed a 4-layer IoT architecture for the health area; the layers are sensors, network, services and, applications; they analyzed international standards and applied them to each layer according to the function, responsibility and scope of the architecture [2]. The researchers highlighted the security requirements in confidentiality, integrity and availability, for this they recommended using an ISO standard; they named open IoT architectures, here they focused on a 7-layer architecture, each called: things, acquisition, network, aggregation, centralization, warehouse and application; among its advantages it has authorizations, encryption and identification [3]. For minimize human errors, avoid data breaches and reduce vulnerabilities on health services; the researchers proposed a governance framework to adopt cybersecurity on health, the areas it covers are platforms, storage, software, data and people; cybersecurity approach in a hospital is based on laws, availability of services, recovery and adaptation for staff [4]. The authors made guidelines to obtain the cost and benefits of processes applied in cybersecurity; this model reviews the practices, standards and quantitative risk analysis, presents the impact on hardware and software assets [5]. To guarantee cybersecurity in organizations, the authors proposed an audit model; it serves to verify the strategy adopted to minimize risks, it also evaluates the security policy [6]. The interviewed experts affirmed that Confidentiality, Integrity and Availability are important and transversal concepts on cybersecurity [7]. It was reviewed 3, 4 and 5 layer IoT architectures; the most used layers are sensors or perception, network, services or middleware, application or business; they described the applications on health, transportation and smart domains; they also described the standards that are used on IoT cybersecurity [8]. The authors' recommendations were to update the health system, educate staff, connect the public health sector with the private, attach importance to the security and privacy of health data [9]. The benefits of hybrid blockchain were used on health data proposal, under the standard of protection and regulations with notification rights, access to information, transparency and data portability for patients [10]. The University of Texas produced a list of requirements to apply cybersecurity on IoT network; they applied it on a farm and in the parking area of the university campus; they affirm the need for time to identify connected devices, work on a governance model,

the development of a standard architecture for the organization, the use of NIST as a framework for cybersecurity [11]. IoT Cybersecurity in IoT domain was proposed; for this they used a 3-layer IoT architecture; the first perception layer captures the sensors, Gateway, guest computers; the second network layer includes Wireless Fidelity (WiFi), Global System for Mobile Communications (GSM), 2G, 3G and, other access connections; the third application layer comprises smart environments [12].

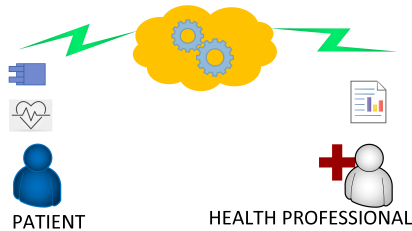


Figure 1: Architecture IoT for a patient

The researchers created an electrocardiography device for the service and monitoring of the patient's health in the IoT environment (Figure 1); the data is stored in the cloud of a provider, in addition the provider applies HIPAA for health data security [13].

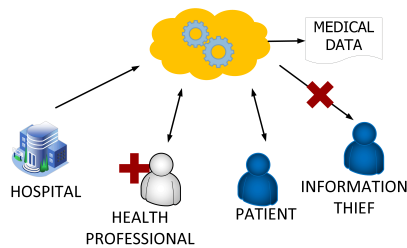


Figure 2: Architecture IoT for hospital

The researchers reviewed the privacy and security of various works, described various recording models for health data; the study presented a health architecture that has hospitals, patients and doctors, and also deposits its data in a cloud (Figure 2); they stated the challenges in the cloud are interoperability standards, expensive models, performance, privacy and, security [14]. In a process of construction, execution and monitoring on IoT domain, cybersecurity was applied to supervise and generate early alarms; the objective was to anticipate problems or attacks in the implementation through metrics [15]. The researchers analyzed cybersecurity standards applied to IoT, the standards are ISO 27000 series, GSM Association, Open Web Application Security Project (OWASP), Publicly Available Specification (PAS), machine to machine (M2M); several organizations applied IoT Cybersecurity based on law, consulting, technology companies, research centers, commercial organizations [16]. The researchers proposed a medical IoT architecture to safeguard information with characteristics such as: scalable, confidentiality keeping, general and efficient transmission; these characteristics are in areas such as: health files, medical systems, imaging systems, sensors, and information systems, and several medical systems have been tabulated [17]. The author analyzed the security correlation of medical devices that are used and interconnected in medical centers; she disclosed incidents, threats and vulnerabilities in medical devices, and analyzed the ISO, International

Electrotechnical Commission (IEC), Medical Device Regulation (MDR), NIST standards [18].

Table 1: Measurement of proposals on cybersecurity

Ref.	Proposal	Process	Model metrics
[1]	Detection method in IoT	Detect cyberattacks on Smart city nodes through learning algorithms to detect attack behavior	Performance 98%
[2]	Health system in IoT	Protects information with international NIST, ISO, PHI and HIPAA standards, uses 4-layer architecture	There are no metrics
[3]	Health system in IoT	It is used 7-layer architecture, security mechanisms and big data for data analysis	There are no metrics
[4]	Governance framework	Cybersecurity approach in a hospital is based on laws, availability of services, recovery and adaptation for staff	There are no metrics
[5]	Socio economic model	Measure the cost and benefit of cybersecurity, quantitative analysis, audits and standards.	There are no metrics
[6]	Audit model	Contains 18 domains, ranking formats	There are no metrics
[8]	Layered IoT models	Descriptions and types of attacks at each layer, types of cybersecurity standards by layer	There are no metrics
[10]	Hybrid blockchain	Data string with permissions to update information, use HIPAA	There are no metrics
[11]	Security framework on IoT domains	Uses NIST for governance and control of wireless and mobile communications	There are no metrics
[12]	Analysis of IoT 3-Layer Architecture	Cybersecurity application to protect authentication, information privacy	There are no metrics
[13]	Health IoT device	Generate and send data to the cloud, use data sending protocols and HIPAA	There are no metrics
[15]	Monitoring code	Processes for instantiating a cybersecurity infrastructure	There are no metrics
[17]	Architectures	Analysis of cybersecurity architectures in health domains	There are no metrics

The references with their models, processes and measurement are summarized for better understanding (Table 1), only one proposal presents the performance of the model, they carried out tests and others only present the models, other proposals use HIPAA and NIST to secure the information.

2.2. Methods

2.2.1. Scope of proposal

Applying the NIST Framework Core are cybersecurity activities, results, standards and best practices that are frequent in critical sections of the infrastructure; these activities pay off across the organization from the operational, middle and executive levels; the framework has five simultaneous functions: Identify, Protect,

Detect, Respond and Recover; these activities facilitate a high-level strategic approach to managing cybersecurity risks in an organization [19].

- Adopt NIST Cybersecurity Framework applied in [2] and [11];
- Adopt HIPAA used in [2] and the framework described in [19];
- The architecture oriented to strong Confidentiality, Integrity and Availability properties through ISO / IEC 27001 applied in [12];
- Adopt cybersecurity standards for layers [8];
- Establish a 6-layer architecture;
- The generated data is saved, processed and retrieved on cloud;
- Establish three user profiles: patients, medical professionals and hospital administration;
- Only hospital medical devices are considered.

2.2.2. Cybersecurity attacks

The types of security attacks were summarized from references (Table 2):

Table 2: Cybersecurity attacks

Ref.	Information collection attacks	Database attacks	Website attacks, Middleware or Application	Operation device attacks, sensing or network
[8]	Information damage level	Malicious scripts, unauthorized	Malicious insider, under infrastructure, virtualization threat, phishing, virus, trojan	Replay attacks, timing attacks, node capture, routing information,
[12]	Not considered	Not considered	Physical attacks, Malicious code injection, Spear-Phishing attack, Sniffing Attack.	Routing attacks, DoS, Data transit attacks
[17]	Operating System vulnerability, Open SSH vulnerability	Password intrusion, Vulnerability intrusion, SQL Injection	Cross-site scripting, cross-site request forgery, Cross-heterogeneous network attacks	Dropbear SSH Server, DoS
[18]	Unauthorized access	Uncontrolled distribution of passwords	Malware on systems	Malware on devices, DoS, software update

Attacks are independent of an organization, they are internal or external, they occur on devices, network, user access, software, any hardware, protocols, physical, logical; Attacks can deny or interrupt service.

2.2.3. Cybersecurity IoT elements

There is a relationship between health, IoT and cybersecurity; between IoT and Health the relationship is the application of security standards for devices and communication; between health and cybersecurity the relationship is the laws for the protection of patient information; between cybersecurity and IoT the relationship is the application of data protection standards.

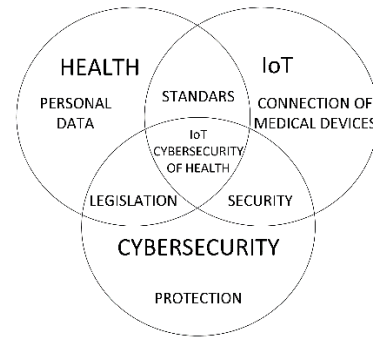


Figure 3: Elements to adapt IoT cybersecurity to health

As show in Figure 3 the connection between health, IoT and cybersecurity; in health we have the personal data of the patients; IoT includes medical devices that capture data over a network, in cybersecurity are the data protection standards.

2.2.4. Ecuador data

According to National Institute of Statistics and Censuses of Ecuador [20]: There were 175 public hospitals and 490 private hospitals in 2016; in 2017 there were 179 public hospitals and 466 private hospitals; in 2018 there were 183 public hospitals and 451 private hospitals.

In public hospitals: Attention to people in 2016, 2017 and 2018 was 752000, 780208 and 807245 respectively; in availability of spaces or beds in the years 2016, 2017 and 2018 was 12300, 13400 and 14144 respectively.

In private hospitals: Attention to people in 2016, 2017 and 2018 was 376000, 364000 and 357000 respectively; in availability of spaces or beds in the years 2016, 2017 and 2018 was 10600, 10100 and 9700 respectively.

In 2018, the provision of spaces or beds in the Ministry of Public Health is 40.47%; Ministry of National Defense is 2.30%; Social Security is 15.86%, other public is 1.97%, private non-profit is 10.6% and private for-profit is 29.03%; at the country level, the main areas of care in descending statistical order are: general services, medicine, gynecology, surgery, pediatrics, neonatology, traumatology, psychiatry, cardiology, urology, infectology, gastroenterology, otorhinolaryngology, ophthalmology and other services.

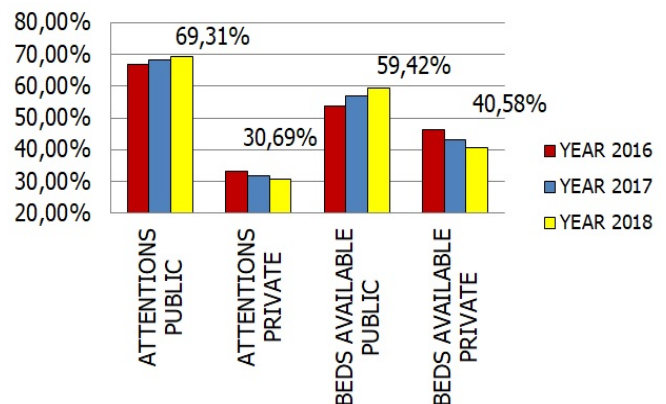


Figure 4: Attentions and public / private availability

In 2018 (Figure 4), public hospitals attended 69.31% of cases and a capacity of available beds of 59.42%; private hospitals attended 30.69% of cases and a capacity of available beds of 40.58%; the public sector each year has the greatest burden on medical care and spaces.

3. Results

The following results were obtained:

- Cybersecurity and competitiveness survey of hospitals in Ecuador
- Model conceptual prototype of IoT Cybersecurity for a public or private hospital
- Architecture prototype of IoT Cybersecurity for a public or private hospital
- Algorithm prototype of cybersecurity for IoT architecture

3.1. Cybersecurity and competitiveness survey of hospitals in Ecuador

The Latin American cybersecurity indices for 2017 [21], the cybersecurity indices for 2018 [22] and the competitiveness indices for 2017-2018 [23] were tabulated; to understand the impact and connection between these indices (Figure 5).

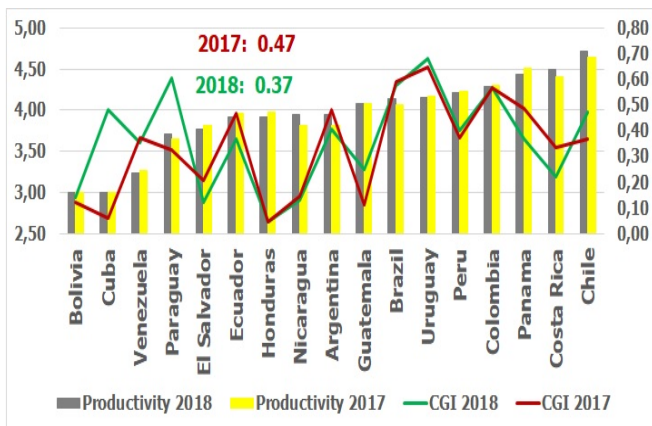


Figure 5: Cybersecurity Index vs. Competitiveness

2017 and 2018 of the Global Cybersecurity Index (CGI) and Global Competitiveness Report were considered; the first are the indicators of responsibility of the countries in the matter of cybersecurity, it was issued by the United Nations; the second is the set of institutions, policies and factors that establish the level of productivity, it was issued by the Economic Forum.

Cybersecurity values are between 0.01 and 0.99 on the secondary Y axis; the CGI of Ecuador in 2017 was 0.47 and in 2018 the CGI was 0.37; Ecuador is below 0.50; the CGI of Uruguay in 2017 was 0.65 and in 2018 the CGI was 0.68; Uruguay is highly committed to implementing cybersecurity.

Competitiveness values are between 0 and 5 on the main Y axis; Ecuador's competitiveness in 2017 was 3.96 and in 2018 it was 3.91; Chile's competitiveness in 2017 was 4.64 and in 2018 it was 4.71; Chile, Costa Rica and Panama are the first countries with the best productivity in 2018; Uruguay, Paraguay and Brazil are the first countries with the best cybersecurity application in 2018.

It is evident that the levels of competitiveness are not directly linked to the levels of cybersecurity.

Data from the World Health Organization (WHO) observatory were tabulated in the category of health information systems, among Latin American countries from 2008 to 2016; from each country (Figure 6) were obtained the score for the application of international health guidelines (HR), the average percentages of: reliability of the information systems in the Civil Registry for cause of death (CRCD), Civil Registry for births (CRBI) and data integrity due to death (INCD).

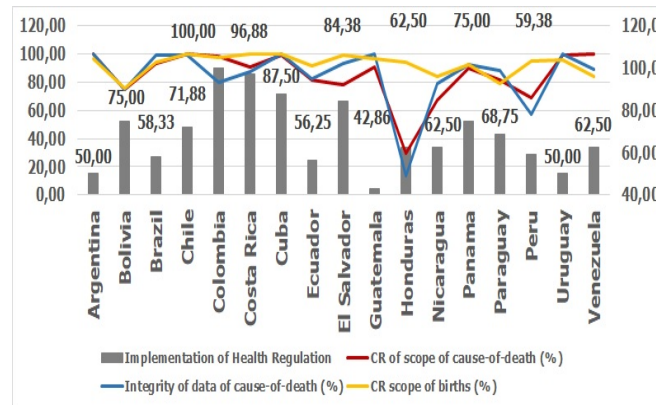


Figure 6: Health Regulations vs. Computer Registry

Colombia has 100 points in HR, in CRCD it has 98.06%, in INCD it has 80% and in CRBI it has 96.97, it follows that this country applies ICTs at an excellent level; Costa Rica has 96.88 points in HR, in CRCD it has 90.87%, in INCD it has 87% and in CRBI it has 99.66%, it follows that this country applies ICTs at an excellent level; Ecuador has 56.25 points in HR, in CRCD it has 81.77%, in INCD it has 82% and in CRBI it has 91.59%, it follows that this country applies ICTs at a good level.

In descending order by CRCD, Ecuador is in eleventh place; in descending order by INCD, Ecuador is in twelfth place; in descending order by CRBI, Ecuador is in thirteenth place; in descending order by HR, Colombia and Costa Rica are the first countries to apply health guidelines, Ecuador is in fourteenth place.

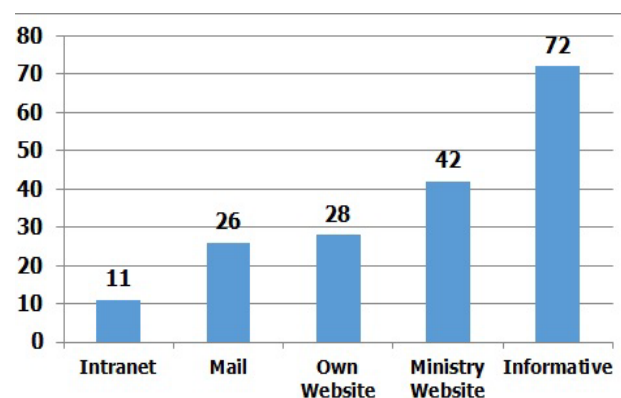


Figure 7: ICT in public hospitals of Ecuador

Review of infrastructure and hospital deficiencies in Ecuador

The web pages of 129 public hospitals were reviewed (Figure 7); it was found that 22 hospitals have a newsletter in pdf format on the website of the Ministry of Public Health of Ecuador (MSP), in addition it was considered that 50 hospitals have the same newsletter, 11 hospitals have their own intranet with institutional mail and their own website; 26 hospitals have institutional mail; 28 hospitals have their own website; 42 hospitals have their information attached to the MSP website; it follows that in 11 hospitals it is possible to implement a Cybersecurity IoT in the short term due to their existing intranet infrastructure.

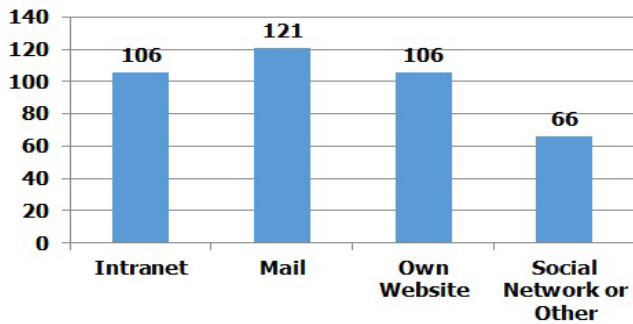


Figure 8: ICT in private hospitals of Ecuador

The web pages of 186 private hospitals were reviewed (Figure 8); here 106 hospitals have their own intranet; 121 hospitals use institutional mail; 106 hospitals have their own web page and 66 use social networks or third party pages for their communications; here it is possible to implement an IoT Cybersecurity to the 106 hospitals in the short term.

In the public sector only 8% of hospitals can apply IoT Cybersecurity to existing networks; 57% of hospitals can apply IoT Cybersecurity in the private sector; in the other hospitals, they must start from the design of the IoT infrastructure for the hospital; each IoT infrastructure depends on the physical infrastructure of the hospital; therefore we do not propose standard IoT network design; in Ecuador, the time and cost of applying IoT cybersecurity to hospitals depends on political, economic, cultural or social factors.

Critical review of existing IoT and cybersecurity measures

The MSP controls and regulates the implementation of the Ecuador National Health System of public and private entities; the public sector is made up of ministry hospitals, hospital of the armed forces, police hospital, social security, municipal care centers; the private sector is made up of a cancer society and private medicine; at the country level, hospitals are classified into 3 levels of hospitals; among the users are: the population without the ability to pay, the population with the ability to pay, members of the armed forces, members of the police, workers affiliated with social security and the population without social insurance [24].

Since 2014, the Public Administration has among its actions the implementation of a technological architecture and information security framework [25], according to references, there is no application of Cybersecurity to IoT in public or private hospitals.

The proposal [26] presented an IT Governance Model between Cobit and ISO 27002 to provide Information Security in Public Hospitals; aligns with IT and Hospital objectives in health data security; Ecuador's investment in Health is \$ 11 billion based on health law; It should be emphasized that Ecuador does not have a standard to safeguard the confidentiality, integrity and availability of health data.

3.2. Model conceptual prototype of IoT Cybersecurity for a public or private hospital

The NIST Framework is applied in organizations of any size and helps to understand cybersecurity risks, risk management and protect information assets; this framework has good practices for managing resources on cybersecurity protection issues.

The ISO 27001 standard has the information security guidelines to maintain the reliability, integrity and availability of the information, allowing the hospital to assess and mitigate risks; it also allows improving the competitiveness and image of the hospital.

HIPAA is used to protect the patient's private medical information and data, which is legislation that provides privacy and security provisions; this law is widely accepted due to the increase in violations of health information.

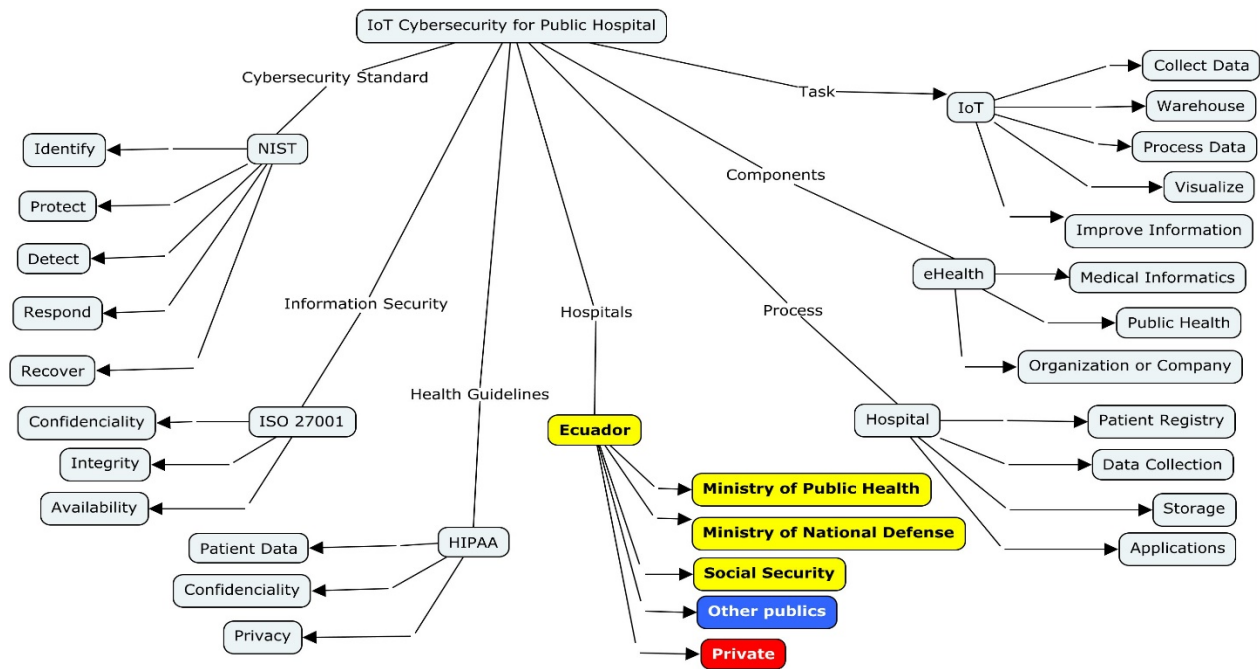
An IoT performs activities such as: Collecting patient data is done through the medical devices that generate the data and have access to the hospital's IoT; the collection is wired or wireless; storing patient data in a repository that has their access validated, may be their own space or a provider's space; process the data that was collected by the devices, the medical result is useful information for the health professional; improving information through correct management supports the health professional in the quality of the service; visualize the information through the applications and personalized systems on health area, indicators and reports are generally presented.

eHealth has components and medical software to manage the knowledge towards health professionals, the objective is to maintain the health care of the population; this care is carried out in a hospital that the government provides care to citizens.

The general information processes in a hospital occur from the patient arrival when taking and saving their personal and medical data; data collection is through the devices connected to IoT; storage guarantees the integrity of information and is presented in personal and medical applications or information systems.

As show in Figure 9 the cybersecurity components that apply to IoT of a public or private hospital, among its components are NIST, ISO 27001 and HIPAA standards; the other IoT, eHealth and hospital components manage the data and convert it into information and valid knowledge for health professionals.

The model assists in the management of medical results such as: diagnosis, prevention, follow-up, prediction, prognosis, treatment or relief of the disease, relief of an injury or disability, updating of physiological or pathological data, results of sample examinations, organ or blood donations.



3.3. Architecture prototype of IoT Cybersecurity for a public or private hospital

Cybersecurity was proposed at the time of designing an IoT for a public or private hospital, that is, not applying late security; each layer with its elements and functions to apply security management from the beginning and continuously; the hardware and software of the hospital and providers must operate and integrate in a reliable way. The architecture adopts standards to apply IoT hardware, software, sensors, control, storage, services and users; reduces implementation costs, reduces delivery times, and increases security levels. A NIST framework is used because it is a common language for dealing with security risks and ISO standard good practices are used.

The following layers were proposed:

3.3.1. First Layer: Medical Devices

This level gathers all the data from the devices through medical equipment, sensors and other equipment, captures the physical world and uploads it to the digital world; at this layer DoS attacks are very likely; according to [18] there are four groups of devices in the health area:

- Implantable medical devices, such as pacemakers, skin sensors and other implantable cardiac devices.
- Portable medical devices, such as portable insulin pumps.
- Mobile devices, such as glucose measuring devices or insulin pumps.
- Stationary medical devices, such as tomography scanners

Security requirements: raise the level of data confidentiality transferred between devices, for integrity devices must use encryption and device authentication to prevent the entry of strangers.

3.3.2. Second Layer: Network

The routing and exchange of data between devices must be in a reliable transmission from the first layer, to transmit data wireless, wired and communication protocols are used; this layer has wireless sensors, access points, and a gateway to transfer data to storage with high reliability; here the data is added, filtered and transmitted between the sensors. Security requirements: Sensors and nodes must be authenticated to avoid strange nodes, confidentiality and integrity are significant in the data.

3.3.3. Third Layer: Services

In this layer the confidentiality, integrity and authenticity of the transferred data are managed through the IoT architecture, the services are called from devices, sensors, servers and systems; it could be vulnerable to internal or underlying attacks; developed services must keep the application safe and transparent of the IoT network.

3.3.4. Fourth Layer: Interface

This layer presents custom applications according to the hospital and patients; it applies protocols and standards with functional systems for users; doctors, patients, administrators and providers access information through interfaces or indicators. Security requirements: User authentication for access, protection of user privacy, defining profiles, data processing and checking software vulnerabilities.

3.3.5. Fifth Layer: Cloud

Cloud Computing provides capabilities to create, store and retrieve patient information, here the hospital, medical care points, doctors and laboratories view / update patient data. The cloud provider provides analytics, access management, data protection, and service integration. It is recommended to use AWS to process, store and transmit health information; this cloud

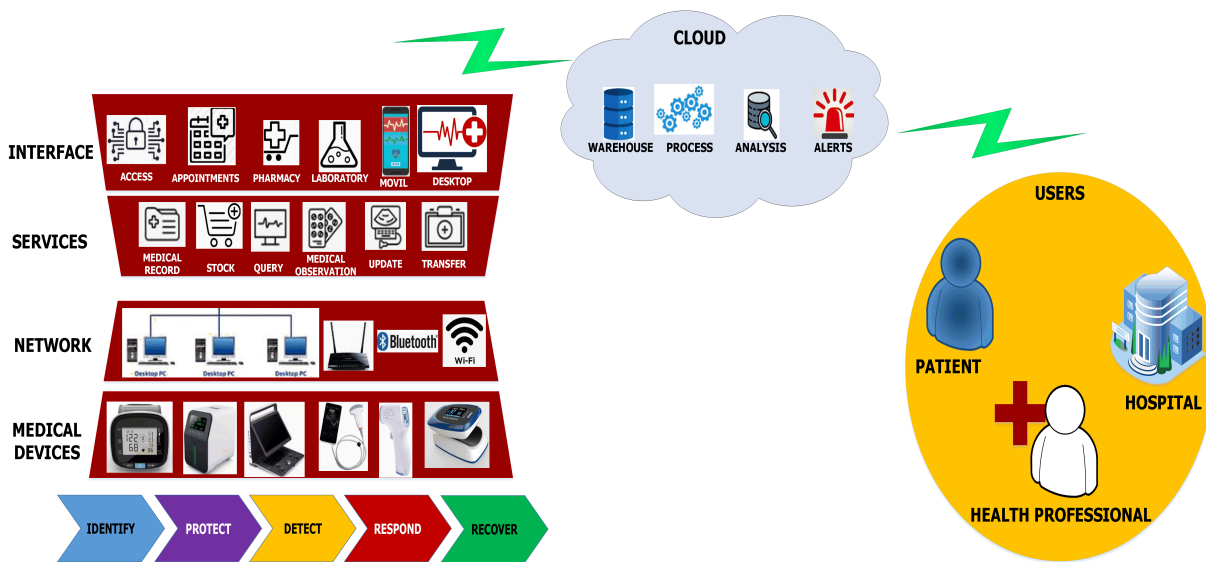


Figure 10: Architecture prototype of IoT Cybersecurity for a public or private hospital

allows HIPAA to be applied in a secure environment; the main characteristics of this cloud are: assignment of individual policies, roles and profiles; encrypted web content, registration and backup data; security groups to limit access to services; management of systems or web applications; encrypted MySQL database, registration, monitoring and alerts.

3.3.6. Sixth Layer: Users

Only three participants were considered in this layer: the patients that generated the data through medical devices, health professionals and the hospital; patients leave control of their health information on cloud servers, this is perceived by the patient as a threat to their privacy; as a requirement of this layer is the data integrity it stored; access to users can be by web, mobile or desktop application because the service layer allows delivery of information to any type of application.

As show in Figure 10 the importance of maintaining the confidentiality, integrity, availability, reliability and authenticity of user data on a public or private or hybrid environment; in addition, there may be people or groups interested in the health information such as laboratories, health providers and others.

Another component of the architecture is NIST Framework Core with its activities in infrastructure design; these activities are detailed below:

Identify: For identification, inventories of physical devices, systems and platforms are taken; data flow, business communication, external information systems also enter in the inventory, all organization resources are given a priority value, in addition third-party roles and responsibilities are identified; in addition to the organization, the mission, vision and objectives are identified, that is, the role of the organization in the health sector, the functions and critical services; in the evaluation, vulnerabilities are identified, information threats are reviewed, internal and external threats are identified.

Protect: The authentication and access category manages the credentials of devices, users and processes; in the training

category, hospital staff, providers, and patients receive cybersecurity education, roles, and responsibilities according to policies; in the data security category, it is managed to protect the confidentiality, integrity and availability of the information; the information assets in any state or transaction are protected; in the protection of information management, security policies are implemented to defend information systems and assets.

Detect: In the anomalies category, the data flow for users and systems is managed, determine impacts and collect data on the attack, in addition to establishing alert guidelines; in the monitoring category, the physical network, people activities, unauthorized codes, providers activities, and unauthorized devices are reviewed to minimize activities against cybersecurity.

Respond: Response procedures are planned for possible cybersecurity incidents; in the communication category, the staff knows the response activities to an event, incident information, passing information according to the plans; others activities are: investigation of the notifications, the impact of the incident, forensic analysis, classification of incidents and vulnerability management are carried out, mitigating incidents, lessons learned.

Recover: Recovery procedures are defined and executed to ensure the replacement of information assets in the event of an incident; these procedures should be improved with the lessons learned; in the communication category, the hospital's relationships and reputation are managed, and recovery activities are also communicated.

The ISO 27001 standard maintains the confidentiality, integrity and availability of information on phases of NIST framework.

Architecture features:

- Interoperability to avoid interruption of operations,
- Scalability to connect with smart devices,
- Storage to support data delivered by devices,
- Communication overload to support communication between nodes,

- Processing overhead to perform algorithms on data and deliver it as information,
- Resistance to failures affecting the network,
- An IoT network solves the limits of growth, access, space, data transfer and data transformation; with the features the health system / environment improves the category of information and services.

IoT cybersecurity applied to the health area supports services such as: medical information, individualization, hospital emergency, monitoring, delivery and supervision of medicines, medical equipment, medical waste tracking, blood management, infection review and others.

Formula to measure the architecture workload

A formula was proposed to evaluate the architecture based on the layers, quantities were established according to the components for each layer, the Equation (1) is:

$$Occ = ABS \left(\sin \left(\sqrt{\frac{(Qd * Qn) + (Qs * Qi)}{Qp * Qu}} \right) \right) \quad (1)$$

Here:

Occ = Occupation or workload of architecture; Qd = Quantity of medical devices; Qn = Quantity of receiving nodes; Qs = Quantity of services available; Qi = Quantity of interfaces; Qp = Quantity of processes in the cloud; Qu = Quantity of final users.

Ten scenarios were simulated with the six parameters of the formula (Figure 11).

These were grouped in pairs as follows: medical devices with receiving nodes at Hardware Level, available services with interfaces at Software Level and cloud processes with final users at Process Level; in the first scenario the hardware level is 80, the software level is 108, the process level is 36 and the workload is 59.38%; in the fifth scenario the hardware level is 135, software level is 128, process level is 9 and the efficiency is 76.78%; in the tenth scenario the hardware level is 80, software level is 110, process level is 195 and the efficiency is 83.44%.

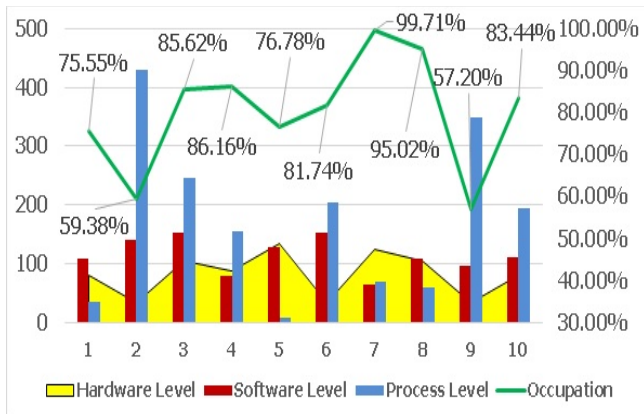


Figure 11: Workload of architecture prototype IoT Cybersecurity

As shown in Figure 11 the cross-line the workload of IoT cybersecurity in ten scenarios, in this simulation the minimum

value was 59.38% and the maximum value was 99.71%; the average workload of this simulation was 80.06%.

We can deduce that the increase in the number of medical devices and data receiving nodes together with the low number of processes influence of the architecture.

An algorithm that applies cybersecurity on design of IoT architecture for a public or private hospital was proposed.

Flowchart techniques (Figure 12) were used to express and apply cybersecurity in the IoT architecture, it consists of two phases called Determine and Apply; the first phase Determine begins with taking inventories at the hospital such as mission, mission, needs, information assets, devices, and network parameters; having the complete record of inventories and parameters, the second phase Apply can be executed; here, the security standards specified in each layer are adopted and applied; otherwise, the inventory and parameters of the hospital must be reviewed again and the algorithm must be run again.

It is necessary to clarify that in case of massive patient assistance to the hospital, the IoT architecture is scalable in terms of the number of devices and storage, in addition this proposal considers the hospital's own devices; regardless of pandemic times or infections such as HIV, Ebola, H1N1, swine flu or COVID-19.

4. Discussion

Principles of Results: The architecture consists of continuous improvement by the functions of the NIST Framework Core in IoT design and adoption of standards.

Relationships of results: the conceptual model has elements that are used on architecture; the architecture has the layers and standards that the algorithm implements through activities.

Exceptions: the architecture only considers smart hospital devices, they are not considered patient-specific devices, BYOD does not apply; it does not consider the carbon footprint produced by the use of devices in the hospital, nor the costs of devices, nor the number of devices in the network, nor the payment values for the cloud.

The results of this research agree with: applying cybersecurity [4], [7], [8]; use HIPAA for health standard [2], [10], [13] and [14]; use NIST for cybersecurity standard [2] and [11]; the benefits of cloud computing are high considering the low cost, storage, access, processing, updating of information and supports the growth of data [14] and [15]; apply IoT cybersecurity to protect information assets [2], [3], [12], [15] and [16]; it use ISO / IEC 27000 series standard [2], [12], and [16].

Theoretical consequence of the research: The modular design of the IoT cybersecurity architecture allows the interconnection of devices, network, services, applications, cloud and users to support growth and adjustment of critical areas of the hospital; the architecture is adapted to the hospital that safeguards the information assets through layers and adopted standards in phases determined by the algorithm.

Possible practical applications: In 2018 Ecuador had 183 public hospitals and 451 private hospitals; the design and

implementation of IoT Cybersecurity in a hospital can be replicated regardless of the dimensions of areas, devices, floors, workers and other characteristics.

IoT recommendations with security measures

We present standard IoT security measures for any public or private health hospital in Ecuador; we emphasize that this proposal is for architecture in hospitals, here IoT devices and hospital information are secured; devices belonging to doctors or patients are not included or considered; the recommendations are in 3 blocks:

In design phase:

- Select upgradeable devices with standard protocols
- Apply network segmentation for connected IoT devices
- Identification and authentication of devices against the IoT network
- Perform device installation and configuration procedures
- Analysis planning and efficiency verification on devices In implementation phase:
- Device management not accessible from the internet
- Configure connection ports only for access to hospital devices
- Verify privacy and confidentiality of data on devices
- Physical and logical evaluation of the devices before their integration into the IoT network
- Restrict or secure access to the cloud interface In Operation phase:
- Continuous change of default credentials of IoT devices
- Continuous app updates on IoT devices
- Disable inactive connectivities
- Network connectivity and device integration
- Use Big Data to monitor the behavior of devices
- Delete of data from unused devices

5. Future work and Conclusions

As future work we proposed an cybersecurity analysis of internet medical things for care centers in Ecuador.

It was concluded that the cybersecurity standards applied to the design of IoT for a public or private hospital generates trust on information assets, preserves the confidentiality, integrity and availability of the information at the operational, tactical and strategic levels; the architecture prototype is between 59.38% and 99.71% of acceptable workload. With the continued growth of health data, there is a market for clinical data validated and verified by health professionals.

Security is adopted from the design of the infrastructure through standards, framework and good practices to minimize risks, ensure each layer and connectivity; the architecture simplifies the distribution and the relationship between the components.

Critical analysis:

Proper understanding of security requirements are important to this IoT cybersecurity recommendation for public and private hospitals in Ecuador; because health information is sensitive, indispensable for the early evaluation and diagnosis of human beings; that referring to the term cybersecurity, the following concepts converge: protection, security and legislation; here the first two terms (protection and security) are covered by Ecuador as

evidence by Figure 6; but the legislative dimension is not very well implemented, so the proposed model covers this deficiency.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors thank to Universidad Politécnica Salesiana del Ecuador.

References

- [1] I. Alrashdi, A. Alqazzaz, E. Loufi, R. Alharthi, M. Zohdy, and H. Ming, "AD-IoT: Anomaly detection of IoT cyberattacks in smart city using machine learning," *2019 IEEE 9th Annu. Comput. Commun. Work. Conf. CCWC 2019*, pp. 305–310, 2019, DOI 10.1109/CCWC.2019.8666450.
- [2] A. Strielkina, O. Illiashenko, M. Zhydenko, and D. Uzun, "Cybersecurity of healthcare IoT-based systems: Regulation and case-oriented assessment," in *2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, 2018, pp. 67–73, DOI 10.1109/DESSERT.2018.8409101.
- [3] D. Minoli, K. Sohraby, and B. Occhiogrosso, "IoT Security (IoTSec) Mechanisms for e-Health and Ambient Assisted Living Applications," *Proc. - 2017 IEEE 2nd Int. Conf. Connect. Heal. Appl. Syst. Eng. Technol. CHASE 2017*, pp. 13–18, 2017, DOI 10.1109/CHASE.2017.53.
- [4] J. Rajamaki, J. Nevmerzhietskaya, and C. Virag, "Cybersecurity education and training in hospitals: Proactive resilience educational framework (Prosilience EF)," *IEEE Glob. Eng. Educ. Conf. EDUCON*, vol. 2018-April, pp. 2042–2046, 2018, DOI 10.1109/EDUCON.2018.8363488.
- [5] P. Rathod and T. Hamalainen, "A Novel Model for Cybersecurity Economics and Analysis," *IEEE CIT 2017 - 17th IEEE Int. Conf. Comput. Inf. Technol.*, pp. 274–279, 2017, DOI 10.1109/CIT.2017.65.
- [6] R. Sabillon, J. Serra-Ruiz, V. Cavaller, and J. Cano, "A comprehensive cybersecurity audit model to improve cybersecurity assurance: The cybersecurity audit model (CSAM)," *Proc. - 2017 Int. Conf. Inf. Syst. Comput. Sci. INCISOS 2017*, vol. 2017-Novem, pp. 253–259, 2018, DOI 10.1109/INCISOS.2017.20.
- [7] G. Parekh et al., "Identifying Core Concepts of Cybersecurity: Results of Two Delphi Processes," *IEEE Trans. Educ.*, vol. 61, no. 1, pp. 11–20, Feb. 2018, DOI 10.1109/TE.2017.2715174.
- [8] Y. Lu and L. Da Xu, "Internet of things (IoT) cybersecurity research: A review of current research topics," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2103–2115, 2019, DOI 10.1109/JIOT.2018.2869847.
- [9] A. F. Klaib and M. S. Nuser, "Evaluating EHR and health care in Jordan according to the international health metrics network (HMN) framework and standards: A case study of hakeem," *IEEE Access*, vol. 7, pp. 51457–51465, 2019, DOI 10.1109/ACCESS.2019.2911684.
- [10] D. Akarca, P. Xiu, D. Ebbitt, B. Mustafa, H. Al-Ramadhani, and A. Albeyatti, "Blockchain Secured Electronic Health Records: Patient Rights, Privacy and Cybersecurity," *Conf. Proc. 2019 10th Int. Conf. Dependable Syst. Serv. Technol. DESSERT 2019*, pp. 108–111, 2019, DOI 10.1109/DESSERT.2019.8770037.
- [11] J. Webb and D. Hume, "Campus IoT collaboration and governance using the NIST cybersecurity framework," *IET Conf. Publ.*, vol. 2018, no. CP740, pp. 1–7, 2018, DOI 10.1049/cp.2018.0025.
- [12] A. Abdullah, R. Hamad, M. Abdulrahman, H. Moala, and S. Elkhediri, "CyberSecurity: A Review of Internet of Things (IoT) Security Issues, Challenges and Techniques," *2nd Int. Conf. Comput. Appl. Inf. Secur. ICCAIS 2019*, pp. 1–6, 2019, DOI 10.1109/CAIS.2019.8769560.
- [13] I. Medvediev, O. Illiashenko, D. Uzun, and A. Strielkina, "IoT solutions for health monitoring: Analysis and case study," *Proc. 2018 IEEE 9th Int. Conf. Dependable Syst. Serv. Technol. DESSERT 2018*, vol. 2015, pp. 163–168, 2018, DOI 10.1109/DESSERT.2018.8409120.
- [14] S. Chentharu, K. Ahmed, H. Wang, and F. Whittaker, "Security and Privacy-Preserving Challenges of e-Health Solutions in Cloud Computing," *IEEE Access*, vol. 7, pp. 74361–74382, 2019, DOI 10.1109/ACCESS.2019.2919982.
- [15] J. Diaz, J. E. Perez, M. A. Lopez-Pena, G. A. Mena, and A. Yague, "Self-Service Cybersecurity Monitoring as Enabler for DevSecOps," *IEEE Access*, vol. 7, pp. 100283–100295, 2019, DOI 10.1109/access.2019.2930000.
- [16] I. Brass, L. Tanczer, M. Carr, M. Elsdén, and J. Blackstock, "Standardising

a moving target: The development and evolution of IoT security standards,” *IET Conf. Publ.*, vol. 2018, no. CP740, pp. 1–9, 2018, DOI 10.2139/ssrn.3437681.

- [17] A. Razaque et al., “Survey: Cybersecurity Vulnerabilities, Attacks and Solutions in the Medical Domain,” *IEEE Access*, vol. 7, pp. 168774–168797, 2019, DOI 10.1109/ACCESS.2019.2950849.
- [18] I. M. Skierka, “The governance of safety and security risks in connected healthcare,” *IET Conf. Publ.*, vol. 2018, no. CP740, pp. 1–12, 2018, DOI 10.1049/cp.2018.0002.
- [19] M. Barrett, “Framework for improving critical infrastructure cybersecurity,” *Proc. Annu. ISA Anal. Div. Symp.*, vol. 535, pp. 9–25, 2018.
- [20] INEC Instituto Nacional de Estadísticas y Censo, “Registro Estadístico de Camas y Egresos Hospitalarios-Ecuador,” *Regist. estadístico y Egr. 2018*, 2018.
- [21] U. T. International, *Global Cybersecurity Index (GCI) 2017*, no. November. 2017, DOI 10.1111/j.1745-4514.2008.00161.x.
- [22] International Telecommunication Union, *Global Cybersecurity Index (GCI) 2018*, 2018, DOI 10.1111/j.1745-4514.2008.00161.x.
- [23] K. Scwab, *The Global Competitiveness Index Report 2017-2018*, no. 31. 2018.
- [24] R. Lucio, N. Villacrés, and R. Henríquez, “Sistema de salud de Ecuador,” *Salud Publica Mex.*, vol. 53, no. SUPPL. 2, pp. 177–187, 2011.
- [25] D. R. L. Pulles, D. M. C. Urquizo, and Ms. A. C. MBA, “The Present Situation of e-Health and mHealth in Ecuador,” vol. 4, no. 3, pp. 261–267, 2017.
- [26] D. Pillo-Guanoluisa and R. Enríquez-Reyes, “Gobierno de TI con énfasis en seguridad de la información para hospitales públicos,” *Maskana*, vol. 8, no. 0, pp. 42–55, 2017.

Risk Management: The Case of Intrusion Detection using Data Mining Techniques

Ruba Obiedat*

King Abdullah the Second School of Information Technology, The University of Jordan, 11942, Jordan

ARTICLE INFO

Article history:

Received: 02 May, 2020

Accepted: 14 June, 2020

Online: 25 June, 2020

Keywords:

Risk Management

Intrusion Detection

Data Mining

Machine Learning

ABSTRACT

Every institution nowadays relies on their online system and framework to do businesses. Such procedures need more attention due to the massive amount of attacks that occurs. These procedures have to go first through the management team of the institution, in order to prevent exploits of the attackers. Thus, the risk management can easily control and identify the risk that occurs. One of these risks is an intrusion, which is an action or an act that the attacker invades someone's privacy to steal or damage their information. Various techniques have been proposed to prevent these actions in the literature. This research proposed an intrusion detection model to distinguish the most recent attacks using data mining techniques. Three machine learning classification models have been applied namely, J48, Random Forst and REPTree to improve the detection rate. Furthermore, a Feature Selection method has been applied in order to improve the effectiveness of the classifier and also overcome the high dimensionality which presents one of the main technical problems facing the intrusion detection systems and come up with the most important intrusion features affecting the system. These features can be very useful in protecting the systems from attackers. The results identify the top 11 effective features. The best results achieved by the J48 with a 76.271% accuracy rate.

1. Introduction

In the recent years, people all around the world become more depending on the information technology in all kinds, notably, with the expansion of the internet and automated businesses, and procedures from different fields [1]. People utilized computers and applications in order to acquire information about several things; such as stock costs, news, and online trade. On the other hand, others save the information of patient's medical records, credit card, and other personal data on their systems, either offline or online. Many organizations, for example, have a web presence as a fundamental structure of their businesses. Controlling and secure these critical assets and information that come from the decision of the management team [2, 3].

Consequently, without a good plan and scheme, risk can occur regularly. This kind of risk can harm the company for example, in a severe way, especially controlling the intrusion and detect each attack that happens [2]. The availability and integrity of the systems must be ensured against various threats, such as hacking or damaging, in which eventually can hurt the image of the institution. Hence, the secure information and the communication

turned out to be indispensably vital. Furthermore, the need to detect privacy breaches and information security demands of a robust intrusion detection and prevention systems (IDPSs) are more necessity [2-5].

Several techniques have been proposed to prevent such vulnerability. The most recent and effective one is "Data Mining", which is a method of understanding and finding the pattern of the data [6, 7]. This data usually collected from different sources, each one of them portrays a case that occurs on different scenarios. In real life, data can portray as stones and sand, and mining these gravels to extract the jewelry (useful information). Therefore, extracting such benefit information can lead us to prevent attacks and leaking information in a more effective way. In our case, the dataset of these logs is collected, where the instances of each attack and non-attack portrayed as a row, while the columns are the characteristics of each instance. These characteristics called features. Another critical process can be used to help us knowing and understanding the pattern, and hidden information is called Feature Selection. It is a way to remove redundancy and not important features from the dataset and keep the most important ones without affecting the essential information of the data, it relies on identifying the features which are independent of each other but

* Ruba Obiedat, Email: r.obiedat@ju.edu.jo

highly relevant to the output at the same time using the goodness metrics. This technique can increase the accuracy of the classifier, utilize time and resources, reduce complexity, and support generalization [5, 8, 9].

This research presented a model using Data Mining techniques to identify these intrusion attacks. We used a public-online dataset that depicts these attacks. Then examined the data on different classifiers for evaluation and improving the detection rate of the Intrusion Detection System (IDS), which are; Decision Tree, RepTree, and Random Forest. And since IDS deals with massive amount of data [8], we then applied features selection to see the most important features that can help detect each attack. Feature selection method presents one of the most useful and commonly used techniques in data preprocessing step for the Intrusion Detection System (IDS). This technique is essential in our case here since the number of features in any Intrusion Detection System (IDS) is very huge whatever size of networks you work with and we need to reduce it. In addition, applying the feature selection technique can optimize the classifiers results and increase detection rate by removing unimportant features and help identifying the most relevant features affecting the system which present valuable information for the security teams as well.

The rest of this paper is organized as follows: section 2 presents the background of the intrusion detection history. The methodology produced in Section 3. Finally, section 4 and 5 introduces the experiments and results, and the conclusion, respectively.

2. Background

In order to define Intrusion Detection and Prevention Systems (IDPSs), it's important to know the events prior the detection procedures. The term attack is an action that exploits the vulnerability of a system to cause a loss of an information or access. Therefore, intrusion is considered a kind of attack, in which the intruder tries to obtain access or damage the system. In short, it is a process of observation, detecting and analyzing the performance that known as a violation of the information security policies. [10] Illustrates the intrusion detection as a way of detecting the cyber-attacks that take place on the computer networks using a certain framework known as Intrusion Detection System (IDS). This framework usually detects any abnormal patterns of the system.

There are two types of an attack, outsider attack, which is an attack that comes from the previous external origins. The other one is the Insider attack, an unauthorized/authorized inner user attempts to gain and abuse the non-authorized/ authorized entrance privileges to the system. Thus, we need to prevent such acts.

Intrusion Detection is one of these acts and known as a process of observing of the computer or networks for any unauthorized access, activity or modification [11, 12]. Intrusion Detection System helps to detect any intrusions on systems. Also, it helps to catch any unusual behaviors in different ways that covers and displays warnings when logging into a system. Another important procedure is that Intrusion Prevention System (IPS) detects system warnings in real time and prevents unauthorized access to the system by using specific software. Moreover, it can help to prevent any potential harmful events [11]. Intrusion Prevention System can detect threats in various ways; it can work with combination of

access control (firewall/router) or any other security controls in the systems to prevent planned attacks. It can block the malicious network activity and configure or reconfigure privacy controls in browser settings to block these attacks [2].

However, shutdown the prevention characteristics in IPS cause them to operate as ID systems. Although IPS and IDS both monitor and analyze the network flow, IPS is recognized to be an expansion of IDS. IPS and IDS both are implemented to find unionize movement. The major duty of the IDS is to alert the unauthorized movements that are taking place. On the other hand, IPS is used for more effective protection by improving IDS and other common security solutions more. A powerful risk management process is essential part of security. Risk management has to determine the necessary security controls to decrease the danger to a suitable level by organizations. For effective IDPS design, it is necessary to obtain appropriate risk management-based requirements [2, 13].

2.1. Risk Management

Application, and network service of any computer and communication systems, usually behave normally and expected to be guarded against misuse, through a combination of privacy and safety and so on to be available and accessible. Similarly, these functions should provide a secure and trusted data transmission services to the end-users. Risk management means to control the risk and find a way to reduce the risk to an acceptable level. [14] determined the risk management as the procedure enables the managers of IT to adjust the financial as well as operational expenses of protection and to satisfy the mission capacity by securing the Information Technology frameworks and information that eventually help their associations' missions [13, 14].

Security concerns can rapidly disintegrate client certainty and conceivably diminish the appropriation rate and rate of a degree of profitability for strategically critical items or administrations. Risk management procedure is a critical part of a successful security system. Main objective of an organization's risk management procedure must take into consideration the organization and its capacity to accomplish its central goal and missions, instead of essentially its IT resources and assets. Risk management shouldn't be treated as only a technical issue, but as a fundamental critical management issue inside organizations. Risk protection plans are described by understanding, identifying and accepting the leftover risks related to the use of information systems. In order to improve protection of the organizations from serious and increasingly threats of information systems, organizations should utilize a risk strategy alongside IDPSs, as a whole system of protection to secure the CIA triangle (Confidentiality, Integrity, Availability) of information systems [15, 16].

2.2. Intrusion Detection and Prevention Systems

Intrusion Detection System is used to observe any usual movements on a computer system or network. Moreover, it shows alerts on any possible violations, threats or attacks on computer system or network. Intrusion Detection System is both hardware and software application that work together to monitor and control computer system as well as computer network for any security policy abuse, furthermore, it reacts to any unusual activities by sending a message or alerts to the admin of the system in different ways [16]. Intrusion Prevention means to operate intrusion

detection and try to stop any potential harmful events. Moreover, Intrusion Prevention System has the same capabilities of IDS of stopping potential accidents; it is designed to secure systems from any abuses or troubles. However, the main difference between IPS and IDS is that IDS warns in case of any attack attempt while IPS tries to block the attack from happening [17].

2.3. Important of the Intrusion Detection

Each organization utilizing computer systems must take into consideration the important of security. The fact that Information Security is an area which depends on specialists to enhance the safety of the enterprise most valuable resource can't be underestimated. The majority of businesses exclusively carry out high standard security policy arrangements, despite the fact that the largest threats are from inner sources. Furthermore, organizations perform network security arrangements which are intended to keep network assets safe. IDPSs improve the safety of InfoSec and network with denying the any unauthorized access of discovered attacks. Moreover, it can provide valuable clues on the best defense for the attack [8]. Unfortunately, there is nothing called "absolutely secure system", any system is vulnerable toward violation by insiders who misuse they system for their own benefits. In order to understand the requirements on focus and improve advanced IDPSs, an overlook into the numbers of incidents and data breaches must be studied [18, 19].

In the last decade, statistics record shows that the technological progress becomes more complex and security concerns grow more and more as shown in Figure 1 [20]. The incidents are limited for the danger and risk of privacy violations, security breaches, and malicious intrusions. Concurrently, with the increasing number of the computer systems as well as the huge expansion in the total number of threats, it has become even more aggressive than ever before, especially with the continuous increasing in size and speed of the networks. It is obvious that enterprises require intelligent InfoSec techniques such as IDPSs which can detect the recent formed attacks to limit the loss of data and keep up the privacy of data with a quick response.

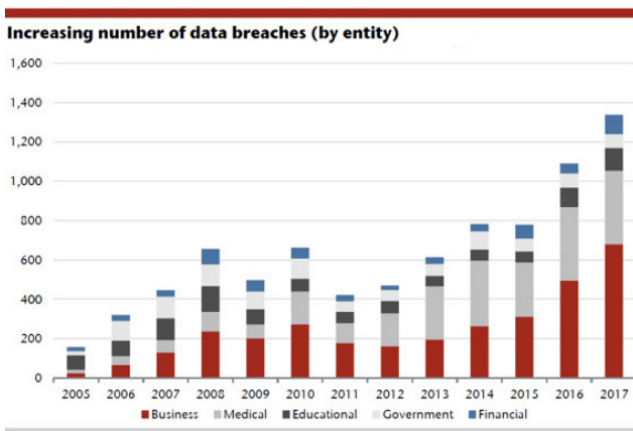


Figure 1: Numbers of data breaches

3. Related Work

Some researchers such as [21] introduced a framework which works with actual Intrusion Detection issues in classifying network data analysis into abnormal and typical behaviors. His paper suggests a "multi-level hybrid intrusion detection model" which www.astesj.com

uses learning machine as well as support vector machine to enhance efficiency of recognizing unknown and known attacks. Also, he added an updates K-means algorithm in order to develop a strong training dataset in order to improve the classifiers performance. The modified K-means is utilized to make new short training datasets to the whole original training dataset, essentially decrease the training time of classifiers, and enhance the performance of IDS. Analysis of the different techniques based on the similar dataset, the proposed model reveals and show high efficiency in accuracy and attack detection.

Furthermore, [6] proposed an "Internal Intrusion Detection and Protection System (IIDPS)", he suggested that in order to discover attacks, we need to utilize a forensic and data mining techniques. IIDPS generate users' individual profiles to remain track of users using their behavior as a forensic feature and decides if a user is the account owner or is not, by examining through user present PC usage practices with the patterns gathered in the account owner's personal profile. The results show that the user identification accuracy rate was 94.29%, while the reaction time is under 0.45 s, suggesting that it can manage to secure the system from insider attacks efficiently and effectively.

In addition, [22] presented a system to defeat menaces each time a detection arrangement was demanded because of highly spread in networks. Within the development of the arrangement, attackers have become stronger and each one damage network protection. Therefore, a need of the Intrusion Detection arrangement appeared to be a vital and necessary instrument in network security. Detection of such attacks loud intrusions normally rely upon the strength of Intrusion Detection Arrangement (IDS). In his approach, a number of elements have been coordinated for utilizing the methods; these systems have their very own advantages and disadvantages. In his paper, he focuses on various classification methods.

A detailed analysis and investigation of different machine learning methods have been developed, for locating the cause or problems related to different machine learning strategies in detecting the intrusive actions [23]. The results identified with low-frequency attacks utilizing network attack logs and dataset are as well discussed and viable techniques are recommended for development. Machine learning methods have been compared and analyzed in terms of their exposure ability for detecting the different classification of attacks. Limitations related to every classification of them are as well discussed. Different data mining devices for machine learning have as well been carried in his paper.

As can be seen in the mentioned works, many researchers focused on machine learning methods side to improve the detection rate of IDS but few had tried to investigate the impact of the features of the datasets. However, in this work, we study the impact of these features and what the implication of them, especially, the most important ones using feature selection method. Thus, improve the Detection performance.

4. Methodology

In this section, a description of the proposed approach processes is presented. As shown in Figure 2, five main processes are defined, namely, Data Mining, Classification Models, Data Description, Feature Selection methods and Evaluation.

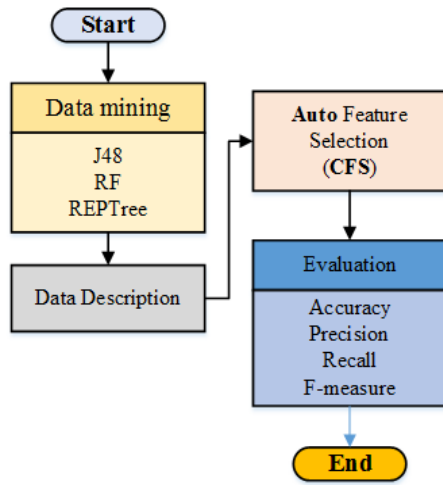


Figure 2. Methodology Steps

4.1. Data Mining

Data Mining reveals an important pattern, connection and directions through studying and analyzing each data using different machine learning techniques, data visualization methods, statistics, and artificial intelligence. This procedure effectively referred implicit, previously unknown, and possibly helpful information found in the data, therefore empowering the detection of patterns, the generation, and examination of hypotheses, and the creation of visualizations [24].

4.2. Decision Trees C4.5 (J48)

C4.5 is one of the algorithms related to decision tree which create trees from preparing data utilizing information gain and entropy. It is considered as an extension of the “ID3 algorithm” that overcomes most of its weakness. At every node, the algorithm checks the information gain and selects the attribute with the highest gain as splitting node, dividing the samples into subsets collection, the most significant characteristic value will settle on the choice then the procedure is taking place again upon smaller size of subsets [25].

4.3. Random Forest (RF)

This technique is based on the shell of a lot of decision trees, utilizing a stochastic procedure over the base of C4.5 algorithms. It was suggested by [26] and intends to make independent and uncorrelated trees dependent on the various and random input vectors attached by similar distribution. The outcome is the mean of the created trees through the process. RF is easy to implement, robust and able to handle large amount of data.

4.4. Reduced Error Pruning Tree (REPTree)

This algorithm is used to deal with decision tree noise. REPTree applies regression tree logic. It was presented in the post-pruning method founded on the concepts of [27, 28]. [29] defined the logarithmic procedure as: the initial step, where the data is separated into two groups. The first is the developing set, generated or created into a realization of learning algorithms, while the second is the pruning set that is generated through removing the subtree rooted at that node, until the outcomes in a predictive accuracy drop measured through the pruning set.

4.5. Data Description

In this work, the experiments applied on the intrusion-detection dataset that publicly available on the Kaggle repository/website [30]. The dataset consists of 20000 instances and 42 features. These features have different characteristics and methods in order to extract them, to name a few, duration, protocol-type, num-failed-logins, service, num-compromised, and so on. All features have numeric values except the protocol-type feature where it is a category-based.

4.6. Feature Selection methods

Feature selection (FS) is a method to choose the best group of features in order to use them in the developed model development. Overall, FS is designed by separating the redundant or irrelevant features without maintaining any loss of information while concentrating on the only relevant ones. It is one of the most vital goodness metrics to select features; generally, we say that a model is good if it is relevant to the output but not redundant with the other relevant features at the same time. The objective of the procedure is to reduce the time of training period of the classification models. Moreover, this procedure helps in developing the capacity of the model by decreasing the overfitting and to encourage researchers to facilitation their models and make them simpler to understand [31].

Feature selection techniques is an essential and effective step in data mining and has been adopted by many researchers in different data mining fields such as pattern recognition, network security and IDS. There are two types of feature selection approach; a Wrapper feature selection and Filter based feature selection. Wrapper method employs the performance of the machine learning algorithm to evaluate the usefulness of features, while Filtering approach relies on the characteristics of the training data itself to identify the redundant or irrelevant data [32]. In this research we use one of the most common filter based feature selection method; which is Correlation-based Feature Selection.

4.7. Correlation-based Feature Selection

Correlation-based Feature Selection (CFS) is a heuristic type feature selection method that uses a search algorithm and evaluation criteria to select a subset of features. It is the most popular measure for the dependency between two variables. CFS measures each feature’s goodness by predicting the usefulness of individual features to predict the class label and their intercorrelation level, and then come up with the optimal subset of features relevant to the class with no redundancy. In short, the explanation can be summarized: "Good feature subsets contain features highly correlated with the class, yet uncorrelated with each other" [32]. Therefore, the selection criteria in CFS select the best subset of features automatically. The following equation computes the merit of feature subset S with k number of features.

$$Merit_s = \frac{k\bar{r}_{cf}}{\sqrt{k+k(k-1)\bar{r}_{ff}}} \quad (1)$$

5. Evaluation

The Confusion matrix is a table that contains a summary of the prediction results for the classification system. Table 1 presents a confusion matrix for a binary classifier. It is used in order to assess

the intrusion detection models of the current research and will be pointed out to as a primary source for the evaluation.

Table 1: Confusion Matrix

		Predicted Class	
		Positive	Negative
Actual Class	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

Accuracy: determined by ratio of correctly classified instances of the two classes divided by the total number of instances.

$$Accuracy = \frac{TP+TN}{TR+TN+FP+FN} \tag{2}$$

Precision: the ratio of classes classified as attackers that actually are attackers

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

Recall: the ratio of correctly classified intrusion attack divided by the total number of instances classifies as attackers.

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

F-measure is determined as a weighted mean of the recall and precision.

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{5}$$

6. Experiments and Results

In this section, we examined the data on several classification models. Named; Decision Tree, RepTree, and Random Forest Then we applied a feature selection technique to select the best subset of features for the classifiers and improve the accuracy of the results. In sum, first, we examine the data without applying the feature selection method. Then, we apply the feature selection before the examination on the classification model.

As shown in Table 2, the best accuracy obtained by J48 with 75.424%, while the second and third achieved by RepTree and RF, with 71.186% and 70.339%, respectively. For the recall measure, the highest results demonstrated by RepTree and RF at the same rate, 0.960%, and the J48 come third with 0.940. In precision and F-measure, J48 classifier has the best results with 0.671% and 0.783%, followed by RepTree with 0.623% and 0.756. The last classifier was RF with 0.600% for precision and 0.738% for the F-measure.

Table 2: Results of J48, RepTree and RF without Feature selection.

Data Without FS	All features			
	Acc	Recall	Precision	F-Measure
J48	75.424	0.940	0.671	0.783
RepTree	71.186	0.960	0.623	0.756
RF	70.339	0.960	0.600	0.738

The original data had 41 features and decreased to 11 features after applying the feature selection method as shown in Table 3. Table 3 contains the top 11 features relevant to the attacks with a brief description about each feature taken from the Kaggle repository website. These data can be very useful in identifying important intrusion features and protecting the system from attackers. After finding the optimal subset of features we examine the classifiers again against these 11 features. As shown in Table 4, most of the results shows a brief increase, where the best classifier in the original data has a nearly 0.847 increase in accuracy with 76.271% rate — the second best results achieved by RF with 74.576%, unlike the original data where the second rank obtained by RepTree. The RF shows a notable increase with 4.237% which is more improved than the previous experiments, the least accurate of all classifier demonstrated by the RepTree with 72.034%. For the F-measure and precision, again the J48 had the highest results with 0.734% and 0.825%, the second and third classifier were RF and RepTree for both measures, respectively.

Table 3: Top selected features.

#	Features	Description
1	Service	Network service on the destination, e.g., http, telnet, etc.
2	Flag	Normal or error status of the connection
3	src_bytes	Number of data bytes from source to destination
4	dst_bytes	Number of data bytes from destination to source
5	Urgent	Number of urgent packets
6	logged_in	1 If successfully logged-in; 0 otherwise
7	srv_error_rate	% Of connections that have "SYN" errors (same-service connections)
8	same_srv_rate	% Of connections to the same service (same-host connections)
9	diff_srv_rate	% Of connections to different services (same-host connections)
10	dst_host_srv_diff_host_rate	Diff_host_rate for destination host
11	dst_host_error_rate	Error_rate for destination host

Table 4: Results of J48, RepTree and RF with Feature selection.

Data WithFS	Selected features			
	Acc	Recall	Precision	F-Measure
J48	76.271	0.940	0.734	0.825
RepTree	72.034	0.940	0.644	0.764
RF	74.576	0.940	0.653	0.770

In sum, the feature selection method increases the results for all measures except the recall measure. Decreasing the features from 41 to 11, not only shows a notable increase on the results but also, help us to know the best-selected features that can identify the intrusion detection attack from the non-attack class. As a result, we can conclude that the experimental results and the selected features show that the data mining techniques as classifiers and feature selection method can play an important role in improving risk management, helping the security teams to build their security strategies and prevent exploits of the attackers, which keeps their networks safe and secured. Finally, the risk factors can be further analyzed to study how they affect the intrusion detection as a case study of risk management.

7. Conclusion and Future Work

The current computer network is a risky domain, packed up with attackers that have millions of hours available to operate against the most grounded of security strategies. The best way to control them is to know when they are trying an attack and counter their efforts. Choosing the right intrusion detection system is the key to ensuring that an enterprise's networks and systems stay secure. While security cases turn out to be more numerous, network intrusions become a significant threat, as a result IDPS is becoming increasingly important and necessary. These intelligent IDPSs should utilize several intelligent methods from the matter fields of data mining, machine learning, and artificial intelligence to support them to figure out what qualifies as an intrusion. This paper proposed an intrusion detection model to prevent and decrease the intrusion attacks. Three machine learning classifiers were applied; which are J48, RepTree and RF to check the detection rate. Moreover, as the network traffic is becoming very huge with massive amount of data, leading to increased number of redundant and irrelevant features which decrease the IDS detection rate, consuming more resources, slowdown the detection process and increase complexity. Consequently, a feature selection method has been applied using the Correlation-based Feature Selection in order to find out the most relevant features to the detection process and remove unnecessary or redundant one. This approach is not only able to reduce the time consuming of the experiments but also increase the results and lead to higher accuracy. The best obtain accuracy achieved by the J48 with 76.271% using the top 11 features identified by the feature selection method. For future work, it is aimed to increase the number of the classifiers and uses different data to help us study the features in detail using a different technique of the features selection.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] Atasoy, H., "The effects of broadband internet expansion on labor market outcomes", *ILR Review*, 66(2), 315-345, 2013. <https://doi.org/10.2139/ssrn.1890709>
- [2] Michael E. Whitman, Herbert J. Mattord, *Management of Information Security*, Cengage Learning, 5th edition, 2016.
- [3] Sennewald, C. A., & Baillie, C., *Effective security management*, Butterworth-Heinemann, 2020.
- [4] Lewis, T. G., *Critical infrastructure protection in homeland security: defending a networked nation*, John Wiley & Sons, 2019.
- [5] Amiri F, Yousefi M, Y, Lucas C, Shakery A, and Yazdani N., "Mutual information-based feature selection for intrusion detection systems", *Journal of Network and Computer Applications*, 34(4), 1184-1199, 2011. <https://doi.org/10.1016/j.jnca.2011.01.002>
- [6] FY Leu, KL Tsai, YT Hsiao, CT Yang, "An internal intrusion detection and protection system by using data mining and forensic techniques", *IEEE Systems Journal*, 11(2), 427-438, 2015. <https://doi.org/10.1109/jsyst.2015.2418434>
- [7] Leskovec, J., Rajaraman, A., & Ullman, J. D., *Mining of massive data sets*, Cambridge university press, 2020.
- [8] Yang Li, Jun-Li Wang, Zhi-Hong Tian, Tian-Bo Lu, Chen Young, "Building lightweight intrusion detection system using wrapper-based feature selection mechanisms", *Computers & Security*, 28(6):466-475, 2009. <https://doi.org/10.1016/j.cose.2009.01.001>
- [9] Ayman I. Madbouly Amr M. Gody ,Tamer M. Barakat, "Relevant feature selection model using data mining for intrusion detection system", *International Journal of Engineering Trends and Technology (IJETT)*, 9 (10), 2014. <https://doi.org/10.14445/22315381/ijett-v9p296>
- [10] D.E. Denning, "An intrusion-detection model", *IEEE Transactions on software engineering*, (2), 222-232, 1987. <https://doi.org/10.1109/tse.1987.232894>
- [11] Lee, W., & Stolfo, S. J., "Data mining approaches for intrusion detection" in *1998 Proceedings of the 7th USENIX Security Symposium*, San Antonio Texas, USA, 1998. <https://doi.org/10.21236/ada401496>
- [12] Vidal, J. M., Monge, M. A. S., & Monterrubio, S. M. M., "Anomaly-Based Intrusion Detection: Adapting to Present and Forthcoming Communication Environments", In *Handbook of Research on Machine and Deep Learning Applications for Cyber Security*, pp. 195-218, IGI Global, 2020. <https://doi.org/10.4018/978-1-5225-9611-0.ch010>
- [13] S. Kamiya, J. Kang, J. Kim, A. Milidonis & R. Stulz, "Risk management, firm reputation, and the impact of successful cyberattacks on target firms", *Journal of Financial Economics*, 2020. <https://doi.org/10.1016/j.jfineco.2019.05.019>
- [14] Chichakli, R., "Information systems risk management", 2009.
- [15] Malik, M. F., Zaman, M., & Buckby, S., "Enterprise risk management and firm performance: Role of the risk committee", *Journal of Contemporary Accounting & Economics*, 16(1), 100178, 2020. <https://doi.org/10.1016/j.jcae.2019.100178>
- [16] Whitman, M. E., & Mattord, H. J., "Readings and cases in the management of information security: Law and Ethics", *Information Security Professional*, 2005.
- [17] Martin, C., What is IPS and how intrusion prevention system work, available online on: <https://www.paloaltonetworks.com/cyberpedia/what-is-an-intrusion-prevention-system-ips>, 2009.
- [18] Beal, V., *Intrusion detection (IDS) and prevention (IPS) systems*, 2005.
- [19] Sundaram, A., "An introduction to intrusion detection", *Crossroads*, 2(4), 3-7, 1996. <https://doi.org/10.1145/332159.332161>
- [20] Reklaitis, V., "How the number of data breaches is soaring - in one chart", Retrieved April 20, 2020, from <https://www.marketwatch.com/story/how-the-number-of-data-breaches-is-soaring-in-one-chart-2018-02-26>, 2018.
- [21] Al-Yaseen, W. L., Othman, Z. A., & Nazri, M. Z. A., "Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system", *Expert Systems with Applications*, 67, 296-303, 2017. <https://doi.org/10.1016/j.eswa.2016.09.041>
- [22] Sree, S. B., *Kernel Based Intrusion Detection Using Data Mining Techniques*, 2018.
- [23] Mishra, P., Varadharajan, V., Tupakula, U., & Pilli, E. S., "A detailed investigation and analysis of using machine learning techniques for intrusion detection", *IEEE Communications Surveys & Tutorials*, 21(1), 686 - 728, 2019. <https://doi.org/10.1109/comst.2018.2847722>
- [24] Hand, D. J., "Principles of data mining", *Drug safety*, 30(7), 621-622, 2007. <https://doi.org/10.2165/00002018-200730070-00010>
- [25] Quinlan, J. R., *C4. 5: programs for machine learning*, Morgan Kaufmann Publisher, 2014.

- [26] Breiman, L., "Random forests", *Machine learning*, 45(1), 5-32, 2001.<https://doi.org/10.1023/a:1010933404324>
- [27] Pagallo, G., & Haussler, D., "Boolean feature discovery in empirical learning", *Machine learning*, 5(1), 71-99, 1990, <https://doi.org/10.1007/BF00115895>.
- [28] Quinlan, J. R., "Simplifying decision trees", *International journal of man-machine studies*, 27(3), 221-234, 1987, [https://doi.org/10.1016/S0020-7373\(87\)80053-6](https://doi.org/10.1016/S0020-7373(87)80053-6).
- [29] Fürnkranz, J., & Widmer, G., "Incremental reduced error pruning" in 1994 Proceedings of the Eleventh International Conference, Rutgers University, New Brunswick, NJ, USA, 1994. <https://doi.org/10.1016/b978-1-55860-335-6.50017-9>
- [30] Intrusion detection, Retrieved 29 April 2020, from <https://www.kaggle.com/what0919/intrusion-detection>, 2020.
- [31] Hee-su Chae, Byung-oh Jo, Sang-Hyun Choi & Twae-kyung Park, "Feature selection for intrusion detection using NSL-KDD", *Recent advances in computer science*, 184—187, 2013.<https://doi.org/10.1109/icomitee.2019.8920961>
- [32] Chen, Y., Yang Li, Xue-Qi Cheng, & Li Guo., "Survey and taxonomy of feature selection algorithms in intrusion detection system", In 2006 International Conference on Information Security and Cryptology, pages 153–167, 2006.https://doi.org/10.1007/11937807_13

Business Process Design for Widuri Indah School Management System with the Support of Cloud Computing

Yulyanty Chandra*, Roy Willis, Calvin Windoro, Sfenrianto

Information Systems Management Department, BINUS Graduate Program-Master of Information Systems Management, Bina Nusantara University, Jakarta, Indonesia 11480

ARTICLE INFO

Article history:

Received: 02 May, 2020

Accepted: 14 June, 2020

Online: 25 June, 2020

Keywords:

Cloud computing

Business Process

School Management System

Information system

ABSTRACT

This study aims to design business processes or school management solutions that are right for the school. Widuri Indah School is private school that has not been integrated with the school management system. Qualitative method such as observation and interview are conducted to gain insights about main activities and support activities in Widuri Indah school business process. Furthermore, designing the use case diagram to construct the business process. The design will be able to produce business process that can integrating between users at school and school management systems in a cloud-based process It can be used as a guide in the procurement of school management system based on cloud computing. The design is believed to bring improvement such as interoperability, flexibility, data management, and efficiency as a result.

1. Introduction

The world of Education has not yet fully raised all its problems in Indonesia. Various efforts and activities are always carried out to improve the innovation and quality of the system that carries the name of education has been carried out by the government, until now. The world of education should mean the availability of various solutions to support educational programs to create an education system that can support and enhance the activities of users at school. However, the use of IT in Indonesia has just entered the stage of studying various possibilities for the development and application of IT for education. When examined in detail in this regard, the government has sought to improve the quality of Indonesia's education, what exactly is at the core of the problem in the world of education, may be far more difficult. Schools in Indonesia are not yet connected with technology in the process, have not yet reached the mind power and adequate sources of information. The solution to all the problems above is an academic information system that is used to monitor all educational activities and is able to produce information that is needed by the school, government, students, guardians and the public at large and up to date. The system must be accessible whenever and wherever and be able to generate information automatically without the need for manual calculations that are error prone and result in incorrect information, and certainly can

support quick decision making. Widuri Indah School's system has not utilized technology as the main base for automation. All things especially administration for payment of school fees is still manually by queuing making payments in school administration. Information such as report card grades, student absences and attendance and submission of teacher and employee leave are still manual, and student assignment information has not been connected to any technological media. This study aims to produce design needs in terms of technological infrastructure in Indonesia in the form of catalogs, diagrams, and matrices that will be needed by Widuri Indah Schools in developing their school management.

2. Literature Review

A good education in the world should have a good business system and process, where the ultimate goal itself is not only to support teaching and learning in schools, but also that an integrated system can support the internal operational performance of schools. A well-integrated system will be able to make decisions faster than using a manual or traditional system.

2.1 Business Process

According to [1], in his book "Business Process Change" (2003), the definition of a business process is a series of activities carried out by a business which includes the initiation of input, transformation of information, and producing output. The output can be valuable for business or market customers, it can also be

* Yulyanty Chandra, Email: yuliyantychandra@binus.edu

valuable for other processes (within the organization). A business process can be broken down into several processes, each of which has its own role to achieve its goals. Sub-processes can be broken down into activities, which are the smallest processes that can consist of one or more steps that must be included in a business process [2] [3].

Business process development is needed by an educational institution with the aim of improving internal quality management [4]. Therefore, the business processes that run on educational institutions are needed to be analyzed in accordance with the needs of management and institutions [5]. Use case is one of the approaches used to design business processes [6].

2.2 School Management System

School Management system is an information system used by educational institution for specific purposes. For example, the Smart School Management System (SSMS) which is used to assist learning activities and Information systems management of schools in Malaysia [7]. While Smart school management in Indonesia, one of them is edConnect as an information system that simplifies administrative work, improves information accessibility, facilitates intensive communication, and provides data visibility for all aspects of the school.

For its own definition, School Management Information System is a design that adjusts the structure, task management, learning processes and the special needs of schools by utilizing information systems management [8]. The purpose of the existence of a school management information system is to support management and educational activities by school managers by utilizing information [9]. It can be concluded that the school management information system is a structural design, task management and special needs of schools that are both direct and indirect learning processes that are used by school managers by utilizing information systems management.

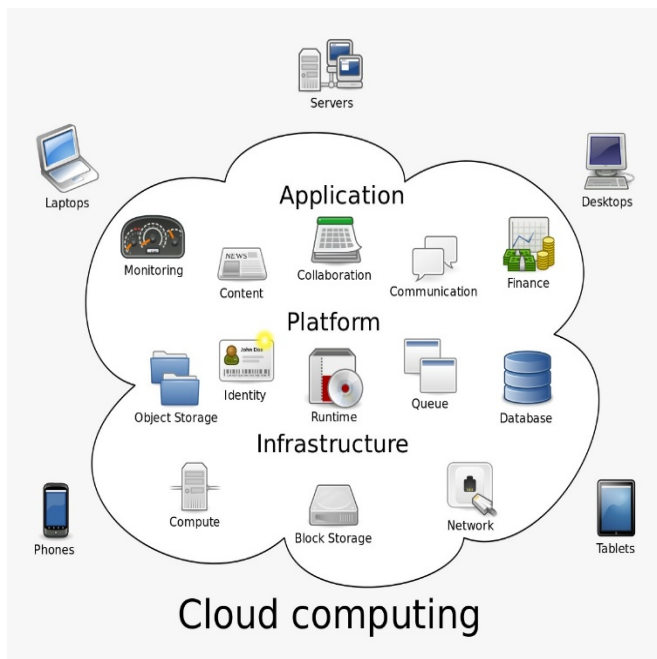


Figure 1: Cloud Computing

2.3 Cloud Computing

Figure 1 shows the three main components used to support cloud computing, namely Infrastructure, Platforms and Applications [10]. There are several infrastructures from clout computing such as server, network, storage, and digital technology that are needed.

Cloud Computing (Cloud computing) is a combination of computer technology and internet networks [11]. Cloud (Cloud) is part of the internet, as the cloud is in the picture of an organization's internet network pattern. Apart from being like a cloud in a computer network diagram, the cloud (cloud) in cloud computing is also an integrated network infrastructure [10].

The term cloud has been widely used in the development of the Internet world, because the Internet can be said to be a large cloud. Cloud computing itself is a computing model, where its resources such as processor / computing power, storage, network, and software become abstract and are provided as services on the network / internet using remote access patterns.

There are many issues that can be solved by cloud. For instance, If the company not sure where to store information, a cloud server is a perfect way to reduce confusion. Cloud infrastructure has evolved tremendously and is now safer and more robust than conventional on-site technologies. Therefore, cloud can also be the answer in solving security issues. Cloud computing decreases or removes the requirement for businesses to buy hardware and develop and run data centers, hence will lead to cost efficiency.

Analytics, AI, and the prospect of safe communication beyond the corporate area give an incentive to implement a cloud infrastructure. Also, one of the benefits of cloud infrastructure for companies is how convenient it is for team leaders to operate from anywhere. Businesses have traditionally been connected to wherever their hardware is based, as that is where they need to get access to all their records. The cloud, though, allows users to carry their data with them everywhere they go.

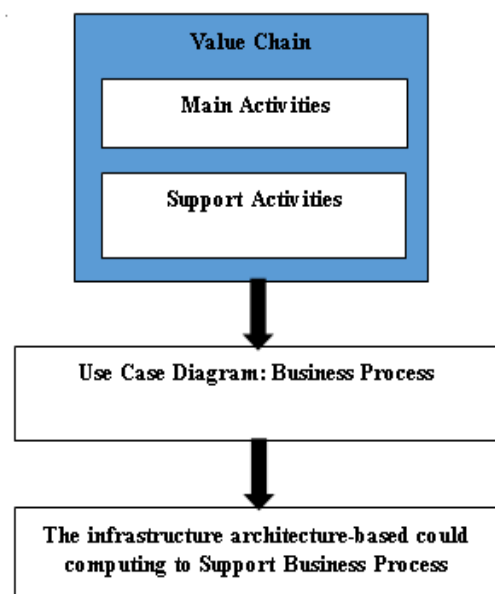


Figure 2: Research Design Stages

3. Methodology

The method used is a qualitative method that is observation and interview. Figure 2 shows research design stages. The design begins by explaining the main activity and supporting processes that exist in the management of the Widuri Indah school's business process. Modeling that will be used is value chain which is a tool to describe main activities and support activities.

Then defining the parts of the business process, determine the relationship of information systems with business processes from Widuri Indah School's management, and detailed the specific business process by using use case diagram to make a clear related part to the system. The final design is making the infrastructure architecture-based cloud computing for the steps internal structure for the school.

4. Result and Discussion

This section presents the result of study for design business process bases on cloud computing to support Widuri Indah School Management System. The following value chain explains the process of main activities and supporting activities (see Figure 3).

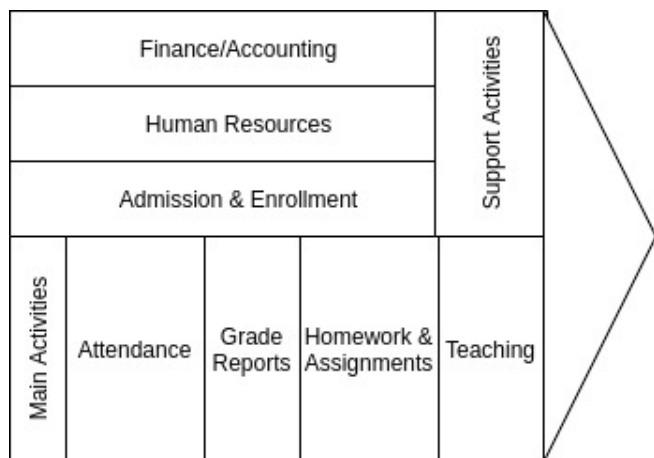


Figure 3: The Value Chain of main activities and support activities

Based on figure 3, the main activities are attendance, grade reports, homework and assignments, and Teaching. The Attendance is a data collection of the presence of students every morning at school and attendance data of teachers, and employees at the school. Grade reports are the grades given from the teacher to the students in the form such as; PH (Daily assessment), PTS (Middle semester assessment), PAS (End of semester assessment) and PAT (End of year assessment).

Then, homework and assignments are given to students every day to provide exercises to keep in mind the lessons given on that day, and train students to be able to think independently in doing assignments. The final main activity is Teaching, that provided by the teachers every day at school, providing daily teaching material for students, where the material taught is in accordance with the curriculum applied at school.

While for the supporting activities are finance and accounting, human resources, and admission & enrollment. For finance and accounting, work to record the financial statements of students, especially for reports income from school fees every month, uniform and book sales reports, and annual school budget data

collection for various needs such as school operational costs, budget for providing school equipment, and others.

Then, human resource is recording the recap of teacher attendance data, along with the application for leave or permit and record the data of teachers in schools that relate to official data in Dapodikdasmen. While admission & enrollment is tasked to accept the registration of new students at the school along with all the completeness documents that must be fulfilled by parents of prospective students.

The next design is to define the business process parts which are the details of each business process that can be seen in the use case diagram below (see Figure 4).

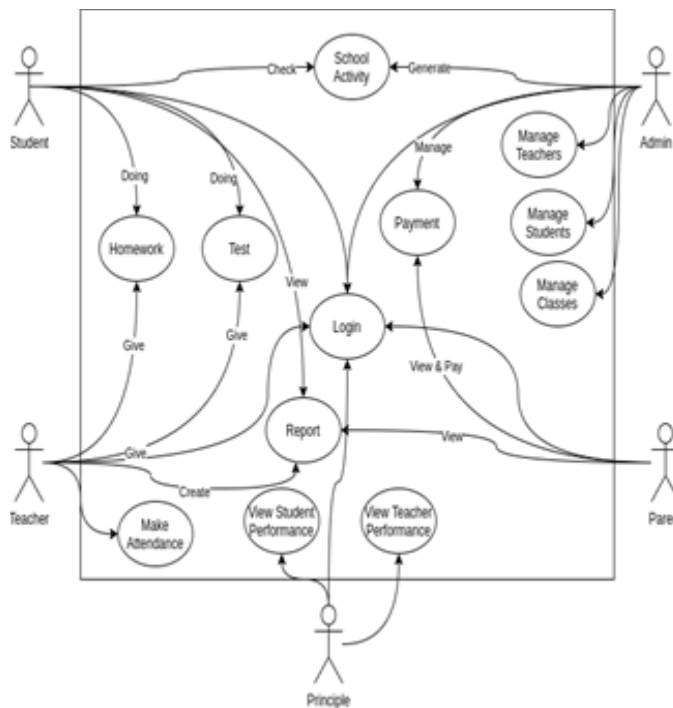


Figure 4: Use Case Diagram business process

From the Use Case Diagram above consist of five actors, they are student, teacher, parent, principle, and admin. They all have the access to the school system management. This use case diagram covers all main activities and support activities in Figure 3.

The principle only has the access to look after student and teacher overall performance as information. To maintain students' grades and keep the teacher on track and achieving their key performance indicator is the main principle's work to be done.

The admin has the access to manage teachers, students, and classes scheduling. Schedule management is crucial to ensure the student and the teachers neatly and precisely assigned to the same class with the subject. The admin also manages the any fees might needed for school operational such as monthly school fees, administration fees, etc. Any kind school activities will be updated by the admin and can directly inspected by the student.

The students and the teachers have interrelation process. The teachers will give assignments and the test to the students. The students will also complete their assignments or test's answer in reply. The teacher generates the report which can be viewed by the students. The teacher specifically has the authority of

attendance checking.

The parents usually only accomplish their duty in terms of payment and checking the report created by the teacher according to students' learning outcomes, especially homework and examination results.

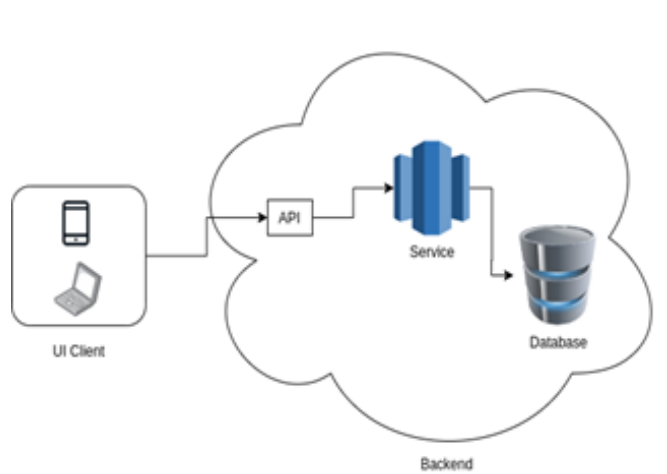


Figure 5: Infrastructure Architecture-Based Cloud Computing

Figure 5 shows a cloud-based architecture. Cloud architecture was chosen over architecture on premise because schools are public places, and it is feared that there will be physical damage or loss. With cloud-based architecture, it can scale their computing solutions as they need. When on the holiday, the student very rare to open the school management system, so the school can decrease the computing resources to make it more efficient. This solution also enables a flexible workplace. Because all information is stored in the cloud, the student and teacher can continue their activity from anywhere, especially when they cannot come to school. With all this advantage, cloud-based architecture can increase productivity.

5. Conclusion

Based on the results of the design that has been done and explained, it can be concluded that by using business processes and cloud technology, teaching and learning processes and other supporting processes can run more optimally. Improvements can be seen in interoperability, flexibility, data management, and efficiency. Apart from that, using cloud was chosen compared to on premise because of physical security reasons from cloud architecture far better than architecture on premise.

References

- [1] Harmon, P. (2003). *Business process change: a manager's guide to improving, redesigning, and automating processes*. Morgan Kaufmann. ISBN-13: 978-1558607583, ISBN-10: 1558607587
- [2] Harmon, P., & Trends, B. P. (2010). *Business process change: A guide for business managers and BPM and Six Sigma professionals*. Elsevier. Paperback ISBN: 9780123741523, eBook ISBN: 9780080553672
- [3] Nurhayati, L., & Setiadi, D. (2017). Pemodelan Proses Bisnis (Studi Kasus PD. Simpati Sumedang). *Infoman's: Jurnal Ilmu-ilmu Manajemen dan Informatika*, 11(1), 40-50. DOI: [10.33481/infomans.v11i1.39](https://doi.org/10.33481/infomans.v11i1.39)
- [4] Sapunar, D., Grković, I., Lukšić, D., & Marušić, M. (2016). The business process management software for successful quality management and organization: a case study from the University of Split School of Medicine. *Acta medica academica*, 45(1), 26. DOI: 10.5644/ama2006-124.153
- [5] Putra, D. M. D. U., & Welda, W. (2019). *Business Process Analysis and Modeling Using the Business Process Improvement Framework at the*

Internal Quality Assurance STMIK STIKOM Indonesia. ACSIE (International Journal of Application Computer Science and Informatic Engineering), 1(2), 75-86. DOI <https://doi.org/10.33173/acsie.53>

- [6] Röglinger, M., Seyfried, J., Stelzl, S., & zur Muehlen, M. (2017, September). Cognitive computing: what's in for business process management? An exploration of use case ideas. In *International Conference on Business Process Management* (pp. 419-428). Springer, Cham. DOI: 10.1007/978-3-319-74030-0_32
- [7] Hussein, S. M., Shariff, S. A., & Mantoro, T. (2017). Parental Involvement and Awareness Using Malaysian's Smart School Management System. *Advanced Science Letters*, 23(2), 712-716. ISSN 1936-6612
- [8] Moshe, T. (1999). A case study of the impact of school administration computerization on the department head's role. *Journal of Research on Computing in Education*, 31(4), 385-401. <https://doi.org/10.1177/0042085998033004004>
- [9] Demir, K. (2006). *School Management Information Systems in Primary Schools*. Online Submission, 5(2). ISSN: 1303-6521 volume 5 Issue 2 Article 6
- [10] Pyar, K., & Khaing, M. M. (2019). Clout Computing Basic : Feature and Service., *International Journal of Trend in Scientific Research and Development (IJTSRD)*, Vol. 3 Issuu 5, ISSN: 2456-6470, Volume-4 | Issue-3 , April 2020
- [11] Waloeyo, Y. J. (2012). *Cloud computing–Aplikasi Berbasis Web yang Mengubah Cara Kerja dan Kolaborasi Anda Secara Online*. Andi Offset, Yogyakarta. ISSN: 978-979-29-2302-5

Efficiency Enhancement of p-i-n Solar Cell Embedding Quantum Wires in the Intrinsic Layer

Nahid Akhter Jahan¹, M. Mofazzal Hossain^{2,*}

¹Department of Electrical and Electronic Engineering, Southeast University, Dhaka 1212, Bangladesh

²Department of Electrical and Electronic Engineering, University of Liberal Arts Bangladesh, Dhaka 1209, Bangladesh

ARTICLE INFO

Article history:

Received: 22 May, 2020

Accepted: 14 June, 2020

Online: 25 June, 2020

Keywords:

Quantum wire

Solar cell

Conversion efficiency

InAs/GaAs

Simulation

ABSTRACT

A high efficiency InAs/GaAs quantum wire solar cell is modelled embedding periodic array of InAs quantum wires (QW) in the intrinsic layer. The promising low dimensional heterostructure such as Quantum Wells, Quantum Wires, Quantum Dots or Dashed (elongated Dots) based intermediate-band-gap solar cells are recently being grasped the attention for ongoing third generation solar cell studies. In this particular work, we contrive, design and thereby simulate the solar cell incorporating QWs with a view to magnify the efficiency. After implementation of 15 layers of InAs QWs conjugated within the intrinsic layer and with the adaptation of AM1.5 solar irradiance, the proposed cell structure ensued a V_{oc} of 1.26 V, J_{sc} of 32.83 mAcm^{-2} and a fill factor of 89.4%, which eventuates an overall cell efficiency of 37%. We achieved an efficiency of 26.98% with the same materials and dimensions without QWs in intrinsic layer. Therefore, it may be optimistically appealed that the insertion of QWs in the intrinsic layer has an affirmative impact on the efficiency of the cell.

1. Introduction

It cannot be over-emphasized that the want for the development of sustainable energy sources are becoming paramount as the ubiquitous demand for energy supply is ever increasing with the constraint or the limitation of fossil fuel reserves which is indeed exhaustible. Our fixation for fossil fuels date back to industrial revolution and the unbridled consumption of these fuels, have been deliberately forgoing the earth toward climatological cataclysm. In many ways, we ceaselessly pushing our planet's terrestrial in catastrophic debacle and whacking the earth's eco-system by these fuel's burning. Eventually, we ended up making an ozone contraction, causing our climate an ever-hotter with ever-rising sea levels, a toxic muddle of nature, recurrent droughts, and alarming global crisis all owing to the gushing of lethal greenhouse gases (GHG) and air pollutants ensued from combustion of fossil fuels [1]. Therefore, there is no gainsaying the fact that we instead must start switching to non-pollutant sustainable sources of renewable energy and the scientists from all over the world are reiterating the world-leaders to patronize the technologies of non-polluting green-energy harness. Corollary, for combating this staggering climate crisis, solar photovoltaic (PV) cells have been prevalently

considered as one of the most promising sustainable green-energy technologies. Depending on the underlying technology, PV cells can be categorized into three generations. The first-generation PV is based on single or multi-crystalline silicon p-n junction cells. Second-generation solar cells are called thin-film solar cells as the semiconductor materials are only a few micrometers thick. However, for single junction solar cells either first- or second-generation, the thermodynamic efficiency limit of sunlight to electric power conversion efficiency (η) is 32.9% [2]. This limitation is well known as the Shockley-Queisser limit [3] which states that the photons with energies less than the bandgap energy are not absorbed and also the photons with energies greater than the bandgap energy release the additional energy in the form of heat which eventuates into thermalization loss. Therefore, the aim of third-generation solar cells is to achieve conversion efficiency which will surpass the Shockley-Queisser limit [4]. In this context, low dimensional hetero-structures (nano-structures), such as Quantum Wells (QWL), Quantum wires (QW), Quantum dots (QD), and elongated Quantum dots (Quantum dash) based solar cells have shown very promising prospect to achieve this expectation. It is worth stating that the QWL assures confinement of excitons in 1-dimension, whereas the QW assures s confinement of exciton in 2-dimension and QDs/dashes establish confinement in all 3-dimensions. Theses nano-structures, specially QW, and

* M. Mofazzal Hossain, Dhaka-1209, Bangladesh, +8801796587888, mofazzal.hossain@ulab.edu.bd

QDs exhibit outstanding opto-electronic properties owing to their ability to create veritable confinement with discrete and disparate energy states above the energy gaps of the bulk constituents [5-7].

They possess a unique property that is by changing the diameter of the QDs one can actually modulate their bandgap (the exciton energy-gap) which provides tunability of absorption spectrum [8]. It is seen that as the quantum diameter decreases, the breach or separation between the energy level increases. In line with this, it has been observed that as the separation between adjacent QDs or Quantum dashes is reduced, it results in the formation of coupled confined exciton states among these nano hetero-structures and thereby acts as a channel of charge-carriers transfer from high exciton-energy QDs to low exciton-energy QDs [9, 10]. All these phenomena establish that the exciton-bandgap and the dynamics of charge carriers are the function of size, proximity and positioning for all these QWL, QW, and QDs. Because of these properties, fabricating or positioning of QWL, QW, and QDs inside a semiconductor renders these structures as one of the potential candidates for realizing intermediate band solar cells (IBSCs) [11] as their discrete energy states act exactly like intermediate band (IB). IB utilizes sub-bandgap absorption which can thereby generate additional photocurrent from low-energy photons of the solar spectrum [4,12]. Convinced by this suitability of QDs, the p-i-n InAs/GaAs quantum dot embedded solar cell model was first modelled and proposed by [13]. In his model, by inserting QDs in intrinsic region the cell efficiency was increased from 19.5% (without QDs) to 25%. Since then quite a lot of studies have been done on InAs/GaAs QD solar cells to improve its efficiency compared to GaAs solar cells using the same structure [14-19]. Those reports demonstrate the achievement of circa 17% conversion efficiency in InAs/GaAs QD solar cell containing five stacks of QDs in intrinsic (i)-region by doping the InAs QDs with Si during the fabrication process.

In this work, we particularly intended to study the effect of Quantum Wire (QW) on the performance of intermediate band solar cell (IBSC). We investigated the plausible chance of increasing the efficiency of InAs/GaAs QW embedded intermediate band solar cell (IBSC) by stacking multiple InAs QW layers within the intrinsic layer. The modelling and simulation was performed using Silvaco TCAD software. Initially, the simulation was conducted without QWs in the intrinsic layer with an intention to compare this structure with the structure of similitude including QWs in intrinsic layer (InAs QWs buried in intrinsic-layer). Thereafter, secondly we carried out the simulation of InAs/GaAs QWs embedded p-i-n solar cell in the same software and adopting similar dimensions of the cell as above.

Here, we will depict our analysis and comparison between the results that upholds our insinuation and expectancy of efficiency enhancement for the structure having QWs inside. The associated design-calculation of band structure by determining exciton-confinement energies of InAs/GaAs QW is quantitatively performed and thereby included to corroborate our interpretation of the results. We, therefore, present an approach and deliver the proof to state that solar cell with Quantum Wires in the intrinsic region indeed improves the cell efficiency conceivably by capturing multiple photons through multiple discrete excitons states encompassing wider range of solar spectrum.

2. Structure and Modeling of p-i-n InAs/GaAs QW Solar Cell

2.1. Schematic Elucidation of the Proposed p-i-n InAs/GaAs QW Solar Cell

The schematic structure of proposed p-i-n InAs/GaAs solar cell is depicted precisely in Figure 1. In this structure of QW embedded p-i-n solar cell, we used GaAs as n-type (3.1 μm thick) and p-type (540 nm thick) layers at the top and bottom of the intrinsic region (112 nm thick) respectively and multiple layers of InAs QWs are buried or sandwiched within the GaAs barriers in the intrinsic region. Figure. 1 is an exemplar of 7-layers of QWs. The dimension of the InAs QW has a height of 5 nm with a width of 10 nm and the spacing between adjacent QWs are kept 10 nm throughout. The barrier height of GaAs beneath the InAs QW is chosen as 6.5 nm and the barrier height of GaAs as capping above the QW is chosen as 4.5 nm. The total thickness of each InAs QW embedded inside the GaAs-barriers (InAs QWs/GaAs-Barriers) becomes 16 nm (6.5 nm + 5 nm + 4.5 nm). The collective thickness of intrinsic layers becomes 112 nm (16 x 7 nm) for 7-layers of InAs QWs/GaAs-Barriers. The Aluminum and Molybdenum are used as anode and cathode, respectively, where the length of Aluminum is specified as 20 nm. Owing to the opacity of Aluminum, the active solar insolation region is quantified as 80 nm as seen in the Figure. 1. Therefore, each embedded layer of InAs QWs is comprised of four periodic QWs alongside. Furthermore, a highly doped p⁺ layer of 200 nm thick InGaP is introduced amid the back contact and the 540 nm thick p-type GaAs which is called the back-surface field (BSF) layer. BSF layer is commonly introduced to reduce the loss due to recombination and thus enhancing the conversion efficiency. At the very top of the solar cell, 498 nm thick ZnO is used as a transparent conductive oxide (TCO) layer to reduce the series resistance of the cell that mostly results from our thick n-type window layer, thereby increasing the open circuit voltage. Eventually, the net thickness of the semiconductor (p-i-n) becomes 3.852 μm , whereas including TCO-layer the net thickness becomes 4.35 μm . It is worth noting that, in the window layer a wider band-gap of 50 nm thick n-type InGaP (1.8 eV at 300 K) is inserted with a plausible chance of achieving meteoric rise in the absorption of higher energy photons even above the GaAs energy-gap (1.43 eV at 300 K) limit. For simplicity, during simulation, the top and bottom contacts are assumed to be Ohmic rather than Schottky.

The energy band diagram of the proposed p-i-n InAs/GaAs QW solar cell is illustrated in Figure 2 from which a comprehensive understanding of the operation of the proposed p-i-n InAs/GaAs QW solar cell is possible. Although practically the conduction and valance band offsets between barrier and QWs are not equal, however, for simplicity we sketched the band diagram considering the band offsets same for both conduction and valance band region. The effects of strain on the valance band has been neglected for this consideration. The engineering of energy band diagram at the intrinsic region due to the incorporation of

QWs as low dimensional structure will ensure more effective utilization of solar spectrum, wherein photons matched with multiple confinement energies of different sized QWs will be effectively absorbed. In this way it will be possible to utilize the wider range of solar spectrum for efficient conversion in comparison with typical single junction solar cells. Thereby this proposed structure is credibly suitable to overcome the Shockley-Queisser limit. The calculation of 1-dimensional (1-D), 2-dimensional (2-D) exciton energies and band-diagram for 1-D confinement of InAs/GaAs QW hetero-structure is represented in the next sub-section.

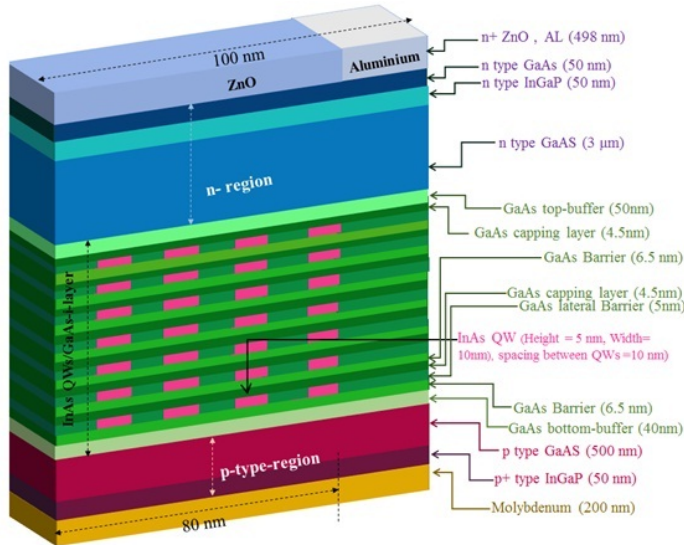


Figure 1. Structure of the proposed p-i-n InAs/GaAs QW solar cell

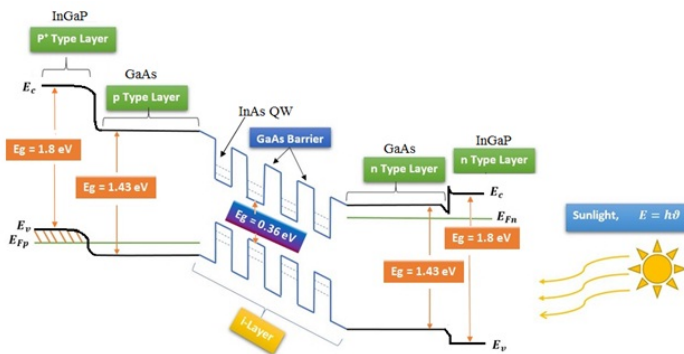
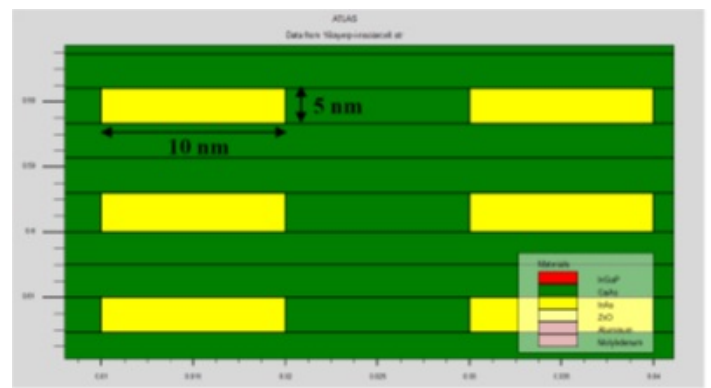
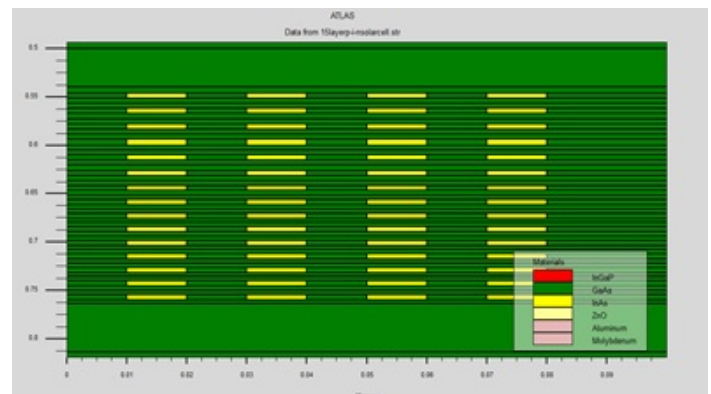
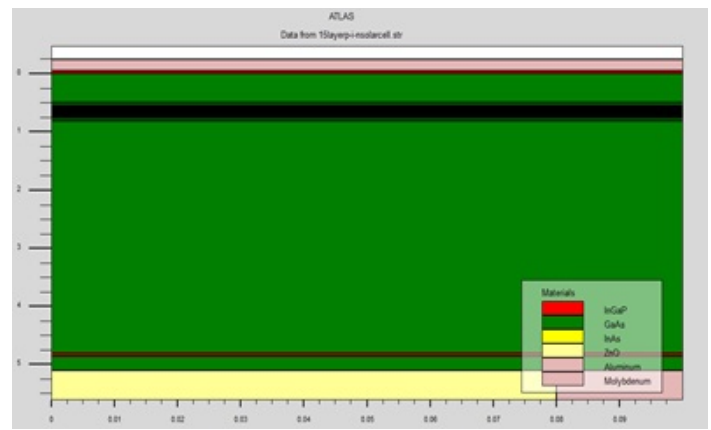


Figure 2. Energy band diagram of proposed p-i-n InAs/GaAs QW solar cell

Figure 3 (a) shows the proposed structure of InAs/GaAs QWs embedded solar cell generated in Silvaco environment. The zoomed-in view of InAs QW layers inside the intrinsic region produced in Silvaco is also shown in Figure 3 (b). The information regarding the dimension of QWs and buffer layers is illustrated in Figure 3 (c). In this structure we considered 15 layers of similar sized QWs. For all of our simulated structures including QWs, the dimension of each QW is consistently kept 10 nm×5 nm. The doping concentrations of n⁺, n, p, and p⁺ regions have been assumed as 2.5×10²⁰ cm⁻³, 1×10¹⁸ cm⁻³, 1×10¹⁷ cm⁻³, and 2×10¹⁸ cm⁻³ respectively.



(c)

Figure 3. (a) Structure of InAs/GaAs QW solar cell (x-axis and y-axis dimension are in microns) generated in Silvaco, (b) The zoomed-in view of 15 layers of QDs in intrinsic region generated in Silvaco, and (c) The extended zoomed-in view for observing the size of InAs QWs in intrinsic region (x-axis and y-axis dimension are in microns)

2.2. Calculation of Exciton Energies and Construction of Band-Diagram for 1-D Confinement of InAs/GaAs QW Hetero-Structure

For comprehensive understanding of the formation of intermediate energy bands, confinement energies of excitons (electrons and hole pair) in InAs QW have been calculated by solving the Schrödinger equation for finite-potential barriers (GaAs-barriers on both side of QW). Our calculated exciton energies (for one dimensional confinement) are illustrated in

Figure 4 and summarized in Table: 1 for different thickness of QWs. Relied on the accessible evidence, the QW confinement-energies calculation have been performed from the solution of the empirical Schrödinger equation by assuming the conduction band offset as 70% (or valence band offset of 30%) [20].

noted that for the case of InAs/InP hetero-structures, the band offset is nearly equally distributed to the both bands (conduction and valence), whereas the band offset is found mostly contained to the conduction band (87%) in our previous study of antimony based Quantum Well [22].

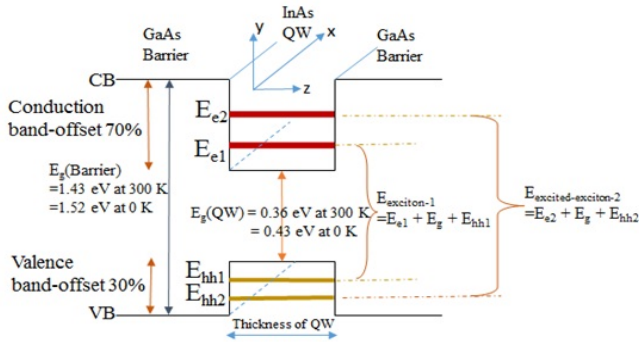


Figure 4. Calculated ground state and excited states confinement energies of excitons (Electron-hole pairs) and construction of energy-band-diagram for 1-dimensional confinement of InAs/GaAs QW hetero-structure

The physical pertinent parameters, e.g., energy bandgap, electron and heavy hole effective masses, the lattice constant, the magnitudes of stiffness constants, and deformation potentials for strain calculation were all adopted from the established values for InAs and GaAs [21]. The associated equations and details of this calculation can be found in our previous work performed for InGaSb/AlGaSb Quantum Well [22]. For most of the cases III-V hetero-structures such as for InAs/GaAs, the band offset is nearly localized to the conduction band. In this case, for confinement energies calculation and band structure development, we assumed the valence band offset as 30%. This lesser valence band offset is analogous to the formerly published standards [20]. It is to be

As we can see that for a QW of height 5 nm, the ground-state exciton energy at 0 K is achieved as— $E_{exciton-1} = E_{e1} + E_g(QW) + E_{hh1} = 774$ meV. At 300 K the ground-state exciton energy will be reduced to $E_{exciton-1} = 704$ meV, because of the temperature dependent bulk E_g of InAs (360 meV at 300 K). Whereas for a QW of 10 nm thick, the ground-state exciton energy at 0 K is determined as— $E_{exciton-1} = E_{e1} + E_g(QW) + E_{hh1} = 552$ meV, and at 300 K this exciton energy will be reduced to $E_{exciton-1} = 482$ meV. Alongside the calculation ensued the excited-state exciton energy as— $E_{excited-exciton-2} = E_{e2} + E_g(QW) + E_{hh2} = 1026$ meV at 0 K, and 956 meV at 300 K. Where, E_{e1} is the energy of ground-state electron confinement, E_{hh1} is the energy of ground-state heavy-hole confinement, E_{e2} is the energy of excited-state electron confinement, E_{hh2} is the energy of excited-state heavy-hole confinement, and $E_g(QW)$ is the energy gap of InAs. (Since the strain contribution for InAs/GaAs is not more than a few meV, it has been neglected).

In our proposed structure, the dimension of our QW is selected as 10 nm × 5 nm (width in y × thickness in z). Therefore, the confinement is assured in 2-dimensional, i.e. the exciton is confined in both y and z direction while along the x-direction is free. Thus the total ground-state exciton energy will be — $E_{exciton-1}$ (2-D confinement) = $E_{e1(z)} + E_{e1(y)} + E_g(QW) + E_{hh1(z)} + E_{hh1(y)} = (307 + 112 + 360 + 37 + 10) = 826$ meV. If we subtract the exciton binding energy of InAs (1.8 to 2 eV [23]), the resultant

Table 1: Electron and heavy-hole confinement of InAs/GaAs

Thickness of QW, L_z (nm)	E_{e1} (meV)	E_{e2} (meV)	E_{hh1} (meV)	E_{hh2} (meV)	$E_g(QW)$ (meV) At 0 K	$E_g(QW)$ (meV) At 300 K	$E_{exciton-1} = (E_{e1} + E_{hh1} + E_g(QW))$ (meV) at 0 K	$E_{exciton-1} = (E_{e1} + E_{hh1} + E_g(QW))$ (meV) at 300 K	$E_{excited-exciton-2} = (E_{e2} + E_{hh2} + E_g(QW))$ (meV) at 0 K	$E_{excited-exciton-2} = (E_{e2} + E_{hh2} + E_g(QW))$ (meV) at 300 K
2	598		164		430	360	1192	1122		
2.5	541		122		430	360	1093	1023		
3	486		92		430	360	1008	938		
3.5	434		71		430	360	935	865		
4	387		56		430	360	873	803		
4.5	345		45		430	360	820	750		
5	307		37		430	360	774	704		
5.5	274		31		430	360	735	665		
6	245		27		430	360	702	632		
6.5	220	763	23	118	430	360	673	603	1311	1241
7	198	753	20	103	430	360	648	578	1286	1216
7.5	179	727	17	91	430	360	626	556	1248	1178
8	162	694	15	81	430	360	607	537	1205	1135
8.5	147	660	14	72	430	360	591	521	1162	1092
9	135	624	13	63	430	360	578	508	1117	1047
9.5	124	586	12	54	430	360	566	496	1070	1000
10	112	550	10	46	430	360	552	482	1026	956

$E_{exciton-1}$ (2-D confinement) becomes 824 meV which is equivalent to 1500 nm of wavelength. Similarly the total excited-state exciton energy will be $-E_{excited} - exciton-2$ (2-D confinement) = $E_{e2(z)} + E_{e2(y)} + E_g(QW) + E_{hh2(z)} + E_{hh2(y)} = (0 + 550 + 360 + 45 + 0) = 953$ meV. Subtraction of exciton binding energy for excited states can be eliminated, because exciton in ground state are more confined than that of the excited states. From the aforementioned analysis it can be inferred that the proposed structure of QW provides an intermediate band (IB) of energies from 826 to 956 meV in the i-layer of p-i-n diagram, which corresponds to a wavelength range of 1290 to 1500 nm. Therefore, our QW-based IBSC can afford multiple lower-energy photons absorption covering this wavelength range (1290-1500 nm) of insolation together with the absorption of higher-energy photons in p and n-GaAs/InGaP of window and base layers. Typically, the maximum photons that can be absorbed by GaAs is around 875 nm [24] and by InGaP is around 688 nm.

3. Results and Discussion

In this work, we considered a one-sun solar irradiance of AM 1.5 for simulation. At first the simulation was carried out for p-i-n solar cell without QWs and achieved an efficiency of 26.9%. Then, simulation was further carried out to investigate the effects of stacking multiple layers of InAs quantum wires in the intrinsic region and to find out how it impacts the conversion efficiency for this proposed cell. For this purpose, the simulation is conducted particularly for three distinct cases; structure's having 7, 10 and 15-layers of QWs in the intrinsic layers. Table 2 shows the simulation results which include the density of short circuit current (J_{sc}), open circuit voltage (V_{oc}), fill factor (FF), and conversion efficiency (η).

Table 2: Characteristic parameters of InAs/GaAs solar cell for various number of QW layers

Number of QD layers	J_{sc} (mA/cm ²)	V_{oc} (Volts)	FF (%)	η (%)
0	28.2	1.08	88.4	26.9
7	32.2	1.26	89.4	36.3
10	32.5	1.26	89.5	36.6
15	32.9 \cong 33	1.26	89.4	37

Now if we compare and analyse the simulation results presented in Table 2, it can be realized that with the increasing number of embedded QW layers inside the intrinsic region the short circuit current density increases, which ultimately indicates the generation of additional excitons (electron-hole pairs). However, with the increasing number of QW layers the open circuit voltage remains completely constant. This is very likely and expected because the upper limit of open circuit voltage predominantly depends on the energy gap, band offsets of the QWs/barriers and the confinement energies of QWs, not on the surplus creation of excitons. Figure 5 shows the current density vs voltage curve of InAs/GaAs quantum dot solar cell exemplarily for 0 and 15 layers of QWs. When comparing the short circuit current densities of InAs/GaAs solar cells without QWs and with 15 layers

of QWs, it is found that there is a relative augmentation of about 16.2%.

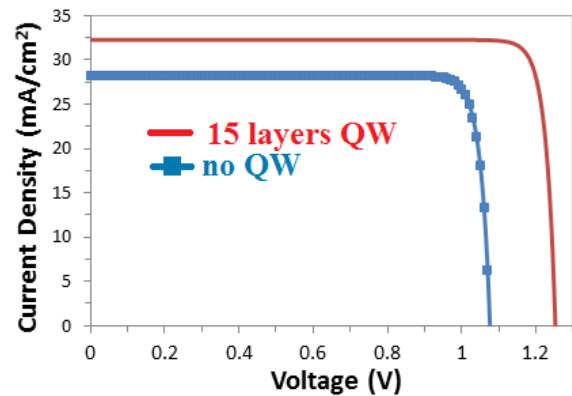


Figure 5. Current density vs voltage characteristics of InAs/GaAs QWSC (QWs layers = 0, 15)

The conversion efficiency of a solar cell indicates the probability that an incident photon would generate an exciton (electron-hole pair). With the increasing number of quantum dot layers in the intrinsic region, the conversion efficiency significantly increases as shown in Figure 6. For InAs/GaAs QW solar cell, 37% conversion efficiency is achieved for 15 layers of QWs whereas for InAs/GaAs solar cell without QWs it is only 26.9%. As we quantitatively calculated the exciton energies and band-diagram for 1-D confinement of InAs/GaAs QW hetero-structure in our previous section (b. sub-section of section 2), we observed that the proposed structure of QW can afford an intermediate band (IB) of energies from 826 to 956 meV in the i-layer, which corresponds to a range of wavelength 1290 to 1500 nm. Thereby it can absorb multiple lower-energy photons covering this wavelength range (1290-1500 nm) of light together with the absorption of higher-energy photons by the window and base layers (GaAs and InGaP). As previously discussed, low dimensional hetero-structure-based solar cells are considered to be a promising candidate in mimicking the characteristics of intermediate band solar cells. The advantage of using QW in the intrinsic region is the absorption of lower energy photons having longer wavelengths which eventually leads to greater conversion efficiency. For a typical GaAs solar cell, the maximum photons that can be absorbed is around 875 nm [24]. Also owing to the inclusion of InGaP in both n and p-regions ensures photons absorption in around 688 nm. Compared to a typical GaAs solar cell the absorption range would be much wider for this QW solar cell due their ability to absorb even lower energy photons. From our analysis, it can be inferred that, our proposed quantum dot embedded structure is indeed helpful for the realization of solar cells to absorb photons with lower energy having wavelengths from 1290 nm to 1500 nm.

The simulation is also conducted to study the effects of thickness of n-type window layer on the cell efficiency while the thickness of all the other layers remained unchanged as shown in Figure 1. It is evident from Figure 7 that there is an almost linear increase in conversion efficiency with the increase in thickness of n-type window layer. This phenomenon can be understood by the fact that larger the n-layer thickness smaller the series resistance of the cell. However, it is noticed that when the thickness of n-type

window layer is increased from 2 μm to 4 μm , the conversion efficiency increases insignificantly. Moreover, the effect of concentration of n-type window layer on efficiency is also studied. Figure 8 depicts the effects of dopant concentrations in the n-type window layer and it is observed that there is a very negligible increase in conversion efficiency with an increase in dopant concentration.

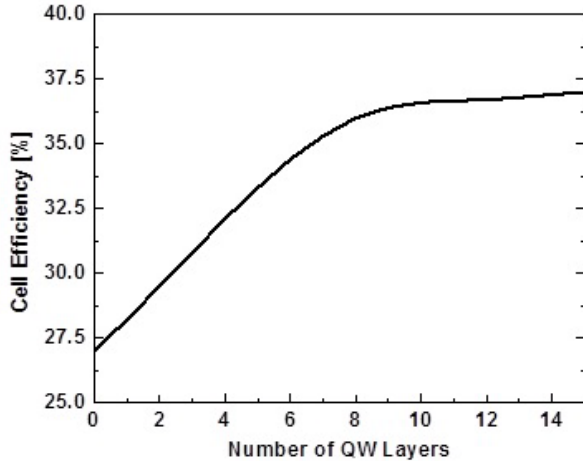


Figure 6. Evolution of cell efficiency with the number of QW layers

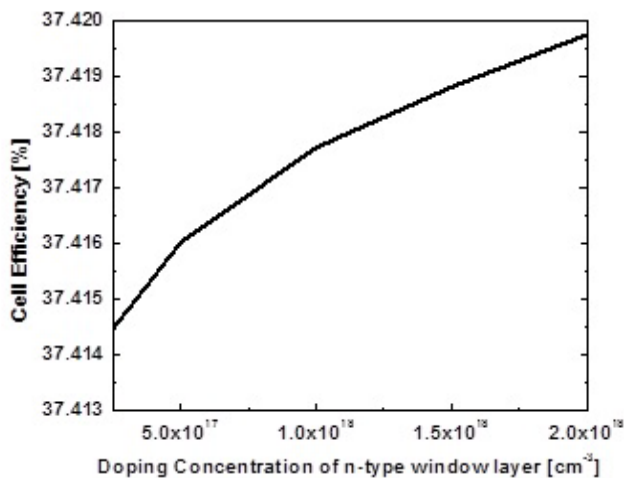


Figure 7. The dependence of Cell efficiency on the thickness of n-type window layer

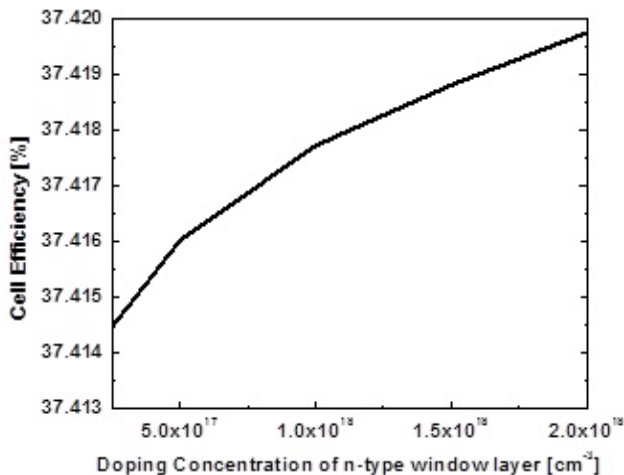


Figure 8. Evolution of cell efficiency with the dopant concentration of n-type window layer

4. Conclusion

InAs/GaAs solar cell is a distinctive example of III-V heterostructure to achieve higher conversion efficiency thereby can be considered as one of the prospective candidates for the realization of solar cell having greater performance. By embedding QWs in the intrinsic region, higher conversion efficiency for third generation solar cell can be realized. In this work, the impacts on the characteristic parameters of a solar cell due to the insertion of stacked QW layers in the intrinsic region are studied. We proposed 7 to 15-layers of QWs based solar cell structure that reveals the highest reachable conversion efficiency of 37% whereas p-i-n solar cell without QWs shows a conversion efficiency of only 26.9%. The insertion of QWs in the intrinsic region improves the photon absorption range by absorbing lower energy photons in the IB having longer wavelength from 1290 nm to 1500 nm. However, the technical limitation of this work is that the achieved conversion efficiency is exclusively based on computer simulation, we could not able to fabricate the cell to verify the achieved conversion efficiency.

References

- [1] A. Nishimura, Y. Hayashi, K. Tanaka, M. Hirota, S. Kato, Masakazu Ito, K. Araki, and E. J. Hu, "Life Cycle Assessment and Evaluation of Energy Payback Time on High-Concentration Photovoltaic Power Generation System," *Applied Energy*, 87(9), pp.2797-2807, 2010. <https://doi.org/10.1016/j.apenergy.2009.08.011>
- [2] Andrew S. Brown and Martin A. Green, "Detailed Balance Limit for The Series Constrained Two Terminal Tandem Solar Cell," *Physica E: Low-Dimensional Systems and Nanostructures* 14 (1-2), pp.96-100, 2002. [https://doi.org/10.1016/S1386-9477\(02\)00364-8](https://doi.org/10.1016/S1386-9477(02)00364-8)
- [3] W. Shockley and Hans J. Queisser, "Detailed Balance Limit of Efficiency of p-n Junction Solar Cells" *Journal of Applied Physics* 32 (3): 510, 1961. <https://doi.org/10.1063/1.1736034>
- [4] Antonio Luque and Antonio Martí, "Increasing the Efficiency of Ideal Solar Cells by Photon Induced Transitions at Intermediate Levels," *Physical Review Letters* 78(26), pp.5014, 1997. <https://doi.org/10.1103/PhysRevLett.78.5014>
- [5] Polina O. Anikeeva, Jonathan E. Halpert, Mounqi G. Bawendi and Vladimir Bulović, "Quantum Dot Light-Emitting Devices with Electroluminescence Tunable over the Entire Visible Spectrum," *Nano Letters* 9 (7), pp. 2532–2536, 2009. <https://doi.org/10.1021/nl9002969>
- [6] Jdira, Lucian, Peter Liljeroth, Eric Stoffels, Daniël Vanmaekelbergh, and Sylvia Speller, "Size-dependent Single-Particle Energy Levels and Intersubband Coulomb Interactions in CdSe Quantum Dots Measured by Scanning Tunneling Spectroscopy," *Physical Review B* 73 (11): 115305, 2006. <https://doi.org/10.1103/PhysRevB.73.115305>
- [7] An, J. M., A. Franceschetti, S. V. Dudyi, and Alex Zunger, "The Peculiar Electronic Structure of PbSe Quantum Dots," *Nano Letters* 6 (12): 2728–2735, 2006. <https://doi.org/10.1021/nl061684x>
- [8] Takagahara, Toshihide, and Kyozauro Takeda, "Theory of the Quantum Confinement Effect on Excitons in Quantum Dots of Indirect-Gap Materials" *Physical Review*, V 46(23):15578, 2010. <https://doi.org/10.1103/PhysRevB.46.15578>
- [9] Nahid A Jahan, Claus Hermannstädter, Jae-Hoon Huh, Hirotaaka Sasakura, Thomas J Rotter, Pankaj Ahirwar, Ganesh Balakrishnan, Kouichi Akahane, Masahide Sasaki, Hidekazu Kumano and Ikuo Suemune, "Temperature dependent carrier dynamics in telecommunication band InAs quantum dots and dashes grown on InP substrates," *J. Appl. Phys.* 113: 033506 2013. <https://doi.org/10.1063/1.4775768>
- [10] C. Hermannstädter, N. A. Jahan, J.-H. Huh, H. Sasakura, K. Akahane, M. Sasaki, and I. Suemune, "Inter-dot coupling and excitation transfer mechanisms of telecommunication band InAs quantum dots at elevated temperatures," *New Journal of Physics*, vol. 14, p.023037 (2012). <https://iopscience.iop.org/article/10.1088/1367-2630/14/2/023037/meta>
- [11] Antonio Luque and Antonio Martí, "The Intermediate Band Solar Cell: Progress Toward the Realization of an Attractive Concept." *Advanced Materials* 22 (2): 160-174, 2010. <https://doi.org/10.1002/adma.200902388>
- [12] Jiang Wu, Y. F. M. Makableh, R. Vasan, M. O. Manasreh, B. Liang, C. J. Reyner, and D. L. Huffaker, "Strong Interband Transitions in InAs Quantum Dots Solar Cell." *Applied Physics Letters* 100: 051907, 2012. <https://doi.org/10.1063/1.3681360>

- [13] V. Aroutiounian, S. Petrosyan, and A. Khanchatryan, "Quantum Dot Solar Cells." *Journal of Applied Physics* 89(4): 2268, 2001. <https://doi.org/10.1063/1.13392>
- [14] R. B. Laghumavarapu, M. El-Emawy, N. Nuntawong, A. Moscho, L. F. Lester, and D. L. Huffaker, "Improved Device Performance of InAs/GaAs Quantum Dot Solar Cells with GaP Strain Compensation Layers," *Applied Physics Letters* 91, 243115, 2016. <https://doi.org/10.1063/1.2816904>
- [15] P. Lam, J. Wu, M. Tang, D. Kim, S. Hatch, I. Ramiro, V. G. Dorogan, M. Benamara, Y. I. Mazur, G. J. Salamo, J. Wilson, R. Allison, and H. Liu, "InAs/InGaP Quantum Dot Solar Cells with an AlGaAs Interlayer," *Solar Energy Materials and Solar Cells*, 144: 96-101, 2016. <https://doi.org/10.1016/j.solmat.2015.08.031>
- [16] Alonso-Álvarez, D., A. G. Taboada, J. M. Ripalda, B. Alén, Y. González, L. González, J. M. García, F. Briones, A. Martí, A. Luque, A. M. Sánchez, and S. I. Molina, "Carrier Recombination Effects in Strain Compensated Quantum Dot Stacks Embedded in Solar Cells." *Applied Physics Letters* 93: 123114, 2008. <https://doi.org/10.1063/1.2978243>
- [17] N. S. Beattie, G. Zoppi, P. See, I. Farrer, M. Duchamp, D. J. Morrison, R. W. Miles, and D. A. Ritchie, "Analysis of InAs/GaAs quantum dot solar cells using Suns- V_{OC} measurements." *Solar Energy Materials and Solar Cells* 130: 241-245, 2014. <https://doi.org/10.1016/j.solmat.2014.07.022>
- [18] Ayami Takata, Ryuji Oshima, Yasushi Shoji, Kouichi Akahane and Yoshitaka Okada, "Fabrication of 100 layer-stacked InAs/GaNAs strain-compensated quantum dots on GaAs (001) for application to intermediate band solar cell." *35th IEEE Photovoltaic Specialists Conference, Honolulu, HI, USA, November 001877-001880, 2010.* <https://doi.org/10.1109/PVSC.2010.5616186>
- [19] Xiaoguang Yang, Kefan Wang, Yongxian Gu, Haiqiao Ni, Xiaodong Wang, Tao Yang and Zhanguo Wang, "Improved Efficiency of InAs/GaAs Quantum Dots Solar Cells by Si-doping." *Solar Energy Materials and Solar Cells* 113: 144-147, 2013. <https://doi.org/10.1016/j.solmat.2013.02.005>
- [20] Yin Wang, Ferdows Zahid, Yu Zhu, Lei Liu, Jian Wang, and Hong Guo, "Band offset of GaAs/Al_xGa_{1-x}As heterojunctions from atomistic first principles" *Appl.Phys.Lett.* 102, 132109 (2013). <https://doi.org/10.1063/1.4800845>
- [21] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for III-V compound semiconductors and their alloys", *Journal of Applied Physics* 89, 5815 (2001); <https://doi.org/10.1063/1.1368156>
- [22] Nahid A. Jahan, Claus Hermannstadter, Hirotaka Sasakura, Thomas J. Rotter, Pankaj Ahirwar, Ganesh Balakrishnan, Hidekazu Kumano, and Ikuo Suemune, "Carrier dynamics and photoluminescence quenching mechanism of strained InGaSb/AlGaSb quantum wells", *Journal of Applied Physics* 113, 053505 (2013). <https://doi.org/10.1063/1.4789374>
- [23] P. J. P. Tang, M. J. Pullin, and C. C. Phillips "Binding energy of the free exciton in indium arsenide", *Phys. Rev. B* 55, 4376 – Published 15 February 1997. <https://doi.org/10.1103/PhysRevB.55.4376>
- [24] Christopher G. Bailey, David V. Forbes, Stephen J. Polly, Zachary S. Bittner, Yushuai Dai, Chelsea Mackos, Ryne P. Raffaele, Seth M. Hubbard, "Open-Circuit Voltage Improvement of InAs/GaAs Quantum-Dot Solar Cells Using Reduced InAs Coverage," *IEEE Journal of Photovoltaics* 2 (3): 269-275, 2012. <https://doi.org/10.1109/JPHOTOV.2012.2189047>

Efficient Discretization Approaches for Machine Learning Techniques to Improve Disease Classification on Gut Microbiome Composition Data

Hai Thanh Nguyen^{*1}, Nhi Yen Kim Phan¹, Huong Hoang Luong², Trung Phuoc Le², Nghi Cong Tran³

¹College of Information and Communication Technology, Can Tho University, Can Tho city, 900100, Vietnam.

²Department of Information Technology, FPT University, Can Tho city, 900000, Vietnam.

³National Central University, Taoyuan, 320317, Taiwan, R.O.C.

ARTICLE INFO

Article history:

Received: 10 May, 2020

Accepted: 18 June, 2020

Online: 28 June, 2020

Keywords:

Personalized medicine

Bacterial composition

Classic machine learning

Discretization

Species abundance

Read counts

Disease prediction

Deep learning

ABSTRACT

The human gut environment can contain hundreds to thousands bacterial species which are proven that they are associated with various diseases. Although Machine learning has been supporting and developing metagenomic researches to obtain great achievements in personalized medicine approaches to improve human health, we still face overfitting issues in Bioinformatics tasks related to metagenomic data classification where the performance in the training phase is rather high while we get low performance in testing. In this study, we present discretization methods on metagenomic data which include Microbial Compositions to obtain better results in disease prediction tasks. Data types used in the experiments consist of species abundance and read counts on various taxonomic ranks such as Genus, Family, Order, etc. The proposed data discretization approaches for metagenomic data in this work are unsupervised binning approaches including binning with equal width bins, considering the frequency of values and data distribution. The prediction results with the proposed methods on eight datasets with more than 2000 samples related to different diseases such as liver cirrhosis, colorectal cancer, Inflammatory bowel disease, obesity, type 2 diabetes and HIV reveal potential improvements on classification performances of classic machine learning as well as deep learning algorithms. These binning approaches are expected to be promising pre-processing techniques on various data domains to improve the performance of prediction tasks in metagenomics.

1 Introduction

This paper is an extension of work originally presented in The 11th IEEE International Conference on Knowledge and Systems Engineering (KSE) 2019 in Da Nang, Vietnam [1].

Recent years, the field of health care has been receiving great attention from the world. Many services and high-tech applied equipment are manufactured for medical use. Medical results require high accuracy and meet a wide range of diseases, so it is necessary to deploy diagnostic methods and treatments with new technologies. Also, dangerous diseases such as Liver Cirrhosis, Colorectal, HIV, etc. have an increasing rate due to lifestyle, way of life, diet. Liver Cirrhosis is a disease of concern due to the increasing trend each year. The main cause of cirrhosis is the living environment, using a lot of alcohol, toxic chemicals. The level and duration of consumption is an important determinant of the

development of liver pathology. In 2015, cirrhosis was the 12th leading cause of death in the United States, with a total of 42,443 deaths 2,494 compared to 2014 [2]. Colorectal cancer is the third most common disease in the United States. The main object of the disease is elderly but the proportion of young people with the disease tends to increase. According to statistics in early 2020, an estimated 53,200 deaths (28,630 males and 24,570 females) are due to the disease, but there is an increased incidence in young people. Although the incidence decreases by 3.6% per year from 2007 to 2016 in adults 55 years and older, they increase 2% per year in adults under 55 years. This year, colorectal cancer is estimated to be the fourth most commonly diagnosed cancer in the United States for men and women aged 30 to 39 [3]. Heart failure is a common complication of Cardiometabolic diseases (CMD). When the pumping action of the heart weakens, the amount of blood pumped out

*Corresponding Author: Hai Thanh Nguyen, Email: nthai@cit.ctu.edu.vn, nthai.cit@ctu.edu.vn

is insufficient for the body to make it difficult for the person to feel short of breath, or chest pain, which is called heart failure. Genetic diversity has not been considered in the diagnosis and prognosis of the disease. We apply a single method of treatment to all patients with a similar diagnosis. The results showed that some patients' health did not improve, and the rest gradually recovered. This shows that many methods need to be applied for an effective treatment for each patient. Advances in data processing technology make us understand the importance of metagenomic to human health and explore the diversity of genetics. Deep learning has provided many algorithms to help scientists propose models, methods of diagnosis and treatment.

Modern techniques in healthcare are still developing at a great speed. One of them is Personalized Medicine which defines the impact goals that will work for a patient based on the patient's environmental factors, genes, etc. and is used on a group of patients. Today, scientists have studied numerous methods for Personalized Medicine and metagenomic is one of them. As we have known, metagenomics is a method of sequencing and analyzing the DNA of microorganisms collected from the environment without culturing them. We are looking at the human gut environment. Bacteria are often very diverse, they are classified into seven basic types: domain, kingdom, phylum, class, order, family, genus, and species. This diversity helps to provide more information about diseases to support more effective diagnosis and prognosis. The diseases under consideration are complex and we only have a limited number of observation data samples, so the prediction tasks are produced in inconsistent results with comparable diseases.

To test and propose models for healthcare services, Machine learning algorithms have been strongly researched and developed to solve metagenomics-related problems, mainly prediction genes [4, 5], Operational Taxonomic Unit clustering [6–9], comparative metagenomics [10–13], binning, taxonomic profiling and assignment [14]. All of the issues mentioned are given in [14].

The authors presented the basic principles of machine learning in [15] and a typical pipeline was introduced [16, 17]. This model is clustered or classified based on pre-processed results and feature extraction. A machine learning program consists of three basic components, such as data (or experience), task (formed by the output of the algorithm) and target (possibly in the form of measuring the efficiency of the output). MetAML was introduced in [17] by Edoardo et al. which is a computational framework that learns about the presence of specific species signs and the relative abundance of species, these two collectively called are quantitative microbiome configurations. Using machine learning to independently evaluate the accuracy of models on large metagenomic datasets, thereby analyzing and comparing the practical microbiome usage strategies that were recommended in [18]. With the task of classifying metagenomics, Ph-CNN [19] was introduced to the OTU hierarchical structure and it was also compared to other technologies in machine learning such as random forests. PopPhy-CNN [20] is introduced by D. Reiman et al. PopPhy-CNN is deep learning framework, using embedded information based on a phylogenetic tree to predict diseases from metagenomic data. PopPhy-CNN has superior performance compared to random forests, support vector machines, LASSO and the basic 1D-CNN model built with bacterial vectors. In addition to retrieving information microbiological

classifications from trained CNN models, PopPhy-CNN also visualizes phylogenetic tree classifications. Several deep learning algorithms have been evaluated as a feasible approach to speed up DNA sequencing [21]. To identify viruses by deep learning with metagenomic data, J. Ren et al. [22] proposed using DeepVir Downloader (a reference-free and alignment-free machine learning method) to improve accuracy and support virus research. We will present 1D representations through packaging and expansion methods, and demonstrate the effectiveness of applying Multi-layer Perceptron (MLP), traditional artificial neural networks to perform predictive tasks.

In the study [23], the authors presented the MSC algorithm with the goal of classification to detect the circulating rate and estimate their relative level. MSC is a metagenomic sequence classification algorithm that has accuracy, memory and runtime and gives an approximate estimate of abundance over other algorithms. Jolanta Kawulok [24] presented research on the environmental classification of metagenomic data to build a microbiome fingerprint. Another study by Lo Chieh et al. [25] proposed the MetaNN model, this is a model of host phenotypes classification from metagenomic data using neural networks. The results show that MetaNN is superior to the exact classification team for synthetic and real metagenomic data compared to other models, contributing to the development of microbiome-related disease treatments.

In this research, we have investigated and implemented a variety of binning methods to improve predictive performance. We have proposed different binning methods including binning based on frequency, width, the proposed breaks and combination between scaler and binning. After applying binning approaches, the data are fetched into machine learning algorithms including both classical machine learning and deep learning. We present the results which are produced with more running times in [1] for deeper comparison and include p-values for finding significant results. Additionally, we include another dataset, Crohn disease, in the experiments, for a complete comparison with the state-of-the-art. The results of [1] run by only MultiLayer Perceptron, we extend to run with a variety of machine learning algorithms including classical algorithms such as Random Forest, Linear Regression and deep learning (Convolutional Neural Network (CNN)) technique. The number of bins also carried out to consider in this work for choosing the number of bins for metagenomic data binning approaches. The work is expected to provide robust pre-processing methods to enhance the performance of machine learning algorithms applying to metagenomic data. Results, datasets and scripts for the experiments are uploaded to the public GitHub repository. To sum up, the contributions of this work include:

- The study presents various binning methods. The width of all bins can be equal or the width of each bin is conducted from the frequency of values. We also consider binary bins to determine whether the feature exists in the considered sample or not. The proposed methods as shown results can improve the performance comparing the original data.
- A combination of scaler algorithms and binning methods is also introduced for the comparison. Some scalers such as logarithm calculations and quantile transformation reveal good performance on some datasets.

- Methods are evaluated by disease prediction tasks on a variety of diseases including liver cirrhosis, colorectal cancer, IBD, obesity, HIV, Type 2 diabetes. Considered data types include species abundance and counts at other taxonomic ranks such as genus, family, etc.
- Several machine learning techniques including both classic machine learning and deep learning are investigated with Classification tasks on metagenomic data.
- The proposed framework, namely Metagenomic-To-Bins (**Met2Bin**), including scripts, results and datasets is published at <https://github.com/thnguyencit/met2bin>.

In the next sections of the paper, we describe 8 metagenomic datasets used in the experiments including the total number of features, the total number of samples, both the number of disease and non-disease (Section 2). In Section 3, we introduce the metagenomic data binning with various approaches along with scaler algorithms. Section 4 describes the empirical results and Section 5 provides insightful remarks of the study.

2 Metagenomic data benchmarks

To evaluate the performance of classifiers, we run the prediction tasks on a variety of datasets (8 metagenomic datasets) including species abundance datasets and read counts related to different specific diseases, such as Liver cirrhosis (CIR), colorectal cancer (COL), Crohn’s disease, Human Immunodeficiency Virus (HIV) infection, Inflammatory Bowel Disease (IBD), Obesity (OBE), Type 2 Diabetes (T2D and T2W). Each dataset consists of 4 main parameters: (1) the number of features, (2) the number of samples, (3) the number of samples affected by the disease, (4) the number of healthy samples.

CIR dataset comprises 542 features with 232 samples including 118 patients and 114 healthy individuals. COL dataset consists of 121 individuals with 48 patients. The number of patients affected by Crohn’s disease is 663 out of 975 people were considered. For HIV dataset, the number of positive cases is 129 out of 155. IBD dataset includes 253 samples of which 164 are affected by the disease, OBE dataset consists of 174 non-obese and 170 obese individuals. T2D and WT2 datasets include 344 samples and 96 samples, respectively.

The HIV, Crohn’s datasets contain 155 and 975 samples, respectively, which have values greater than 1, evaluated using the recommended method [26] with the number of reads for microbial taxa at the levels which are higher than species. Crohn’s disease is a type of inflammatory bowel disease (IBD). This disease can affect any segment of the gastrointestinal tract from the mouth to the anus. The features in two these datasets can be genus counts, family counts, or order counts. We bring read counts datasets from the analysis in [27] to compare to our method.

For species abundance datasets, each sample, species abundance is a relative proportion and it is revealed as a real number that has the total abundance of all species summing to 1 (The details are shown in Table 1).

Let D be the set of considered datasets, $D = \{d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8\}$, with $d_1 = CIR$, $d_2 = COL$, $d_3 = Crohn$, $d_4 = HIV$, $d_5 = IBD$, $d_6 = OBE$, $d_7 = T2D$, $d_8 = WT2$, $d = 1..8$

- $F_i = \{f_1, f_2, \dots, f_m\}$ includes m features corresponding to d_i
- $S_i = \{s_1, s_2, \dots, s_n\}$ includes n samples corresponding to d_i
- $P_i = \{p_1, p_2, \dots, p_k\}$ includes k patients who affected by diseases corresponds to d_i
- $C_i = \{c_1, c_2, \dots, c_k\}$ includes x controls / healthy individuals that correspond to d_i

$$Matrix(C) = \begin{pmatrix} d_1 & F_1 & S_1 & P_1 & C_1 \\ d_2 & F_2 & S_2 & P_2 & C_2 \\ d_3 & F_3 & S_3 & P_3 & C_3 \\ d_4 & F_4 & S_4 & P_4 & C_4 \\ d_5 & F_5 & S_5 & P_5 & C_5 \\ d_6 & F_6 & S_6 & P_6 & C_6 \\ d_7 & F_7 & S_7 & P_7 & C_7 \\ d_8 & F_8 & S_8 & P_8 & C_8 \end{pmatrix} = \begin{pmatrix} CIR & 542 & 232 & 118 & 114 \\ COL & 503 & 121 & 48 & 73 \\ Crohn & 48 & 975 & 663 & 312 \\ HIV & 60 & 155 & 129 & 26 \\ IBD & 465 & 253 & 164 & 89 \\ OBE & 572 & 344 & 170 & 174 \\ T2D & 381 & 96 & 53 & 43 \\ WT2 & 443 & 110 & 25 & 85 \end{pmatrix}$$

The read counts of each feature in HIV and Crohn datasets can be greater than 1 while total species abundance of all features in one sample of other species abundance datasets is sum up to 1:

$$\sum_{i=1}^k f_i = 1$$

with:

- k is the number of features for a sample.
- f_i is the value of the i -th feature.

The next section, we will introduce pre-processing methods based on binning approaches on these metagenomic datasets.

3 Metagenomic data binning

Data binning or Data Discretization is a data processing method which transforms continuous value into discrete value. To discretize continuous values into “bins”, we need to determine “breaks” where indicates which bin these values belong to. “Breaks” are real values which can be 0.1, 0.35, etc. that are considered as “boundaries” of bins. Let say, we have an array of values including 0.000012, 0.02, 0.56, 0.92. We would like to divide 10 bins which own an equal width for each bin on a considered value range from 0 to 1. The width of each bin or interval width, in this case, is $\frac{1-0}{10} = 0.1$. The value range of the first bin is from 0 to 0.1, the second bin is from 0.1 to 0.2, etc. In our study, we do not consider the values of 0 (zeros), so values which are greater than 0 and lower than 0.1 (such as values of 0.000012, 0.02) will belong to the first bin while the second bin contains values which are greater than or equal to 0.1 and lower than 0.2. For other values with the computation as above, 0.56 will belong to the 6th bin while the last bin (10th bin) contains 0.92. The breaks as the example mentioned consist of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9. We also add 0 to the breaks to

Table 1: Information on eight considered datasets.

	CIR	COL	Crohn	HIV	OBE	T2D	WT2	IBD
#Features	542	503	48	60	465	572	381	443
#Samples	232	121	975	155	253	344	96	110
#Patients	118	48	663	129	164	170	53	25
#Controls	114	73	312	26	89	174	43	85

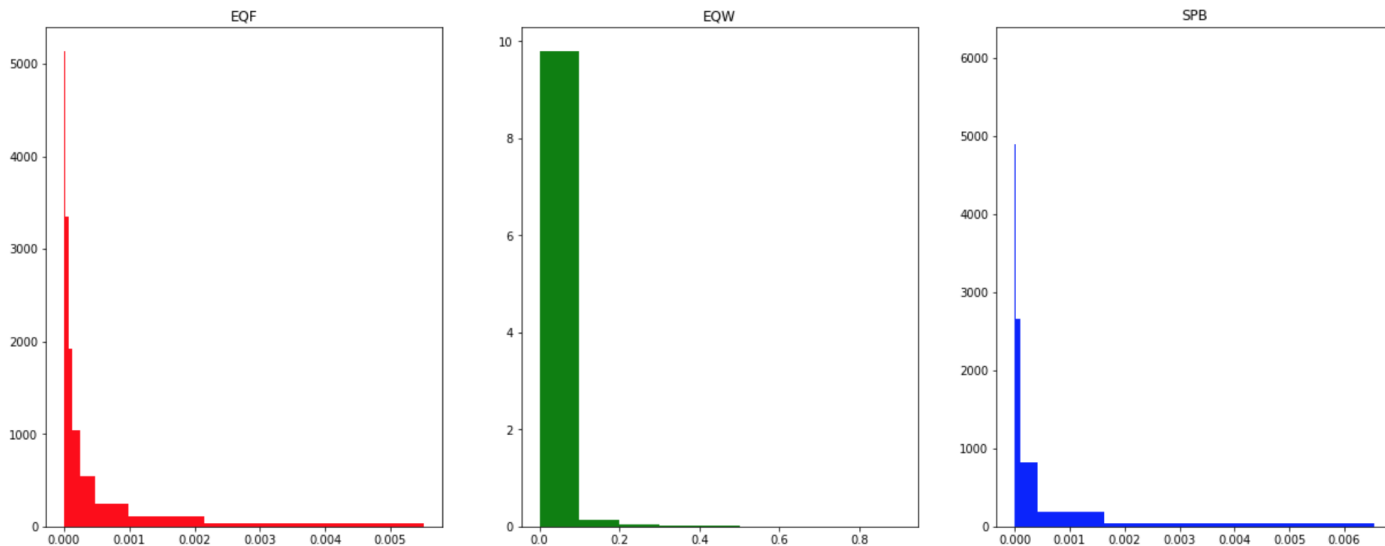


Figure 1: Data density with various binning methods on Liver Cirrhosis bacterial species abundance dataset using the same 10 bins. X-axis shows a value range of abundance.

compare whether the values are greater than 0 to distribute values to bins. Metagenomic data can exist outliers that cause many wrong prediction results by learning machine algorithms. Binning approaches are expected to improve the performance by reducing the effects of minor observation errors and to get rid of noise in the data.

This section will present various binning approaches which can be binning with **equal width** or **equal frequency** of values or basing on **species abundance distribution** of several considered species abundance datasets or simply only considering whether the feature exists in the sample or not ($value > 0$), namely **Binary Binning**. Some methods combining between binning approach and transformation with scaler algorithms are also presented.

3.1 Equal Width binning

Equal Width binning (EQW) divides and delivers continuous values to bins which have equal width. Each bin has the equal width which is computed by $\frac{Max_value - Min_value}{number\ of\ bin}$ in the range of [Min, Max] of the data. For instance, we would like to deliver original values to 5 equal width bins ($k = 5$) using a range of [Min=0,Max=1], then width of each bin is 0.1 ($w = 0.1$). The interval boundaries include $Min + w, Min + 2 \times w, \dots, Min + (k - 1) \times w$. The idea is simple but this method show improvements in prediction tasks.

3.2 Equal Frequency binning (EQF)

Equal frequency binning method cuts the data into n parts (bins) which each part contains approximately the same number of values.

The breaks are identified using the training set so the performance in the testing phase will be poor if the training set cannot reflect exact general data distribution of the considered disease. Breaks depend totally on data distribution so the width of each bin can vary significantly.

3.3 Binning based on species abundance distribution

Species Bins (SPB) is extended from EQW combining species abundance distribution conducted from 6 species abundance datasets in [1]. Authors in [1] presented breaks including $0, 10^{-7}, 4 \times 10^{-7}, 1.6 \times 10^{-6}, 6.4 \times 10^{-6}, 2.56 \times 10^{-5}, 0.0001024, 0.0004096, 0.0016384, 0.0065536$ for Species Bins. The first break ranges from 0 to 10^{-7} which is the smallest value of species abundance known in six species abundance datasets of CIR, COL, IBD, OBE, T2D, and WT2 [1]. The width of each bin is equivalent to a 4-fold increase from the previous bin.

3.4 Equal Width binning combining scaler algorithms

Some transformation algorithms applying to original data can be useful for binning. Standardization method is a widely-used technique for numerous machine learning algorithms to resolve the problem of different data distributions. Quantile Transformation (**QTF**), MinMaxScaler (**MMS**), and logarithmic computations scalers are considered to convert data before binning.

Quantile Transformation is implemented to combine with EQW in these experiments. QTF is considered as a robust pre-processing

technique because it can reduce the effect of the outliers. Samples in test and validation sets which are smaller or larger than the fitted range then will be assigned to the bounds of the output distribution. Another algorithm illustrated in this study is MinMaxScaler, to make a comparison with QTF and logarithmic computations. MinMaxScaler converts each feature to a given range by (1) and (2) formulas:

$$X_{std} = \frac{X - X.min}{X.max - X.min} \tag{1}$$

$$X_{scaled} = X_{std} * (max - min) + min \tag{2}$$

Functions which perform the transformation as above are now available in scikit-learn library.

As described in [1], metagenomic data usually follow the zero-inflated distribution. Data scaled with the methods of transformation based logarithm calculation reveal more normally-distributed. In this study, we use logarithm computation base 4 and base 100 for comparison.

3.5 Binary Bin

Binary Bin (**B2**) which can be considered as the one-hot encoding method, also is brought to compare. B2 indicates whether a feature is present or absent in a sample. If values are greater than 0, Bin 1 contains them. Otherwise, they are delivered to Bin 0 (with all values=0).

3.6 Data distribution Visualization of binning methods and scalers

Figure 1 shows various binning approaches on CIR dataset. The breaks of EQF include $0, 4 * 10^{-07}, 2.47 * 10^{-05}, 6.2 * 10^{-05}, 1.27 * 10^{-04}, 2.466 * 10^{-04}, 4.788 * 10^{-04}, 9.783 * 10^{-04}, 2.1484 * 10^{-03}, 5.5149 * 10^{-03}$. These breaks are approximate to SPB. We note that some first bins (with EQF) own high density with the width of these are rather small ($4 * 10^{-07}$ for the first bin) while the width of the 9th bin is about $3.4 * 10^{-03}$. Similar results are exhibited for SPB. Width of each bin with EQW is equal, so we can see that the first bin contains most of the data.

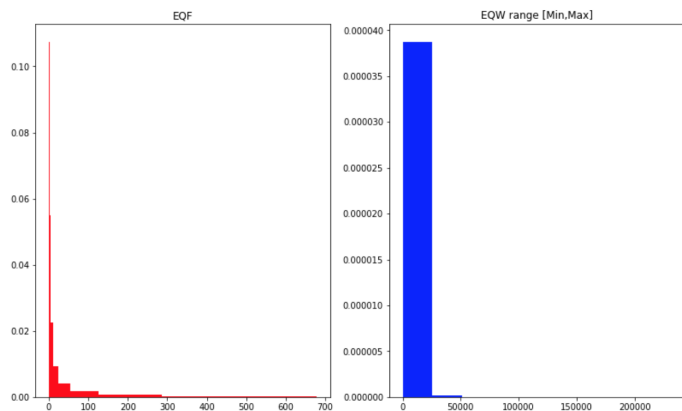


Figure 2: Data density on Crohn Read Counts dataset. X-axis shows a value range of counts with breaks using EQF and EQW on Min-Max range of training set.

For data type of counts, the values in features can be greater than 1, so SPB and EQW considering in a value range from 0 to

1 are not efficient for this type. EQW with a range between Min and Max values in the training set and EQF can work in this situation (Figure 2). Observed and conducted from Figure 2 for EQW method, we see that data distribution of metagenomic is the zero-inflated distribution, no matter what data is abundance or counts. However, a transformation with logarithm enables data to be more normally-distributed (Figure 3).

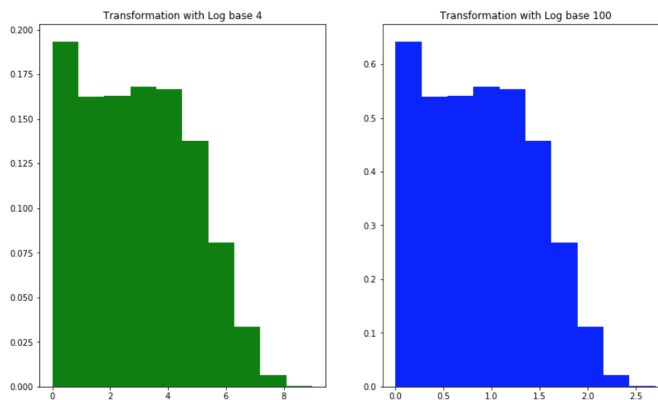


Figure 3: Data density on Crohn Read Counts dataset after transformation with logarithm base 4 and 100. The X-axis shows a transformed value range

The performance of the proposed binning methods will be evaluated in the next section (Section 4).

4 Experimental Results on Metagenomic data binning Approaches

To exhibit the efficiency of binning approaches on various machine learning algorithms, we present the results with a classic machine learning algorithm (Random Forest), Linear Regression and a famous deep learning technique that is Convolutional Neural Network on 1D data (CNN1D).

OPERATION		DATA DIMENSIONS	WEIGHTS (N)	WEIGHTS (%)
Input	#####	542	1	
InputLayer		-----		
	#####	542	1	0 0.0%
Conv1D	\\	-----		
relu	#####	540	64	256 0.7%
Flatten		-----		
	#####	34560		0 0.0%
Dense	XXXXX	-----		
sigmoid	#####	1	34561	99.3%

Figure 4: CNN1D architecture.

The algorithms of **CNN1D**, **Linear Regression** are both implemented using Adam optimization function, a learning rate of 0.001, and using an overall epoch of 500 along with a batch size of 16 and binary cross-entropy loss function. To reduce overfitting issue, we use "Early Stopping" with the number of epoch patience of 5. The learning will stop if the Loss is not improved after 5 consecutive epochs. As conducted from [1], we should use a shallow deep learning architecture instead of deeper architectures, so the proposed CNN1D architecture includes a convolutional layer consisting of 64

filters of the size of 3, following by a max-pooling of size 2, and an activation function of ReLU [28]. The details of CNN1D are visualized in Figure 4.

Random Forests algorithm is a robust learning algorithm and widely-used in numerous studies related to bioinformatics tasks. In this study, Random Forest algorithm is deployed with 500 trees, nodes are expanded until all leaves are pure or until all leaves contain less than $min_samples_split = 2$ where $min_samples_split$ is the minimum number of samples required to split an internal node.

The performance of each classifier is measured by an average of Area Under the Curve (AUC) and an average Accuracy (ACC) on 10-stratified-fold-cross validation repeated 5 times. The same folds are used for all classifiers, i.e. training and test sets were identical for each classifier. Besides, results are visualized by Boxplot to exhibit graphically depicting groups of numerical data through their quartiles. Our results are compared to state-of-the-arts including *MetAML* [18] on 6 species abundance datasets and *Selbal* [27] on 2 read counts datasets. *MetAML* [18] is a framework for metagenomic data analysis running on species abundance with classic machine learning algorithms such as SVM and Random Forest. *MetAML* performed the best with Random Forest; hence in comparison with our methods, we also run the classification tasks using Random Forest with the same parameters with *MetAML*. *Selbal* uses balance score to find good features, then fetching the features into Linear Regression algorithms for the prediction tasks.

We need to specify the value range to divide bins for binning approaches. In this study, the considered value range can be either $[0,1]$ or $[Min,Max]$ to divide bins for data. $[Min,Max]$ means we consider the range covered by the minimum value and the maximum value of all features in the training set to bin the data.

We present the experimental results as followings. First, we show the disease prediction performance of all considered binning approaches (Section 4.1). Next, we evaluate and compare the differences in performance when we change the number of bins. Then, promising methods are compared with the state-of-the-art including *MetAML* [18] and *Selbal* [27].

4.1 Evaluation on different data pre-processing methods for metagenomic data

We compare the performance of various pre-processing methods based on binning approaches with three different widely-used machine learning algorithms including Random Forests (Figure 5), Convolutional Neural Network (Figure 6) and Linear Regression (Figure 7).

Figure 5 shows the prediction performance of the considered binning methods performed by Random Forests on 8 considered datasets. We compare the efficiency of different binning approaches (B2, EQF, EQW, SPB) and various scalers (Logarithms, Min-Max Scaler, Quantile Transformation). Except for Crohn dataset, there are not too significant differences in the performances of Random Forests algorithm with different approaches. In the chart, NA (“Not Available”) means the model running on the original data without using binning, so the value range for binning is also NA (for example, NA_log4_NA, NA_none_NA, etc.). For CIR and OBE datasets, EQF without using scaler achieves the best performance. The average AUC for predicting CIR, OBE datasets using EQF_none are

0.95582, 0.68238, respectively. For samples from COL dataset, the highest result is with QTF scaler. The datasets of Crohn’s disease, HIV, IDB achieve the best results with EQW binning combined with QTF scaler on the value range of $[Min,Max]$. The best AUC for Crohn’s disease dataset is 0.86976 and IDB obtains the best value at 0.88888 while HIV dataset obtains the best at 0.72438. The remaining 2 datasets including T2D and T2W using the binning method of EQW on the range of $[Min,Max]$ without using scaler, reach the AUC best at 0.76286 and 0.80868, respectively. Crohn dataset shows worse results on the value range of $[0,1]$. It seems more appropriate because this dataset uses read count where the values are either equal 0 or greater than 1.

Figure 6 exhibits the results of CNN model on the considered datasets. As seen from the figure, binning approaches can outperform other methods. CIR dataset has the highest AUC value of 0.95986 while IDB dataset owns the best AUC value of 0.90894. Both two those results are evaluated with EQF without using any scalers. COL dataset obtains the best results with using the EQW combined Min-Max scaler on the value range of $[Min,Max]$ of the training sets with AUC of 0.83732. Crohn’s gets the best result using QTF scaler with AUC of 0.85698. The prediction results on HIV disease using EQF without scaler on the value range of $[Min,Max]$ peak at the best AUC value of 0.72788. OBE dataset has the highest AUC when we use the SPB approach. Two datasets of T2D, T2W running with EQW binning on the range of $[0,1]$ exhibit the best AUCs of 0.75746, 0.80238, respectively.

We also present performances of Linear Regression algorithm in Figure 7. As exhibited, the results are rather similar to mentioned previous two algorithms. CIR dataset obtains the best AUC value of 0.95870 with using EQF binning on the range of $[0,1]$ while we achieve the best AUC of 0.83990 on COL with EQW. Original data of Crohn dataset being run by QTF scaler reaches the highest AUC value of 0.86884. Some results on other datasets are similar to CNN’s results.

From the shown experimental results, we notice that CNN, in general, achieves better results than using Random Forest and Linear Regression. Binning approaches appear to be more efficient to enhance significantly the performance with CNN and Linear Regression.

The methods of EQW on the value range $[min, max]$ of training sets, EQF binning and scalers algorithms appear to be appropriate methods for the prediction tasks of HIV and Crohn’s disease where the values of features can be greater than 1. Comparing to the performance of the original data (NA_none_NA), these methods can give significant improvements.

4.2 Number of bins for Metagenomic binning

A comparison among the numbers of bins for binning EQW approach is presented in Figure 8. Average AUCs on each number of bins applying to 8 considered datasets are calculated to compare. The numbers bins of 5, 10 give average AUCs (on 8 datasets) of 0.8022450, 0.8011900, respectively while using 100 and 255 bins reveal AUCs of 0.7621925, 0.7471600, respectively. The binary bin approach reaches the average AUC of 0.7958750. As observed, the numbers bins of 5 and 10 obtain significantly better results compared to 100 and 255 bins. As shown from the average performance on the all considered datasets, data discretization with 5 bins

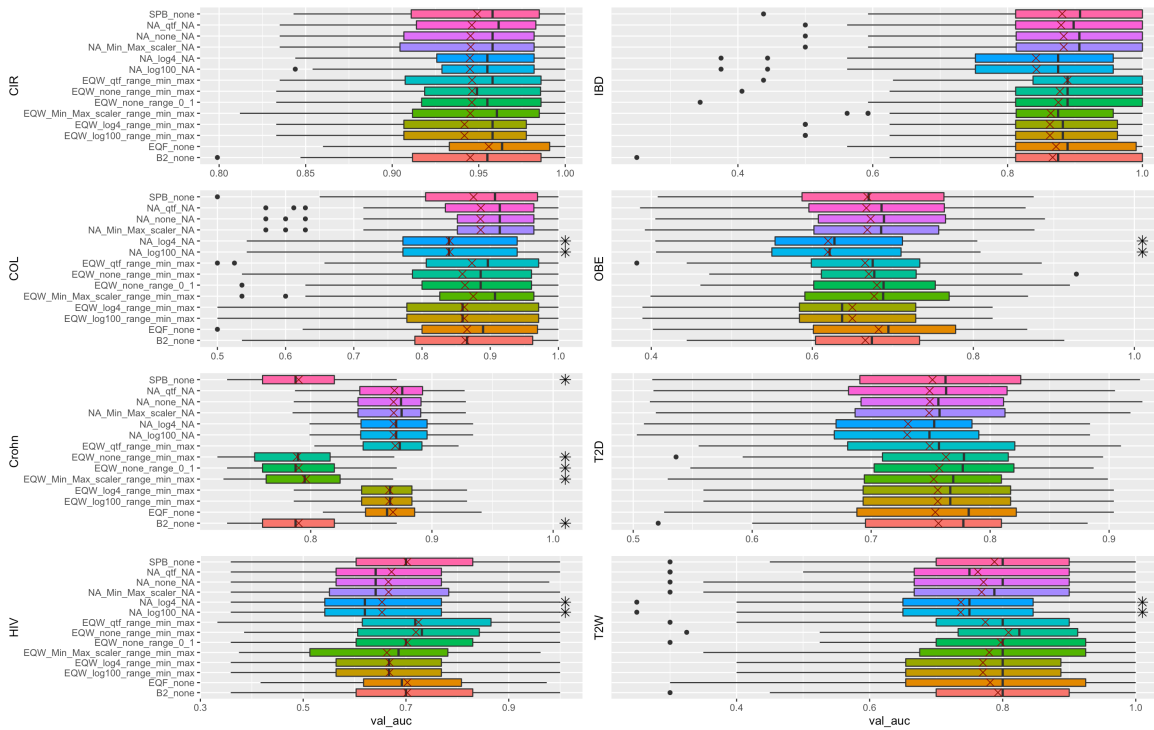


Figure 5: Different Methods Comparison using Random Forest. "*" reflects significant differences compared to the best result on each dataset. "X" reveals average performance on 10-fold cross-validation repeated 5 times. Methods names denote The binning method combining Scaler and data range for binning. For example, EQW_Min_Max_scaler_range_min_max denotes that we performed MinMax scaler and binning by EQW on the range of [Min,Max] of values in training set while NA_none_NA means no binning method or scaler is applied. Black Dots exhibit outliers in results.

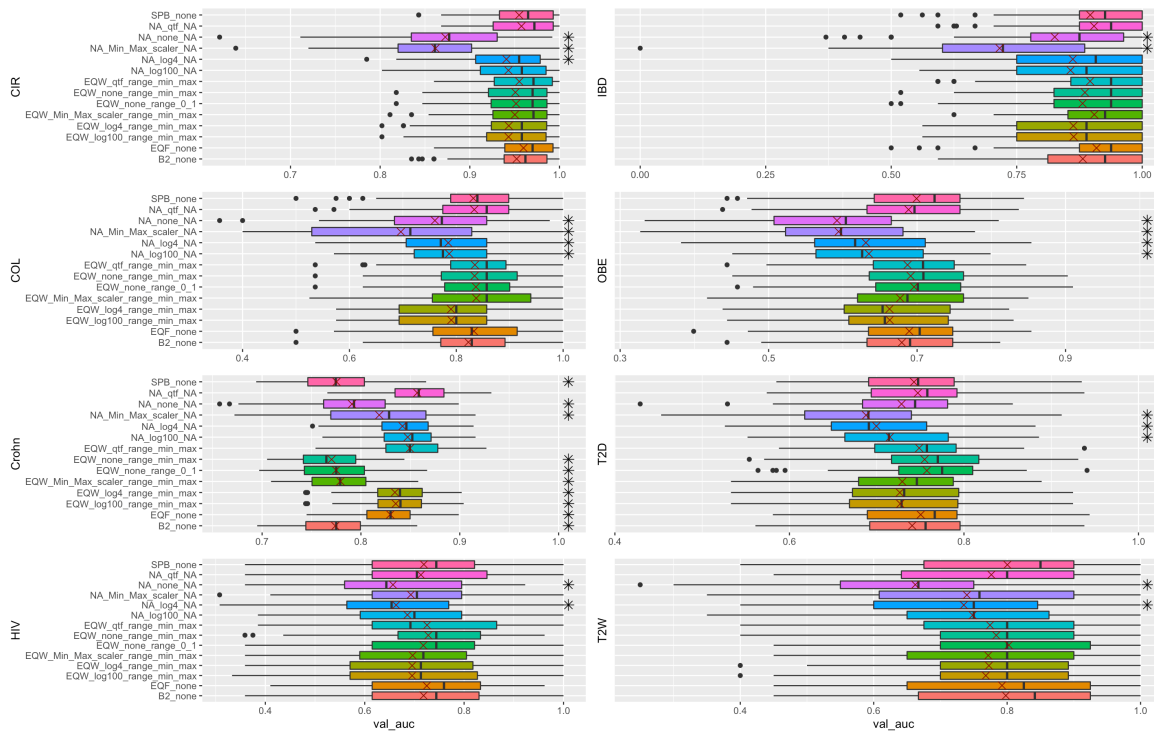


Figure 6: Different Methods Comparison using CNN. All symbols and stickers in the chart are the same as Figure 5.

achieves the best.

We see that CIR dataset obtains the best AUC value is 0.9522 while IBD dataset achieves the AUC value of 0.88162 with 2 bins.

Additionally, The Crohn dataset reaches the best AUC of 0.77488 with 100 bins. Three datasets of COL, OBE and T2W reveal the best AUC values with 5 bins with AUCs of 0.83670, 0.69594, 0.80238,

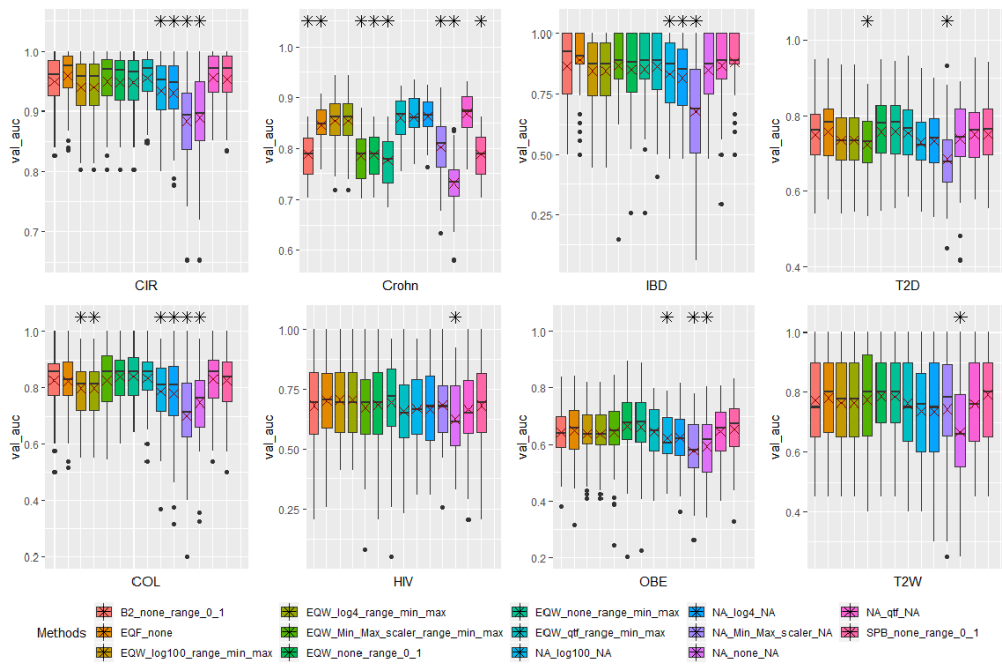


Figure 7: Different Methods Comparison using Linear Regression. All annotations in the chart are the same as Figure 5.

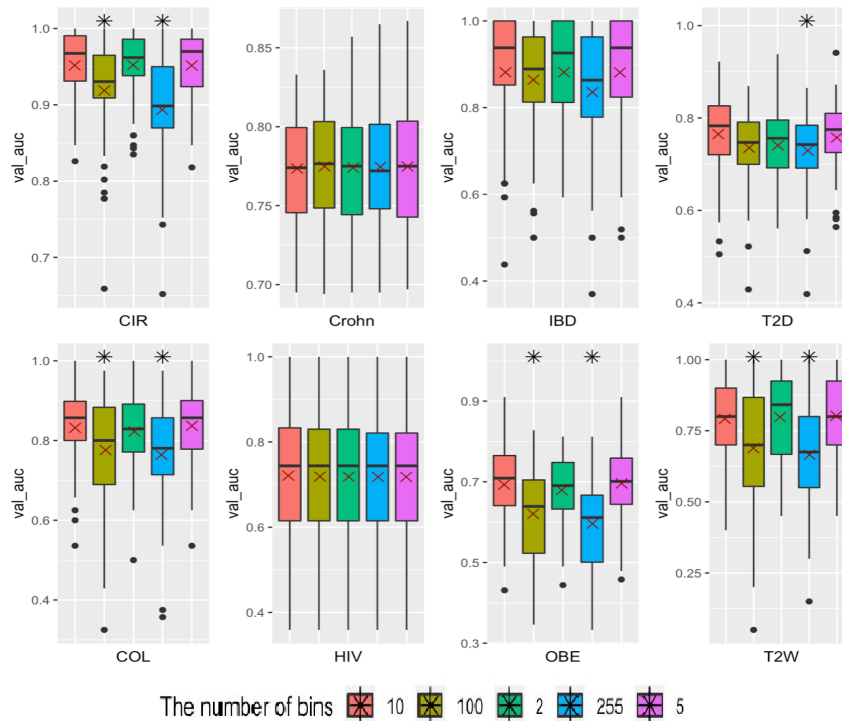


Figure 8: Performance of the various number of EQW bins using the CNN model. “*” reflects significant differences compared to the best result on each dataset. “X” reveals average performance on 10-fold cross-validation repeated 5 times.

respectively. Otherwise, prediction tasks on the diseases of HIV and T2D with 10 bins give the best results.

4.3 State-of-the-art comparison

To reflect the efficiency of binning approaches on metagenomic data, we compare the proposed binning approaches to some state-of-the-

art including MetAML [18] and selbal [27].

In Table 2, we display the results using binning approaches of EQW, EQF, SPB with a total of 5 bins and comparing state-of-the-art in average ACC and average AUC on 10-fold cross-validation repeated 5 times. Three datasets including CIR, OBE and T2W all had better results than state-of-the-art. COL disease has only better

Table 2: Results Comparison of robust binning methods (the number of bins is 5 for EQW, EQF and SPB) and state-of-the-art (Selbal on HIV, Crohn’s disease and MetAML on other datasets) in average ACC and average AUC on 10-fold cross-validation repeated 5 times. The results formatted in bold text are better compared to the state-of-the-art.

Datasets			CIR	COL	Crohn	HIV	IBD	OBE	T2D	T2W	AVG
State-of-the-art		val_acc	0.877	0.805	NA	NA	0.809	0.644	0.664	0.703	0.750
EQF	CNN	val_acc	0.906	0.792	0.783	0.833	0.829	0.680	0.655	0.722	0.775
EQF	RF	val_acc	0.887	0.808	0.813	0.816	0.810	0.657	0.675	0.720	0.773
EQW	CNN	val_acc	0.883	0.785	0.739	0.827	0.851	0.668	0.667	0.707	0.766
EQW	RF	val_acc	0.886	0.791	0.745	0.809	0.807	0.645	0.680	0.708	0.759
SPB	CNN	val_acc	0.903	0.790	0.737	0.828	0.830	0.672	0.650	0.727	0.767
SPB	RF	val_acc	0.883	0.785	0.739	0.827	0.851	0.668	0.667	0.707	0.766
State-of-the-art		val_auc	0.945	0.873	0.820	0.674	0.890	0.655	0.744	0.762	0.795
EQF	CNN	val_auc	0.960	0.833	0.830	0.725	0.909	0.689	0.750	0.792	0.811
EQF	RF	val_auc	0.956	0.866	0.868	0.703	0.872	0.682	0.754	0.781	0.810
EQW	CNN	val_auc	0.952	0.837	0.775	0.718	0.881	0.696	0.757	0.802	0.802
EQW	RF	val_auc	0.946	0.863	0.790	0.703	0.875	0.680	0.757	0.797	0.801
SPB	CNN	val_auc	0.955	0.832	0.775	0.718	0.896	0.699	0.742	0.801	0.802
SPB	RF	val_auc	0.952	0.837	0.775	0.718	0.881	0.696	0.757	0.802	0.802

ACC value when binning with EQF combined with Random Forest model (ACC value is 0.868). In Crohn’s disease, when performing the binning method with EQF combined with 2 models CNN (ACC value is 0.830) and Random Forest (ACC value is 0.868) both give better results than state-of-the-art. AUC values for HIV disease are better than state-of-the-art in all models and methods, but there is no ACC result in this disease higher than state-of-the-art. IBD has 5 good results when done with ACC values but only 2 good results for AUC. T2D has most of the results better than state-of-the-art, only when doing SPB method with CNN model (both ACC and AUC values) and when binning with SPB, CNN model with ACC measurements are lower than state-of-the-art.

5 Conclusion

In this study, we presented **Met2Bin** with various binning approaches using Equal with binning, Equal frequency binning, Species bins and binary bins to reduce the effects of minor observation errors and to get rid of noise in the metagenomic data. Scaler with QTF and logarithm transformation also show potential improvements in the data type of reading counts. In most cases, binning approaches and scaler algorithms can improve performance for machine learning algorithms.

The binning and scaler approaches are examined on a vast of datasets including different diseases and various data types (species abundance and read counts at genus or family or order levels). We can see that the proposed method can work on any value ranges. This research only takes into account unsupervised binning methods. Considerations on the labels of samples should be carried out in further studies.

As revealed from the performance of disease prediction, we can predict Liver cirrhosis, IBD with high accuracy while Obesity, HIV and T2D diagnosis are still challenges. Further research should investigate to improve those diseases.

In general, CNN produces better results than classic machine learning. However, the considered CNN architecture in this work

is rather small and modest but its performance exhibits promising results. Further investigations on the CNN architectures should be considered to improve the performance. We also do not consider the labels of samples when we build the breaks for data binning. In the future, the research should consider and investigate the supervised binning approaches to evaluate whether those can be efficient or not on metagenomic data.

The binning methods are potential methods so that we can use such bins for converting numeric data and showing them in 2D images. A bin which represents the magnitude of value can be shown in the image with a specific colour. Binning techniques enable us to visualize bio-markers in images as well as to leverage advancements in deep learning algorithms for images to do prediction tasks.

The results and other materials of this work can be downloaded from <https://github.com/thnguyencit/met2bin>.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] Thanh Hai Nguyen, Jean-Daniel Zucker. Enhancing Metagenome-based Disease Prediction by Unsupervised Binning Approaches. The 2019 11th International Conference on Knowledge and Systems Engineering (KSE-IEEE), ISBN: 978-1-7281-3003-3, pp 381-385. 2019. <https://doi.org/10.1109/KSE.2019.8919295>
- [2] Colorectal Cancer: Statistics Approved by the Cancer.Net Editorial Board. 2020. <https://www.cancer.net/cancer-types/colorectal-cancer/statistics>
- [3] Young-Hee Yoon, Chiung M. Chen; Liver cirrhosis Mortality in the United States: National, State and Regional Trends, 2000-2015. 2018. <https://pubs.niaaa.nih.gov/publications/surveillance111/Cirr15.htm>.
- [4] C. Mathee et al. ;SURVEY AND SUM- MARY: Current methods of gene prediction, their strengths and weaknesses; p. 4103-4117.ISSN 0305-1048. 2002. <https://doi.org/10.1093/nar/gkf543>

- [5] Z. Wang, Y. Chen and Y. Li; A Brief Review of Computational Gene Prediction Methods; p. 216-221. ISSN 1672-0229. 2004. [https://doi.org/10.1016/S1672-0229\(04\)02028-5](https://doi.org/10.1016/S1672-0229(04)02028-5)
- [6] N.P. Nguyen et al.; A perspective on 16S rRNA operational taxonomic unit clustering using sequence similarity; ISSN 2055-5008. *Nature*. <https://doi.org/10.1038/npjbiofilms.2016.4>
- [7] M. Arumugam et al.; Enterotypes of the human gut microbiome; 473, p. 174-180. ISSN 1476-4687. 2011 <https://www.nature.com/articles/nature09944>
- [8] S. Park et al.; hc-OTU: A Fast and Accurate Method for Clustering Operational Taxonomic Units Based on Homopolymer Compaction; 15, p. 441-451. ISSN 1545-5963. 2016. <https://doi.org/10.1109/TCBB.2016.2535326>
- [9] Wei ZG et al.; A Dynamic Multi-Seeds Method for Clustering 16S rRNA Sequences Into OTUs. *PubMed*. 2019. <https://doi.org/10.3389/fmicb.2019.00428>
- [10] D. H. Huson, D. C. Richter, S. Mitra, A. F. Auch and S. C. Schuster; Methods for comparative metagenomics; 10, p. S12. ISSN 1471-2105. 2009. <https://doi.org/10.1186/1471-2105-10-S1-S12>
- [11] L.-x. Chen, M. Hu, L.-n. Huang, Z.-s. Hua, J.-l. Kuang, S.-j. Li and W.-s. Shu; Comparative metagenomic and metatranscriptomic analyses of microbial communities in acid mine drainage; 9, p. 1579-1592. ISSN 1751-7370. 2015. <https://doi.org/10.1038/ismej.2014.245>
- [12] S. Nayfach et al.; Toward Accurate and Quantitative Comparative Metagenomics; 166, p. 1103-1116. ISSN 0092-8674. <https://doi.org/10.1186/1471-2105-10-S1-S12>
- [13] S. M. Dabdoub et al.; Comparative metagenomics reveals taxonomically idiosyncratic yet functionally congruent communities in periodontitis; 6, p. 38993. ISSN 2045-2322. <https://doi.org/10.1038/srep38993>
- [14] K. Sedlar et al.; Bioinformatics strategies for taxonomy independent binning and visualization of sequences in shotgun metagenomics; 15, p. 48-55. ISSN 2001-0370. 2016. <https://doi.org/10.1016/j.csbj.2016.11.005>
- [15] H. Soueidan et al.; Machine learning for metagenomics: methods and tools; *Metagenomics* 1. 2017. <https://doi.org/10.1515/metgen-2016-0001>
- [16] G. Ditzler et al.; Forensic identification with environmental samples; *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2012. <https://doi.org/10.1109/ICASSP.2012.6288265>
- [17] G. Ditzler et al.; MultiLayer and Recursive Neural Networks for Metagenomic Classification; *IEEE Transaction Nanobioscience* 14, p. 608-616. 2015. <https://doi.org/10.1109/TNB.2015.2461219>
- [18] E. Pasolli et al.; Machine Learning Meta-analysis of Large Metagenomic Datasets: Tools and Biological Insights; *PLoS Comput*. 2016. <https://doi.org/10.1371/journal.pcbi.1004977>
- [19] D. Fioravanti, Y. Giarratano, V. Maggio, C. Agostinelli, M. Chierici, G. Jurman and C. Furlanello; Phylogenetic convolutional neural networks in metagenomics; 19, p. 49. 2018. ISSN 1471-2105. <https://doi.org/10.1186/s12859-018-2033-5>
- [20] D. Reiman, A. A. Metwally and Y. Dai; PopPhy-CNN: Attention Neural Networks for Metagenomic Phylogenetic Tree Embedded Architecture for Convolution. 2018. <http://biorxiv.org/lookup/doi/10.1101/257931>
- [21] F. Celesti, A. Celesti, J. Wan and M. Villari; Why Deep Learning Is Changing the Way to Approach NGS Data Processing: a Review; ISSN 1937-3333. 2018. <https://doi.org/10.1109/RBME.2018.2825987>
- [22] J. Ren et al.; Identifying viruses from metagenomic data by deep learning. 2018. <http://arxiv.org/abs/1806.07810>
- [23] Subrata Saha, J. Johnson, S. Pal, G. M. Weinstock, S. Rajasekaran, MSC: a metagenomic sequence classification algorithm, *Bioinformatics*; vol.35, 17p. 2932-2940. 2019. <https://doi.org/10.1093/bioinformatics/bty1071>
- [24] Jolanta Kawulok, M. Kawulok and S. Deorowicz; Environmental metagenome classification for constructing a microbiome fingerprint. *Biol Direct* 14. 2019. <https://doi.org/10.1186/s13062-019-0251-z>
- [25] Lo Chieh, Marculescu, Radu; MetaNN: accurate classification of host phenotypes from metagenomic data using neural networks. *CM University. Journal contribution*. 2019. <https://doi.org/10.1186/s12859-019-2833-2>
- [26] K. Sedlar et al.; Bioinformatics strategies for taxonomy independent binning and visualization of sequences in shotgun metagenomics; 15, p. 48-55. ISSN 2001-0370. 2016. <https://doi.org/10.1016/j.csbj.2016.11.005>
- [27] Rivera-Pinto J, Egozcue JJ, Pawlowsky-Glahn V, Paredes R, Noguera-Julian M, Calle ML. Balances: a New Perspective for Microbiome Analysis. *mSystems*. 2018;3(4):e00053-18. Published 2018 Jul 17. <https://doi.org/10.1128/mSystems.00053-18>
- [28] Abien Fred Agarap. Deep Learning using Rectified Linear Units (ReLU). 2018. <https://arxiv.org/abs/1803.08375>

Dynamics Model and Design of SMC-type-PID Control for 4DOF Car Motion Simulator

Pham Van Bach Ngoc^{1,*}, Bui Trung Thanh²

¹Space Technology Institute, Vietnam Academy of Science and Technology, Hanoi, Vietnam

²Hung Yen University of Technology and Education, Hung Yen, Vietnam

ARTICLE INFO

Article history:

Received: 30 March, 2020

Accepted: 07 June, 2020

Online: 26 June, 2020

Keywords:

SMC-type-PID

SMC

GA Optimization

Lyapunov

4DOF

Motion simulator

ABSTRACT

4-DOF car motion simulator helps to simulate real-life experiences that drivers do not have the opportunity to access the real environment. The dynamics equation of 4-DOF car motion simulator is a very complex problem with many uncertain parameters, so it requires intelligent control algorithms. Sliding mode controller (SMC) can achieve good tracking performance and robustness to the disturbances, but SMC has worse stability and reliability than PID controller. As a most widely used controller, PID controller has many obvious advantages, but it has poorer tracking performance than SMC. In this paper, a control method of sliding mode type PID controllers is proposed to fully combine the advantages of the two controllers. In this study, based on the dynamics equation of 4-DOF car motion simulators the author develop two algorithms to control sliding mode control type PID (SMC-type-PID) and sliding mode controller type PID with GA optimization for 4-DOF car motion simulator. Firstly, the authors used Lyapunov theory to prove the stability of the system, next presenting the simulation results of two control algorithms with different uncertain components and comparing them to find and demonstrate the effectiveness of the new control method applied to the 4-DOF car motion simulator.

1. Introduction

Parallel mechanisms have been researched and applied in many areas such as motion simulators, which simulate the experience of being in a car, plane, tank, and in a virtual reality environment or create motions to serve different goals. Parallel mechanisms have outstanding advantages compared to serial machines such as: high rigidity, high load bearing capacity, ability to change position and spiritual orientation, activation accuracy, high stability, high loading capability. The simulation model has 6-DOF [1] based on Stewart-Gough structure, this structure has good load capacity, however it needs to use 6 driving mechanisms leading to complex control.

For motion simulation model for devices such as cars, tanks moto bike...in fact we only need 4-DOF is enough. These car motion simulators consist rotating and translating along the vertical OZ axis, rotating around the OX axis and OY axis. In this research the authors focus on the hybrid mechanism including parallel and serial mechanism, which generate 4 motions (4-DOF) in the space. The control problem for 4-DOF car motion simulators is difficult because of many non-linear parameters. Currently, a

number of control methods have been developed and applied to control problems for a defined model or a model with constant parameter uncertainty in [2-7]. However, the control problem for a driving practice model with indefinite parameters has always paid much attention to further improving the kinetic quality and dynamics of the driving platform [8]. Sliding mode controller is used in nonlinear systems. It is robust and efficient in maintaining stability for nonlinear dynamic systems. Later this method was more interested by scientists because of its stability and stability even when there is the impact of noise as well as changes of model parameters. However, the chattering phenomenon remains the main disadvantage of this robust control. As solution to this problem, a PID sliding mode approach is exposed and tested in this paper. Sliding mode controller with type PID is used to eliminate the oscillation around the sliding surface when the amplitude of the slider control law changes greatly. The gain before the sign function in the sliding mode control law is calculated according to the Lyapunov stability criterion introduce in [9], [10] [11]. This gain, unless properly selected, is likely to cause oscillation. In this study, the authors presented the 4-DOF car motion simulators with the serial and parallel mechanism, kinematics and dynamic equations. Base on dynamics equations

*Pham Van Bach Ngoc, 18 Hoang Quoc Viet Street, pbnngoc@imech.vast.vn

authors focused on building intelligent control algorithms for the 4-DOF car motion simulators model. Specifically, the author will develop an SMC-type-PID and compared the results with the SMC-type-PID with GA Optimization with uncertain parameters.

2. Kinematics and dynamics model

In this section, the model of the 4-DOF car motion simulator shown in fig 1 would be constructed. The car motion simulators mechanism model with the global axis is chosen. The motion of the upper plate is due to the movement of three vertical linear actuator and the rotation of the revolute joint attached below the lower plate. The piston's movement makes changes in vertical movement, rotate around the OX axis and OY axis of the mobile platform. The revolute joint rotates the moving parts around the OZ axis. Therefore, the car motion simulators has 4-DOF. The systems variables need to control to track the desired trajectory are the length of three robot's legs and the rotation angle around the OZ axis of the revolute joint. Fig.1 illustrates the car motion simulator model with the movements that are translation along the OZ axis, rotation along the OZ, rotation along OX axis, and rotation along the OY axis and they are defined by $P_z, \gamma, \alpha,$ and β .

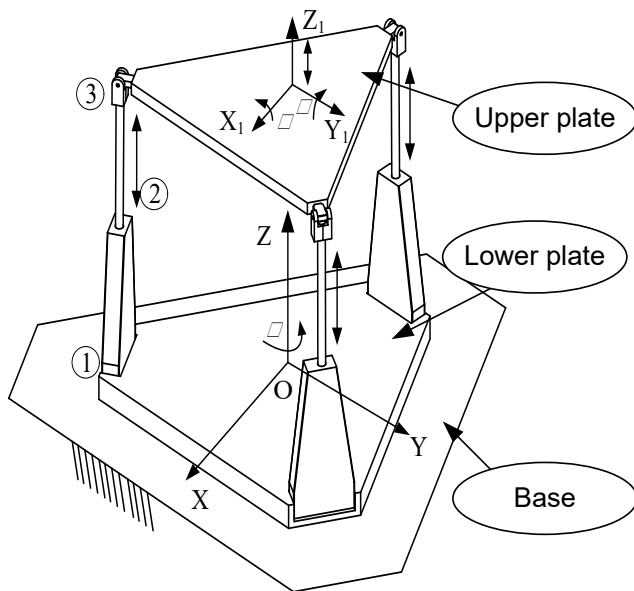


Figure 1: 3D model of the 4DOF car motion simulator

2.1. Inverse Kinematics of 4-DOF car motion simulators

In this part author briefly present the kinematics of 4-DOF car motion simulator [12], from the trajectory of the center point of upper plate we need to find the limb lengths l_i ($i = 1, 2, 3$) and the angle of rotation about OZ axis γ . The kinematic parameters are demonstrated in the vector diagram in Fig. 2 with a is the radius of an upper plate, b is the radius of a lower plate, in this research assumption as $a = b$.

The vector loop equation for each limb of 4-DOF car motion simulator [12], [13] can be written as:

$$\overline{A_i B_i} = \overline{OP} + \overline{PB_i} - \overline{O_A A_i} - \overline{OO_A} \quad (i = 1, 2, 3) \quad (1)$$

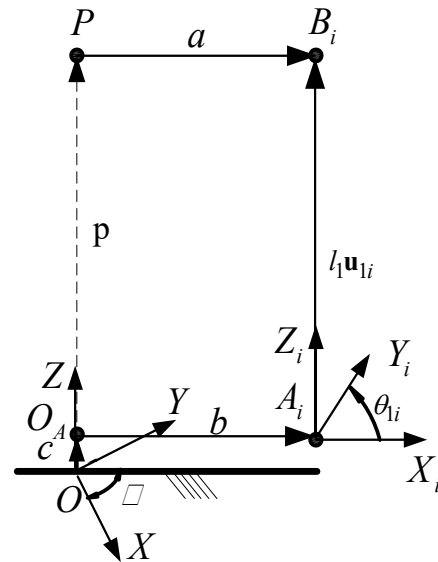


Figure 2: Vector diagram of the 4DOF car motion simulator

where positions of $A_1, A_2, A_3, B_1, B_2, B_3$ are given by:

$$A_1 = \begin{bmatrix} a \times \sin(\frac{\pi}{6}) & a \times \cos(\frac{\pi}{6}) & 0 \end{bmatrix}^T, A_2 = \begin{bmatrix} -a & 0 & 0 \end{bmatrix}^T,$$

$$A_3 = \begin{bmatrix} -a & 0 & 0 \end{bmatrix}^T; B_1 = \begin{bmatrix} a \times \sin(\frac{\pi}{6}) & a \times \cos(\frac{\pi}{6}) & 0 \end{bmatrix}^T,$$

$$B_2 = \begin{bmatrix} -a & 0 & 0 \end{bmatrix}^T, B_3 = \begin{bmatrix} a \times \sin(\frac{\pi}{6}) & -a \times \cos(\frac{\pi}{6}) & 0 \end{bmatrix}^T.$$

The positions of center of the lower plate and upper plate can be written as:

$$O_A = \begin{bmatrix} 0 & 0 & a_z \end{bmatrix}^T \text{ and } P = \begin{bmatrix} 0 & 0 & p_z \end{bmatrix}^T$$

The coordinate of A_i and B_i are obtained as:

$${}^0 A_i = {}^0 R_A \times A_i + O_A \text{ and } {}^0 B_i = {}^0 R_p \times B_i + P$$

with ${}^0 R_A$ and ${}^0 R_p$ are transfer matrices.

Solving the vector equations the length of limbs of the 4DOF car motion simulators are computed as

$$l_i = \sqrt{A_i B_i^T \times A_i B_i} \quad (2)$$

where $A_i B_i = {}^0 B_i - {}^0 A_i$ and we obtain the vector q .

2.2. Dynamic Model of 4-DOF car motion simulators

In the section, the author using Euler-Lagrange to establish the dynamic model of the car motion simulator, the equation is described by [13],

$$M\ddot{q} + C(\dot{q}, q) + G(q) = F \tag{3}$$

where $q = [l_i, \gamma]^T$ (for $i = 1, 2, 3$);

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} & 0 \\ m_{21} & m_{22} & m_{23} & 0 \\ m_{31} & m_{32} & m_{33} & 0 \\ 0 & 0 & 0 & I_{pz} \end{bmatrix} \tag{4}$$

$$m_{11} = \frac{m_p (\sin \theta_{11})^2}{9} + m_2 + \frac{I_{py} (\cos \theta_{11})^2}{(a^2 - (a - l_1 \cos \theta_{11}))^2}$$

$$m_{12} = \frac{m_p \sin \theta_{12} \sin \theta_{11}}{9};$$

$$m_{13} = \frac{m_p \sin \theta_{13} \sin \theta_{11}}{9}$$

$$m_{21} = \frac{m_p \sin \theta_{11} \sin \theta_{12}}{9};$$

$$m_{23} = \frac{m_p \sin \theta_{13} \sin \theta_{12}}{9}$$

$$m_{22} = \frac{m_p (\sin \theta_{12})^2}{9} + m_2 + \frac{I_{px} (\cos \theta_{12})^2}{(a^2 - (a - l_2 \cos \theta_{12}))^2}$$

$$m_{31} = \frac{m_p \sin \theta_{13} \sin \theta_{11}}{9};$$

$$m_{32} = \frac{m_p \sin \theta_{13} \sin \theta_{12}}{9}$$

$$m_{33} = m_2 + \frac{m_p (\sin \theta_{13})^2}{9}$$

with $m_1 (kg)$, $m_2 (kg)$, $m_{dc} (kg)$ and $m_p (kg)$ are the mass of the cover of pistons, the mass of pistons, the mass of motors and the mass of the upper plate, respectively.

$$C = [C_1 \quad C_2 \quad C_3 \quad 0]^T \tag{5}$$

$$C_1 = \frac{1}{(a^2 - (a - l_1 \cos \theta_{11}))^2} \begin{pmatrix} 2I_{px} l_1 \dot{\theta}_{11} \sin \theta_{11} \cos^2 \theta_{11} (a - l_1 \cos \theta_{11}) \\ -I_{py} (\cos \theta_{13})^3 (a - l_1 \cos \theta_{11}) \dot{l}_1^2 \\ -2I_{py} \dot{\theta}_{11} \sin \theta_{11} \cos \theta_{11} \end{pmatrix} + \frac{2}{9} m_p \begin{pmatrix} \dot{\theta}_{11} \dot{l}_1 \cos \theta_{11} \sin \theta_{11} + \dot{\theta}_{12} \dot{l}_2 \cos \theta_{12} \sin \theta_{11} \\ + \dot{\theta}_{13} \dot{l}_3 \cos \theta_{13} \sin \theta_{11} \end{pmatrix}$$

$$C_2 = \frac{1}{a^4 (a^2 - (a - l_2 \cos \theta_{12}))^2} \begin{pmatrix} 2I_{px} l_2 \dot{\theta}_{12} \sin \theta_{12} \cos \theta_{12} (a - l_2 \cos \theta_{12}) \\ -2\dot{\theta}_{12} I_{py} \sin \theta_{12} \cos \theta_{12} \\ -I_{py} (\cos \theta_{12})^3 (a - l_2 \cos \theta_{12}) \dot{l}_2^2 \end{pmatrix} + \frac{2}{9} m_p \begin{pmatrix} \dot{\theta}_{11} \dot{l}_1 \cos \theta_{11} \sin \theta_{12} + \dot{\theta}_{12} \dot{l}_2 \cos \theta_{12} \sin \theta_{12} \\ + \dot{\theta}_{13} \dot{l}_3 \cos \theta_{13} \sin \theta_{12} \end{pmatrix}$$

$$C_3 = \frac{2}{9} m_p \begin{pmatrix} \dot{\theta}_{11} \dot{l}_1 \cos \theta_{11} \sin \theta_{13} + \dot{\theta}_{12} \dot{l}_2 \cos \theta_{12} \sin \theta_{13} \\ + \dot{\theta}_{13} \dot{l}_3 \cos \theta_{13} \sin \theta_{13} \end{pmatrix}$$

$$D = m_2 g [D_1 \quad D_2 \quad D_3 \quad 0]^T \tag{6}$$

$$D_1 = \frac{m_p \sin \theta_{11}}{9} \begin{pmatrix} -l_3 \dot{\theta}_{13}^2 \sin \theta_{13} - l_2 \dot{\theta}_{12}^2 \sin \theta_{12} - l_1 \dot{\theta}_{11}^2 \sin \theta_{11} \\ + l_1 \ddot{\theta}_{11} \cos \theta_{11} + l_2 \ddot{\theta}_{12} \cos \theta_{12} + l_3 \ddot{\theta}_{13} \cos \theta_{13} \end{pmatrix} + g \left(\frac{m_2}{2} + \frac{m_p}{3} \right) \sin \theta_{11} - \frac{I_{py} l_1 \ddot{\theta}_{11} \cos \theta_{11}}{(a^2 - (a - l_1 \cos \theta_{11}))^2} \begin{pmatrix} l_1 \ddot{\theta}_{11} \sin^2 \theta_{11} (a - l_1 \cos \theta_{12}) \\ + (\ddot{\theta}_{11} \cos \theta_{11} + \sin \theta_{11}) \end{pmatrix}$$

$$D_2 = \frac{m_p \sin \theta_{12}}{9} \begin{pmatrix} -l_3 \dot{\theta}_{13}^2 \sin \theta_{13} - l_2 \dot{\theta}_{12}^2 \sin \theta_{12} - l_1 \dot{\theta}_{11}^2 \sin \theta_{11} \\ + l_1 \ddot{\theta}_{11} \cos \theta_{11} + l_2 \ddot{\theta}_{12} \cos \theta_{12} + l_3 \ddot{\theta}_{13} \cos \theta_{13} \end{pmatrix} + \frac{m_2 g \sin(\theta_{12})}{2} - \frac{\cos \theta_{12}}{(a^2 - (a - l_2 \cos(\theta_{12}))^2)} \begin{pmatrix} I_{px} l_2^2 \dot{\theta}_{12}^2 \sin^2 \theta_{11} (a - l_2 \cos \theta_{12}) \\ + I_{py} l_2 (\dot{\theta}_{12}^2 \cos \theta_{12} + \ddot{\theta}_{12} \sin \theta_{12}) \end{pmatrix}$$

$$D_3 = \frac{m_p \sin \theta_{13}}{9} \begin{pmatrix} -l_3 \dot{\theta}_{13}^2 \sin \theta_{13} - l_2 \dot{\theta}_{12}^2 \sin \theta_{12} - l_1 \dot{\theta}_{11}^2 \sin \theta_{11} \\ + l_1 \ddot{\theta}_{11} \cos \theta_{11} + l_2 \ddot{\theta}_{12} \cos \theta_{12} + l_3 \ddot{\theta}_{13} \cos \theta_{13} \end{pmatrix} + g \sin \theta_{13} \left(\frac{m_p}{3} + \frac{m_2 g}{2} \right)$$

with $g = 9.8 (m/s^2)$ is the gravity coefficient.

$$F = [F_1 \quad F_2 \quad F_3 \quad \tau_\gamma]^T \tag{7}$$

is defined as a control signal vector.

3. Design controller for 4dof car motion simulator

3.1. Sliding mode controller with PID

The sliding mode control (SMC) is strongly requested due to its robustness against the disturbances. To ensure the convergence of the system to the wished state, a high level switching control is requested which generates the chattering phenomenon. In this way, a PID sliding surface with a saturation function will be proposed in this paper to solve this problem. Slide control law are designed based on the sliding function. In this case, the author has chosen the sliding function of the following form to eliminate the oscillation around the sliding surface when the amplitude of the slider control law changes greatly.

$$S = \dot{e} + K_p e + K_I \int_0^t e(t) dt \quad (8)$$

where K_p, K_I are PID parameters

The trajectory deviation is defined by $e = q - q_d$ with q_d is the desire trajectory of q . The goal of the controller design is to control q following the desire trajectory q_d with the small e . Select the sliding surface of controller as:

$$S = \dot{e} + \lambda_1 e + \lambda_2 \int_0^t e(t) dt \quad (9)$$

Sliding derivative \dot{S} can be obtained as:

$$\begin{aligned} \dot{S} &= \ddot{e} + \lambda_1 \dot{e} + \lambda_2 e = \ddot{q} - \ddot{q}_d + K_p \dot{e} + K_I e \\ &= M^{-1} (F - C(\dot{q}, q) - G(q)) - \ddot{q}_d + K_p \dot{e} + K_I e \end{aligned} \quad (10)$$

The control signal in SMC design consisting of equivalent control and switching control where the control action is corresponding with the sliding phase and reaching phase. The equivalent control is determined when $s(t)=0$, while the switch control is described when $s(t) \neq 0$. With $\dot{S} = 0$, equivalent control can be defined as:

$$F_{eq} = - (M (-\ddot{q}_d + \lambda_1 \dot{e} + \lambda_2 e) - C(\dot{q}, q) - G(q)) \quad (11)$$

We use the condition of the slider controller, $\dot{S} < 0$, F_{sw} is written in form:

$$F_{sw} = -M (k_1 S + k_2 \text{sign}(S)) \quad (12)$$

From there we have the control signal of the system like this:

$$F = F_{eq} + F_{sw} \quad (13)$$

Select the Lyapunov function: $V = \frac{1}{2} S^T S$, and the derivative of the function V :

$$\dot{V} = S^T \dot{S} = S^T (M^{-1} (F - C(\dot{q}, q) - G(q)) - \ddot{q}_d + \lambda_1 \dot{e} + \lambda_2 e) \quad (14)$$

$$\dot{V} = -S^T k_1 \dot{S} - S^T k_2 \text{sign}(S) \quad (15)$$

with control signal F is calculated in eq.7 above, and k_1, k_2 is selected as positive definite diagonal matrices, so we have the system of stability and $e \rightarrow 0$ lead to $q \rightarrow q_d$.

3.2. Design of Sliding Mode Controller type PID with GA Optimization

In this controller, K_p, K_I, k_1, k_2 are the constant parameters existing in sliding surfaces and control laws determine the overall performance. Hence, it is necessary to find the optimal values of them using optimization algorithm. Genetic algorithm (GA) is one of the fundamental evolutionary stochastic optimization algorithms. It mimics the process of natural selection and uses biological evolution to develop a series of search space points toward an optimal solution.

The goal of SMC-type-PID is to achieve accuracy trajectory tracking for 4-DOF car motion simulators; that is, the smaller the trajectory errors are, the more effective the controller is. Those parameters to be optimized are relevant to trajectory error; hence the fitness function is defined as follows:

$$f(\lambda_1, \lambda_2, k_1, k_2) = \sum_{i=0}^{\infty} \|e(i)\|^2 \quad (16)$$

with the fitness function, those parameters can be found with the minimization of tracking errors by using the designed control law with the minimization of tracking errors. The flowchart of GA Optimization is shown in figure 3 and figure 4.

After 15 generations, we have the following results: $k_1 = -25$; $\lambda_1 = 36$; and $k_2 = 64.1$; $\lambda_2 = 751.9$;

3.3. Numerical simulate and compare the results of two designed control algorithm.

The simulation parameters of 4-DOF car motion simulator are shown in Table 1. At the initial time, the position of the robot is denoted by $p_z=0.7(m)$, $\alpha=\beta=\gamma=0(rad)$, $l_1=l_2=l_3=0.65(m)$.

The control results of the proposed controller are evaluated through numerical simulation using Matlab/Simulink tool. Reference trajectory equations are described by a ternary function [13]:

$$q(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 \quad (17)$$

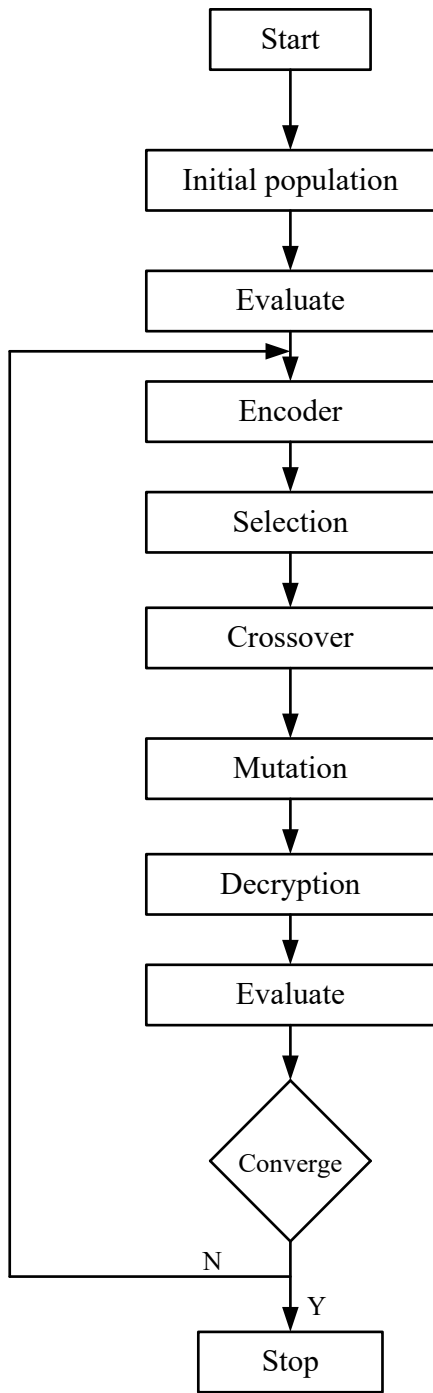


Figure 3: Genetic algorithm flowchart

Table 1: 4-DOF Car Motion Simulator Specifications

Items	Value
Base radius	500 [mm]
Moving radius	500[mm]
Leg strut	300[mm]
c	50[mm]
mp	15(kg)
m1	0.5(kg)
m2	10(kg)
mdc	3(kg)
k_1 (SMC-PID)	14.79
k_2 (SMC-PID)	14.79
λ_1 (SMC-PID)	15.1
λ_2 (SMC-PID)	14.79

From the motion of the center point of upper plate we can calculate the length of legs l_i ($i = 1, 2, 3$) using inverse kinematic equation of 4-DOF car motion simulator, the length of each limb shown in figure 5.

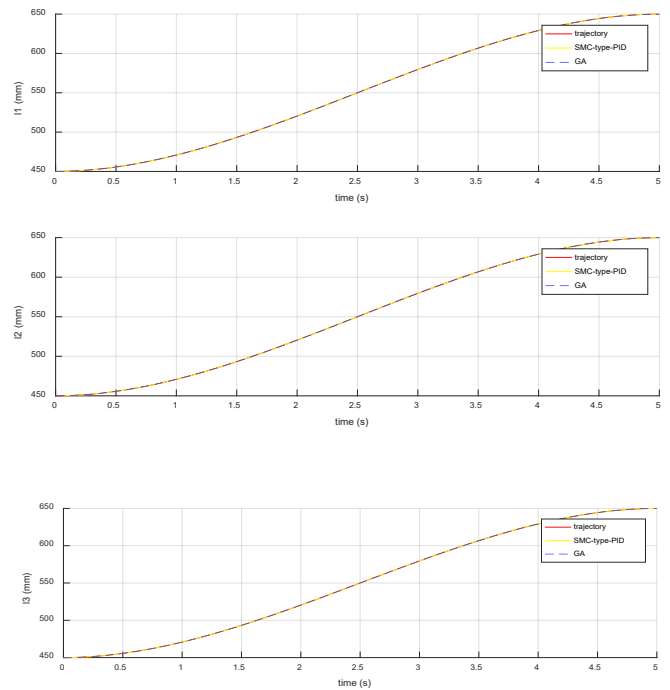


Figure 5: The trajectory responses of 3legs l_i ($i = 1, 2, 3$)

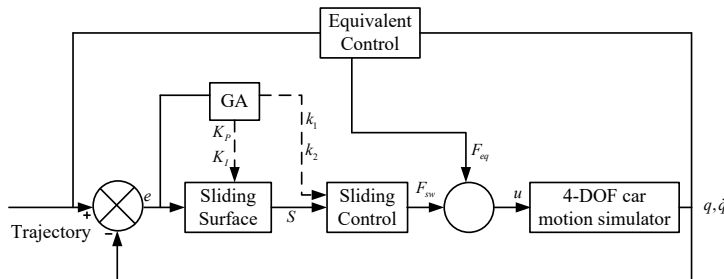


Figure 4: Flowchart of SMC-type-PD controller with a genetic algorithm optimisation

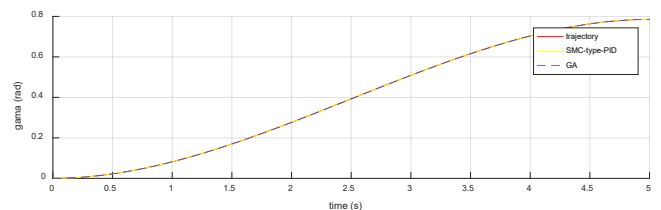


Figure 6: The trajectory responses of trajectory angle gamma

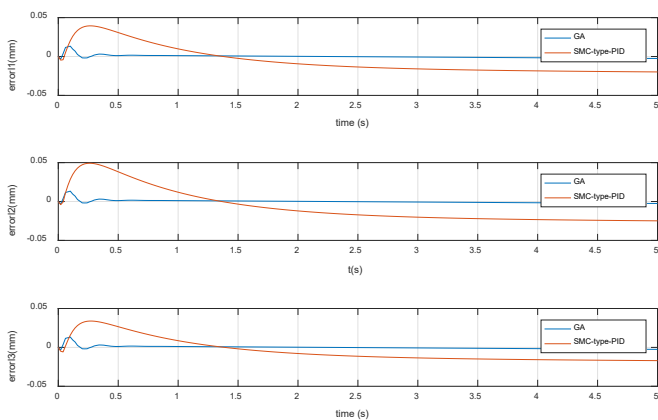


Figure 7: The trajectory error of the 3legs $l_i (i = 1, 2, 3)$

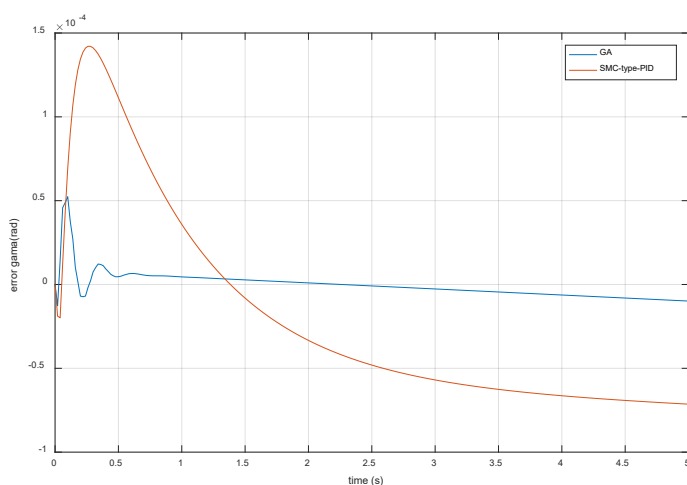


Figure 8: The tracking error of the angle trajectory

The simulation results show that the adaptive controller constructed by combining the SMC-type-PID with the GA optimization is able to ensure the stability for the car driving simulator system and the tracking error converge to zero rapidly. In general, the SMC-type-PID control with the GA Optimization is achieved accuracy trajectory tracking; that is, the smaller the trajectory errors are, the more effective the controller is.

4. Conclusions

This paper presented an adaptive control method for the 4DOF car motion simulator. The control method design is based on the sliding mode controller type PID structure combined with the sliding mode controller type PID with GA optimization for the system. In the simulation results we can see that the SMC-type-PID with the saturation functions give us good results, stability and accuracy of the process output and control evolution. The parameters selected by the genetic algorithm give good results and smaller errors. The stability of the system is proven by using Lyapunov theory, the simulation results show the proposed controller can be ensure the stability and tracking performance for the system. Through the simulation results, the author found that the sliding mode controller type PID with GA optimization gives better results because the deviation of the set value with the actual value is very small, approximately 0. Therefore, the sliding mode

controller type PID with GA optimization can find the optimal values using optimization algorithm and it be recommended for nonlinear systems request for high accuracy as 4-DOF car motion simulators. Further and conduct experiments on real models in the near future.

Acknowledgment

This work is funded by VAST project under grant number VAST01.06/19-20.

References

- [1] J. Pearce, "Research and Development of a 6 Degrees of Freedom Electric Motion Platform," Bachelor's Thesis, 2016.
- [2] Y. Qiu, X. Liang, and Z. Dai, "Backstepping dynamic surface control for an anti-skid braking system," *Control Eng. Pract.*, vol. 42, pp. 140–152, 2015. <https://doi.org/10.1016/j.conengprac.2015.05.013>
- [3] G. Sun, D. Wang, X. Li, and Z. Peng, "A DSC approach to adaptive neural network tracking control for pure-feedback nonlinear systems," *Appl. Math. Comput.*, vol. 219, no. 11, pp. 6224–6235, 2013. <https://doi.org/10.1016/j.amc.2012.12.034>
- [4] T. Zhang, M. Xia, and Y. Yi, "Adaptive neural dynamic surface control of strict-feedback nonlinear systems with full state constraints and unmodeled dynamics," *Automatica*, vol. 81, pp. 232–239, 2017. <https://doi.org/10.1016/j.automatica.2017.03.033>
- [5] S. Qi, D. Zhang, L. Guo, and L. Wu, "Adaptive Dynamic Surface Control of Nonlinear Switched Systems with Prescribed Performance," *J. Dyn. Control Syst.*, vol. 24, no. 2, pp. 269–286, 2018. <https://doi.org/10.1007/s10883-017-9374-7>
- [6] K. A. Semprun, L. Yan, W. A. Butt, and P. C. Y. Chen, "Dynamic surface control for a class of nonlinear feedback linearizable systems with actuator failures," *IEEE Trans. neural networks Learn. Syst.*, vol. 28, no. 9, pp. 2209–2214, 2016. <https://doi.org/10.1109/TNNLS.2016.2572205>
- [7] Y.-J. Liu, J. Li, S. Tong, and C. L. P. Chen, "Neural network control-based adaptive learning design for nonlinear systems with full-state constraints," *IEEE Trans. neural networks Learn. Syst.*, vol. 27, no. 7, pp. 1562–1571, 2016. <https://doi.org/10.1109/TNNLS.2015.2508926>
- [8] X. Ji, X. He, C. Lv, Y. Liu, and J. Wu, "A vehicle stability control strategy with adaptive neural network sliding mode theory based on system uncertainty approximation," *Veh. Syst. Dyn.*, vol. 56, no. 6, pp. 923–946, 2018. <https://doi.org/10.1080/00423114.2017.1401100>
- [9] M. Kim, B. Jung, B. Han, S. Lee, and Y. Kim, "Lyapunov-based impact time control guidance laws against stationary targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 2, pp. 1111–1122, 2015. <https://doi.org/10.1109/TAES.2014.130717>
- [10] A. Wolf et al., "Determining Lyapunov exponents from a time series," vol. 16, no. 3, pp. 285–317, 2017. [https://doi.org/10.1016/0167-2789\(85\)90011-9](https://doi.org/10.1016/0167-2789(85)90011-9)
- [11] S. K. Y. Nikraves, *Nonlinear systems stability analysis: Lyapunov-based approach*. CRC Press, 2018.
- [12] Kiem Nguyen Tien, Duyen Ha Thi Kim, Tien Ngo Manh, Cuong Nguyen Manh, Ngoc Pham Van Bach, Hiep Do Quang, "Adaptive Dynamic Surface Control for Car Driving Simulator based on Artificial Neural Network", 2019 International Conference on Mechatronics, Robotics and Systems Engineering (MoRSE), 2019 <https://doi.org/10.1109/MoRSE48060.2019.8998749>
- [13] L. W. Tsai, *Robot Analysis: The Mechanics of Serial and Parallel Manipulators*, Wiley, 1999.

Promotion of the Research Activities at the Image Processing Research Laboratory (INTI-Lab) of the UCH as Knowledge Management Strategy

Avid Roman-Gonzalez^{*1,2}, Natalia Indira Vargas-Cuentas¹

¹*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, 15834, Peru*

²*Business on Engineering and Technology (BE Tech), 15076, Lima, Peru*

ARTICLE INFO

Article history:

Received: 08 April, 2020

Accepted: 22 June, 2020

Online: 26 June, 2020

Keywords:

Knowledge management

INTI-Lab

Research assessment

Research promotion

Research achieves

ABSTRACT

In Peru, approximately since 2013, a necessary change has begun in the importance given to research, science, technology, and technological innovation. Likewise, in 2014, a new University Law was approved that among other aspects also promotes research production in universities. Against this context, the universities begin to improve with more emphasis activities related to research. The Universidad de Ciencias y Humanidades creates different research centers, one of them being the Image Processing Research Laboratory (INTI-Lab). Through INTI-Lab research projects are generated both with own resources and through external financing sources. INTI-Lab helps to increase the number of papers published and indexed in SCOPUS, increase the score in Researchgate platform, improve the position on the Webometrics ranking, and brings students closer to the research activity. In this context, a Knowledge Management strategy is essential. In the present work, an analysis of the different strategies and results will be presented. The analysis shows that the knowledge management strategies adopted by INTI-Lab contribute to increase the scientific production of the UCH.

1. Introduction

The Peruvian government, through the National Council of Science, Technology and Technological Innovation (CONCYTEC for its acronym in Spanish) since about 2013 have begun the implementation of some policies that encourage the research activities oriented to solve real problems of society. These actions include contest funds for the development of postgraduate studies, conducting research stays, financing for research projects, financing for scientific equipment, funding for the organization of technological events, incentives for the publication of scientific articles, tax incentives for companies that invest in science and technology [1].

On the other hand, the Ministry of Education of Peru, through the National Superintendency of Higher Education (SUNEDU for its acronym in Spanish) has been performing the licensing of the universities that meet the Basic Conditions of Quality (CBC for its acronym in Spanish) [2].

Faced with this promising context for research, fostered by the policies described above, universities have increased the recruitment of researchers to increase scientific production.

The Universidad de Ciencias y Humanidades (UCH), in this context, could not be left behind. The UCH began its academic activities in 2008, its headquarters are located in Los Olivos district, Lima, Peru. UCH currently has ten professional schools and provides service to more than 3000 students. The UCH obtained the SUNEDU license in November 2017.

The UCH on the subject of research has made a substantial investment to promote the development of research projects and increase scientific production [3]. This investment includes the creation of 3 research centers: e-Health (created in 2014), a research center that focuses on the subject of health sciences and is related to the professional school of nursing. CIICS (created in May 2016), an interdisciplinary research center focused on social and economic issues, related to professional schools of initial education, primary education, accounting, psychology, marketing, and administration. INTI-Lab (created in June 2016), a research center focused on the application of engineering to solve real

*Avid Roman-Gonzalez, Av. Universitaria 5175, Los Olivos, Lima, Peru, avid.roman-gonzalez@ieeee.org

problems of society, related to the professional schools of Systems Engineering, Electronic Engineering, and Industrial Engineering.

Likewise, the UCH promotes formative research in its undergraduate students through 3 programs that are developed throughout the five years of study of each professional school [4].

All these mentioned actions have allowed the UCH to improve positions in different rankings based on research such as those presented in [5, 6]. Taking the described situation into account, a good knowledge management strategy is necessary to maintain achievements and continue to improve.

Different authors know knowledge management (KM) as a fundamental strategy for companies in this contemporary situation [7, 8]. The KM plays a pivotal role in the competitiveness of large firms and small and medium enterprises [9] due to different strategies for conserving, passing, and continuing the organization strategy regardless if there are changes in the personal resources.

The present work continues as follows: Section II presents the Image Processing Research Laboratory (INTI-Lab), its creation and its research topics, as well as its members. Section III shows the different achievements obtained by INTI-Lab since its establishment and the Knowledge Management strategy applied in INTI-Lab. Finally, Section IV presents the analysis and conclusions of the results.

2. Image Processing Research Laboratory (INTI-Lab)

The Image Processing Research Laboratory (INTI-Lab for its acronym in Spanish) is one of the three research centers of the UCH. The acronym was chosen due to INTI means Sun in Quechua, Quechua is the mother language of INCAS (pre-Spanish culture in Peru), and Sun was considered a God in the INCAS culture. INTI-Lab was created on June 23, 2016, through Resolution No. 090-2016-CU-UCH. INTI-Lab is a research center dedicated mainly to the development of projects related to the signal and image processing, aerospace technology, rehabilitation engineering, smart cities, ICT, biotechnology, development of electronic systems, computer systems and artificial intelligence.

The main objective of INTI-Lab is to contribute to the generation and growth of research production at the Universidad de Ciencias y Humanidades, both at the professor level and at the student level, developing research projects that can help to solve real problems of our society.

INTI-Lab has principally six research topics that are:

Electronic Circuits and Communication Systems.

- Computing Systems and Computer Science.
- Industrial Applications
- Engineering in Medicine and Biology
- Aerospace systems
- ICT Management

At its beginning, INTI-Lab was created with only one member, in 2017 INTI-Lab go to 12 members, in 2018, 11 members, and currently, INTI-Lab has 15 members.

In Figure 1, one can see de logo of the INTI-Lab.

www.astesj.com



Figure 1: INTI-Lab Logo

3. Achievements and Knowledge Management Strategy of INTI-Lab in the Research Field

Since its creation in June 2016, INTI-Lab has had several achievements among which we can mention:

- First place in the COPROING 2016.
- First and second place in the 1st Conference of Sciences and Humanities (2016).
- Third place in the Science and Humanities Fair 2017.
- First Place in the II Conference of Sciences and Humanities (2017).
- 5 United Nations grants.
- First place in the II Creaton UCH 2018.
- One of our members won a scholarship to pursue a master's degree in Russia.
- One member won the OHB SE Competition (2018).
- Best paper award in ETCM 2018.
- Five grants for research project development.
- Emerging Space Leader - ESL 2019
- Young Space Leader - YSL 2019
- 8 research grants (international and national)

3.1. Invited Lectures

Part of the research work is to be able to publicize the results of research, not only at the technical level through the publication of scientific articles, but also doing scientific dissemination tasks for the general public through talks. In that sense, in Figure 2, one can observe the evolution of lectures given by the members of INTI-Lab in different years.

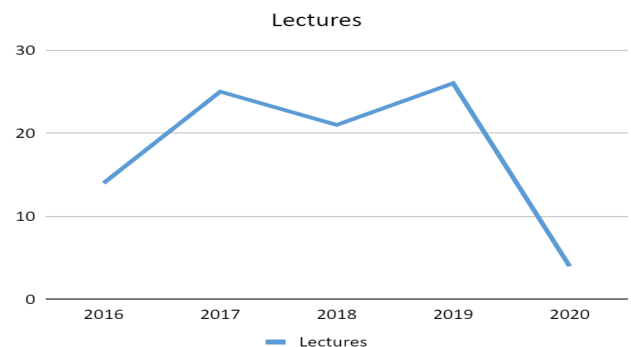


Figure 2: Lectures given by INTI-Lab members (collected data at 20/06/2020)

3.2. Press

The works developed in INTI-Lab not only reach the public through published scientific articles, or through the given talks, but also attract the attention of the press, could be written, radio, and/or television. In this sense, since 2016, different members of INTI-Lab have had press appearances giving the results of the developed projects. INTI-Lab projects had press coverages in 2016, 2017, 2018, 2019 and so far in 2020.

In Figure 3, one can see a sample of the different press coverages made to the INTI-Lab projects in 2016, 2017, 2018, and so far in 2019.



Figure 3: INTI-Lab in press

3.3. Publications

The publication of scientific articles is one of the final tasks of researchers. The idea is to be able to share the results of the research with the scientific community of the whole world. Whatever the topic one is investigating, let's be sure that other research groups are working on similar topics. The publication of articles allows us to share results; it can open the possibility of working in groups, sign agreements, make academic exchanges, among other benefits.

Likewise, there are many rankings of universities that take into account the number of published articles. One of these rankings is the one developed by the Scimago group [10].

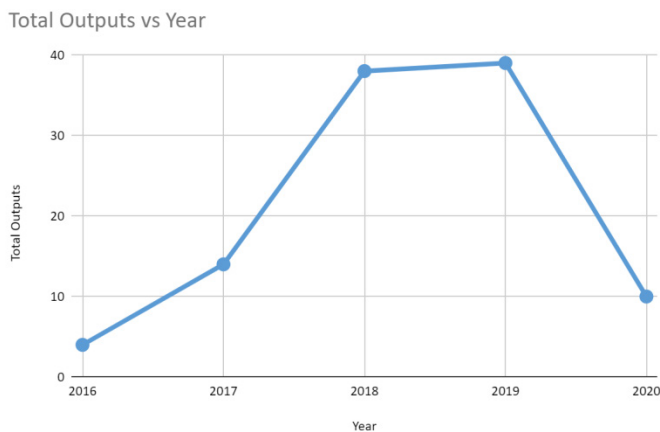


Figure 4 : Scientific articles published by INTI-Lab members (collected data at 20/06/2020)

Figure 4 shows the evolution in terms of the number of published papers by INTI-Lab; these data include indexed and non-indexed documents.

3.4. Agreements

An essential part of the work carried out in INTI-Lab are the agreements signed with other national and international institutions. These agreements allow us to develop projects together, the use of technological tools, the ability to carry out academic exchanges, research stays, among other benefits.

The institutions with which INTI-Lab currently has an agreement are the following:

- INICTEL-UNI (Peru)
- Universidad Peruana Cayetano Heredia (Peru)
- Universidad de Malaga (Spain)
- Beihang University (China)
- Universitatea Tehnica Gheorghe Asachi (Romania)
- Peruvian Space Agency (Peru)
- Open Cosmos (United Kingdom)

In Figure 5 one can see the logos of the institutions with which INTI-Lab has an agreement.



Figure 5: Logos of the institutions with which INTI-Lab has an agreement (collected data at 20/06/2020)

3.5. Scientific production of the UCH

The UCH, thanks to its investment in research, has increased its scientific production, which can be seen in Figure 6.

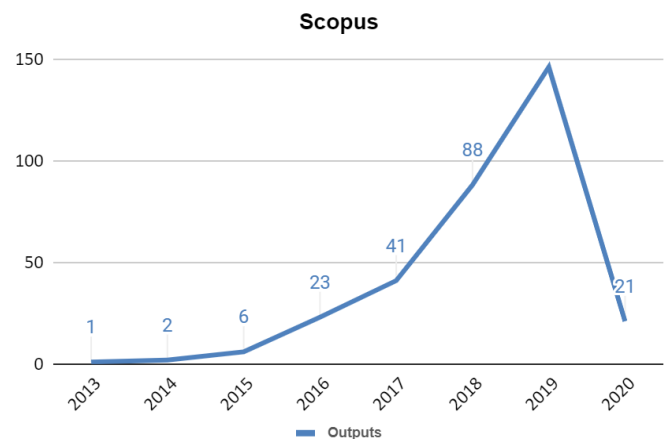


Figure 6: The number of UCH scientific articles indexed in SCOPUS over the years (collected data at 20/06/2020)

This increase has achieved a better positioning of the UCH in different rankings since different rankings are depending on the research. One ranking using metrics based on the web is

Webometrics [11]; the performance of the UCH in the Webometrics ranking can be seen in Figure 7.

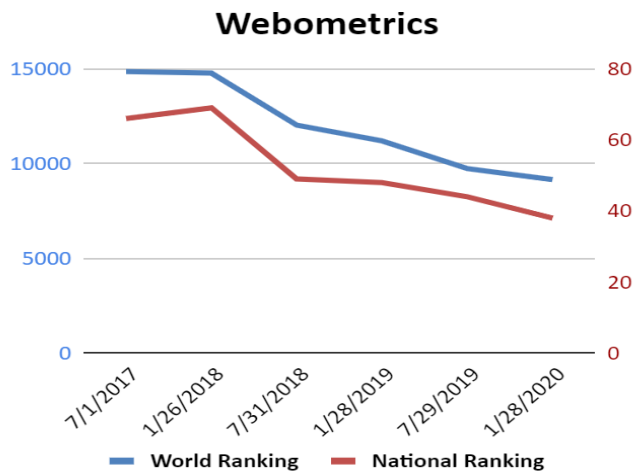


Figure 7: Progress of the position of the UCH in the ranking developed by Webometrics. In the gray curve and the left vertical axis, one can see the place worldwide. In the curve in blue and the right vertical axis, the position at the national level is observed. (collected data at 20/06/2020)

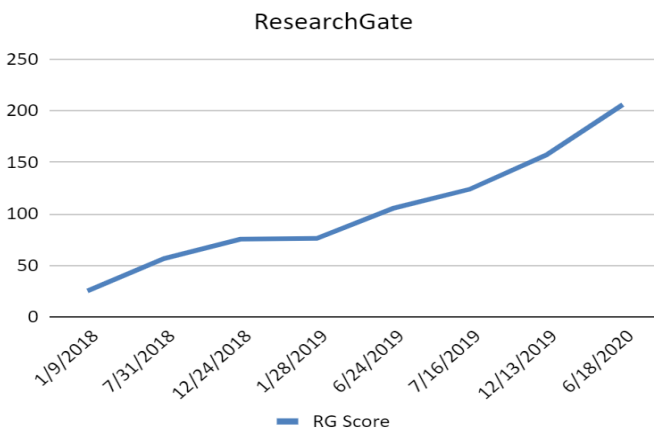


Figure 8: Evolution of the UCH concerning the ResearchGate Score. (collected data at 20/06/2020)

But the benefits are not only reflected in the ranking but also in the ResearchGate scores as can be seen in Figure 8. ResearchGate is a kind of social network for researchers in which what is shared are scientific articles, thesis, dissertations, posters, books, etc. [12].

3.6. Contribution of INTI-Lab to the scientific production of the UCH

The UCH has different dependencies among which one can mention three research centers (e-Health, CIICS, and INTI-Lab), ten professional schools (electronic engineering, systems engineering, industrial engineering, nursing, psychology, accounting, administration, marketing, primary education, and initial education), and the general studies department. Each of these dependencies, organizing in faculties (Science and Engineering Faculty - SEF, Health Science Faculty - HSF, Accounting, Economics, and Financial Faculty - AEFF, Humanities and Social Sciences Faculty - HSSF, and General Studies), contributes to the total scientific production of the UCH.

In Figure 9, one can see the contribution in the number of papers of each unit. It can be seen that the number of documents of INTI-Lab varies concerning what is shown in Figure 2 since, in the case of Figure 9, only the papers indexed in SCOPUS are shown. For the scientific production of the UCH, the analysis is done only taking into account the documents indexed in SCOPUS since they are the ones that count to elaborate the Scimago ranking.

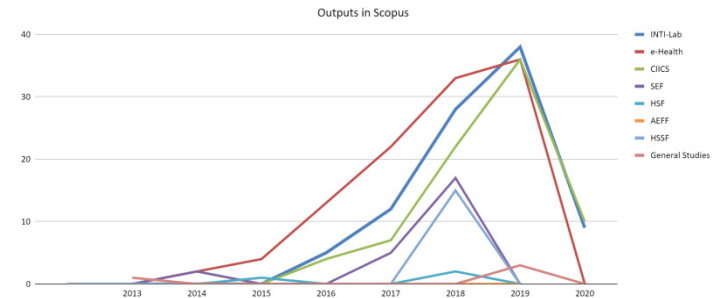


Figure 9 Scientific productions in SCOPUS of the different dependencies of the UCH. (collected data at 20/06/2020)

INTI-Lab being the newest of the three research centers of the UCH presents an essential contribution to the number of papers of the UCH

3.7. INTI-Lab Scientific Tuesday

Knowledge management is essential in any institution to transmit knowledge. In INTI-Lab, knowledge management is vital to continue growing in the indicators mentioned above, regardless of the change of members. At INTI-Lab, as a knowledge management strategy, "INTI-Lab Scientific Tuesdays" have been implemented. These scientific events are days in which the members of the research team (in turn, according to a pre-established schedule) share their progress on the projects that they are developing. These sessions take place once a month. Thanks to this strategy, each member knows and can contribute to the work of others. In Figure 10, one can observe a session of the "INTI-Lab Scientific Tuesday".



Figure 10: INTI-Lab Scientific Tuesday.

4. Discussion and Conclusion

According to what is presented in Section 3, it can be observed that the contribution of INTI-Lab to the scientific production of the

UCH is essential. One can see in Figure 11 that this contribution represents 27.96% of the total scientific output of the UCH throughout the different years.

The contribution of INTI-Lab is only below e-Health, which is the oldest research center of the university. INTI-Lab since its creation has occupied the second place and has been shortening distances with e-Health. For 2019, INTI-Lab go to the first place.

In each of the aspects measured, whether publications, invited lectures, agreements, contests, and grant; the growth of INTI-Lab has been exponential.

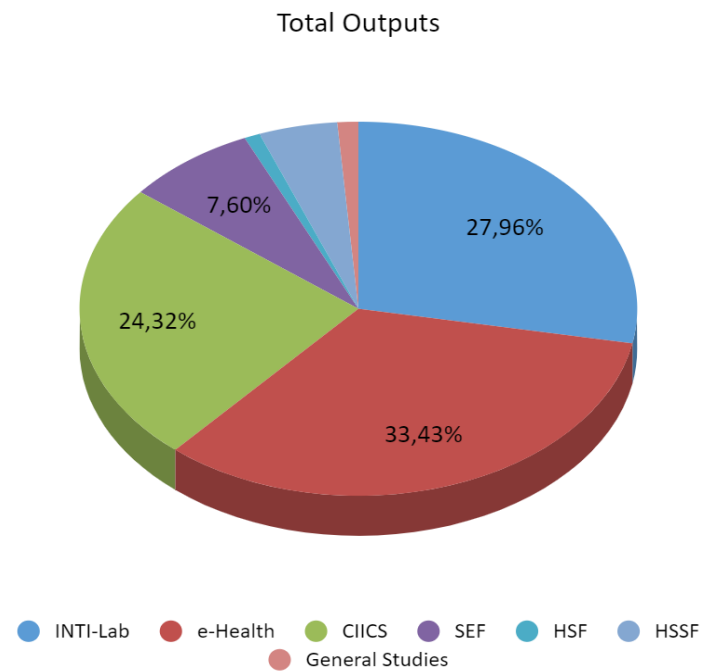


Figure 11: Contribution of the different dependencies to the scientific production of the UCH.

INTI-Lab has always tried to work hand in hand with students. So it is the UCH research center with the most significant number of papers whose first author is a student, reaching 17.14% of the total number of articles published by INTI-Lab.

Unlike other places where publication with students is done mainly with postgraduate students, INTI-Lab works with undergraduate students, since the UCH does not yet have a post-graduate school. The fact of achieving a good cup of publications with undergraduate students is another aspect to highlight.

In addition to the statistics already mentioned, one can suggest other achievements of INTI-Lab that highlight the promotion of research at the UCH. For example, one of the members of INTI-Lab won a scholarship to do postgraduate studies in Russia. Another INTI-Lab member was recognized as a researcher by CONCYTEC, becoming the first CONCYTEC researcher entirely UCH (UCH student, UCH internship, and currently working at UCH).

On the other hand, the Knowledge Management (KM) must be fundamental in every company, also in startups where the learning curve must be shorter. KM is a differentiating factor. Universities have the main focus of generating knowledge, and one has to design the model or strategy that is needed from what KM means

for the university. Patents, scientific articles, as indicators or metrics, can be used. In companies with high capital investments and with significant operational risks (such as companies related to space missions), the KM is essential because an error can cause considerable losses. These losses could be in infrastructure or lives.

The INTI-Lab strategies described in this work contribute to the objective on the research activities at university

References

- [1] Lastra, Javier Ulises Solis. "Tax Incentive to Promote Research Development and Innovation in Peru." 2018 IEEE Sciences and Humanities International Research Conference (SHIRCON). IEEE, 2018. DOI: 10.1109/SHIRCON.2018.8593197
- [2] López, Sonia F. Paredes, José Clemente Flores Barboza, and Franks Paredes Rosales. "El licenciamiento de la Universidad Ricardo Palma. La experiencia exitosa del trabajo en equipo." *Tradición, segunda época* 16 (2018): 97-103. DOI: <https://doi.org/10.31381/tradicion.v0i16.1440>
- [3] B. B. Gonzales, I. Iraola-Real, D. Llulluy, B. Meneses-Claudio, and A. Roman-Gonzalez, "Relationship Between Research Production and Funding at the Universidad de Ciencias y Humanidades", III IEEE World Engineering Education Conference – EDUNINE 2019, March 2019, Lima - Peru. DOI: 10.1109/EDUNINE.2019.8875754
- [4] U. Lapa-Asto, G. Tirado-Mendoza, and A. Roman-Gonzalez, "Impact of Formative Research on Engineering Students", III IEEE World Engineering Education Conference – EDUNINE 2019, March 2019, Lima - Peru. DOI: 10.1109/EDUNINE.2019.8875842
- [5] Roman-Gonzalez, Avid, and Natalia I. Vargas-Cuentas. "Scientific Production in the 50 First Universities Licensed by SUNEDU." 2018 IEEE Sciences and Humanities International Research Conference (SHIRCON). IEEE, 2018. DOI: 10.1109/SHIRCON.2018.8593133
- [6] Roman-Gonzalez, Avid, et al. "Ranking of the Licensed University of Peru Based on the Active Research Grade (ARG)." 2019 IEEE 39th Central America and Panama Convention (CONCAPAN XXXIX). IEEE, 2019. DOI: 10.1109/CONCAPANXXXIX47272.2019.8977117
- [7] Grant, Robert M. "The resource-based theory of competitive advantage: implications for strategy formulation." *California management review* 33.3 (1991): 114-135. DOI: <https://doi.org/10.2307/41166664>
- [8] Sokolov, D., and E. Zavyalova. "Knowledge management strategies, HRM practices and intellectual capital in knowledge-intensive firms." (2018).
- [9] Cerchione, Roberto, and Emilio Esposito. "Using knowledge management systems: A taxonomy of SME strategies." *International Journal of Information Management* 37.1 (2017): 1551-1562. DOI: <https://doi.org/10.1016/j.ijinfomgt.2016.10.007>
- [10] De-Moya-Anegón, Félix; Herrán-Páez, Estefanía; Bustos-González, Atilio; Corera-Álvarez, Elena; Tibaná-Herrera, Gerardo (2018). *Ranking Iberoamericano de instituciones de educación superior. SIR Iber 2018*. Barcelona, España: Ediciones Profesionales de la Información SL. ISBN: 978 84 09 03911 1. DOI: <https://doi.org/10.3145/sir-iber-2018>.
- [11] Webometrics, "Ranking de Universidades", Edition 2018 2.1.2, <http://www.webometrics.info/>.
- [12] Yu, Min-Chun, et al. "ResearchGate: An effective altmetric indicator for active researchers?." *Computers in human behavior* 55 (2016): 1001-1006. DOI: <https://doi.org/10.1016/j.chb.2015.11.007>

A Fuzzy-PID Controller Combined with PSO Algorithm for the Resistance Furnace

Trinh Luong Mien^{1,*}, Vo Van An², Bui Thanh Tam³

¹University of transport and communications, 03 Cau Giay, Lang Thuong, Dong Da, Hanoi, Vietnam

²Binh Duong economics and technology university, 530 Binh Duong avenue, Thu Dau Mot, Binh Duong, Vietnam

³Thu Dau Mot university, 06 Tran Van On, Thu Dau Mot, Binh Duong, Vietnam

ARTICLE INFO

Article history:

Received: 05 April, 2020

Accepted: 08 June, 2020

Online: 26 June, 2020

Keywords:

PID

PSO

Fuzzy-PID

Fuzzy logic

Resistance furnace

Temperature control

ABSTRACT

The paper presents a novel control strategy applying the particle swarm optimization (PSO) algorithm to optimize the scaling weights coefficients of the fuzzy-PID controller for the resistance furnace temperature control system, called PSO-based fuzzy-PID controller/algorithm. The proposed PSO-based fuzzy-PID controller in this paper consist of the fuzzy-PID controller and the PSO algorithm. The proposed fuzzy-PID controller is combination of the advantage of PID control and fuzzy logic control. Firstly, the paper presents the mathematical model of the resistance furnace by identification method, based on the experimental data. Then, the design of the fuzzy-PID controller is given in this study. And then, the paper presents the design of the temperature control board using PIC16f with the installed PSO-based fuzzy-PID algorithm. Finally, the simulation and experimental results proved the stability of the proposed PSO-based fuzzy-PID controller with the disturbance, improved the furnace temperature control quality, through obtained major control criteria, such as overshoot, steady-state error, settling time, rising time.

1. Introduction

In industry and transportation, the resistance furnace is a very commonly device. Its temperature control has the characteristics of the one-direction temperature rise, large inertia, time delay and time-varying parameters. In the temperature control system, it is difficult to establish accurate models and determine parameters by mathematical methods, especially when it is disturbed.

The traditional control methods based on the exact object mathematical model, such as feedback control, PID control [1-3] can meet the system performance requirements, and has the advantage of eliminating the steady-state error, but its performance depends on the tuning of the parameters.

Fuzzy logic controller (FLC) has the advantage of inaccurate object model, rapidity and small overshoot, but the control process will have steady-state error. The research works [4-7] used only fuzzy logic for controlling temperature of the resistance furnace and achieved limited results. The quality of the temperature fuzzy logic control system depends on expert experiment and FLC's parameters. These works only stopped at the computer simulation, did not solve the effects of disturbance.

The research works [8-16] were combined the advantages of the PID controller with fuzzy logic to design the intelligent controller for the resistance furnace. However, all most of them are still stopped in simulation on computer, but have not given the hardware control board with the intelligent control algorithm on microprocessor.

The research work [17] introduced the fuzzy PID based on genetic algorithm for the resistance furnace. This work [17] indicated that the temperature control system dynamic quality and stable precision is improved, but has not been conducted experimentally on physical equipment and has not assessed the impact of disturbance.

This paper presents a novel fuzzy-PID controller combined with PSO algorithm to control temperature of the resistance furnace with disturbance: from theory to practical experiments. The fuzzy-PID is designed based on the combination of PID control and fuzzy logic control, in which the basic parameters of PID are calculated by fuzzy logic control, and the parameters of fuzzy-PID are determined by multiplying the basic parameters of PID with the scaling weights. These multiplied scaling weights of the fuzzy-PID controller are calculated based on PSO algorithm, so that the dynamic response of the control system is better and the

*Trinh Luong Mien. Tel: +84904684595. Email: mientl@utc.edu.vn

steady-state error is eliminated. The part 2 of the paper presents the resistance furnace identification, determining the PID basic parameters. In part 3, a new PSO-based fuzzy-PID controller is given, here PSO algorithm is used to optimized the scaling weights of the fuzzy-PID controller. And then the simulation model of the resistance furnace temperature control system is built on Matlab in part 4 in order to verify the feasibility of the proposed PSO-based fuzzy-PID controller. Then, the design of the temperature control board using PIC16f microcontroller with PSO-based fuzzy-PID algorithms is introduced in part 5. Finally, part 6 is conclusions. The simulation and experimental results show that the proposed PSO-based fuzzy-PID controller improve the resistance furnace temperature control system quality.

2. The resistance furnace mathematical model by identification method

The resistance furnace, the studied object in this paper, has dimensions L375xW255xH345mm. The furnace uses resistance wire, powered by 220VAC/50Hz, maximum power of 1000W, maximum temperature of 100°C. In this system, a 2-wire K-type thermocouple sensor with the setting temperature range 0-100°C corresponding to the electrical signal 0-5VDC is adopted. Controlling the AC power supply capacity for the thermistor wire changes the furnace temperature here, the authors use the TRIAC BTA10-800B 10A 600V [18].

For resistance furnace identification, the control board here only plays the role of opening 100% the TRIAC to supply 220VAC power to the resistance wire. When the TRIAC was active with 100% capacity of the resistance furnace, at the time of measurement the authors realized that the furnace temperature was not changed immediately, it took a certain amount of time to convert electricity into heat energy, heat transfer in the furnace - delay time of the furnace. The temperature in the furnace will gradually increase to the maximum value, corresponding to the maximum power of the furnace.

Table 1. The experimental temperature values of the furnace

No.	Time [minute]	Temperature [°C]	Note
1	1	27.8	ambient temperature
2	2	29.1	
3	3	30.5	
4	4	32.0	
5	5	33.5	
6	6	35.9	
7	7	38.5	
8	8	41.3	
9	9	44.7	
10	10	47.8	
11	11	51.2	
12	12	54.9	
13	13	58.2	
14	14	61.6	
15	15	65.0	
16	16	68.2	
17	17	71.1	
18	18	74.2	
19	19	76.8	
20	20	79.4	
21	21	81.7	
22	22	84.0	
23	23	86.0	
24	24	87.7	
25	25	88.9	
26	26	89.9	
27	27	90.5	
28	28	91.0	
29	29	91.4	
30	30	91.4	

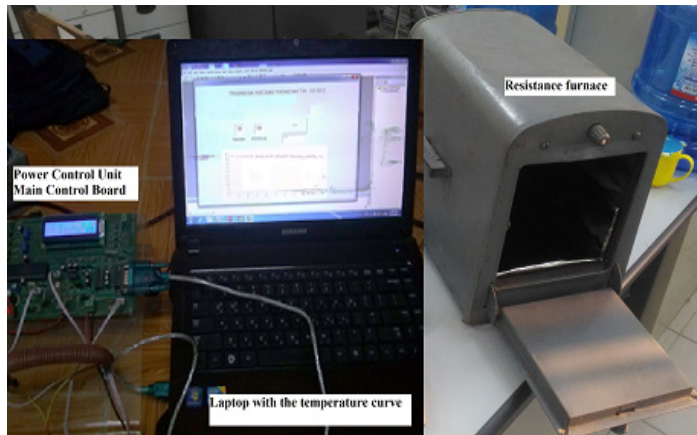


Figure 1. The resistance furnace identification process

Based on this experimental method, the resistance furnace mathematical model can be identified. The identification process was carried out at the ambient temperature about 27.8°C. The measured steady temperature was about 91.4°C. The time to reach the steady output temperature of the furnace was about 30 minutes.

The measured temperature curve of the resistance furnace as shown in Fig.2.

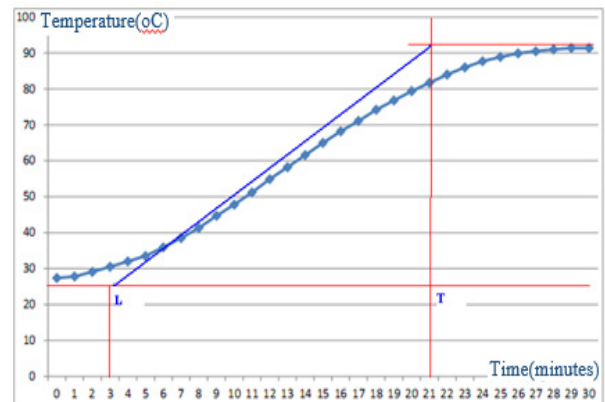


Figure 2. Measured furnace temperature curve

The measured furnace temperature curve has the S-shape, so the furnace mathematical model is described as below [1]:

$$P(s) = \frac{K}{1 + sT} e^{-Ls} \tag{1}$$

where L is the time delay, period time that the output signal does not respond immediately; K is the transfer coefficient, is the limit output value as $t \rightarrow \infty$; T is the inertial time.

Based on the furnace temperature characteristic curve, the authors draw the tangent line with the inflection point of this curve, the parameters of the furnace as follows:

$$K = \frac{91.4 - 27.8}{100} = 0.64 \quad (2)$$

It is easy to identify $L=225$ second and $T=1230$ second.

Therefore, we obtain the transfer function of the furnace as:

$$G_{obj}(s) = \frac{0.64e^{-225s}}{1230s + 1} \quad (3)$$

It is easy to see that the mathematical model of the resistance furnace is nonlinear, large time delay.

According to Ziegler Nichols-1 (ZN1) [1], the PID controller parameters for the furnace, are determined as equations (5).

$$G_{PID}(s) = K_p(1 + \frac{1}{T_I s} + T_D s) = K_p + \frac{K_I}{s} + K_D s \quad (4)$$

$$K_{p0} = \frac{1.2T}{KL}, T_{I0} = 2L, T_{D0} = 0.5L \quad (5)$$

Therefore, the initial PID parameters are chosen as below:

$$K_{p0}=10.256; T_{I0}=450, T_{D0}=112.5$$

3. Design of PSO-based Fuzzy-PID controller

3.1. Proposed PSO-based Fuzzy-PID controller

To improve the performance efficiency of the fuzzy-PID controller for the resistance furnace, this work proposed a novel control approach applying the PSO algorithm to tune the scaling weights of the fuzzy-PID controller. The structure diagram describing the incorporate between the fuzzy-PID controller and PSO algorithms is presented in Figure 3.

The fuzzy-PID controller is synthesized based on the structure of the PID with K_p , T_I and T_D parameters, determined according to the fuzzy logic calculated blocks P, I, D corresponding to defuzzified output values K_{pf} or T_{If} or T_{Df} and then multiply with the corresponding scaling weights, i.e. K_a , K_b , K_c , following formula as below:

$$K_p = K_{pf} K_a; T_I = T_{If} K_b; T_D = T_{Df} K_c \quad (6)$$

The PSO algorithm is used to optimize three scaling weights, i.e. $K_a \rightarrow K_a^*$, $K_b \rightarrow K_b^*$, $K_c \rightarrow K_c^*$. So that after applying the PSO algorithms, the parameters' Fuzzy-PID is optimized as below:

$$K_p^* = K_{pf} K_a^*; T_I^* = T_{If} K_b^*; K_D^* = T_{Df} K_c^* \quad (7)$$

For this proposed PSO-based fuzzy-PID controller, the parameters of the fuzzy-PID are continuously adjusted in a specified range. However, the scaling weights' fuzzy-PID is optimized by PSO algorithm. Therefore, the furnace temperature control quality will be greatly improved

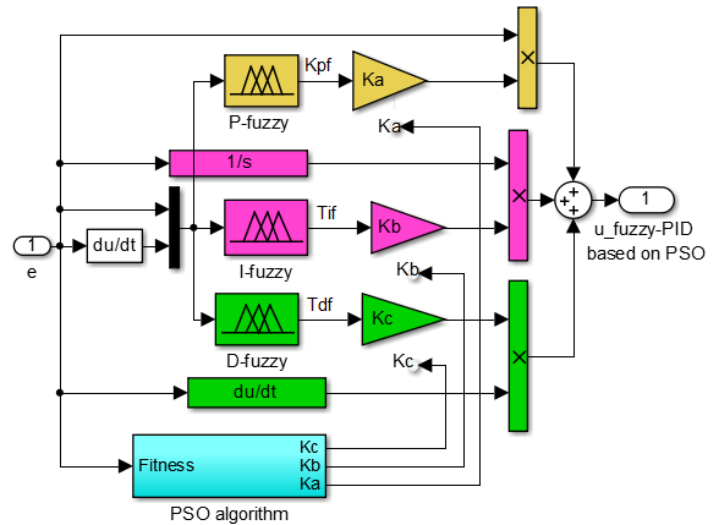


Figure 3. Structure diagram of a PSO-based Fuzzy-PID

3.2. Design of the fuzzy logic calculated blocks

Each fuzzy logic calculated block has two inputs and one output: first input as temperature error ($e=T_d^p-T^p$), in which T_d^p -setpoint temperature, T^p - measured temperature, and second input as temperature error derivation (de/dt), the output corresponding to P-fuzzy, I-fuzzy, D-fuzzy are the K_{pf} , T_{If} , T_{Df} .

The P, I, D - fuzzy logic calculated blocks, corresponding to K_{pf} , T_{If} , T_{Df} , are presented in Fig.4.

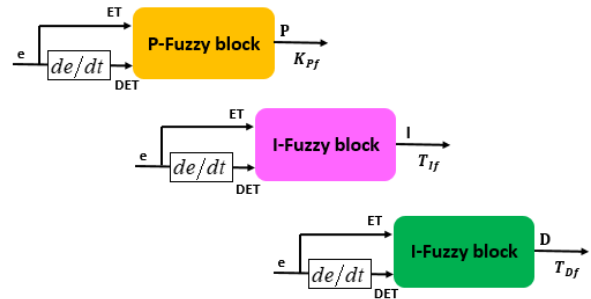


Figure 4. Structure diagram of the fuzzy calculated blocks

Two input fuzzy variables of each P-fuzzy, I-fuzzy, D-fuzzy are ET and DET, corresponding to e and de/dt . The output fuzzy variables of the P-fuzzy, I-fuzzy, D-fuzzy blocks are P, I, D, corresponding to parameters K_{pf} , T_{If} , T_{Df} .

The physical value range of the input variables and the output variables are as follows: $ET \in [-20.0, 20.0]$, $DET \in [-2.0, 2.0]$, $P \in [0, 20]$, $I \in [0, 0.05]$, $D \in [0, 2000]$. For D-fuzzy calculated block, $ET \in [-100.0, 100.0]$, $DET \in [-10.0, 10.0]$

In these fuzzy blocks, membership functions are selected triangles, and setting the input variables by 5 fuzzy sets: $ET = \{HQ, HD, HI, HV, HL\}$; $DET = \{TA, TZ, TI, TV, TL\}$, and output variables are equal to 5 fuzzy sets, corresponding to the language variables P, I, D we have $P = \{VD, VG, VI, VV, VL\}$; $I = \{VD, VG, VI, VV, VL\}$; $D = \{VD, VG, VI, VV, VL\}$.

With the number of fuzzy sets of five inputs and two fuzzy inputs for each fuzzy calculated block K_{pf} , T_{If} , T_{Df} , so we have the total of $5 \times 5 = 25$ fuzzy rules for each fuzzy calculated block. Based on the resistance furnace specification and the principle of

PID parameters adjustment, the authors built the fuzzy rules table for the fuzzy block of calculated K_{P_f} , T_{I_f} , T_{D_f} , as Table 2.

Table 2. Fuzzy rules of the P,I,D-fuzzy calculated blocks

P I D	ET					
	HQ	HD	HI	HV	HL	
DET	TA	VD	VD	VD	VD	VD
	TZ	VD	VG	VD	VD	VD
	TI	VD	VD	VI	VI	VV
	TV	VD	VD	VI	VV	VL
	TL	VD	VD	VV	VL	VL

These P-fuzzy, I-fuzzy, D-fuzzy calculated blocks in this study use Max-Min inferential law and defuzzification according to the centroid point method, the output clear values of the P-fuzzy, I-fuzzy, D-fuzzy calculated blocks, corresponding to K_{P_f} , T_{I_f} , T_{D_f} , are defined.

3.3. Building the PSO algorithm

The particle swarm optimization is one of the best effective optimization techniques. Firstly, all particles are assigned initial position and velocity values. Based on optimal value of fitness function, new position and velocity of the particles are updated. There are two fitness value are required in the update process of particles, that are p_{best} (personal best) and g_{best} (global best). The p_{best} value is trace in every iteration of particles and g_{best} value is computed among the best solution in p_{best} value [19]. In the PSO algorithm, the velocity value and position values of particles are continuously updated by following formulas.

$$v_{i,d}^{j+1} = \alpha v_{i,d}^j + C_1 r_1 (p_{best} - x_{i,d}) + C_2 r_2 (g_{best} - x_{i,d}) \quad (8)$$

$$x_{i,d}^{j+1} = x_{i,d}^j + v_{i,d}^{j+1} \quad (9)$$

where $v_{i,d}^{j+1}$ - new updated velocity of i -particle in $(j+1)$ iteration. $v_{i,d}^j$ -the velocity of i -particle in j -iteration; $x_{i,d}^{j+1}$ - new updated position of i -particle in $(j+1)$ iteration. $x_{i,d}^j$ -last position of i -particle in j -iteration; C_1 and C_2 are the acceleration factors for updating the particle velocity; r_1 and r_2 are initially random positive values which should be less than one; α is the assigned inertial weight coefficient to maintain the particles last velocity while the accelerations factors drive the flow of particles towards optimum solutions.

It is easy to see that the particles position is mainly depended upon the update velocity, and this update velocity is handled by the acceleration factors and assigned inertial weight coefficient. During the update process, the velocity and position of particles must satisfy the constraints, as shown below.

$$V_{d_min} \leq v_{i,d}^{j+1} \leq V_{d_max} \quad (10)$$

$$X_{d_min} \leq x_{i,d}^{j+1} \leq X_{d_max} \quad (11)$$

In this study, PSO algorithm is used to determine three the scaling weights of the Fuzzy-PID controller, i.e. K_a , K_b , K_c . And

the fitness function can be selected based on the combination of three goals, i.e. settling time – T_{STL} , quality index of integral of the absolute magnitude of the error - E_{IAE} , and peak overshoot - M_{POT} .

$$J(\theta) = \gamma_1 E_{IAE} + \gamma_2 M_{POT} + \gamma_3 T_{STL} \rightarrow \min \quad (12)$$

$$E_{IAE} = \int_0^T |e(t)| dt \quad (13)$$

where $\theta = [K_a, K_b, K_c]$ are parameters to optimized. T denotes the simulation time. $\gamma_1, \gamma_2, \gamma_3$ are the assigned weights, here $\gamma_1 = 1, \gamma_2 = 1, \gamma_3 = 0.5$.

The search boundary for the scaling weights of the Fuzzy-PID controller is assigned as follows:

$$0 \leq K_a \leq 30; 0 \leq K_b \leq 0.1; 0 \leq K_c \leq 10 \quad (14)$$

The PSO algorithm continuously varies the values of Fuzzy-PID controller parameters, following equation (6), until the objective function J is minimized to J_{min} . The implementation of PSO algorithm involves the following steps:

Step 1. Initialize size of particles $N=10$, dimension search-space $D=3$ corresponding to K_a, K_b, K_c , maximum iteration $M=30$, inertial weight $\alpha=0.5$, acceleration factors $C_1=2, C_2=2$,

Step2. Initialize two velocity and position values vectors, and then compute the initial fitness p_{best} values vector for each particle.

Step 3. Minimize g_{best} value from the p_{best} values vector, following the formulas (12), (13), (14).

Step 4. Update new velocity and position values into the velocity and position vectors of particles in step 3, following the formulas (8), (9).

Step 5. Compute the values of fitness p_{best} and then g_{best} for new update position if these fitness values are improved the replaced last values with these new values (less than the previous fitness values).

Step 6. Repeat the step 3, 4, 5 until the last iteration or the desired value of the fitness function is satisfied.

4. Simulating the furnace temperature control system using the proposed PSO-based Fuzzy-PID controller

4.1. Modeling fuzzy-PID controller on Matlab

Based on Mamdani fuzzy method, the proposed fuzzy-PID in this study, was built on Matlab as show in Figure 5.

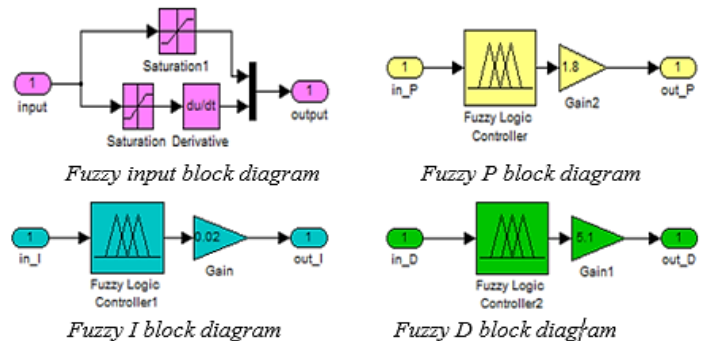


Figure 5. Modeling the new-type fuzzy-PID on Matlab

The triangle membership functions were selected. The fuzzy rules were built based on Table 2. The fuzzy inference mechanism was chosen Max-Min method and the defuzzification according to the centroid point method were adopted. The relational surface of the P-fuzzy, I-fuzzy, D-fuzzy blocks are described in Figure 6.

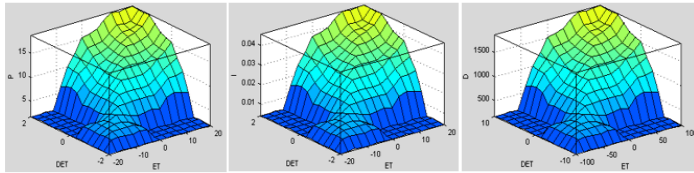


Figure 6. The relational surface of the P, I, D-fuzzy blocks

The simulation diagram of the heater temperature control system is built on Matlab, as shown Fig.7. We conducted the quality assessment of the control system with three designed controllers above: PID-ZN1, PID-CHR and new-type fuzzy-PID.

4.2. Determining the optimal parameters of the fuzzy-PID by the PSO algorithm

The initial scaling weights of the Fuzzy-PID controller, i.e. K_a, K_b, K_c , in simulation are selected randomly based on the initial parameters of PID, following the formular (5), so as to improve the quality of the furnace temperature control system. Here, the initial values K_a, K_b, K_c are as follows $K_a=10, K_b=0.02, K_c=5.5$

When applying the proposed PSO algorithm in part 3.3 to determine optimal parameters of fuzzy-PID controller, at this point, the optimal values K_a, K_b, K_c are as follows $K_a=22.2, K_b=0.021, K_c=5.1$

4.3. Simulation results on Matlab

The simulation scenarios are carried out with 3 controllers: PID controller with the parameters determined by Ziegler-Nichols 1, denoted PID-ZN1; Fuzzy-PID controller with the unoptimized parameters, denoted FPID; and Fuzzy-PID controller with the optimal parameters, denoted PSO-FPID.

When there is not disturbance, the response of the system corresponding to three controllers as shown Figure 7.

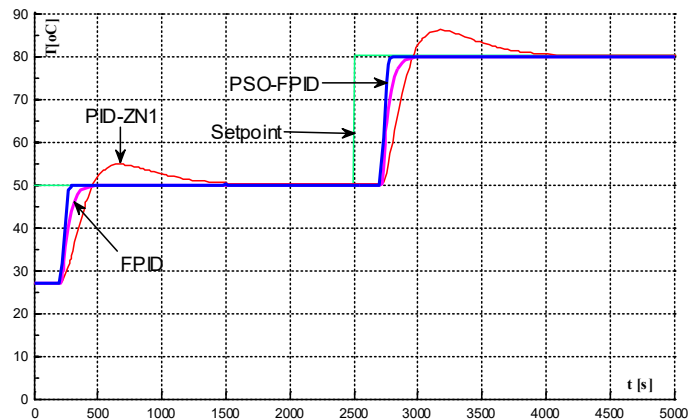


Figure 7. Response of the control system without disturbance

The overshoot of the control system with the proposed fuzzy-PID controller (FPID and PSO-FPID) is better than conventional PID controller (PID-ZN1). Specifically, when

applying the PSO-based fuzzy-PID and the fuzzy-PID controller then the control system is not overshoot, meanwhile with PID-ZN1 controller then the control system is an overshoot of about 10%. The settling time of the system (included the time delay, 225 seconds) at the desired temperature value, $T_d=50^{\circ}\text{C}$, for the proposed PSO-based fuzzy-PID controller is so rapid, about 278 seconds, and for the proposed fuzzy-PID controller is about 401 seconds. The rising time of the system with the proposed fuzzy-PID controller is faster than PID controller. The steady-state error of the system with the proposed fuzzy-PID controller is almost eliminated.

Table 3. The system quality indexes without disturbance

Quality indexes	PID	Fuzzy-PID	PSO-based fuzzy-PID
Rising time (sec)	117.6	71.8	38.8
Settling time (sec)	1243.6	401.5	278.4
Overshoot (%)	<10	0	0
Steady-state error (%)	<1	~0	~0

Therefore, the two controllers proposed in this study (FPID, PSO-FPID) can bring improved quality of the furnace temperature control system.

When the sine disturbance impacts on the system: In fact, the operation environment is not ideal. The system operation is always affected by external disturbance, so that it makes to change the actual signals. Assume that the sine disturbance has the mixed frequency in range of 0.01-50Hz, and maximum amplitude of 50. At this time, the response of the furnace temperature control system corresponding to three controllers PID-ZN1, FPID, PSO-FPID, are shown in Figure 8.

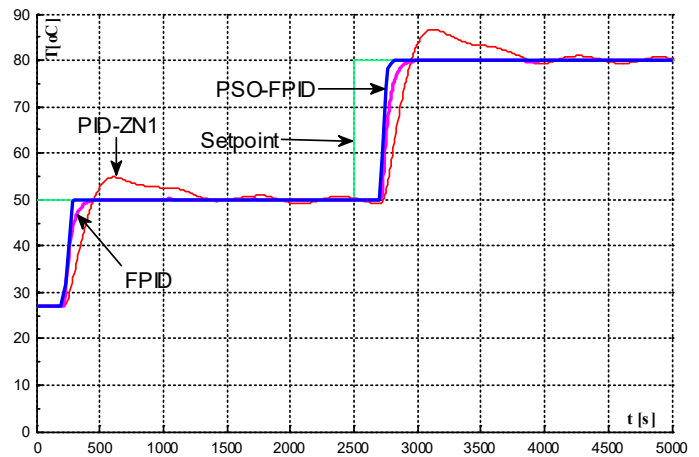


Figure 8. Response of the control system with sine disturbance

Figure 8 indicated the traditional PID controller (PID-ZN1) causes the system to oscillate around the set-point temperature value. This oscillated amplitude depends on the disturbance amplitude. Meanwhile, applying the proposed fuzzy-PID or PSO-based fuzzy-PID, the system quality is quite good, ensuring the system stability.

If the disturbance amplitude is so large, the furnace temperature control system may be unstable with the traditional

PID controller, but still ensure stability of the system with the proposed PSO-based fuzzy-PID.

Table 4. The system quality indexes with disturbance

Quality indexes	PID	Fuzzy-PID	PSO-based fuzzy-PID
Rising time (sec)	110.5	65.5	36.4
Settling time (sec)	1255	380.1	265.1
Overshoot (%)	<15	<1	~0
Steady-state error (%)	-2 ÷ 2	-0.5 ÷ 0.5	~0
Oscillated response	Yes	Small	No

5. Design of the temperature control board

5.1. Block diagram of temperature control board

In this section, the authors present the design of the PIC16F control board for controlling temperature of the resistance furnace. This temperature control board is designed by using PIC16F877A microcontroller [20], as shown in Figure 9.

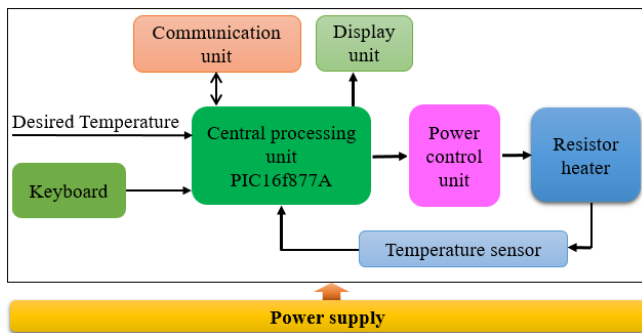


Figure 9. Block diagram of the temperature control board

The power supply block is responsible for supplying power to the entire system. The control object here is the resistance wire, wrapped around the copper pipe, causing the heat in the copper pipe to rise temperature in the furnace, it is controlled via the voltage signal of the power control unit. The temperature in the heater is measured by a temperature sensor. The display unit is responsible for displaying the actual temperature and set-point temperature. The keyboard is responsible for entering the set-point temperature. The central processing unit converts from analogue signal to digital values and performs the functions of fuzzification block, fuzzy-rules law table, defuzzification block and generating signals to command the power unit, thereby stabilizing the furnace temperature according to the set values. The power unit is responsible for controlling the AC power supplied to the resistance wire. The communication unit, using USB/RS232 port, is responsible for communicating between the board and the computer.

5.2. Layout circuit of the temperature control board

The principle schematic circuit diagram of the temperature control board includes the schematic circuit diagram of the PIC16f central processing unit, the LCD display block, the power control unit, the USB/RS232 communication block and the AC/DC converter block. The detail schematic circuit of the temperature control board is presented in [18].

Based on the principle schematic circuit, the layout circuit diagram is given in Figure 10 and the realistic layout circuit of the control board as shown in Figure 11.

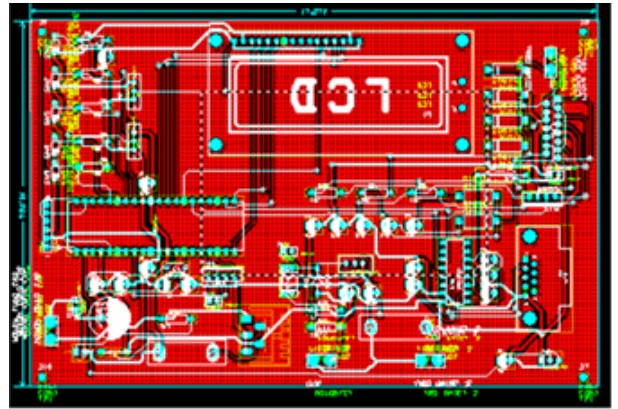


Figure 10. Layout circuit of the temperature control board



Figure 11. Realistic layout circuit of the control board

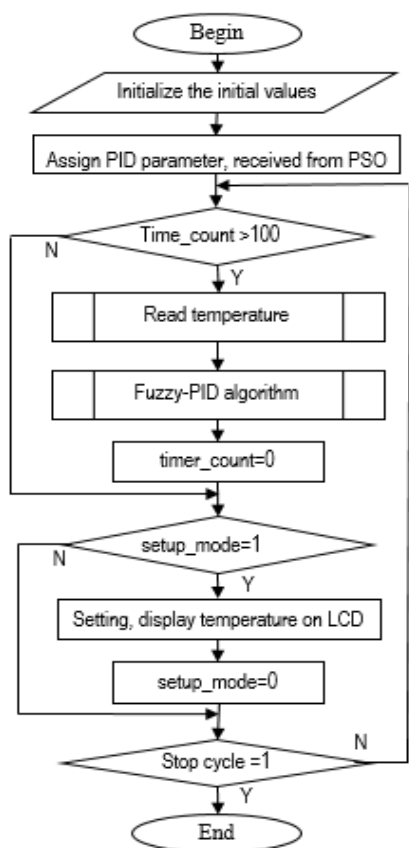


Figure 12. The main program flowchart for controlling the furnace temperature using PSO-based fuzzy-PID

5.3. Temperature control algorithm on PIC16f chip

The furnace temperature control algorithm is installed on the PIC16f microprocessor as described in Fig.12.

The detail of algorithm flowcharts corresponding to each subprogram, are presented in Fig.13.

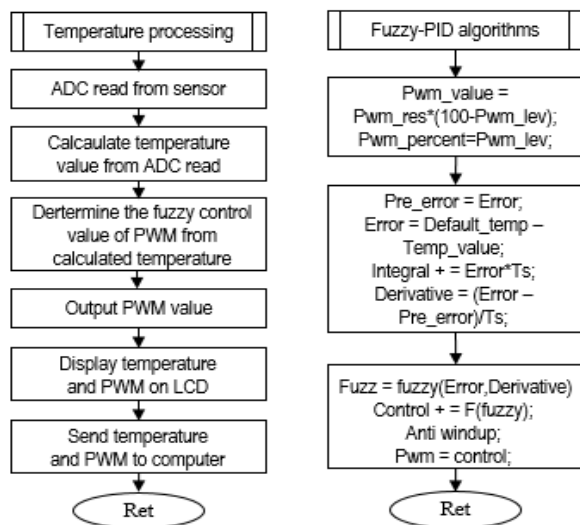


Figure 13. Algorithm flowchart of subprograms

5.4. Experimental control for the heater in the laboratory

Experiment of the resistance furnace temperature control system, using PIC16F control board with the proposed PSO-based fuzzy-PID algorithm, is presented in Figure 1.

The monitoring control interface based on Visual C# is developed on the computer to obtain the actual temperature value of the resistance furnace. From this interface, the set-point temperature value and the parameters of the PSO-based fuzzy-PID controller also can be assigned remotely and easily.

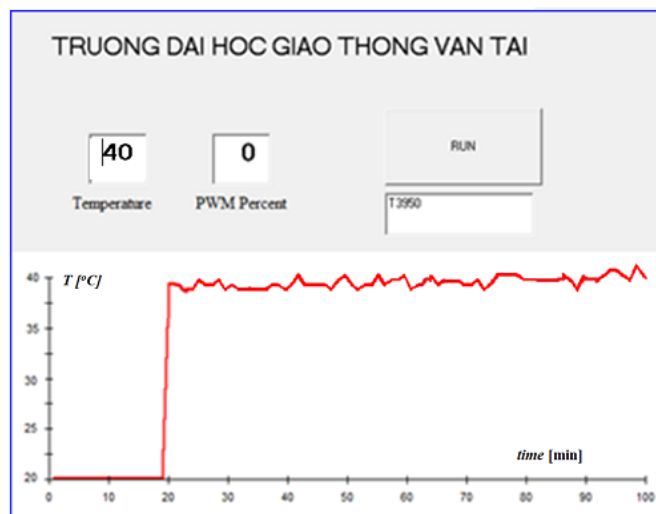


Figure 14. Real temperature response at setpoint 40°C

At the beginning of the experiment, the control board with installed PSO-based fuzzy-PID algorithm, will create the 100% PWM output signal to supplying the suitable AC power values to the resistance wire, after the delay time, the temperature in the furnace changes from environmental temperature $T_0=20^{\circ}C$ up to $39^{\circ}C$, and then reduce PWM signal to maintain the actual temperature value at setpoint value of $T_d=40^{\circ}C$. When the furnace temperature exceeds $40^{\circ}C$, the PIC16f microcontroller interrupts the PWM signal so that the temperature will reduce to the setpoint temperature value.

The experimental results in laboratory showed that the proposed PSO-based fuzzy-PID controller brings good control quality for the resistance furnace temperature control system. The proposed PSO-based fuzzy-PID controller with the optimal parameters usually gives better quality the proposed fuzzy-PID controller, because the parameters of the fuzzy-PID controller was not optimized. However, these two proposed controllers offer better quality than the traditional PID controller, especially when the disturbance impacts on the system.

6. Conclusion

This paper presented a novel control approach applying the PSO algorithm to tune the scaling weights coefficients of the fuzzy-PID controller which is combination of the advantage of the traditional PID control and fuzzy logic control. The fuzzy-PID controller is designed based on the typical fuzzy rules, and combined with the PID control, but the fuzzy-PID three scaling weights are tuned by the PSO algorithm. And then the simulation model of the resistance furnace temperature control system

without/with disturbance using the proposed controller was built on Matlab. The simulation results have confirmed the effectiveness of the two proposed controllers for the resistance furnace temperature control system, i.e. fuzzy-PID controller and PSO-based fuzzy-PID controller. Then the proposed PSO-based fuzzy-PID algorithm was implemented on the temperature control board which made of PIC16f microprocessor. Experimental results showed that proposed PSO-based fuzzy-PID algorithm provided better control quality than the traditional PID algorithm. In addition, the proposed PSO-based fuzzy-PID controller is consider to be an efficient control strategy applying for a class of nonlinear and uncertain complicated control objects.

From this research, the future work focusing on optimization structure of fuzzy logic calculated blocks, and apply adaptive control to tuning online/adaptive parameters of the fuzzy-PID controller. The future work also continues to develop the temperature control board using PIC16f in order to complete the equipment, and put the equipment into the actual application to control the temperature of materials in industry and transportation.

Conflict of interest

The authors declare no conflict of interest.

References

- [1] Aidan O'Dwyer, Handbook of PI and PID Controller Tuning Rules, Imperial College Press, 2003.
- [2] Yu Feng Zhang, Ming Li and Jing Min Dai, PID Heating and Temperature Control Method Based On Dynamic Assignment, 5th International Conf. on Mechatronics, Materials, Chemistry and Computer Eng., 2017. <https://doi.org/10.2991/icmmce-17.2017.208>
- [3] V.Kabila, Glan Devadhas, Comparative Analysis of PID and Fuzzy PID Controller Performance for Continuous Stirred Tank Heater, Indian Journal of Science and Technology, Volume 8, Issue 23, Pages: 1-9, 2015, <https://doi.org/10.17485/ijst/2015/v8i23/85351>
- [4] L.X. Wang, A course in Fuzzy Systems and Control, Prentice Hall International, Inc, 2002.
- [5] Z.R. Radakovica et al., Application of temperature fuzzy controller in an indirect resistance furnace, Applied Energy, Volume 73, Issue 2, 2002. [https://doi.org/10.1016/S0306-2619\(02\)00077-6](https://doi.org/10.1016/S0306-2619(02)00077-6)
- [6] Moh Syahrir Bin Ribuan, Development Of Fuzzy Logic: Temperature Controller Implementation - Microcontroller Based, University Teknikal Malaysia Melaka, 2009.
- [7] P. Singhala, D. N. Shah, B. Patel, Temperature Control using Fuzzy Logic, International Journal of Instrumentation and Control Systems, Vol.4, No.1, 2014, <https://doi.org/10.5121/ijics.2014.4101>
- [8] Nguyen Chi Ngon, Fuzzy PI controller: From design to application, Science magazine, Can Tho University, 2011.
- [9] Shi Dequan, et al., Application of Expert Fuzzy PID Method for Temperature Control of Heating Furnace, 2012 International Workshop on Information and Electronics Engineering, Procedia Engineering, Volume 29, 2012, <https://doi.org/10.1016/j.proeng.2011.12.703>
- [10] Sneha S Patole1, Shailendra K Mittal, Fuzzy PID Controller Design for Heating Control System, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 6, Issue 6, 2007, <https://doi.org/10.15662/IJAREEIE.2017.0606140>
- [11] Qingjie Yang, Guohou Li, Xusheng Kang, Application of fuzzy PID control in the heating system, 2008 Chinese Control and Decision Conference, 2008, <https://doi.org/10.1109/CCDC.2008.4597814>
- [12] Jiangjiang Wang, Dawei An, Chengzhi Lou, Application of Fuzzy-PID Controller in Heating Ventilating and Air-Conditioning System, 2006 International Conference on Mechatronics and Automation, June 2006, <https://doi.org/10.1109/ICMA.2006.257656>
- [13] L. D. Vijay Anand, Lakshmi Asok., Design of Fuzzy PID with Expert Control for a Temperature Process, International Journal of Engineering Research & Technology, Volume 03, Issue 02, 2014.
- [14] Guofang G., Zhengzhong L., Design of Heating Furnace Temperature Control System Based on Fuzzy-PID Controller, Information Engineering and Applications. Lecture Notes in Electrical Engineering, Volume 154. Springer, https://doi.org/10.1007/978-1-4471-2386-6_196
- [15] Trinh Luong Mien, Design of Fuzzy-PID Decoupling Controller for the Temperature and Humidity Process in HVAC System, International journal of engineering research & technology, Vol. 05, Issue 01, January-2016. <https://doi.org/10.17577/IJERTV5IS010473>
- [16] Y.Zhang, et al., Application and Research of Fuzzy PID Control in Resistance Furnace Temperature Control System, IOP Conf. Series: Earth and Environmental Science, Vol.186, Issue 5, 2018. <https://doi.org/10.1088/1755-1315/186/5/012051>
- [17] Liu Jing Yan, The Fuzzy PID Control of Resistance Furnace Temperature System Based on Genetic Algorithm, Applied Mechanics and Materials Vol. 273, 2013, <https://doi.org/10.4028/www.scientific.net/AMM.273.678>
- [18] Trinh Luong Mien, Vo Van An, Lecture notes: Design of the temperature controller based on fuzzy logic using PIC16f microprocessor, University of transport and communications, 2018.
- [19] A.H. Al-Mter and S. Lu, A Particle Swarm Optimization Algorithm Based On Uniform Design, International Journal of Data Mining & Knowledge Management Process, Vol. 6, No. 2, 2016. <https://doi.org/10.5121/ijdkp.2016.6203>.
- [20] Microchip, PIC16F/LF1825/1829 Data Sheet DS41440A, Microchip Technology Inc., 2010.

Warehouse Relocation of a Company in the Automotive Industry Using P-median

Zarate-Zapata, Aldo Cesar *, Garzón-Garnica, Eduardo Arturo, Cante-Mota, Román, Olmos-Álvarez, Fernando, Martínez-Flores, José Luis, Sánchez-Partida, Diana

Logistics and Supply Chain Management Department, Universidad Popular Autónoma del Estado de Puebla, 72410, México

ARTICLE INFO

Article history:

Received: 02 April, 2020

Accepted: 03 June, 2020

Online: 26 June, 2020

Keywords:

Logistics

P-median

Risk management

ABSTRACT

To have enough information on time can be helpful when companies try to reduce costs and operate more efficiently. An international company that supplies parts for the automotive industry is currently testing its new facilities in Mexico. The relocation of the raw materials and finished goods warehouses were tested using a P-Median model. The operating costs and risk factors were included in the model to provide a better solution and improve the operation of the warehouses and production lines. The research results compared different scenarios and indicated that the proposed better location isolates the forklift routes, mainly for finished products, and minimizes the cost of moving both raw materials and finished products to and from warehouses.

1. Introduction

The logistics activities have always been important for companies, because they represent various concepts, principles, and methods of different areas, like marketing, production, accounting, warehouse management, among others. With logistics, the chain value for the customers is increased, as well as for suppliers and stakeholders. An efficient logistics management could represent a decrease in costs, and continuous improvement in production [1].

In a warehouse, typically products arrive packaged and the warehouse personnel reorganizes, classifies, sorts and repackages such products. This function is of vital importance in any warehouse because it is needed to break down large chunks of products and redistribute them in small quantities. So, the supply chain is the sequence of the different processes in the organization within which the products move from their origin toward the customers [2].

In this paper, we propose a solution to a warehouse problem of a company in the automotive industry, using a p-median model to minimize the travel distance from the warehouses to all of the production stations.

2. Research objectives and research methodology

2.1. Research Objectives

This document aims to implement the p-median problem modeling with Lingo® software, which is a tool designed to solve

linear, nonlinear, integer, stochastic, and optimization programming models as well as concentrate and model exact algorithms. It is done to explain the company's problems and decrease risk in the industrial process to obtain an exact solution that does not require model validation.

2.2. Research Methodology

The research methodology adopted for the research work has been represented as a theoretical concept. The application of the mathematical model is based on the p-median problem. Such a problem has been implemented in several documents, and by different authors, as cited in the literature review. Such literature review is based on studies carried out with the application of exact algorithms using the p-median problem as the basis.

The following criteria were used to select the literature:

- Literature published on risk industries, supply chain, and cases of study using the p-median problem.
- Literature published from 1975 to 2017.
- Articles published in refereed scholarly journals, working papers, and thesis.
- Journals explain p-median problem algorithms in their editorial scope
- Keywords used in the article: risk, logistics, p-median, risk management

3. Literature Review

In [3] it is mentioned that the p-median problem is useful to model many real-world situations such as the location of public or

* Zarate-Zapata, Aldo Cesar, aldocesar.zarate@upaep.edu.mx

industrial facilities, warehouses, and others. The p-median problem differs from the Uncapacitated Facility Location Problem in two respects:

1. There are no costs for opening facilities
2. There is an upper bound on the number of facilities that should be opened.

It models the problem of finding a minimum cost clustering and belongs to the class of NP-hard problems.

In [4], a set L of m facilities (or location points), a set U of n users (or customers or demand points), and a $n \times m$ matrix D with the distances traveled (or costs incurred) d_{ij} for satisfying the demand of the user located at i from the facility located at j , for all $j \in L$ and $i \in U$ are considered. The objective is to minimize the sum of these distances or transportation costs.

$$\sum_{(i \in U) m} j \in J dij$$

Where $J \in L$ and $|J| = p$. The p-median can be defined as a purely mathematical problem: given an $n \times m$ matrix D , select p , columns, or D so that the sum of minimum coefficients in each line within these columns is the smallest possible.

The p-median can also be interpreted in terms of cluster analysis; locations of users are then replaced by points in an m-dimensional space [5].

Besides this combinatorial formulation, the p-median is also an integer programming problem. Let us define two sets of decision variables: (i) $y_j = 1$ if a facility is opened in $j \in L$, and 0, otherwise; (ii) $x_{ij} = 1$, if customer i is served from a facility located in $j \in L$, and 0, otherwise. Then the integer programming formulation is as follows:

$$\text{minimize} = \sum_{i=1}^n * \sum_{j=1}^m d_{ij} x_{ij} \tag{1}$$

Subject to

$$\sum_{j=1}^m x_{ij} = 1, \forall i \tag{a}$$

$$x_{ij} \leq y_j, \forall i, j \tag{b}$$

$$\sum_{j=1}^m y_j = p, \tag{c}$$

$$x_{ij}, y_j \in \{0, 1\} \tag{d}$$

The objective function of the p-median model seeks to minimize the sum of the distances d_{ij} for each client. Constraint (a) is the one that assures that each client is assigned to a facility. Constraint (b) makes sure that each client is attended by precisely one server. Constraint (c) indicates that p facilities should be assigned. Equation (d) indicates that the total number of open facilities is set to p by constraint.

The research in [6] shows that the academic research on safety in logistics has mainly focused on transportation, and in particular, on safety concerning motor carriers [7]. Among others, studies research characteristics of professional drivers (e.g., personality, health, attitude), stress factors they face (time pressure, fatigue, stress) and how these relate to safety behavior and/or accidents [8], [9], [10], [11], [12]. Recently, warehouse safety has started to gain attention. For instance, In [13] it is analyzed which factors impact warehouse safety.

In addition, in [14] it is mentioned that the risk in an environment is very significant [15], and the companies that measure it are characterized by several risk factors. The term *supply chain vulnerability* [16] has been used to describe the dependence and risks that exist among organizations as they rise to the challenge of better, faster, cheaper. Two aspects connected to the risk assessment in supply chains can be found:

- 1) Risk exists inside the company and at the network level.
- 2) Risk evaluation is subjective because each people perceives differently a risk and the nature of the upstream and downstream relationships of such risk.

According to [17], the success of risk management will be based on the understanding of various categories of risks, that affect projects in an organization. Achieve an analysis of their mutual dependence, categorize them according to their importance, and develop strategies for risk management based on prioritization.

The physical risk can be further measured in terms of the potential damages in physical spaces such as warehouses and production lines. These measures, called *physical risk indexes*, are derived from inspection procedures by external experts. The risk measures are monitored from both the logistics and supply chain perspective and the top-management control perspective [18]. Furthermore, in the supply chain the objectives are considered in terms of the organization's goals concerning the final customer. Thus, the risk evaluation, in terms of weights and importance of indicators, should be guided by an awareness of the nature and importance of the market objectives [14].

4. Problem Description

An International Company of the Automotive Industry has a facility in Mexico; its primary purpose is the manufacturing of automotive cooling components, like radiators and heat exchangers. Currently, this company has only one production line in its planning stage. Their forecasted demand is very significant, and it is expected to grow further. Therefore, they will require a total of four production lines. The design of such lines is already available, as well as the preliminary data of the required and projected production levels.

The raw materials warehouse, as well as the finished goods warehouse, are currently operating on a small scale, working only at floor level, without racks. As the plant is not yet operating in full production mode, it is still possible to relocate the materials and finished goods warehouses, to a certain extent.

The cost of relocating the warehouse; at this stage, is minimal because production lines are not yet established. Still, not all of the building's spaces can be used for warehouses, as power, gas, water, and overhead lines for production lines are already established, close to the places where the machines would be installed.

The use of the current layout, or relocation to a new position within the plant, should consider the operating cost of the warehouses, but also the risk that comes with the location of those warehouses. If materials must be carried through the plant using a forklift, the cost is higher than if a manual cart can be used. Nevertheless, when using a cart, if the warehouse operators must cross the forklift's pathway, there is a higher risk of an accident

happening than if the route is free of such crossings. The risk of damaging finished goods with the forklift, or by crossing its roads, is also noted. The relationship between cost and risk should then be considered for the solution.

The company leases its forklifts, so the cost of ownership does not change whether they are used or not, except for the cost of fuel. If the distance traveled by the forklift can be kept small enough to meet the need for full production with the fewest number of forklifts, some savings can be made by that means. A shorter distance traveled by the forklift could mean that all production lines are serviced by just one forklift, saving on the fixed cost of the rent.

5. Methods and Procedures

The problem for this company was modeled as a p-median problem, using a distance matrix calculated from the warehouses to each production station.

The distances were computed as follows:

A map of the whole plant was used as a base, and divided, on the scale, to a square grid of 1 meter by side. The location of each working station, and those of the existing warehouses were marked in the map

The possible alternate locations of the warehouses were selected simply by finding the spaces within the plant that were currently not in use, and where the warehouses would fit. Such a method resulted in three possible locations for the warehouses, with little or no movement of the production lines from the current plant layout.

The distance from the warehouse entrance, in each of the proposed locations, to each one of the working stations, whether current or projected, was computed by counting the meters in the map, as if the cart or forklift would have traversed the route from the warehouse to the working station. The same process was done for the reception of raw materials and the finished goods.

If one of the warehouses proposed locations required the relocation of one or more production lines, each one of the distances was computed again, considering the new location of the production lines and warehouses.

As stated, three different scenarios were found possible, and those were evaluated to obtain a comparison against the current situation. The first scenario is the current situation. The second one involves moving the raw materials warehouse to a top-right position, in a space currently unused. The third scenario considers moving the raw materials warehouse to a down-center position, very close to the place where the production lines start, but this last scenario involves the location of the two middle production lines, a few meters offset to the top. Still, those production lines were not operating yet, so a small relocation was still possible.

The considerations to the problem at hand are as follows:

1. The cost of relocating the warehouses is minimal, because there is no hardware, like racks, in the warehouse at this point.
2. Although production lines 2-4 are not yet set up, some of the needed infrastructures are already prepared. It prevents such

lines from being set anywhere. They must remain close to the originally intended position.

3. Forklift streets, walkways, isles, etc., can be relocated easily. However, there is no unlimited space.

Figure 1 shows the current layout of the plant, including only one production line, which is the current state of the company. However, they foresee the need to upscale their production levels to a full mode, needing four lines, in at most two years from now.

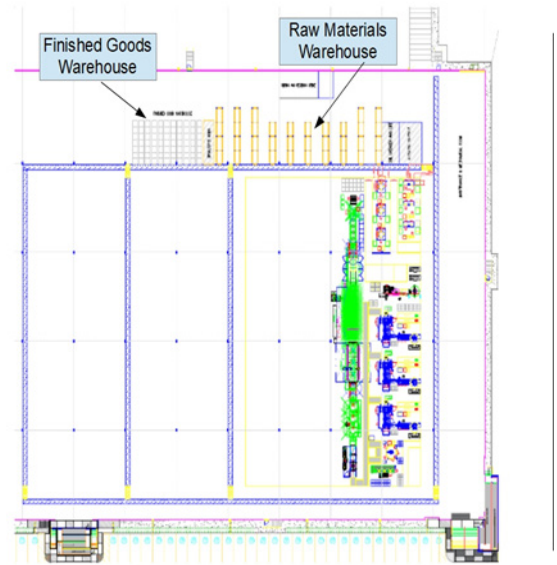


Figure 2 shows the possible layout of the plant when all four lines are in place, considering the current location of both the raw materials and finished goods warehouses. Two possible alternatives to the warehouse location have been found in the current plant layout.

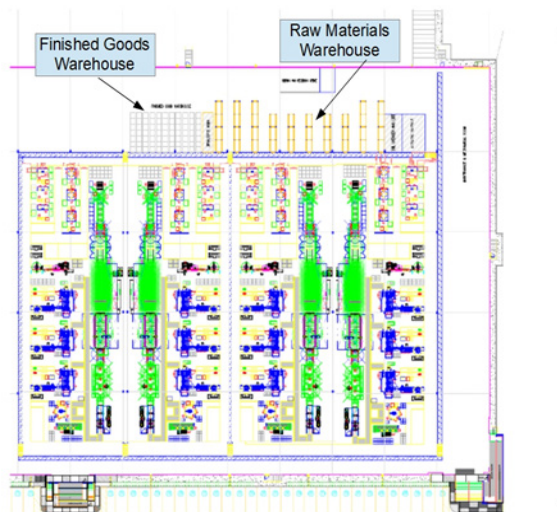


Figure 2. Current layout with no change in the warehouse locations.

Figure 3 shows the possible layout with all four lines in place, but with the warehouses relocated to a place currently unused in the plant, without the need to relocate any of the production lines.

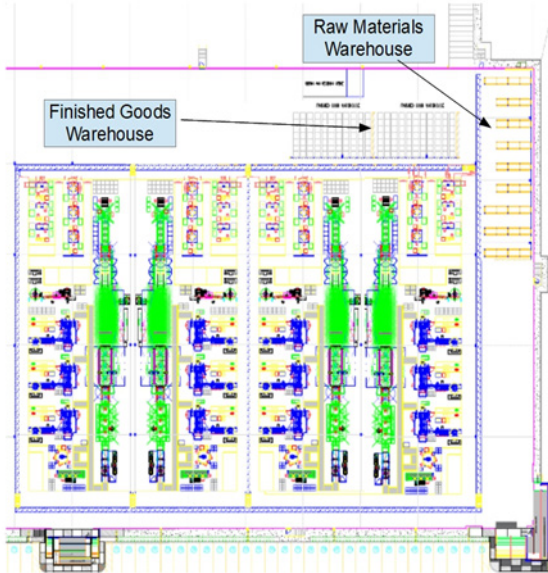


Figure 3. Proposed layout with warehouses in the top right corner.

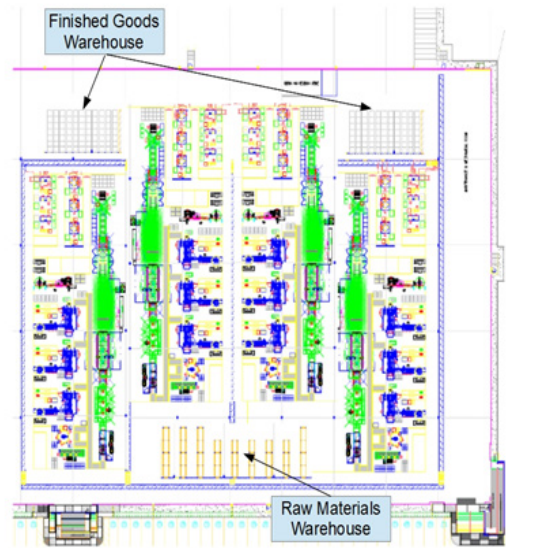


Figure 4. Proposed layout with raw materials warehouse in the bottom center.

6. Data Acquisition and Preparation

The primary way to compute the cost was to obtain the distance from the warehouses to each working station. More distance traveled means the operator would take more time to carry the items, whether by forklift or by manual cart. It was considered that for one production run in each production line, one shipment of raw materials would be required for each working station.

Table 1: Extract of the data table.

Equipment	Raw Material 1			Finished Goods 1			Raw Material 2			Finished Goods 2			Raw Material 3			Finished Goods 3			Arrival Factor
	Distance	Risk	Forklift	Distance	Risk	Forklift	Distance	Risk	Forklift	Distance	Risk	Forklift	Distance	Risk	Forklift	Distance	Risk	Forklift	
0-Raw Material Receive	17	0	1	0	0	0	47	0	1	0	0	0	121	0	1	0	0	0	1
1-Core Builder 1	62	1	1	0	0	0	34	1	1	0	0	0	70	1	1	0	0	0	0.033
1-Core Builder 2	54	1	1	0	0	0	26	1	1	0	0	0	78	1	1	0	0	0	0.033
1-Core Builder 3	46	1	1	0	0	0	18	1	1	0	0	0	86	1	1	0	0	0	0.033
1-Tube Mill	38	1	1	0	0	0	10	1	1	0	0	0	94	1	1	0	0	0	0.125
1-Laser Welding	78	1	0	0	0	0	51	1	1	0	0	0	65	1	0	0	0	0	0.125
1-Paint Flux	83	1	0	0	0	0	56	1	0	0	0	0	58	1	0	0	0	0	0.125
1-Furnace	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.125
1-Flame Brazing 1	14	0	0	0	0	0	21	1	0	0	0	0	129	1	0	0	0	0	0.125
1-Flame Brazing 2	18	0	0	0	0	0	16	1	0	0	0	0	135	1	0	0	0	0	0.125
1-Flame Brazing 3	23	0	0	0	0	0	12	1	0	0	0	0	141	1	0	0	0	0	0.125
1-EOL Pack 1	14	1	0	60	1	0	14	1	0	14	1	0	118	1	0	14	1	0	0.125
1-EOL Pack 2	18	1	0	64	1	0	9	1	0	18	1	0	112	1	0	18	1	0	0.125
1-EOL Pack 3	23	1	0	68	1	0	5	1	0	23	1	0	106	1	0	22	1	0	0.125

As stated, a grid was drawn over the map with the proposed layouts, with each square measuring approximately one meter. The same grid size was used for all the proposed locations. The distance within the warehouses was not computed, because currently there is not a definitive warehouse layout in place. The point of origin was set as the point within the warehouses closest to the forklift or cart roads, in the case of the current warehouse location, such point was the entrance of the warehouse. All the distances were then computed and stored in a single table. Each row in that table represents one working station of one production line. Each column represents one of the possible locations of both the raw materials and finished goods warehouses.

A column was added to each location to consider the cost if the type of material that has to be brought to the working station requires the use of a forklift, a value of 1 was added if the materials can be carried with a manual cart, or by hand, the value was 0. In the case of the risk, if the materials could be carried by hand or cart, but the route needed to cross the forklift pathways, a value of 1 was set.

Another value that had to be computed was the reception of the raw materials. This case had to also consider the cost of bringing the raw materials from the receiving docks to the warehouse. A modifier was added to transform this cost for each production round. The modifier consisted of the number of production lines that could be served with a single raw materials shipment. If an item arrives in a pallet containing 40 boxes, and three boxes would be needed for each production run, the modifier for that distance would be 3/40 boxes, or 0.075. Table 1 below shows an extract of the resulting data table, showing the risk, forklift, and arrival cost modifiers, as well as the distance computed for each scenario.

The model needed to include the risk and cost factors, apart from the distance from/to the warehouse. The model was modified to include those values as a weight. If a forklift works through all the shifts, and all of the weekdays, its cost is close to 8 times that of an employee. So, a cost factor of 8 to 1 is considered if the forklift is needed to bring raw material or finished goods to/from a working station. The risk of an accident would be considered to be 2 to 1, because, if an accident happens, the company would have to send the employee to the hospital, while paying his wages, and hiring an extra employee to cover for the recovery time.

Finally, the cost of receiving raw materials was also considered as the distance from the reception points to the raw materials warehouse. Shipments always arrive on pallets, so they must be moved with a forklift. However, as one pallet contains more than one box that could be sent to the production line, a factor was added, considering how many production orders could be fulfilled with that pallet for each working station. Two main factors were computed: For the Core Builder steps, one raw material shipment could cover 30 days' worth of orders; for the other steps, each shipment would cover eight days. So, factors of 1/30th (0.033) and 1/8th (0.125) were added to the data table.

7. Formulation

It was decided to implement the mathematical model of the p-median problem because the objective of the model is to find the midpoint for a group of clients and their location, so it was

extrapolated to a group of workers in a space delimited by the production lines.

A slight modification had to be made to include all those factors in the model. So that each distance would be modified by the risk, cost, and reception factors. The change was made locally, only by computing the cost factors (risk, forklift use, raw materials) and adding it to the main cost. In general, the model remains as a weighted p-median problem, just as the one found in the literature.

$$\text{minimize } \sum_{i=1}^n * \sum_{j=1}^m x_i * ((d_{ij} (1 + (Cf \times Fl_{ij}) + (Rf \times LC_{ij})) + Rd_i Sf_j Cf + Fd_{ij} (1 + Rf \times FLC_{ij} + Cf \times FFL_{ij}))) \quad (2)$$

Subject to

$$\sum_{i=1}^n x_i = 1 \quad (a)$$

$$\forall i \in \{1..n\} \quad x_i \in \{1,0\} \quad (b)$$

$$\forall i \in \{1..n\}, \forall j \in \{1..m\} \quad Fl_{ij}, LC_{ij}, FLC_{ij}, FFL_{ij} \in \{1,0\} \quad (c)$$

Where

n is the number of different options to choose from.

m is the number of working stations or manufacturing steps.

Cf is the cost factor of the forklift.

Fl indicates if the materials for the working station *must* be carried with a forklift.

Rf is the risk cost of crossing a forklift's lane.

LC indicates if the path to the working station must cross or traverse a forklift's lane.

d is the travel distance from the warehouse to the working station.

Rd is the distance from the reception point to the raw materials warehouse.

Sf is the number of production runs that can be fed by a single pallet.

Fd is the distance from the working station to the finished goods warehouse.

FLC indicates if a forklift's lane must be crossed with finished goods.

FFL indicates if a forklift must be used for finished goods.

x indicates if the option will be used.

Restriction (a) indicates that one option must be selected.

Restriction (b) states that x_i can only have a value of 1 or 0.

Restriction (c) assures that FL_{ij} (Must use a Forklift), LC_{ij} (Lane Crossing), FLC_{ij} (Lane Crossing for Finished Goods), and FFL_{ij} (Use of Forklift for Finished Goods) have only a value of 1 or 0.

The model was coded using Lingo® and solved. The code is the following:

1. Algorithm 1: LINGO ® Coded Model

2. 1. Model:

3. 2. Sets:

- 4. Alternativa: Alt, Rd, x;
- 5. Estaciones: WS, Sf;
- 6. Arco(Alternativa,Estaciones):Dist, Fl, LC, FDist, FFI, FLC;
- 7. EndSets

8. 3. Data:

- 9. Alt = @OLE("C:\Tmp\WHR.xls");
 - 10. Rd = @OLE("C:\Tmp\WHR.xls");
 - 11. WS = @OLE("C:\Tmp\WHR.xls");
 - 12. Sf = @OLE("C:\Tmp\WHdR.xls");
 - 13. Dist= @OLE("C:\Tmp\WHR.xls");
 - 14. Fl = @OLE("C:\Tmp\WHR.xls");
 - 15. LC = @OLE("C:\Tmp\WHR.xls");
 - 16. FDist=@OLE("C:\Tmp\WHR.xls");
 - 17. FFI = @OLE("C:\Tmp\WHR.xls");
 - 18. FLC = @OLE("C:\Tmp\WHR.xls");
 - 19. EndData
-
- 20. **min**=@Sum(Alternativa(i):@Sum(Estaciones(j):x(i)*
(Dist(i,j)*(1+(Cf *Fl(i,j)))+(Rf*LC(i,j)))+(Rd(i) * Sf(i) *
Cf) + (FDist(i,j)*(1+ (Cf*FFI(i,j)))+(Rf*FLC(i,j))))));
 - 21. Cf = 7;
 - 22. Rf = 2;
 - 23. @Sum(Alternativa(i): x(i)) = 1;
 - 24. @For(Alternativa(i): @BIN(x(i)));
 - 25. @For(Alternativa(i): @For(Estaciones(j):
 - 26. @BIN(Fl(i,j));
 - 27. @Bin(LC(i,j));
 - 28. @Bin(FFI(i,j));
 - 29. @Bin(FLC(i,j))
 - 30.);
 - 31. End

8. Results and Discussion

The model considered all the risk, cost, and incoming material cost modifiers. The solution obtained showed that the location that minimized the distance from the warehouse to each of the working stations was location number 2. An extract of the raw results obtained by solving the model is shown in table 2.

Table 2. Raw results of the model solver

Variable	Value	Reduced Cost
CF	7	0
RF	2	0
ALT(1)	1	0
ALT(2)	2	0
ALT(3)	3	0
RD(1)	17	0
RD(2)	47	0
RD(3)	121	0

X(1)	0	22328
X(2)	1	18161
X(3)	0	26156
WS(1)	1	0
WS(2)	2	0
WS(3)	3	0
WS(4)	4	0
WS(5)	5	0
WS(6)	6	0
WS(7)	7	0
WS(8)	8	0
WS(9)	9	0
WS(10)	10	0

9. Conclusions and future research

Logistics within the factory add no value to the finished goods, but wrong logistic management can become costly if accidents happen, whether to people, to the facilities, or the finished goods. The proposed location isolates the forklift pathways, mostly for finished goods, and minimizes the cost of moving both raw materials and finished goods from and to the warehouses. It is expected that, even in full production mode, the need for forklifts is kept below the line that requires more than one device. It is also expected that the manual cart movement is reduced and that the crossing of forklift lanes is also kept at a minimum. Both the company and the authors expect that both the raw materials and the finished goods warehouses can be operated by only one operator, as well as the raw materials receiving process. It would, in turn, keep the operating costs at a level similar to the present, where test runs are being performed, but when in full production mode.

Future work in this area could include more general computing of the costs associated with each physical movement. Nevertheless, the company still needs improving in many areas, like minimizing idle time, reducing cooling and heating of the furnaces to make more efficient use of energy, thus reducing costs, optimizing the use of significant items, reducing even further the need to return big items to the raw materials warehouse, and reorganizing the items within each warehouse.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] R. Ballou, *Logística administración de la cadena de suministro*. México: Pearson Educación de México, S. A. de C. V., 2004.
- [2] J. Bartholdi and S. Hackman, "Allocating space in a forward pick area of a distribution center for small parts", *IIE Transactions*, vol. 40, no. 11, pp. 1046-1053, 2008. <https://doi.org/10.1080/07408170802167662>.
- [3] P. Mirchandani and R. Francis, "Discrete location theory", *Discrete Applied Mathematics*, vol. 36, no. 1, pp. 93-94, 1992. [https://doi.org/10.1016/0166-218x\(92\)90210-2](https://doi.org/10.1016/0166-218x(92)90210-2).
- [4] N. Mladenović, A. Alkandari, J. Pei, R. Todosijević and P. Pardalos, "Less is more approach: basic variable neighborhood search for the obnoxiousp-median problem", *International Transactions in Operational Research*, vol. 27, no. 1, pp. 480-493, 2019. <https://doi.org/10.1111/itor.12646>.

- [5] P. Hansen and B. Jaumard, "Cluster analysis and mathematical programming", *Mathematical Programming*, vol. 79, no. 1-3, pp. 191-215, 1997. <https://doi.org/10.1007/bf02614317>.
- [6] N. Hofstra, B. Petkova, W. Dullaert, G. Reniers and S. de Leeuw, "Assessing and facilitating warehouse safety", *Safety Science*, vol. 105, pp. 134-148, 2018. <https://doi.org/10.1016/j.ssci.2018.02.010>.
- [7] D. Cantor, "Workplace safety in the supply chain: a review of the literature and call for research", *The International Journal of Logistics Management*, vol. 19, no. 1, pp. 65-83, 2008. <https://doi.org/10.1108/09574090810872604>.
- [8] M. Douglas and S. Swartz, "A multi-dimensional construct of commercial motor vehicle operators' attitudes toward safety regulations", *The International Journal of Logistics Management*, vol. 20, no. 2, pp. 278-293, 2009. <https://doi.org/10.1108/09574090910981341>.
- [9] M. Douglas and S. Swartz, "Career stage and truck drivers' regulatory attitudes", *The International Journal of Logistics Management*, vol. 27, no. 3, pp. 686-706, 2016. <https://doi.org/10.1108/ijlm-11-2014-0180>.
- [10] R. Grytnes, H. Shibuya, J. Dyreborg, S. Grøn and B. Cleal, "Too individualistic for safety culture? Non-traffic related work safety among heavy goods vehicle drivers", *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 40, pp. 145-155, 2016. <https://doi.org/10.1016/j.trf.2016.04.012>.
- [11] E. Kemp, S. Kopp and E. Kemp, "Six days on the road", *The International Journal of Logistics Management*, vol. 24, no. 2, pp. 210-229, 2013. <https://doi.org/10.1108/ijlm-08-2012-0080>.
- [12] J. de Vries, R. de Koster, S. Rijdsdijk and D. Roy, "Determinants of safe and productive truck driving: Empirical evidence from long-haul cargo transport", *Transportation Research Part E: Logistics and Transportation Review*, vol. 97, pp. 113-131, 2017. <https://doi.org/10.1016/j.tre.2016.11.003>.
- [13] R. de Koster, D. Stam and B. Balk, "Accidents happen: The influence of safety-specific transformational leadership, safety consciousness, and hazard reducing systems on warehouse accidents", *Journal of Operations Management*, vol. 29, no. 7-8, pp. 753-765, 2011. <https://doi.org/10.1016/j.jom.2011.06.005>.
- [14] B. Gaudenzi and A. Borghesi, "Managing risks in the supply chain using the AHP method", *The International Journal of Logistics Management*, vol. 17, no. 1, pp. 114-136, 2006. <https://doi.org/10.1108/09574090610663464>.
- [15] G. Zsidisin, "Managerial Perceptions of Supply Risk", *The Journal of Supply Chain Management*, vol. 39, no. 1, pp. 14-26, 2003. <https://doi.org/10.1111/j.1745-493x.2003.tb00146.x>.
- [16] G. Svensson, "A conceptual framework of vulnerability in firms' inbound and outbound logistics flows", *International Journal of Physical Distribution & Logistics Management*, vol. 32, no. 2, pp. 110-134, 2002. <https://doi.org/10.1108/09600030210421723>.
- [17] R. Dandage, S. Mantha and S. Rane, "Strategy development using TOWS matrix for international project risk management based on prioritization of risk categories", *International Journal of Managing Projects in Business*, vol. 12, no. 4, pp. 1003-1029, 2019. <https://doi.org/10.1108/ijmpb-07-2018-0128>.
- [18] Kirk, D.L., "Enforcement time nears for German Law," *Business Insurance*, Vol. 33 No. 36, pp. 19-20, 1999.

Solutions for Building a System to Support Motion Control for Autonomous Vehicle

Quach Hai Tho^{1,*}, Huynh Cong Phap², Pham Anh Phuong³

¹University of Arts, Hue University, 530.000, Vietnam

²Vietnam - Korea University of Information and Communication Technologies, Da Nang University, 550.000, Vietnam

³Faculty of Information Technology, University of Education, Da Nang University, 550.000, Vietnam

ARTICLE INFO

Article history:

Received: 12 May, 2020

Accepted: 14 June, 2020

Online: 26 June, 2020

Keywords:

Autonomous vehicle

Model predictive control

Path planning

Motion planning

Intelligent transportation systems

ABSTRACT

With a model predictive control approach including boundary analysis and uncertain prediction of activities of different road participants, this paper proposes solutions that support motion control by steering control and appropriate acceleration to create safe motion trajectories for an autonomous vehicle. The motion control support element is determined by the principle of minimal intervention and can handle complex situations, while building control model to predict real-time operation with speed factors, ability to control driving and limit the long period. The performance of this solution is assessed through simulation, then there are applied research orientations on practical autonomous vehicle accounting.

1. Introduction

Currently, the manufacturers have equipped with standard safety systems in researching the production of autonomous vehicles. However, in the context of traffic in complex environments, with the development of technology, the safety standards for autonomous vehicles have been raised, in which the driver assistance technology is a problem that needs to be cared about creating safe movements for vehicle. This technology can be mentioned by tracking driving conditions such as road conditions, vehicle status information, sensor technology, etc. to contribute to the development of advanced driver-assistance technology. For example, technology that supports the safety system in the adaptive cruise control system helps maintain distance from the vehicle ahead, collision avoidance assistance system helps predict collisions with obstacles with timely braking, as well as the lane keeping system helps the vehicle recognize the separator to adjust the vehicle direction so that the vehicle is always moving in the middle of the lane.

However, the scenario that these systems handle may be relatively simple compared to the diverse and complex situations we often encounter in traffic environment conditions. To solve this problem, we offer a motion control assistance solution to create a safe motion trajectory for the vehicle. The design of this control support system has two main goals: the first is minimal

intervention - that is, applying autonomous control only when necessary, the second is to ensure safety - meaning the vehicle's collision-free state is clearly enforced through optimal constraints.

The control support solution we propose in this article is implemented by predictive control based on the non-linear model predictive control and optimize the steering support system by the steering system along with the acceleration of the vehicle. In the solution assuming the current position of the vehicle, road boundaries, vehicles ahead and uncertain predictions about the future state of the vehicle in the form of a parametrized posterior distribution by their mean and covariance are known. Specifically, for this solution, we will combine the uncertainty changes over time to the mobile obstacle predictions into the optimization problem, and also introduce constraints for boundary limits and moving obstacles while maintaining a vehicle's movement plan for a limited time.

The next section of the paper will introduce the basic principles of predictive control and thereby propose a control solution based on Non-linear Model Predictive Control. Next is the experimental section and conclusion with some suggestions for further research directions for the problem of autonomous vehicles.

2. Building solution to support motion control

Theoretically, in order for an autonomous vehicle to avoid collisions, we need to calculate to find the set of states in which the

vehicle may encounter a collision situation when participating in traffic and then control the movement so that the vehicle does not move into that state set. There are many studies [1]-[5] doing this task that have determined the inevitable collision state set or target motion set. However, these studies do not limit the applicability or make assumptions for simple traffic scenarios, so it is difficult to analyze calculations. In this paper, the solution that we propose is the idea of defining collision state set and thereby determining the set of probability constraints to avoid collisions. At the same time, in this solution, we combine the steering speed control and increase/decrease speed so that the collision avoidance effect can be realized better. The ideas have originated from the studies [6,7,8,9] in the process of determining the difference of the steering angle or the deviation of the front wheel to achieve a safe trajectory.

In many studies, the model predictive control (MPC) has been applied to control autonomous vehicle [7]-[11] as the MPC approach to plan motion based on the vehicle's basic movements and track the road to avoid obstacles, or the MPC approach without constraint conditions and determine the stability of the vehicle with the surrounding environment to provide a safe steering angle at constant speed in a discrete environment. In this article, the MPC approach that we use can handle complex situations with steering control, increase/decrease vehicle speed, and avoid mobile obstacles at a certain extent in an uncertain environment.

The cost function of most MPC methods [7], [11]-[13] often depends on the time and the constraints on the path, so it needs to be pre-optimal or specific time steps or generated out and track the fixed motion of the vehicle, this leads to differences in result from the initial conditions of optimization that could result in invalid linear constraints and unpredictable motion planning. Therefore, we apply the point processing method in [14] and directly solve the nonlinear model predictive control problem by focusing on providing all costs and constraints for the decoding set that does not need to be linear manually.

2.1. Statement of problems

This problem is built on two basic principles: the first is minimal intervention, the second is to ensure safety, meaning the probability of a collision involving the surrounding environment and other traffic objects must be below certain thresholds. And this problem is done in discrete time intervals $k \triangleq t_k$, with $t_k = t_0 + \sum_{i=1}^k \Delta t_i$ (t_0 is the current time, Δt_i is the i time step of the plan).

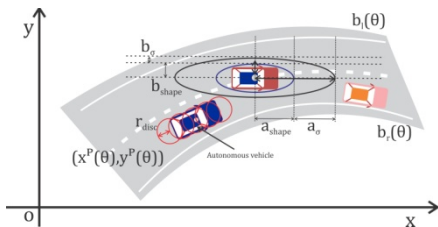


Figure 1. Modeling of autonomous vehicle and other vehicles

At interval k , with values of position $s_k = (x_k, y_k)$, linear velocity v_k , direction of vehicle ϕ_k and steering angle δ_k , the mathematical model of the vehicle is determined by the set $\mathbf{z}_k = [s_k, \phi_k, \delta_k, v_k] \in \mathbb{Z}$.

Assuming $\mathcal{B}(\mathbf{z}_k) \subset \mathbb{R}^2$ is the occupied area of the vehicle at \mathbf{z}_k state (illustrated in Figure 1, with autonomous vehicle modeled

by linked circles), the input values on the control using δ_k steering speed and a_k acceleration is assigned by $\mathbf{u}_k = [u_k^\delta, u_k^a] \in \mathcal{U}$.

Thus, the future state of the vehicle is represented by a discrete dynamic system as follows:

$$\mathbf{z}_{k+1} = f(\mathbf{z}_k, \mathbf{u}_k) \quad (1)$$

With other objects in traffic such as different kinds of car, bicycles, pedestrians, etc. called other objects will be assigned by the index $i = \{1, \dots, n\}$, the input control and their configuration parameters are determined by the values of $\mathbf{z}_k^i \in \mathbb{Z}_i$ and $\mathbf{u}_k^i \in \mathcal{U}_i$.

In order to combine uncertainty, it is necessary to assume that the future state distributions of the objects at m future time intervals is known. They are parameterized to the average state value $\mathbf{z}_{1:m}^i$ and covariance $\sigma_{1:m}^i$. Therefore, the degree of uncertainty in the forecast can be reflected through the covariance value $\sigma_{1:m}^i$.

At each given state, each object will occupy one space $\mathcal{B}^i(\mathbf{z}_k^i, \sigma_k^i, p_\epsilon) \subset \mathbb{R}^2$ having probability greater than p_ϵ (p_ϵ is the acceptable probability of collision that may happen), the model of the object and this occupied space are shown in Figure 1.

The free space defined in this problem is the workspace $\mathcal{W} = \mathbb{R}^2$ and the location map of obstacles $\mathcal{O} \subset \mathcal{W}$ containing static obstacles such as limit of roads and systems of separator, etc., at the same time, the surrounding environment $\mathcal{E}(k)$ is determined to be the state of other objects (including means of traffic, obstacles) at the time k .

In this study, with the set of states $\mathbf{z}_{0:m} = [\mathbf{z}_0, \dots, \mathbf{z}_m] \in \mathbb{Z}^{m+1}$ and input set $\mathbf{u}_{0:m-1} = [\mathbf{u}_0, \dots, \mathbf{u}_{m-1}] \in \mathcal{U}^m$, we will build a general discrete time constraint optimization at m time steps with a time limit $\tau = \sum_{k=1}^m \Delta t_k$.

Thus, the goal of the solution is to calculate the optimal input values $\mathbf{u}_{0:m-1}^*$ for autonomous vehicle with minimizing cost function $\hat{J}_h(\mathbf{u}_{0:m-1}, \mathbf{u}_0^h) + \hat{J}_t(\mathbf{z}_{0:m}, \mathbf{u}_{0:m-1})$.

In which: $\hat{J}_h(\mathbf{u}_{0:m-1}, \mathbf{u}_0^h)$ is the minimum cost to minimize deviations from the input value \mathbf{u}_0^h , and $\hat{J}_t(\mathbf{z}_{0:m}, \mathbf{u}_{0:m-1})$ is the cost depending on the properties of the trajectory according to the motion plan.

This optimum problem follows a set of constraints: the first is to use the vehicle transition model, the second is the non-collision constraint with static obstacles and the third is the probability that no collision will occur p_ϵ with other traffic participants.

And the optimal trajectory of the vehicle is given as follows:

$$\mathbf{u}_{0:m-1}^* = \arg \min_{\mathbf{u}_{0:m-1}} \hat{J}_h(\mathbf{u}_{0:m-1}, \mathbf{u}_0^h) + \hat{J}_t(\mathbf{z}_{0:m}, \mathbf{u}_{0:m-1}) \quad (2)$$

Where:

$$\mathbf{z}_{k+1} = f(\mathbf{z}_k, \mathbf{u}_k)$$

$$\mathcal{B}(\mathbf{z}_k) \cap \mathcal{O} = \emptyset$$

$$\mathcal{B}(\mathbf{z}_k) \cap \bigcup_{i \in \{1, \dots, n\}} \mathcal{B}^i(\mathbf{z}_k^i, \sigma_k^i, p_\epsilon) = \emptyset$$

$$\forall k \in \{0, \dots, m\}$$

$\mathbf{z}_{0:m}^i, \delta_{0:m}^i$ with $i = 1, \dots, m$: are parameters for all other traffic objects, \mathbf{z}_0 is the initial state of vehicle.

2.2. Building a solution

The solution was built to calculate the creation of a safe motion trajectory with a predefined horizon. Optimal problems are constrained by cost values, transition models, maintaining vehicle movement within the boundaries of the lane median, and avoiding collisions with other road participants that must ensure the probability of collision below the value p_ϵ .

In solutions using model predictive control of previous studies [7,11,13], the researchers often use factors such as constant longitudinal speed and small angle assumptions in scenarios avoiding obstacles on the straight road. In this study, we will consider the impact of longitudinal speed to ensure overall safety in complex traffic environments.

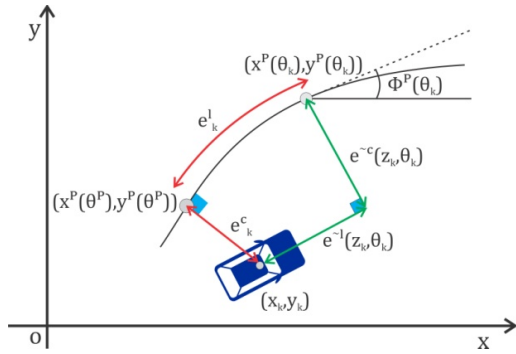


Figure 2. Vehicle dynamics model

The vehicle model in this solution has been introduced with a fixed rear wheel and the motion control is located on the front wheel with \mathbf{z} and \mathbf{u} states set here, and the control of rear wheel movement is via center bridge with distance L and a continuous dynamic model. This motion model is described by a discrete time model with integrals $\mathbf{z}_{k+1} = \mathbf{z}_k + \int_k^{k+\Delta t_k} \dot{\mathbf{z}} dt = f(\mathbf{z}_k, \mathbf{u}_k)$, as follows:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\phi} \\ \dot{\delta} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} v \cos(\phi) \\ v \sin(\phi) \\ \frac{v}{L} \tan(\delta) \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u^\delta \\ u^a \end{bmatrix} \quad (3)$$

Where: $\dot{\mathbf{z}} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\phi} \\ \dot{\delta} \\ \dot{v} \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} u^\delta \\ u^a \end{bmatrix}$

In this model, the limits applied include steering angle $\|\delta\| \leq \delta_{max}$, steering speed $\|u^\delta\| \leq \delta_{max}$ and longitudinal speed $v \leq v_{max}$. The introduction of these limits is consistent with vehicle performance and road traffic rules, as well as a number of restrictions will be set to ensure safety such as speed limits when moving into the corners by limiting the slip coefficient $\|\phi\| \leq \phi_{max}$ and limiting the maximum speed quickly by the acceleration limit $u^a \in [a_{min}, a_{max}]$.

The model predictive control method in this study is based on model predictive contouring control method [15,16,17] and is used to solve problems, in this study, it is not required that the motion trajectory of the vehicle must be exactly according to the reference trajectory, but the motion trajectory must be in the safe range.

On the reference trajectory, the process takes place as follows:

At time k , the vehicle is in position $s_k = (x_k, y_k)$, the process of following the reference path is limited and continuously differentiable in geometric space $(x^p(\theta), y^p(\theta))$ of path parameter θ , with the tangential vector \mathbf{t} and the normal vector \mathbf{n} as follows:

$$\mathbf{t} = \begin{bmatrix} \frac{\partial x^p(\theta)}{\partial \theta} \\ \frac{\partial y^p(\theta)}{\partial \theta} \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} -\frac{\partial y^p(\theta)}{\partial \theta} \\ \frac{\partial x^p(\theta)}{\partial \theta} \end{bmatrix} \quad (4)$$

and the path instructions are described as follows:

$$\phi^p(\theta_k) = \arctan\left(\frac{\partial y^p(\theta)}{\partial x^p(\theta)}\right) \quad (5)$$

the path parameterized according to the arc length $\left(\frac{\partial \theta}{\partial s} = 1\right)$ will allow to estimate the vehicle's progress with v_k on the reference path and the actual path $s = \int v dt$ if during the parameterization of the curves the arc length is negligible and the distance between the points is small compared to the arc length. At the same time, since the motion trajectory of the vehicle will follow a certain path and have a slight deviation from the reference trajectory determined by the road boundary, we can assume that the trajectory remains the same offset, so: $\Delta \theta \approx \Delta s = v \Delta t$, with this additional assumption will generate an approximate process according to the path parameter, as follows:

$$\Delta \theta_{k+1} = \Delta \theta_k + v_k \Delta t_k \quad (6)$$

where $v_k \Delta t_k$ describes the approximation process in time step k .

In the general case, finding the path parameter $\theta^p(x_k, y_k)$ of the nearest point $S(x_k, y_k)$ on the reference path is not feasible and not suitable for fast optimization. Therefore, the value of $\theta^p(x_k, y_k)$ will be approximated according to equation (6).

During the motion planning process, if the actual path deviates from the reference path, the delay error from the first approximation point in time progression to the next point and the position error referenced to the horizontal roads θ_k along the tangential path \mathbf{t}_k are defined as follows:

$$e^{-l}(\mathbf{z}_k, \theta_k) = \frac{\mathbf{t}_k^T}{\|\mathbf{t}_k\|} \begin{bmatrix} x_k - x^p(\theta_k) \\ y_k - y^p(\theta_k) \end{bmatrix} = -\cos \phi^p(\theta_k)(x_k - x^p(\theta_k)) - \sin \phi^p(\theta_k)(y_k - y^p(\theta_k)) \quad (7)$$

If the delay error $e^{-l}(\mathbf{z}_k, \theta_k)$ is small, the process of constructing an approximate path will be as close to the horizontal asymptote as $\Delta \theta \approx \Delta s = v \Delta t$ and $\theta_k \approx \theta^p(x_k, y_k)$. Also, in the process of optimizing the prediction control, the delay error should be actively handled so that the error of the predictive process θ_k is small enough according to the motion planning process.

When projected on the standard path between the actual position and the predicted position, we will determine the contouring error by the deviation of these two positions, as follows:

$$e^{-c}(\mathbf{z}_k, \theta_k) = \frac{\mathbf{n}_k^T}{\|\mathbf{n}_k\|} \begin{bmatrix} x_k - x^P(\theta_k) \\ y_k - y^P(\theta_k) \end{bmatrix} = \sin\phi^P(\theta_k)(x_k - x^P(\theta_k)) - \cos\phi^P(\theta_k)(y_k - y^P(\theta_k)) \quad (8)$$

and the contouring error is a standard for establishing good motion planning when the vehicle is in motion not deviated from a given reference path. Therefore, the cost function of predictive state control is built based on the balance between the contouring error factors $e^{-c}(\mathbf{z}_k, \theta_k)$, delay error $e^{-l}(\mathbf{z}_k, \theta_k)$ and the process of building approximate path v_k to achieve the best combination, as follows:

$$J_{av}(\mathbf{z}_k, \theta_k) = \mathbf{e}_k^T Q \mathbf{e}_k - v_k \quad (9)$$

with path error vector formed from delay error and contouring error as follows:

$$\mathbf{e}_k = \begin{bmatrix} e^{-l}(\mathbf{z}_k, \theta_k) \\ e^{-c}(\mathbf{z}_k, \theta_k) \end{bmatrix} \quad (10)$$

For the performance of the path, all reference paths are parameterized by C^1 - a continuous clothoid path in a system of paths through predetermined points. The clothoid path estimate will be replaced by the third-order spline function of equally spaced nodes and parameterization of the spline functions along the arc length is sufficiently accurate, as well as achieves good performance in the calculation. Because the spline functions provide an analytical parameter about the reference path, the limit of the path and the derivatives needed to solve the nonlinear optimization problem.

At position $S(x_k, y_k)$, from the projection along the normal of the reference path at the actual curvilinear abscissa θ^P , we get the lateral distance $d(\mathbf{z}_k, \theta)$ with the reference path and approx slag equal to θ_k such that $d(\mathbf{z}_k, \theta_k) = e^{-c}(\mathbf{z}_k, \theta_k)$. Just as the movable space of autonomous vehicles at the crossroads θ_k is limited by the left boundary $b_l(\theta_k)$, the right boundary $b_r(\theta_k)$ of the lane and other static obstacles will be parameterized by the third order spline function to allow analytical evaluation and derivation.

As such, the horizontal traverse for the path is limited as follows:

$$b_l(\theta_k) + w_{max} \leq d(\mathbf{z}_k, \theta_k) \leq b_r(\theta_k) - w_{max} \quad (11)$$

where: w_{max} is the upper limit of the vehicle position projection on the standard reference path and this value is greater than $\frac{1}{2}$ of the vehicle width. At the same time, maintaining the validity of w_{max} as a limit just taking the radius of the vehicle's operating range as an upper limit is enough to ensure safety.

Because the relationship between the path and the direction of the vehicle is relative, we need to establish a constraint between the direction of the path $\phi^P(\theta_k)$ and the movement direction of the vehicle ϕ_k , as follows:

$$\|\phi_k - \phi^P(\theta_k)\| \leq \Delta\phi_{max} \quad (12)$$

In the solution of this study, other traffic participants will be described by an orientation ellipse \emptyset with a_{shape} being a semi-major axis and b_{shape} being a semi-minor axis of the ellipse in the longitudinal and horizontal directions of the objects.

The development of future trajectories of these objects is assumed to have some uncertain positions of posterior distribution and is parameterized by the average trajectory $\mathbf{z}_{0:m-1}^i$ and degree uncertainty $\sigma_{0:m-1}$, as follows:

$$\sigma_{k+1} = \sigma_k + \sigma \Delta t_k \quad (13)$$

In the general case, we propose a model for generating indeterminate positions of vehicles with uncertainty $\sigma_k = [\sigma_k^a, \sigma_k^b]^T$ at the time k and $\sigma = [\sigma^a, \sigma^b]^T$ is the uncertainty incurred. Therefore, the value of the variance is determined to approximate to adjust the direction of the vehicle motion aligning the main axis of the surrounding ellipse. The generation of indeterminate positions in the horizontal direction of the vehicle is limited by a maximum value to consider the maximum rationality of the maximum flow of vehicles currently in the current lanes.

The level-set of Gaussian $\mathcal{N}(0, \text{diag}(\sigma_k))$ describe the indefinite position of other road users at p_ϵ degree and ellipses formed with coefficients are set as follows:

$$\begin{bmatrix} a_{\sigma_k} \\ b_{\sigma_k} \end{bmatrix} = \begin{bmatrix} \sigma_k^a \\ \sigma_k^b \end{bmatrix} \left(-2 \log(p_\epsilon 2\pi \sigma_k^a \sigma_k^b) \right)^{1/2} \quad (14)$$

Therefore, we can use the main axis direction and add coefficients to the cross axles of the vehicle to identify the area of the obstacle with the probability of occupancy above the threshold p_ϵ . At the same time, the rectangular area used to determine the position and occupied area of autonomous vehicles will be replaced by a set of circles with radius r_{av} . The use of circles instead of ellipses represents obstacles because the presence of an autonomous vehicle does not need to be aligned in a straight axis, and Minkowski's sum calculation efficiency is not possible if shown by ellipses whose axes do not fit tightly. The Minkowski sum of the sets of circles surrounding the vehicle and the collision constraints shown by the previous displaced ellipses are calculated as follows:

$$c_k^{obs,i}(\mathbf{z}_k) = \begin{bmatrix} \Delta x_j \\ \Delta y_j \end{bmatrix}^T R(\emptyset)^T \begin{bmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{b^2} \end{bmatrix} R(\emptyset) \begin{bmatrix} \Delta x_j \\ \Delta y_j \end{bmatrix} \Bigg|_{k,i} > 1 \quad (15)$$

with $\forall j \in \{1, \dots, 4\}$

where $\Delta x, \Delta y$ are the distances from the circle set covering the area around the vehicle to the center of obstacle i at the time k and $R(\emptyset)$ is the rotation matrix corresponding to the motion direction of the obstacle, and the semi-major axis of the constraint ellipse. The result is calculated as follows:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a_{shape} + a_{\sigma_k} + r_{disc} \\ b_{shape} + b_{\sigma_k} + r_{disc} \end{bmatrix} \quad (16)$$

Thus, we have a collision-free constraint with a higher probability p_ϵ than other means.

As our goal of the solution in this research is to minimize the intervention of the driving, that is, the control system only intervenes with the steering operation when really necessary with the minimum intervention time:

$$J_h(\mathbf{z}_k, \mathbf{u}_k, \mathbf{u}_0^h) = \begin{bmatrix} \mathbf{u}_k^a - \mathbf{a}_0^h \\ \delta - \delta_0^h \end{bmatrix}^T K \begin{bmatrix} \mathbf{u}_k^a - \mathbf{a}_0^h \\ \delta - \delta_0^h \end{bmatrix} \quad (17)$$

where $\mathbf{u}_0^h = [\delta_0^h, \mathbf{a}_0^h]^T$ is the inconsistent value of system state, δ^h is the steering angle value and \mathbf{a}^h is the acceleration value at the time t_k .

In the process of setting values for control inputs, we only determine the steering angle value δ^h and acceleration \mathbf{a}^h without determining the steering speed value $\dot{\delta}^h$. However, in the general case if the front view angle can be controlled, the speed value is still used as the input value, and the control process \mathbf{u}_k maintains the steering speed and acceleration of the vehicle.

The calculation of the trajectory cost will include the cost of the model predictive control calculated by the equation $\mathbf{e}_k = [e^{-l}(\mathbf{z}_k, \theta_k), e^{-c}(\mathbf{z}_k, \theta_k)]$, and at the same time add input control deviations and slip coefficients to create smooth and comfortable trajectories when the vehicle is in motion. The weights R and A allow for sorting based on different priorities, so the trajectory cost is calculated as follows:

$$J_t(\mathbf{z}_k, \mathbf{u}_k, \theta_k) = J_{av}(\mathbf{z}_k, \theta_k) + \mathbf{u}_k^T R \mathbf{u}_k + \dot{\phi}_k A \dot{\phi}_k \quad (18)$$

where $J_{av}(\mathbf{z}_k, \theta_k)$ has transformed the inconsistency from the reference path error into a better direction of motion.

Finally, the solution proposed is the optimization problem. It is done by the minimum factor combining linear between the cost of intervention into the system and trajectory cost, as follows:

$$J_{av}(\mathbf{z}_k, \mathbf{u}_k, \theta_k, \mathbf{u}_0^h) = \beta \omega(t_k) J_h(\mathbf{z}_k, \mathbf{u}_k, \mathbf{u}_0^h) + (1 - \omega(t_k)) J_t(\mathbf{z}_k, \mathbf{u}_k, \theta_k) \quad (19)$$

in which, the weight β and exponential decay function $\omega(t_k) = \exp(-\alpha t_k)$ are used to enhance the input value for the system.

We have chosen the solution to make the weight β reach a high value, so that when moving forward in the predictive model, the system will be able to respond well to inputs but still depend on J_t . By doing so, the solution presented in this study can plan a full implementation without having to predict the planned motion trajectory. Therefore, the nonlinear optimization problem with constraints on state, dynamics, paths and obstacles is constructed as follows:

$$\mathbf{u}_{0:m-1}^* = \arg \min_{\mathbf{u}_{0:m-1}} \sum_{k=1}^m J_{av}(\mathbf{z}_k, \mathbf{u}_k, \theta_k, \mathbf{u}_0^h) \Delta t_k \quad (20)$$

where: $\mathbf{z}_{k+1} = f(\mathbf{z}_k, \mathbf{u}_k)$;

$\theta_{k+1} = \theta_k + v_k \Delta t_k$;

$\mathbf{z}_k \in [\mathbf{z}_{min}, \mathbf{z}_{max}]$;

$\mathbf{u}_k \in [\mathbf{u}_{min}, \mathbf{u}_{max}]$;

$\|\dot{\phi}_k\| < \dot{\phi}_{max}$;

$\|\phi_k - \phi^P(\theta_k)\| < \Delta \phi_{max}$;

$d(\mathbf{z}_k, \theta_k) \in [b_l(\theta_k) + \omega_{max}, b_r(\theta_k) - \omega_{max}]$;

$c_k^{obs,i}(\mathbf{z}_k) > 1, i = \{1, \dots, n\}, \forall k \in \{0, \dots, m\}$

with initial initialization values for the path $(x^P(\theta), y^P(\theta))$, the left boundary $b_l(\theta)$ and the right boundary $b_r(\theta)$ set by the arc and the static obstacles that the vehicle is moving. At the same time, in each loop when the system executes the initial states \mathbf{z}_0 and θ_0 , the input variables \mathbf{u}_0^h and predicts other objects of traffic $\mathbf{z}_{0:m}^i, \sigma_{0:m}^i$ will be given a predictive model for the control system. After that, the optimal solution given by equation (20) and the optimal control \mathbf{u}_0^* will be executed by the system.

3. Experimental results

To test and evaluate the proposed solution, we have conducted empirical simulation of processes in the matlab environment. At the same time, in order to ensure the objectivity and reliability when evaluating, we have conducted simulations with different scenarios and autonomous vehicles controlled moving with steering angle δ_0^h and desired acceleration \mathbf{a}_0^h . Input variables will be handled using model predictive control to ensure generating safe movement, the reference path and the left boundary b_l , right boundary b_r are designed and determined accordingly with the road system. During the experiment, the calculations uses SI measurement system with sampling interval of 0.1s, the trajectory is transferred to I/O controller with simulation time of 0.005s.

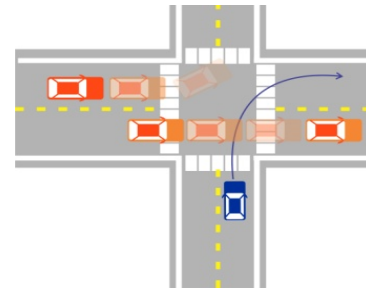


Figure 3. Simulating the scenario 1

Scenario 1: In this scenario, the autonomous vehicle will move into a corner, with the input values for the control system will be able to make the vehicle's direction of travel out of the limit of traffic lanes. At this point, the driver assistance system will perform braking to reduce the vehicle speed to the safe speed limit with the coefficient constraint before the car enters a corner, then increase the vehicle speed after going out of the corner to ensure the progress in the motion planning. With such a driver assistance process, it shows the advantages of controlling the longitudinal force, horizontal force and acceleration the vehicle, as well as implementing the vehicle's movement plan that is complete when entering corners with the brake and increase/decrease speed operations.

Scenario 2: In this scenario, the autonomous vehicle will implement the movement plan from the slip road and turn left to enter the main traffic lane, the initial positions and the speed of

other traffic participants are initialized randomly. To increase the time span of the planning without adjusting the calculations, we will take the approach of changing the distance at each step as follows: The first 40 steps have a distance of $\Delta t_k = 0.1s$ and the next 50 steps have $\Delta t_k = 0.4s$, which leads to a planning time of 24 seconds for all calculations to be performed in real time.

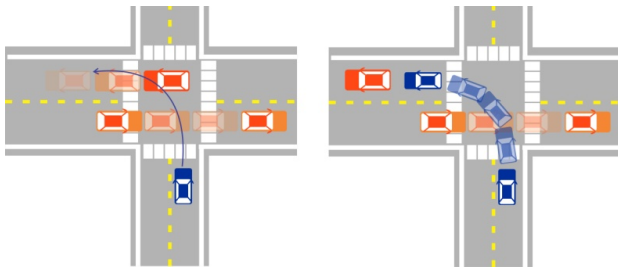


Figure 4. Simulating the scenario 2

The process of following this scenario will be repeated many times with randomly generated input values and there are times when the input values are insecure input to the system, which will lead to accident when the vehicle is moving as planned. However, with the control support system that our solution offers, it is possible to prevent incidents in the most likely cases of accidents.

Specifically: The planning process of movement when the vehicle is moving in a slip road before entering the main road, a collision may occur with the right boundary of the lane. After that, to move into the main lane, the vehicle turns left, at this time on the main lane, other vehicles are moving complicatedly, and unsafe situations may occur. To handle this movement of the vehicle, the control support system will activate the brakes to bring the vehicle's motion status to a stop state so that other vehicles in traffic can move through it, until the safe distance between the autonomous vehicle and other vehicles is large enough, the system activates the vehicle velocity so that the vehicle can move over the cross and merge into the main lane.

During the experiment, we can see that the uncertainty in the predictive of other vehicles is very important because the future states of the predictive model can deviate from the desired predictive point. In the case of omitting uncertainty, the motion planning process needs to provide more precise and specific constraints to ensure the vehicle safety.

4. Conclusions

This paper has proposed a motion control assistance solution to ensure the safety of an autonomous vehicle. The optimal feature of this solution is to shorten the motion planning cycle to minimize deviations from inputs of the prediction while ensuring movement safety. This controller support solution is only implemented in situations where a complex vehicle movement scenario is likely to have a collision with realistic warning elements.

The main idea of this technical solution will support designing an autonomous vehicle with a safe stop whatever the current vehicle control. The experimental simulation with the given scenarios shows that the safety factor can be achieved by calculating to consider all possible possibilities of other vehicles.

In the future, in order that this solution will be more reliable, we have experimented the settings by transferring the simulation to the real environment with experimental vehicle equipped with sensors. When experimenting on reality, it will add a number of factors to analyze the stability of the system so that the behavior of traffic participants is more accurately forecasted. The extensive

implementation of this driver-assistance solution for semi-autonomous or autonomous vehicles in vehicle control systems will be able to minimize the amount of damage and create a safe movement plan for the future.

References

- [1] A. Bautin, L. Martinez-Gomez, T. Fraichard, "Inevitable Collision States: A probabilistic perspective," *IEEE International Conference on Robotics and Automation* 2010, pp.4022–4027,2010,DOI: 10.1109/ROBOT.2010.5509233.
- [2] D. Althoff, M. Althoff, D. Wollherr et al, "Probabilistic collision state checker for crowded environments," *IEEE International Conference on Robotics and Automation - 2010*, pp. 1492–1498, 2010, DOI: 10.1109/ROBOT.2010.5509369.
- [3] D. Hoehener, G. Huang, D. D. Vecchio, "Design of a lane departure driver-assist system under safety specifications," *IEEE 55th Conference on Decision and Control (CDC)-2016*, pp. 2468–2474, 2016, DOI: 10.1109/CDC.2016.7798632.
- [4] M. Forghani, J. M. McNew, D. Hoehener et al, "Design of driver-assist systems under probabilistic safety specifications near stop signs," *IEEE Transactions on Automation Science and Engineering*, vol. 13, no. 1, pp. 43–53, 2016, DOI: 10.1109/TASE.2015.2499221.
- [5] T. Fraichard, H. Asama, "Inevitable collision states. A step towards safer robots?" *Advanced Robotics*, vol. 18, no. 10, pp. 1001–1024, 2003, DOI: 10.1109/IROS.2003.1250659.
- [6] J. Alonso-Mora, P. Gohl, S. Watson et al, "Shared control of autonomous vehicles based on velocity space optimization," *IEEE International Conference on Robotics and Automation (ICRA)-2014*, pp. 1639–1645, 2014, DOI: 10.1109/ICRA.2014.6907071.
- [7] S. Erlien, S. Fujita, J. C. Gerdes, "Shared steering control using safe envelopes for obstacle avoidance and vehicle stability," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 2, pp. 441–451, 2015, DOI: 10.1109/TITS.2015.2453404.
- [8] V. A. Shia, Y. Gao, R. Vasudevan et al, "Semiautonomous vehicular control using driver modeling," *IEEE Transactions on Intelligent Transportation Systems*, vol.15,no.6,pp.2696–2709,2014, DOI: 10.1109/TITS.2014.2325776.
- [9] Y. Gao, A. Gray, A. Carvalho et al, "Robust nonlinear predictive control for semiautonomous ground vehicles," in *2014 American Control Conference*, pp. 4913–4918, 2014, DOI: 10.1109/ACC.2014.6859253.
- [10] A. Gray, Y. Gao, T. Lin et al, "Predictive control for agile semi-autonomous ground vehicles using motion primitives," *American Control Conference (ACC)-2012*, pp. 4239–4244, 2012, DOI: 10.1109/ACC.2012.6315303.
- [11] S. J. Anderson, S. B. Karumanchi, K. Iagnemma, "Constraint based planning and control for safe, semi-autonomous operation of vehicles," *IEEE Intelligent Vehicles Symposium-2012*, pp. 383–388, 2012, DOI: 10.1109/IVS.2012.6232153.
- [12] B. Paden, S. Z. Yong, D. Yershov et al, "A survey of motion planning and control techniques for self-driving urban vehicles," *IEEE Transactions on Intelligent Vehicles*, vol. 1, no. 1, pp. 33–55, 2016, DOI: 10.1109/TIV.2016.2578706.
- [13] S. J. Anderson, S. C. Peters, T. E. Pilutti et al, "An optimal-control-based framework for trajectory planning, threat assessment, and semi-autonomous control of passenger vehicles in hazard avoidance scenarios," *International Journal of Vehicle Autonomous Systems*, vol. 8, no. 2-4, pp. 190–216, 2010, DOI: 10.1504/IJVAS.2010.035796.
- [14] A. Domahidi, J. Jerez, "FORCES Pro," embotech GmbH (<http://embotech.com/FORCES-Pro>), 2014.
- [15] A. Liniger, A. Domahidi, M. Morari, "Optimization-based autonomous racing of 1:43 scale RC cars," *Optimal Control Applications and Methods*, vol. 36, no. 5, pp. 628–647, 2017, DOI: 10.1002/oca.2123.
- [16] D. Lam, C. Manzie, M. C. Good, "Model predictive contouring control for biaxial systems," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 2, pp. 552–559, 2013, DOI: 10.1109/TCST.2012.2186299.
- [17] T. Faulwasser, B. Kern, R. Findeisen, "Model predictive pathfollowing for constrained nonlinear systems," in *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pp. 8642–8647, 2009, DOI: 10.1109/CDC.2009.5399744.