ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 4, 296-313 (2020)

www.astesj.com

Special Issue on Multidisciplinary Innovation in Engineering Science & Technology

ASTESJ

ISSN: 2415-6698

Cluster Centroid-Based Energy Efficient Routing Protocol for WSN-Assisted IoT

Nalluri Prophess Raj Kumar*, Josemin Bala Gnanadhas

Department of Electronics and Communication Engineering, Karunya Institute of Technology and Sciences, Coimbatore, 641114, India

A R T I C L E I N F O

Article history:
Received: 31 May, 2020
Accepted: 19 July, 2020
Online: 28 July, 2020

Keywords:
Internet of Things
Wireless sensor networks
Energy management
Basestation
Clustering
Energy Centroid

A B S T R A C T

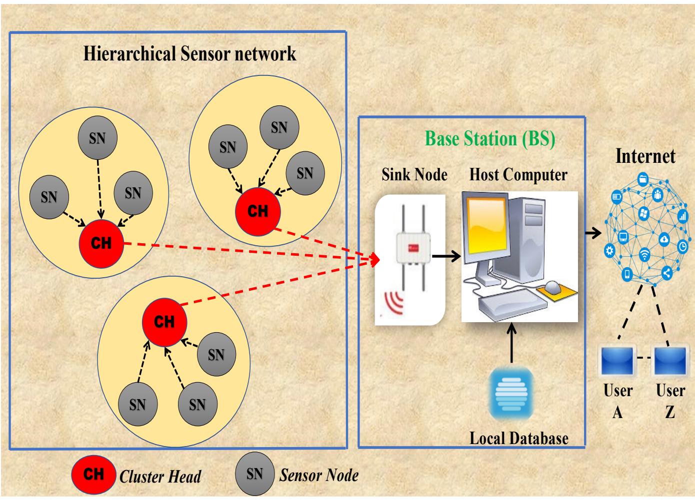
Wireless sensor network is highly resource constrained, where energy efficiency and network lifetime plays a major role for its sustenance. As the sensor nodes are battery operated and deployed in hostile environments, either recharging or replacement of batteries in sensor nodes is not possible after its deployment in inaccessible areas. In such condition, energy is the vital factor for the survival of sensor node in the sensing field. In order to increase the network lifetime and balance the energy consumption, robust routing protocols are required. The proposed network routing has three phases: 1. Network initiation phase to create a zone which enables the communication among local nodes 2. Zone co-ordinator selection phase algorithm to form zone cluster and its re-election procedure and 3. Zone head selection with its replacement phase based on energy centroid positional information and distance to the basestation to distribute load equally among zone co-ordinators, local sensor nodes. The data path between zone heads and basestation is distance centric and is optimized at one hop and dual hop levels to avoid data packet loss at zoneheads. Each zone is designed to own atmost ¼ rth of deployed sensor node count through uniform random deployment. Simulations results when basestation is placed inside sensing field indicates that the proposed network algorithm outperforms when benchmarked against similar protocols like conventional LEACH, Traditional PEGASIS, existing PRRP, ES3 protocols in terms of performance metrics like Network energy consumption, Average energy consumed by sensor node, Packet delivery ratio, Packet loss percentage and Network throughput.

1. Introduction

Internet of Things (IoT) is a concept to interlink the mechanism of the object with the internet and communicate with them to get identified by others. IoT defines the world where a physical object can be connected and communicated in an intelligent way. IoT can build the network comprising physical objects and provide them the capability to gather and share their own information with others. Wireless sensor network is a subset of IoT as it has the technology that can be utilized within an IoT system to meet the user demands. So, a sensor network assisting IoT, in an application using wireless network may have an advantage of lesser cost, friendly deployment of sensor nodes and better scalability. But a main drawback is that, its energy resources are hard to replace as they operate in rugged environment. This makes the energy management a key metric in wireless sensor network assisted IoT applications. A simple architecture of WSN- assisted IoT is shown in Figure 1.

Here a sensor node (SN) gathers information from environmental properties like temperature, pressure or moisture of air and convert it into digital form to send them wirelessly via *cluster* heads to the *base station* (BS). A basestation in contrast to sensor node possesses more computational power, large memory and connected to best source of energy unlike sensor nodes. In an IoT application, a basestation deployed can store, analyze and visualize the *sensed data* collected from cluster heads (CH). The basestation provides *graphical user interface* to interact with users directly or forward sensed data to a *remote server* via *Internet*. Then these *remote servers* will relay sensed data to the authorized users. Also, this sensed data can be saved as *web pages*, so that it can be accessed worldwide via *Internet*.

* Corresponding Author: Nalluri Prophess Raj Kumar, Karunya Institute of Technology and Sciences, +91-8309329926 & prophessms2k17@gmail.com, nalluriprophess@karunya.edu.in



In order to simplify the network management, the *cluster/group/ zone* concept is proposed by various researchers. CH nodes are the managers of grouped sensors nodes. CH has the responsibility to organize sensor nodes in its cluster, framing a routing table, collecting, aggregating and retransmitting the sensed data of the cluster. CH in sensor network will fast deplete their energy due to load imposed on them. To reduce the energy dissipation of CH nodes, the communication distance between CH and the basestation should be low or else information needs to be multi-hop forwarded from CH to basestation with the help of intermediate cluster heads.

Aiming at higher energy efficiency for the sensor network proposed, an extension of work originally presented in an IEEE conference [1] is modified as Cluster Centroid based Energy Efficient Routing (CEER) to balance energy consumption in cluster based WSN-assisted IoT. The main contributions of this research work are as follows:

- Producing a zone-clustering protocol which operates based on energy centroid position. The residual energy of sensor nodes and the communication distance to the basestation is taken into consideration here.
- To optimize the path between the zone head and the basestation, a dual hop or single hop communication routing algorithm is proposed. Zone heads count and distance metric are taken into consideration in a network area of *A* sq.m
- To optimize the path between the local sensor nodes in the zone and the basestation, zone co-ordinators are selected in the zone clusters. One among them acts as the zone head to forward sensed data to avoid local packet loss and long-distance communication among sensor nodes inside the zone.
- The user is provided with an option to opt for either 2 X 2 CEER (mostly suitable for dense environments) or 4 X 4 CEER (mostly suitable for sparse environments) to analyze the results based on their own preferred node deployment type. Provided the basestation should be placed inside the sensing field of the user defined area.

The research contributions reduced the average energy consumption of sensor nodes inside the zone clusters which eventually increased the lifetime of sensor nodes in the network when simulated under a simulation time t = 200 Sec. Moreover, the network lifespan is increased, as the sensor nodes balance the energy consumption among themselves (zone co-ordinator and zone head replacement strategy) to live for a longer period of time and contribute to the network throughput.

The rest of the paper is organized as follows: In section 2, the related work is discussed. The proposed CEER protocol and its phases along with its radio energy dissipation model is explained in section 3. The simulation results and discussions of the proposed protocol in comparison to the existing works is shown in section 4. Section 5 concludes the paper with further proceedings.

2. Related Work

Based on the network architecture, hierarchical routing protocols are classified into three types [2]: 1. Cluster-based; 2. Chain-based; 3. Tree-based protocols.

In *cluster-based* routing, network area is divided into clusters with the assistance of BS or sometimes through grid clustering and so on. A cluster head is selected for each cluster to transfer information collected from all the sensor nodes to the basestation either directly or with the assistance of other cluster heads or sensor nodes outside the cluster. In *chain-based* routing, the nodes are arranged in a chain like structure and only one node will act as the chain head to transmit the entire information of the chain to the basestation directly in the entire network area. In *tree-based* routing, all the sensed information by the leaf node is carried to the parent node (sink node) and at last, the information is sent to the root node i.e., BS.

2.1. Conventional Clustering approach

The first and foremost hierarchical routing protocol designed for wireless sensor networks is the LEACH protocol. In the open system interconnections (OSI) reference model, cluster routing protocols works on the network layer which is connected to both the data link layer and the transport layer. The primary function of the network layer is to permit different networks to be interconnected. It also translates logical network addresses into physical address. So, a network layer forms the major part in forwarding and routing the information across the network.

The LEACH [3-6] protocol is explained in detail in section 4.1.1. Its main characteristic is the local cluster generation, dynamic CH node rotation with data fusion inside the cluster. By generating a random number which is between 0 or 1, all sensor nodes are given provision to be the CH and the threshold set for selection of node is calculated as in Eq. (1)

$$T(n) = \begin{cases} \frac{P}{1-P\left[r*mod\left(\frac{1}{P}\right)\right]} & n \in N \\ 0 & Otherwise \end{cases} \quad (1)$$

The generated random number and $T(n)$ are compared for each sensor node. If the random number generated by a sensor node is less than the calculated threshold value $T(n)$, then that particular sensor node is selected as a cluster head for the current round. In Eq. (1) P is the desired percentage of sensor node inside the local cluster to be selected as cluster head, r is the round number, n is node count and $T(n)$ is the threshold. Number of cluster head nodes are not constant in LEACH due to its random selection of cluster head nature. The drawback of LEACH [7] is that, the different cluster count noted in each round will make number of sensor nodes in each cluster different. Uneven sensor node count in clusters dissipate uneven energy in each round. During CH selection, residual energy of sensor node is not taken into consideration in LEACH protocol.

2.2. Energy based clustering approach

An Energy Based Clustering – Self organizing map (EBC-S) to bring the effectiveness in cluster-based routing algorithms is proposed in [8] for topological clustering and incorporating a topological energy-based clustering technique to achieve extended lifetime and network coverage. The assumption made on BS is that it has no constraints on energy resources. A hierarchical and distance-based clustering technique is proposed in [9] which utilizes a new rank-order distance measure for agglomerative hierarchical clustering. Here, authors generate a rank order list by sorting all other sensor nodes in near neighbors by absolute distance. The distance based on rank order of two sensor nodes is computed using their rank order. The algorithm designed by them by grouping sensor nodes into small number of sub-clusters is similar to CEER networking like the formation of zone clusters inside the zone in CEER. The word *centroid* for wireless sensor networks is used by authors in [10-12] for *k*-means clustering algorithm. This algorithm is mainly based on Euclidean distance of nodes and CH selection based on residual energy of nodes. In this method, sink node/BS collects information about the *identifier, position* and *residual energy* of all nodes and store that information.

The steps in *k*-means clustering algorithm for WSNs are as follows:

- To form *k* clusters of sensor nodes, *k* centroids are to be selected initially at different locations inside the network area *A* sq.m.
- Euclidean distance from each sensor node to the selected centroids is calculated and information of a sensor node is saved in its nearest centroid. Thus, *k* initial clusters are formed in network area.
- After initial round, the position of centroid in each cluster is calculated again to check for position change from the previous one. If it is so, again calculation of Euclidean distance step repeats, else clusters are finalized.

2.3. Mobile sink node clustering approach

Data generated by an individual sensor may not appear significant, but overall data generated by sensor nodes of network area in dense environments is big. So, utilizing sink node's mobility is the concept proposed by authors in [13] to enable data gathering. The proposed technique by authors may reduce energy consumption of sensor nodes, but create additional challenges like determining sink node's trajectory and cluster formation prior to data aggregation. The mobile sink data collection process, cluster head selection problem and *mobile sink* path optimization is discussed in [14]. The mobile sink path optimization is formulated as shortest path finding problem. So, artificial bee colony algorithm is used to find optimal solution and the shortest path of mobile sink to improve data collection efficiency in network area.

2.4. Grid based Clustering approach

Grid clustering for energy optimization of sensor nodes is proposed in [15,16], where size of grid is directly related to the transmission coverage range of the node. So, as the grid size is reduced, the transmission range for sensor node also decrease and thus conserves the energy. But it may lead to more control overhead and consumes extra bandwidth. Moreover, if the grid size is too small, CH has to dissipate more energy to transmit its information to basestation. The similar simulation is done in proposed CEER protocol in the form of 2 X 2 CEER and 4 X 4 CEER protocols. But depending on concentration of network nodes in network area the concerned protocol is implemented in proposed research.

2.5. Centroid (Midpoint) based routing approach

An energy efficient clustering protocol based on K-means algorithm named Energy Efficient Clustering Protocol (EECPK-means) has been proposed by authors in [17] for WSNs where midpoint algorithm is used to improve initial centroid selection procedure. It considers residual energy and Euclidean distance as the parameter metrics for CH selection. A clustering algorithm using spatial correlation is proposed in [18], which groups sensor nodes with similar readings into one cluster and reporting the same, as the reading of the entire group. One node is selected as CH using centroid method here. The sensor node which has minimum distance to cluster centroid point is chosen to be the cluster head of the similar reading sensor group. Similarly, an energy efficient clustering protocol to prevent unbalanced clusters based on firefly and midpoint algorithms is proposed in [19] which uses residual energy and Euclidean distance as the performance metrics. It produced balanced clusters to balance CHs load and increase network lifetime. Similarly, the term *gateway node* is used in [20] to reduce the data load on cluster heads and forward the data to BS. But their assumption that every gateway node should be in the range of its neighbor gateway node in their proposed topology proves high algorithm complexity and limits user defined random sensor node deployment process. The term *energy centroid* for sensor networks is used in [21] where each cluster is designed to own 25% of sensor nodes using *distance centroid* algorithm. Here CH selection is based on *energy centroid* while the communication between CHs and the basestation is distance centric. An *energy centroid* algorithm for WSN-Assisted IoT is proposed by authors in [22] with the EECRP (Energy Efficient Centroid based Routing Protocol) algorithm when the BS is placed inside the network. This algorithm produced better results than conventional LEACH routing algorithm but no proper routing phase is described for the communication between the CHs and the BS. Moreover, the EECRP algorithm finds the CHs, and transfer the information of cluster to the BS directly in single hop. But the proposed finds the help of other ZHs to relay the information of ZH which is far from BS and thus has the provision to opt for dual hop communication to the BS. Most of the existing centroid routing approaches has following drawbacks when choosing the initial centroids randomly for CHs.

- An empty cluster.
- Residual nodes
- User desired clusters as input to algorithm.
- Unbalanced workload on the CHs
- Non-selection of optimal CHs count.
- Unsuitable for Multi-hop routing.

3. Cluster Centroid based Energy Efficient Routing (CEER) protocol

In this section, deployed sensor node network either densely or sparsely distributed in uniform random fashion is utilized to obtain optimal path from source node to basestation/sink using cluster centroid-based routing. Subsequently the radio energy dissipation model is introduced. Moreover, few assumptions are presented for better understanding of our proposed work.

3.1 Assumptions

Wireless sensor network (WSN) in CEER consists of Sink node/Basestation along with other sensor nodes. The communication established between sensor nodes and basestation is multihop based. The routing structure is hierarchy. The sensor network is grouped to zone clusters (Interzonal communication) and each cluster has four zonal quadrants (Intrazonal communication). The *cluster / zonehead* collects information from four *zone co-ordinators* which is collected from each of four *zone-quadrants* or *zone-clusters*. The cluster/zone head communicates with the basestation either in single hop (directly) or through multihop (*relaying zone heads*) communication. In the proposed work the following assumptions are made.

- Distribution of sensor nodes in network is uniform random.
- There is only one sink node located inside the sensing field. (Placed at centre of sensing field).
- Both sink and sensor nodes position is static/fixed i.e., the position of sensor nodes will not change once the sensor network arrangement is done.
- All nodes can adjust their transmission power according to distance of transmission by using free space model or multipath fading model.
- All sensor nodes deployed are isomorphic i.e., equipped with equal energy levels when deployed and has the same processing and communication capabilities.
- Sensor nodes are aware of their locations through some localization techniques i.e., the locational co-ordinates are already fed into them.
- Sensor nodes are capable of transceiving information.
- Energy cost for zone formation is on BS and no control packets for sensor nodes.
- BS has complete knowledge about the energy level and locational co-ordinates of sensor nodes in network area of *A* sq.m.

Once the sensor network formation is done, position of sensor node is not changed. The location co-ordinates of sensor nodes are fed into the node already during network deployment. We assume that every sensor node knows the basestation co-ordinates after initial broadcasting as well as their residual energy at any time. The shape of sensing area in which sensor nodes are distributed is in rectangle. Network region is constructed using cartesian co-ordinate system (xlocation, ylocation). The notations and its explanation are provided in Table 1.

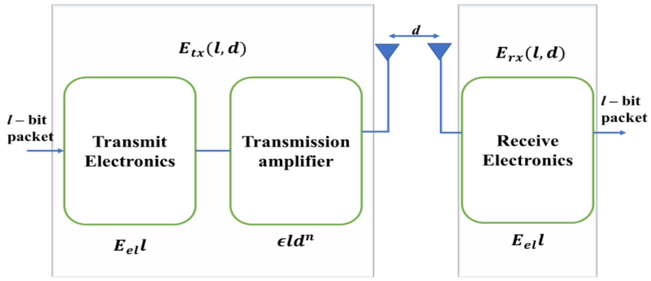
Table 1: Notations in CEER protocol and its explanation

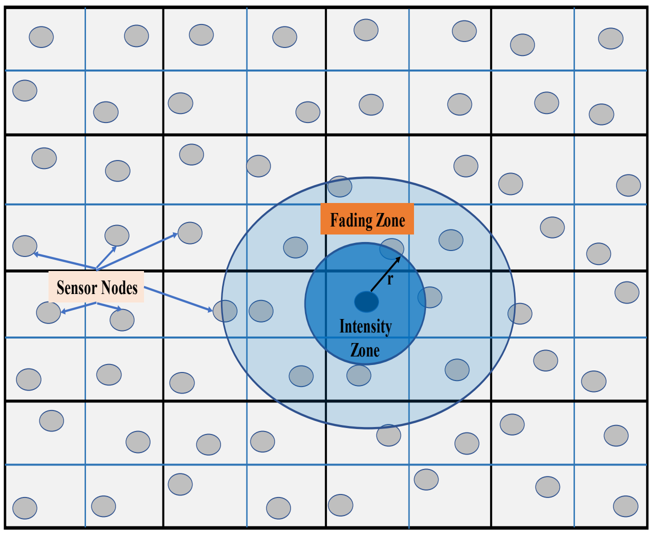
Notation	Explanation
l	*data packet*
d	*distance*
A	*Area of sensing field in Sq.m*
d_{th}	*distance threshold*
E_{tx}	*Energy of transmitter*
E_{rx}	*Energy of receiver*
E_{el}	*Energy dissipated at tx or rx*
ϵ_{fsm}	*Free space model* (d^2 *power loss*)
ϵ_{mpf}	*Multi-path Fading* (d^4 *power loss*)
BS	*Base Station*
ZH	*Zone Head*
ZC	*Zone Co-ordinator*
E_{simR}	*Energy consumed in a Simulation Round*
E_{ZH}	*Energy consumed by Zone Head*
E_{SN}	*Energy consumed by a Sensor Node*
K_i	*I Zone Count*
$E_{ZH\ to\ BS}$	*Energy consumed in transferring data from ZH to BS*
$E_{Reception}$	*Energy consumed by ZH while receiving data from ZC*
$E_{Aggregation}$	*Energy consumed by ZH in transferring l bit data packet to BS*
$d_{ZH\ to\ BS}$	*distance between ZH and BS*
$dist_{MAX}$	*Maximum broadcast distance of BS in Sensor Network of Area A sq.m*
$L_{x,y}$	*Cartesian locational co-ordinates of Sensor Node*
$d_{S\ to\ D}$	*distance from source node to destination node*
NS	*Total network size in Sq.m*
Z_S	*Zone Size in Sq.m*
Z_q	*Zone quadrant/Zone cluster size in Sq.m*
$(\overline{X_C}, \overline{Y_C})$	*Mathematical centroid co-ordinates for sensor field*
$(\overline{X_{ec}}, \overline{Y_{ec}})$	*Energy centroid co-ordinates of CEER network sensor field*

3.2 Radio Energy Dissipation Model of CEER Network

The clustering in CEER is dependent on the energy model implemented below in figure 2.

From the radio energy model shown in figure 2, the required energy to transmit *l* bit data to distance *d* can be formulated as:





$$E_{tx}(l,d) = \begin{cases} l.E_{el} + l.\epsilon_{fsm}.d^2 \text{ if } d < d_{th} \\ l.E_{el} + l.\epsilon_{mpf}.d^4 \text{ if } d \geq d_{th} \end{cases} \quad (2)$$

The energy used in receiving *l* bit packet at the receiver is formulated as Eq. (3)

$$E_{rx}(l,d) = l.E_{el} \quad (3)$$

where E_{el} is per bit dissipated energy at the transmitter or receiver, ϵ_{fsm} and ϵ_{mpf} reflect "*free-space model*" (d^2 *power loss*) and "*Multi-path fading model*" (d^4 *power loss*) conditions. As shown in Figure 3, the data communication process of sensor nodes to the *intensity zones* uses "*free-space model*" and "*Multi-path fading model* to its *fading zones*. d_{th} is the distance threshold set for both models and *d* represents the distance between source sensor to destination sensor. From Eq. (2) distance threshold may be expressed as follows in Eq. (4)

$$d_{th} = \sqrt{\frac{\epsilon_{mpf}}{\epsilon_{fsm}}} \quad (4)$$

Now consider A sq.m sensor network area with N sensors deployed in uniform random fashion and divided into *K* zones or clusters. In hierarchical approach the total utilized energy within a simulation round E_{SimR} is calculated as follows Eq. (5)

$$E_{simR} = \sum_{i=1}^{K} E_{ZH_i} + \sum_{j=1}^{K} E_{SN_j} \quad (5)$$

where E_{ZH_i} is the utilized energy by Zone Head (ZH) when it receives information from Zone Co-Ordinators and relaying information on behalf of other Zone Heads to base station along with its own zonal information to basestation. E_{SN_j} is the energy consumed by sensor nodes in zone quadrants of a zone including zone co-ordinators.

Energy consumed by zoneheads in each zone/cluster is defined as Eq. (6)

$$E_{ZH_i} = E_{(ZH\ to\ BS)_i} + E_{Reception_i} + E_{Aggregation_i} \quad (6)$$

where $E_{(ZH\ to\ BS)_i}$ is the Energy utilized when the ZH of zone *i* transfer information to basestation, $E_{Reception_i}$ is the utilized energy while receiving information from the zone co-ordinators within the zone cluster *i*, $E_{Aggregation_i}$ is the energy consumed to process *l* bit packet to the basestation by ZH. Now $E_{(ZH\ to\ BS)_i}$ is calculated as Eq. (7)

$$E_{(ZH\ to\ BS)_i} = \begin{cases} l.E_{el} + l.\epsilon_{fsm}.d^2_{to\ BS_i} \text{ if } d < d_{th} \\ l.E_{el} + l.\epsilon_{mpf}.d^4_{to\ BS_i} \text{ if } d \geq d_{th} \end{cases} \quad (7)$$

$$E_{Reception_i} = |N_i|.l.E_{el} \quad (8)$$

$$E_{Aggregation_i} = |N_i|.l.Z_{pc} \quad (9)$$

where $|N_i|$ in Eq. (8) is the node count in zone *i* and Z_{pc} in Eq. (9) is the zone path cost of a bit reporting the basestation.

From the assumptions made above, some zone heads operate in free space mode and others in amplification mode to reach basestation with the zonal information. Here zone heads using amplification mode to transmit their information uses relaying zone heads to reach basestation. Let total ZH count in both modes be *m*. Then $\sum_{i=1}^{K} E_{(ZH\ to\ BS)_i}$ is formulated as Eq. (10)

$$\sum_{i=1}^{K} E_{(ZH\ to\ BS)_i} = l\,(K.E_{el} + \epsilon_{fsm}.\sum_{i=1}^{K-m} d^2_{to\ BS_i} + \epsilon_{mpf}.\sum_{i=1}^{m} d^4_{to\ BS_i} \quad (10)$$

So, total consumed energy by all ZH's in sensing area A x A m is given by Eq. (11)

$$\sum_{i=1}^{K} E_{ZH_i} = l\left((K+N).E_{el} + N.Z_{pc} + \epsilon_{fsm}.\sum_{i=1}^{K-m} d^2_{to\ BS_i} + \epsilon_{mpf}.\sum_{i=1}^{m} d^4_{to\ BS_i}\right) \quad (11)$$

Now the Energy consumed by all sensor nodes (N-K) other than K zone heads is given by $\sum_{j=1}^{K} E_{SN_j}$ in Eq. (13)

Energy consumed by single sensor node inside a zone is given by Eq. (12)

$$E_{(SN)_j} = \begin{cases} l.E_{el} + l.\epsilon_{fsm}.d^2_{to\ BS_j} \text{ if } d < d_{th} \\ l.E_{el} + l.\epsilon_{mpf}.d^4_{to\ BS_j} \text{ if } d \geq d_{th} \end{cases} \quad (12)$$

Similar to Eq. (11) let *n* be the node count operating in both free-space and amplification modes. So,

$$\sum_{j=1}^{N-K} E_{SN_j} = l\Big((N-K)E_{el} + \epsilon_{fsm}.\sum_{j=1}^{N-K-n} d_{to\ ZH_j}^2 + \epsilon_{mpf}.\sum_{j=1}^{n} d_{to\ ZH_j}^4\Big) \quad (13)$$

From Eq. (11) and (13), the total energy consumed in a simulation round of *t* sec in proposed hierarchical approach routing is concluded in Eq. (14)

$$E_{simR} = l\Big(2.N.E_{el} + N.Z_{pc} + \epsilon_{fsm}\Big(\sum_{i=1}^{K-m} d_{to\ BS_i}^2 + \sum_{j=1}^{N-K-n} d_{to\ ZH_j}^2\Big) + \epsilon_{mpf}\Big(\sum_{i=1}^{m} d_{to\ BS_i}^4 + \sum_{j=1}^{n} d_{to\ ZH_j}^4\Big)\Big) \quad (14)$$

3.3 Zone or cluster creation and Node deployment

To implement CEER protocol zone formation plays a major role, as it reduces overhead and energy consumption in data transmission within the zone/cluster. By further dividing the zone into 4 quadrants each the energy consumption of zoneheads/cluster heads is reduced. Zone co-ordinators are created within a quadrant for collecting the entire information of quadrant and sending it to zone head. From Figure 4 one can see the formation of a zone, out of *K* zones with four equal quadrants (Q_n, $where\ n = 1,2,3,4$) displayed inside the sensor network. So, inside a 2 X 2 CEER network i.e., a sensor network formed with 2 zones in a row and 2 columns in a sensor area of A x A m, will give better results when CEER is implemented with a greater number of sensors mostly suitable for dense environment. If a CEER has to be implemented in a sensor network with lesser number of nodes then zone formation of 4 X 4 or 8 X 8 finds more suitable for application in sparse environments. Let the intersection point of all the quadrants within a zone be ZQI (Zone Quadrant Intersection) located at co-ordinates (x, y). Now zone quadrant formation algorithm is coded as Algorithm 1

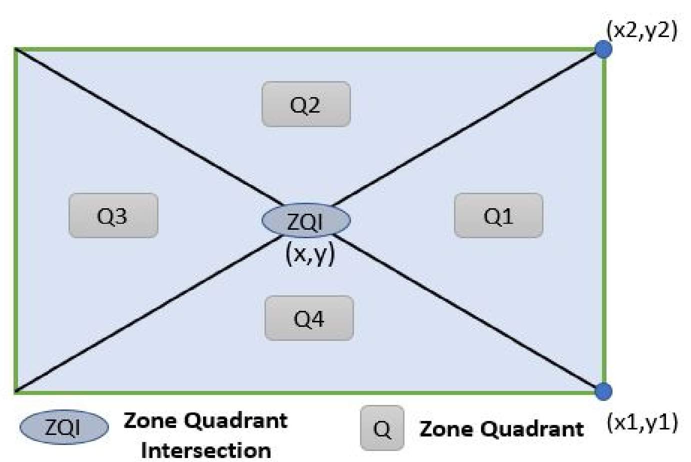


Figure 4: Zone formation in CEER Network

Algorithm 1: Zone Quadrant Formation

Step 1: theta1 = **compute** *angle* (x, y, x1, y1);

theta2 = **compute** *angle* (x, y, x2, y1);

Step 2: compute *angle* (x1, y1, x2, y2); x = x1 - x2; y = y1 - y2;

Step 3: *angle* = atan2(y, x);

Step 4: *angle* = *angle**180/pi;

Step 5: while (*angle* < 0) *angle* += 360;

Step 6: while (*angle* > 360) *angle* -= 360;

Step 7: return *angle*;

Assuming that the sensor nodes are static throughout the simulation round, they are distributed uniform randomly within the zones. In the sensor network, the locational co-ordinates of sensor nodes are obtained using the following Eqs. (15) , (16)

$$\text{X-coordinate } X_c = sensingField_{Length} * random \quad (15)$$

$$\text{Y-coordinate } Y_c = sensingField_{breadth} * random \quad (16)$$

Where $sensingField_{Length}$ and $sensingField_{breadth}$ are the network size locational paramters and *random* denotes positional deployment of node in random fashion inside sensing field and is the number between 0 and 1. Since the simulation is done through Network simulator 2 (NS2 version 2.32), nodes are deployed using *scen*(N) command where N is the Node count in Network area *A* sq.m

The distribution of energy in entire network is done equally among all sensor nodes inside sensing area. Each sensor computes the distance ($d_{S\ to\ D}$) between its own location $L_{x,y}$ to destination sensor or the Base station B_S. The distance ($d_{S\ to\ D}$) is estimated using the Pythagorean Theorem. So, finding the distance between sensors by using the coordinating points of source sensor (X_S, Y_S), and the destination sensor (X_D, Y_D) as shown in Figure 5 and then computing the distance between them is as follows in Eq. (17)

$$(d_{S\ to\ D}) = \sqrt{|X_S - X_D|^2 + |Y_S - Y_D|^2} \quad (17)$$

In R X C CEER network formed with R row zones and C column zones inside A sq.m network field, let total network size be NS = A^2m as represented in fig.5. Consider Z_S as one zone size in sq.m which is split into further equal quadrants of each size Z_q $sq.$m. Now *NS* is formulated from Eq. (18), (19)

$$Z_S = (4 * Z_q)\ \text{sq.m} \quad (18)$$

$$\text{NS} = (R * C * Z_S)\ sq.m \quad (19)$$

Suppose if the 2 X 2 CEER network is implemented in 100 sq.m network field. Z_S is 4 * Z_q = 25 sq.m. Then total network size 2 * 2* 25 = 100 sq.m. from Eq. (19)

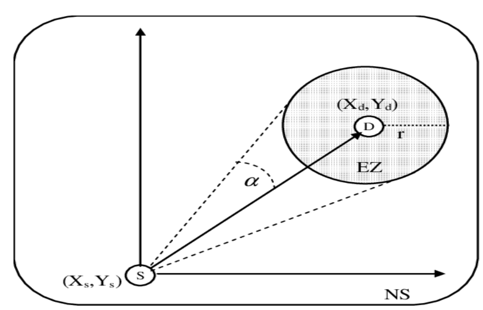


Figure 5: Expected Zone (EZ) formed in radius "r" for the destination sensor w.r.t source sensor locational co-ordinates at an angle α

When the CEER network of network size 2 X 2 is considered i.e., with 2 rows and 2 columns totally comprising of 4 zones. Each zone Z_s will occupy an equal area of 25 sq.m and so total network size NS would be 25 * 2 (Rows of zones formed) * 2 (Columns of zones formed) = 100 Sq.m

3.4 Zone Clustering Algorithm Scheme

The proposed CEER algorithm find the most appropriate zone head (ZH) node for the zone based on energy centroid. The algorithm has 3 phases: 1. Network initiation phase. 2. Zone co-ordinator selection phase and 3. Zone head selection and its replacement phase

3.4.1 Network initiation phase

The aim of this phase is to create mutual message exchange path between sensor nodes and basestation. These messages will have nodes positional, energy level information along with the average energy of deployed network. Also, information of the zone co-ordinators (ZC's) is chosen by the basestation in its first round and the longest transmission distance too. Firstly, sensor nodes will send their *position* message to the *basestation (BS)*. The packet format of *position* message is shown in Figure 6.

msg type	*Sensor ID*	*X-coordinate*	*Y-coordinate*	*Energy Level*

Figure 6: Position message packet format

The *msg type* header shows that the packet has a sensor node locational information. The *Sensor ID* is given to identify the sensor, sending its location update to base station. The position of sensor node in the sensing field is given in cartesian co-ordinate (X, Y). Here *X-coordinate* and *Y-coordinate* of sensor is sent. *Energy level* in the packet format shows the residual energy of sensor node at particular time *t*.

When this location message packet is received from all sensor nodes in the sensing field within the time limit set by timer, basestation will start to estimate the distance from itself to all the sensor nodes which have updated their location message. Consequently, the zone formation will be done with its quadrants, inside the sensing field area. Now the basestation starts identifying the *zone clusters* based on the estimated distances calculated previously. Subsequently, it updates the *node table* with the position information of node and its energy level. Then, the basestation broadcast the *ACK* (Acknowledgement) message specifically to the sensor nodes in one zone. The *ACK* packet format is shown in Figure 7

msg type	$dist_{MAX}$	*ZC ID*	*Average Energy*

Figure 7: Pack format of ACK message

msg type here is used to send the intended information to particular zone cluster nodes from *basestation / sink*. $dist_{MAX}$ field provides the information on maximum broadcast range to each node in particular *zone cluster* and is calculated by basestation and set it as a *communication threshold* for CEER network. Note that, $dist_{MAX}$ is calculated by taking the *Average Energy* from Eq. (2) *ZC ID* is needed here, so that the basestation can send the above-mentioned information to the particular *zone co-ordinator*. The *Average Energy* indicates the CEER network average energy intended for each node. The *ACK* message receiving node will update its routing table according to *ACK* packet content. At the end of this phase, the mutually exchanged information is updated in routing tables of both basestation and sensor nodes memory. Also, the routing table information is updated in real time, as the CEER network enhances to further phases.

3.4.2 Zone co-ordinators selection phase

The main aim of this phase is to select zone *co-ordinator* for every *zone cluster*. After receiving the *position* message and delivering *ACK* message, the basestation selects the node which has the *Energy Level* greater than the *Average Energy*. In the initial round, ZC selection is mostly random as the sensor nodes are equipped with equal energy levels at the deployment stage. In the consecutive rounds as the basestation updates the broadcast information, it is saved by all sensor nodes. After receiving the updated broadcasting information from basestation which has the information on node that can act as the ZC, the sensor node will check its own ID whether it is the ZC. So, if the ID matches then that node claims its election as the ZC for that transmission round. It also acts as a transceiver by receiving the information of all sensor nodes in a *zone cluster* and transmitting it to the ZH from then. Zone Head is elected later based on energy centroid in routing phase as shown in Figure 8. If suppose a sensor node finds its own ID different from BS broadcasted information, it saves the energy by activating its receiving antenna and closing the transmitter antenna to wait for further updates from the basestation through ZC.

3.4.3 Zone head selection and its replacement phase

After selecting the zone co-ordinators, the ZCs' broadcasts the schedule message with their IDs and the positional information to neighbour ZC nodes. Based on ZC ID in *ACK* message and schedule message received by neighbour nodes to determine whether they belong to that zone quadrant/cluster, zone clustering phase is completed. Now the sensor nodes inside each zone cluster will send their positional and energy level information to their ZC. Now zone coordinators which fall under zone K_i (where *i*=1, 2, …., (R*C) as shown in Figure 8) has to find the energy centroid location of the zone K_i. The ZC node which is nearest to that energy centroid will act as ZH for that zone. So, ZH will be one among the four zone coordinators of a selected zone.

By electing the ZH in the above-mentioned way, the CEER network can balance the energy consumption of networking sensor nodes inside the sensing field area. Now the CEER network satisfies the four aspects of our protocol framework. Firstly, in default simulation round, the temporary ZCs' are chosen by basestation itself, thereby gaining an overall picture of CEER network. When the simulation is processing, official ZC is selected in the zone cluster, which shows that CEER network is self-organizing. Also, one among the selected ZCs' of a zone closer to the energy centroid can act as zone head (ZH), which improves coverage of CEER network. The position of energy centroid calculation is described in next section which is done purely based on positional and residual energy information of sensor nodes inside CEER network. So, for a sensor node to become a ZH it should reach the ZC stage first and its energy level must prove worthy to become ZH after considering the closest distance parameter.

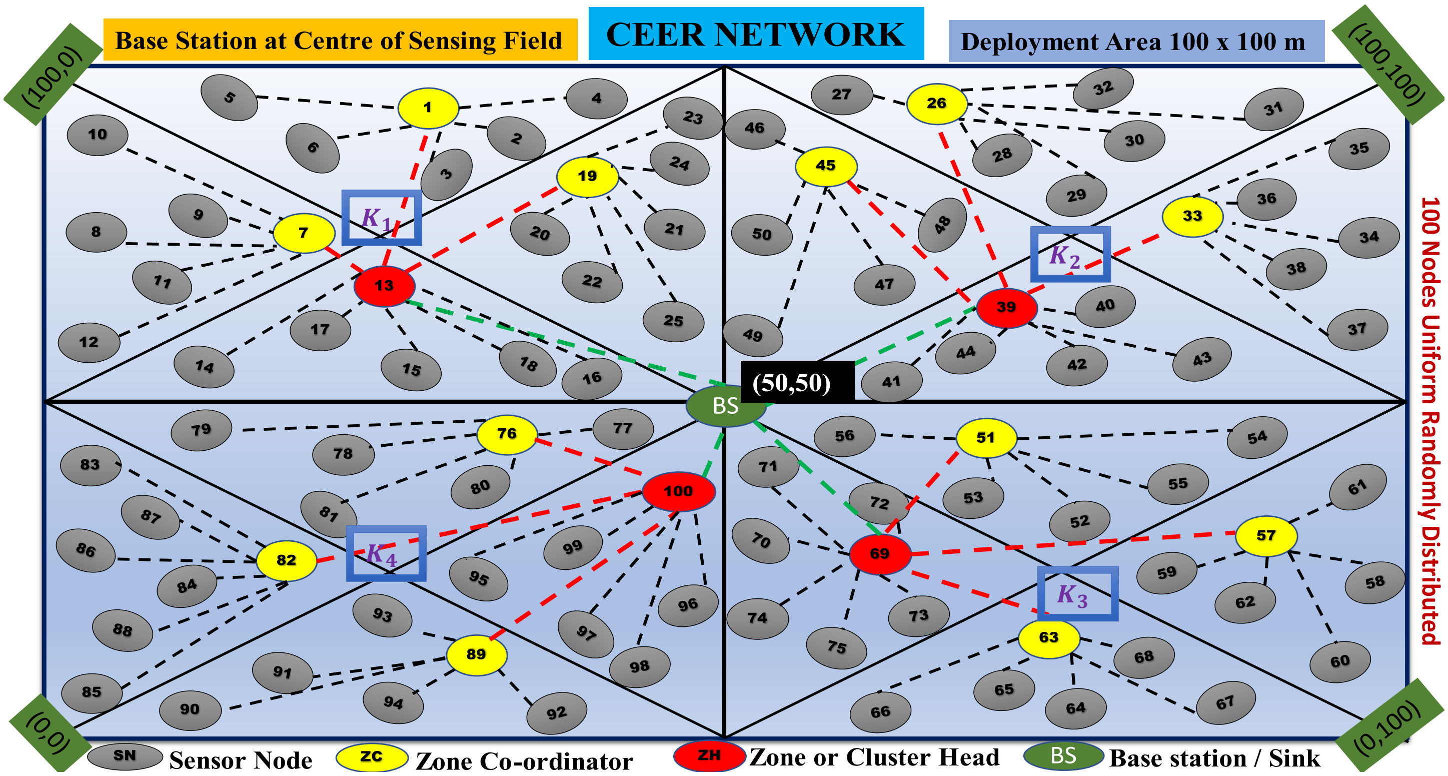


Figure 8: Zone formation and routing process in the CEER network when the basestation (BS) is placed at centre co-ordinates of Network area *A* sq.m

3.5 Zone Energy Centroid Selection Algorithm

In a concept of mathematics, centroid is the point of concurrency. Indirectly, it is the intersection point of all the medians in a geometric figure. Median is the value separating the higher half from the lower half of a data sample i.e., a population or a probability distribution. Now consider the ZC as a median of all the sensor nodes in a zone cluster. These zone coordinators from each zone quadrant are the set of medians. Now all these zone coordinators fulfil their duty to find the centroid position and which would be an imaginary point of mass concentration. Then the ZC near to that point will be selected as ZH for one simulation round. Here the concept is introduced as *Zone Energy Centroid* slightly named different from traditional mathematical centroid for few reasons. Firstly, median in terms of sensor nodes doesn't make sense. Secondly, finding the intersection point of the medians in terms of sensor operation to find zone head is meaningless. Finally, residual energy of sensor nodes is the only metric which changes in the sensor network operation. So, energy centroid in CEER network operation is to display the distribution of remaining energy of entire network.

Here, to calculate the Energy Centroid, the residual energy level and location of zone co-ordinators i is taken into consideration. In the zone cluster, if the weight of ZC is known then position of centroid is calculated using Eq. (20) and (21). In the field of mathematics, calculation for centroid in cartesian co-ordinate system $(\overline{X_C}, \overline{Y_C})$ is calculated as

$$\overline{X_C} = \frac{dM_y}{M} = \frac{\iint_{Az} x.D\, dw}{\iint_{Az} D\, dw} \tag{20}$$

$$\overline{Y_C} = \frac{dM_x}{M} = \frac{\iint_{Az} y.D\, dw}{\iint_{Az} D\, dw} \tag{21}$$

where $\overline{X_C}$ and $\overline{Y_C}$ are the results of centroid in mathematical prospective for CEER network. Here *Az* is the zone geometrical area. *dw* is the centroid weight differential. D is the density of nodes weight. dM_y and dM_x are the static moments of *x* and *y* axis respectively.

In CEER network, when the weight of ZC is unknown i.e., as the term "weight" makes no sense in ZH formation, ZC locational and energy level information is gathered to obtain position of *Energy Centroid*. When its position is obtained, ZC nearer to that *Energy Centroid* can take the responsibility of particular zone by acting as a ZH. Once ZH is selected for that particular simulation round, its locational information is updated in the routing tables of ZH itself, the basestation and the other ZH's in the network area of A sq.m by basestation. The position of centroid is given as $(\overline{X_{ec}}, \overline{Y_{ec}})$ in zone K_i as obtained from Eq. (22) and (23)

$$\overline{X_{ec}} = \frac{\sum_{p=1}^{4} \frac{RE_{ZC_p}}{E_0} .X}{N_j(p)} \tag{22}$$

where RE_{ZC_p} is the Residual Energy of Zone Co-Ordinator p in zone K_i. X and Y are the location co-ordinates of Zone Co-Ordinator p. $N_j(p)$ is the j node count in the zone cluster with zone coordinator p.

When the Zone Heads of *K zones* are ready to send packets to basestation, they estimate their own distance $d_{ZH\ to\ BS}$ to basestation placed at center of network area with $dist_{MAX}$ sent by the basestation in *ACK* packet. If $d_{ZH\ to\ BS} < dist_{MAX}$ then

ZH's will start sending the packets to base station, else ZH's will find their nearest neighbor ZH from pool of (ZH_i) $where\ i = 1,2, \ldots (K-1)$ in terms of distance, to transmit its information in maximum dual hop to the basestation. Then the condition to be checked is whether $d_{ZH_i\ to\ BS} < dist_{MAX}$ or not. If fulfilled then proceed to transmit the information to the nearest ZH and it will send that information to basestation on its behalf, else store the packets in ZH buffer queue and wait for the next round. This process of storing packets in ZH and waiting for consecutive round will be the least case, and if it has to happen there might be packet loss in rare scenario. So, a trading exists in CEER network between energy consumption and packet loss to avoid the long-distance communication of ZH nodes which may lead to more packet loss. So, the energy consumed in a simulation round of *t* sec is calculated through Eq. (14) after updating the Zone path cost Z_{pc}.

Algorithm 2: CEER Network Routing Process

Step 1: Sensor Nodes $SN_j \xrightarrow[PACKET]{\textbf{position}}$ Basestation *BS*

Step 2: Calculate *Average Energy* ***and*** $dist_{MAX}$

Step 3: *BS* $\xrightarrow[PACKET]{\textbf{ACK}} SN_J$

Step 4: ***if*** (Zone Head *ZH* prevail)

Step 5: ***then Begin*** *Zone Clustering* $\xrightarrow{Result}$ *Zone co-ordinator ZC*

Step 6: *ZCs'* $\xrightarrow[Message]{Schedule}$ neighbour *ZC* nodes,

Step 7: Recalculate *Energy Centroid* $\xrightarrow{Result}$ **updated** ZH ID

Step 8: timer is **set,**

Step 9: ***and if*** **(**Node count *N* ≥ *Zone Cluster* **count)**

Step 10: ***and if*** (pre-set **timer** time not reached ***then***

Step 11: $SN_J \xrightarrow{sensed\ data}$ *ZC (p)*, where *p = 1,2,3,4* in a **zone**

Step 12: *ZC (p)* $\xrightarrow[information]{Energy\ Centroid}$ **forward** data to *ZH* K_i

Step 13: ***if*** $(d_{ZH\ to\ BS} < dist_{MAX})$

Step 14: K_i (where *i* = 1, 2, …., (R*C)) $\xrightarrow[Data]{forward}$ *BS directly*

Step 15: ***elseif*** $(d_{ZH_i\ to\ BS} < dist_{MAX})$;

Step 16: ***then*** $\xrightarrow[Data]{forward}$ **nearest** ZH_i where *i* = (1,2, …., (*K -1*)

Step 17: ***return***

Step 18: ***end if***

Step 19: ***return***

Step 20: ***end if***

Step 21: ***return***

Step 22: ***end if***

Step 23: Recalculate *Average Energy* ***and*** $dist_{MAX}$

Step 24: *BS* $\xrightarrow[PACKET]{\textbf{ACK}} SN_J$

Step 25: Select Zone Head *ZH* nodes

Step 26: ***return***

Step 27: ***end if***

Step 28: ***return***

Step 29: ***end if***

As shown in Eq. (14), the energy consumed by CEER network in the initial stage is for the basic information exchange i.e., for sending *position* message and *ACK* message. But when the network enters the stable stage of transferring the sensed information by sensors, the energy consumed by the above messages can be ignored. As it would be very less when the routing proceeds further rounds. We know that hierarchical sensor networks are good at achieving better results in local zone clusters (in our case). So, local sensor nodes i.e., nodes inside a *zone cluster* can only communicate with their ZCs'. Also, the average transmission distance of the local sensor is purely dependent on the location of ZC inside a *zone cluster*. In most of the conventional-clustering protocols like LEACH (Low Energy Adaptive Clustering Hierarchy), the location of sensor node is not considered while choosing the cluster head (CH). As a result, when the CH is chosen it may be located at edges of the cluster and it makes the sensor node inside the cluster to consume most of its energy for communicating with the cluster head. Contrarily, the proposed CEER chooses the ZC and ZH in an entirely different way. Also, the ZC of the next round is chosen by ZC of current round based on metrics encoded in the script. The ZH is chosen as the centre of energy network to optimize energy consumption parameter.

4 Simulation Results and Discussions

The simulation tool used for simulating CEER network is NS-2 version 2.32. Network Simulator (*ns*) is a discrete event simulator targeted at networking research. Based on few sources in NS-2, few assumptions are made for the CEER simulation. The simulation is based on fact that there may be no loss of sensed data from environment by sensor nodes, since the transmission medium used is reliable by default in NS2. Since the omnidirectional antenna is used in simulation, the radio coverage will be in circular direction. The configured node uses free space model /multipath fading model as its propagation *type*.

Here the physical layer is used as wireless channel transmission medium, where the node transmits and receive information. Interface Queue type can be either Drop Tail /Priority Queue based on Routing. If the channel stays busy/channel unavailable i.e., the basestation is far reachable for ZH when $d_{ZH\ to\ BS} > dist_{MAX}$ and even neighbour Zoneheads $d_{ZH_i\ to\ BS} > dist_{MAX}$ in our case, the packet will be saved to the queue. Here the number of packets that can be hold in queue is set to 50 packets. Here the topology loaded is of Flat Grid Type. Two sensor nodes can mutually interchange their data with each other if they lie within their sensing range. But as the topology changes there may be a change in the nodes position so that the node may move out of its neighbours sensing range. Then (GOD) General Operation Director) will monitor Sensor Nodes information in simulation of Wireless Networks. But GOD needs to be re-created for every instance of simulation. GOD creation requires node count to be predefined, so that memory could be reserved for those nodes to capture information. Simulation area of CEER network is 100 x 100 m. The node configuration used for NS-2 simulation of CEER network is shown in Figure 9.

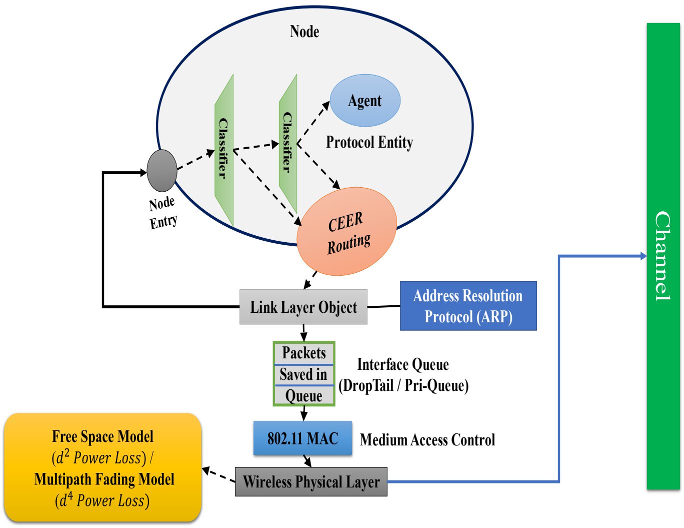


Figure 9: Node configuration of CEER network in NS2 simulator

The "*free space model"* assumes the ideal propagation condition with only one LOS (Line-of-sight) path between transmitter and receiver. It basically represents the communication range as a circle around the transmitter. The receiver receives all packets, if its location is within the circle. This model is best suitable for short distance communication. Otherwise, packets are lost. So, according to *free space model* the received signal power in free space at distance *d* from the transmitter is given as Eq. (24) [23]

$$\epsilon_{fsm} = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L} \tag{24}$$

where P_t is transmitted signal power. G_t and G_r are the transmitter and receiver antenna gain respectively. L ($L \geq 1$) is the system loss and λ is the wavelength. It is common to select $G_t = G_r = 1$ and $L = 1$ in *ns* simulations.

In "*multipath fading model"* both the LOS i.e., direct path and the obstacle reflection path are considered. This model gives more accurate prediction for long distance communication than the *free space model*. In this model the received power at distance *d* is given as Eq. (25) [24]

$$\epsilon_{mpf} = \frac{P_t G_r G_t {h_t}^2 {h_r}^2}{d^4 L} \tag{25}$$

where L = 1 in case of *multipath fading model* model.

P_t is the signal transmitted power, G_r is the receiver antenna gain, G_t is the transmitter antenna gain, h_t and h_r are the heights of transmit and receive antennas respectively.

Since *multipath fading model* doesn't provide good result for short distance communication due to oscillation caused by constructive and destructive combination of two rays, *free space model* best suits the situation when *d* is small. From Eq. (25) as the distance *d* increases the sensor node experience faster power loss. So, to utilize the benefits of both models, we calculate the distance threshold d_{th} from Eq. (4). When $d < d_{th}$ Eq. (24) is used and when $d \geq d_{th}$ Eq. (25) is used. So, distance threshold is calculated as Eq. (26)

4.1 Simulated protocols for comparison with CEER network

The simulation results of proposed CEER protocol is compared with conventional LEACH (Low-Energy Adaptive Clustering Hierarchy), Traditional PEGASIS (Power Efficient Gathering in Sensor Information Systems), Existing PRRP (Position Responsive Routing Protocol) and Pre-proposed ES3 (Energy Efficient Sink Selection Scheme). The results of CEER protocol are taken from two scenarios when R X C is 2 X 2 and 4 X 4 when the BS is placed at center of the sensing field. The evaluated performance metrics in the simulation are Total Energy Consumption of network, Average Energy Consumed by a sensor node, Packet Delivery Ratio (PDR), Dropping ratio of packet and the Network Throughput.

4.1.1 LEACH protocol (Low Energy Adaptive Clustering Hierarchy) [3-6]

Aim: To Improve the lifespan of Wireless Sensor Networks by reducing Energy Consumption in the Network.

- LEACH protocol is a TDMA based MAC Protocol.
- It is a *clustering-based* proactive routing Protocol.

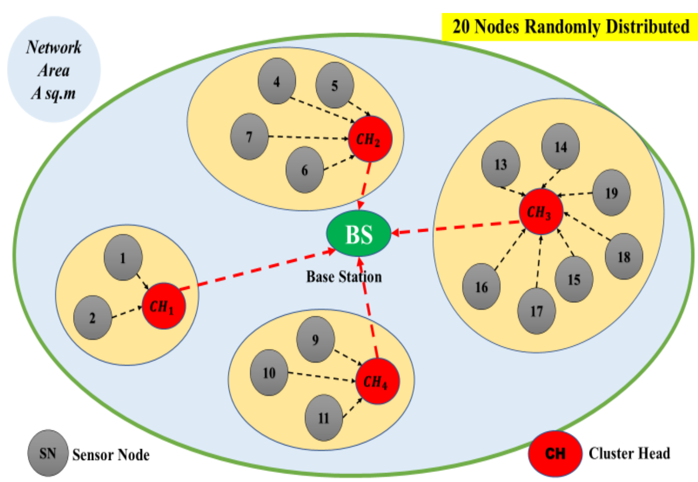


Figure 10: Routing in LEACH Protocol

Protocol Description:

- LEACH is self-organizing and adaptive clustering Protocol.
- In Figure 10 sensor nodes (SN) are organized into local clusters of their choice, where one of them will be a cluster head.
- This randomized approach is used to reduce the energy consumption of one particular sensor node in relaying the information collected from the sensor nodes to the sink.

- The responsibility of cluster head is not only to collect the information from its cluster, but to aggregate the data to be sent to base Station (BS) to reduce the quantity of messages, resulting in less consumption of energy which increases the network lifetime.
- LEACH uses randomized rotation of cluster heads to distribute the energy load among the sensor nodes deployed in the sensing field.

Advantages:

- Cluster Head (CH) aggregates the entire information which reduces the traffic load of the network.
- Since this protocol uses routing with single hop between sensor nodes to cluster heads it saves energy.
- Thus, increasing the lifetime of sensor network.
- Also, geographical location of networking node used to create a cluster is not required.
- LEACH uses no control information from BS and does not require any global knowledge of network.

Drawbacks:

- The protocol does not provide the information on the cluster head count in the network to the Basestation
- Also, if cluster head dies due to any reason the cluster become useless as there is no possibility of that particular cluster info reach the base station.
- Division of clusters is random and results in uneven distribution of clusters.
- E.g. Some clusters have more nodes and some clusters have lesser nodes.
- Some clusters heads are placed at the center of the cluster while some are in the border of cluster.
- This results in an increase in energy consumption and reduces the performance of the network.

4.1.2 PEGASIS protocol (Power Efficient Gathering in Sensor Information Systems)

Aim: To gather the information or data received from sensor nodes and transmit the same to the close neighbors in distance and iterating the leader nodes for transmitting received data to the basestation.

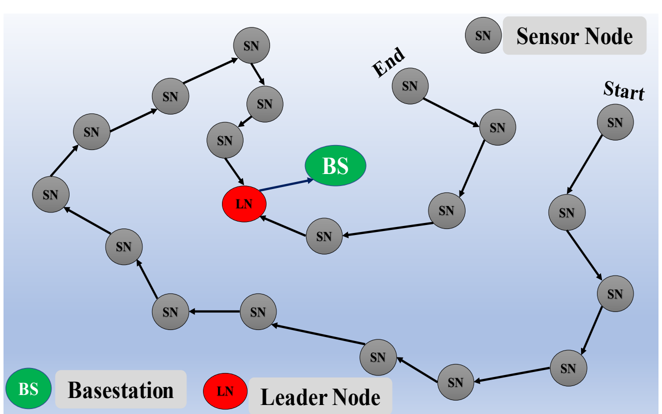


Figure 11: Routing in PEGASIS Protocol

Protocol Description:

- PEGASIS [25,26] is a near optimal chain-based routing protocol as shown in Figure 11. which is an improvement over LEACH.
- It facilitates distribution of energy load evenly among sensor nodes.
- The deployment of sensor nodes in sensing field is random and those sensor nodes arrange themselves into a chain using greedy algorithm.
- Also, the computation of this chain is done by basestation and broadcasts it to all sensor nodes.
- In PEGASIS, every sensor node receives and transmit atleast one packet in every round and be a leader in atleast once in *n* rounds and *n* is also the node count.

Advantages: (Over LEACH)

- Secondly, the leader node can receive atmost two messages from its neighbor nodes which is not the same in the case of LEACH protocol.
- For E.g. In the network of 100 nodes, cluster head (CH) in LEACH receives 20 messages at a time if the 20 nodes transfer their information at a time to the cluster head inside a cluster.
- For every round of communication one node i.e., leader node takes the entire responsibility to transmit the message to the basestation.
- In PEGASIS, data fusion helps in reducing the quantity of data being transmitted between sensor nodes and the basestation.
- Data fusion combines one or more data packets from different sensor measurements to produce a single packet for better inference.
- PEGASIS protocol is used mostly to characterize and monitor the quality of environment.

Drawbacks:

- The distance between the basestation and the leader node is not predefined and so the transmission distance between the leader node and basestation can be far.
- Energy level of the leader node is not considered for its selection.
- Since only one leader node can exist for one instance of communication, it may be the bottleneck for the network in causing delay.
- Redundant transmission (same data or information received more than once) exists as only one node takes the responsibility of transmitting entire chain network information to the basestation.

4.1.3 PRRP protocol (Position Responsive Routing Protocol)

Aim: PRRP protocol [27,28] is mainly designed to reduce the energy consumed by each node by minimizing the time the sensor node is in idle listening state and by reducing the average communication distance over the network.

Protocol Description:

Sink/Basestation builds the tree-based sensor network by broadcasting the control message containing six fields *sender, position, type, level, parent and energy*. The BS creates network tree and each node receives tree information. The protocol finds logically, how many nodes lies in a particular tier T_n as shown in Figure 12.

Energy threshold E_{th} in PRRP is calculated using Eq. (27).

$$E_{th} = lE_{el} + lE_{amp}d^2 + 8lE_{el} \quad (27)$$

Where lE_{el} is the dissipated energy in processing l bits, E_{amp} is the energy of the amplifier electronics used in simulation. Energy consumption of transmitter E_{tx} and receiver E_{rx} are given by Eq. (28) and Eq. (29) respectively.

$$E_{tx} = lE_{el} + lE_{amp}d^2 \quad (28)$$

$$E_{rx} = lE_{el} \quad (29)$$

Where lE_{el} is the dissipated energy in processing l bits, E_{amp} is the Energy of the amplifier electronics used in simulation.

The lifespan of cluster head (CH) in PRRP is calculated in terms of sampling interval number N_s, as in Eq. (30)

$$N_s = \frac{E_{in} - E_{th}}{E_t + lE_r + (T_s - (l+1)T_t)P_s} \quad (30)$$

Where E_{in} denote initial energy of sensor field. E_t is energy consumed in transmitting one data bit. T_s and T_t are sampling time and transit time respectively and P_s is sensor node power.

The sensor nodes are deployed in a random fashion across sensing field as shown in Figure 11 and are assumed that they are aware of its location using GPS or by any other location means. The gateways for the base station are the nodes within tier T_0 . Each tier is having a vertical cross-sectional area on either side of Base station. Each vertical cross-section is provided with *tier-head* which acts as a sink for all the nodes in that vertical cross-sectional tier. *Tier-heads* are selected based on node energy level, its distance from the base station.

Advantage:

- It uses TDMA (Time division multiple access technique) to transfer the data from *tier heads* to the Basestation.
- No delay constraints in data transmission.

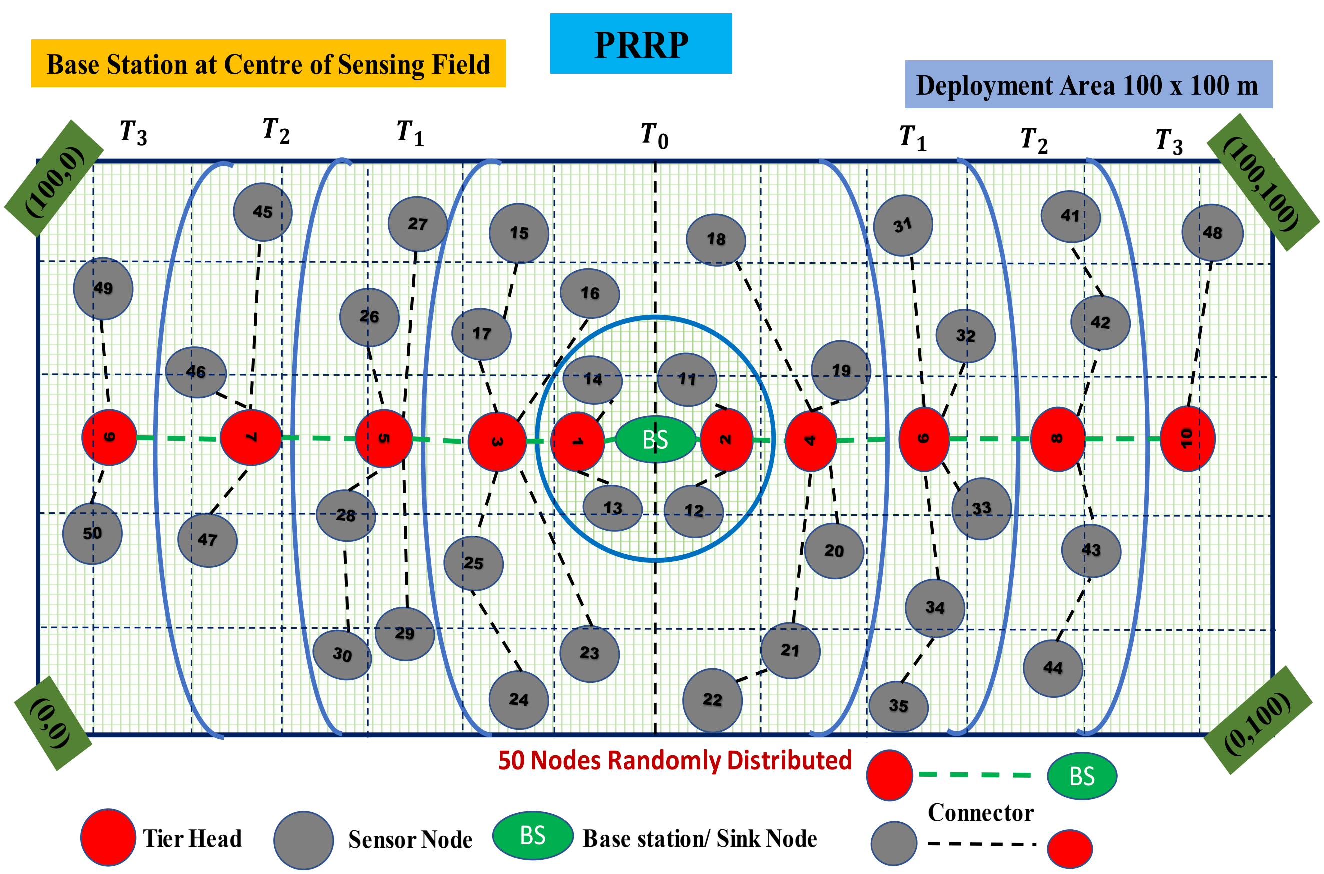


Figure 12: Tier formation and routing process in PRRP (Position Responsive Routing Protocol) when the basestation is placed at the center of the sensing field

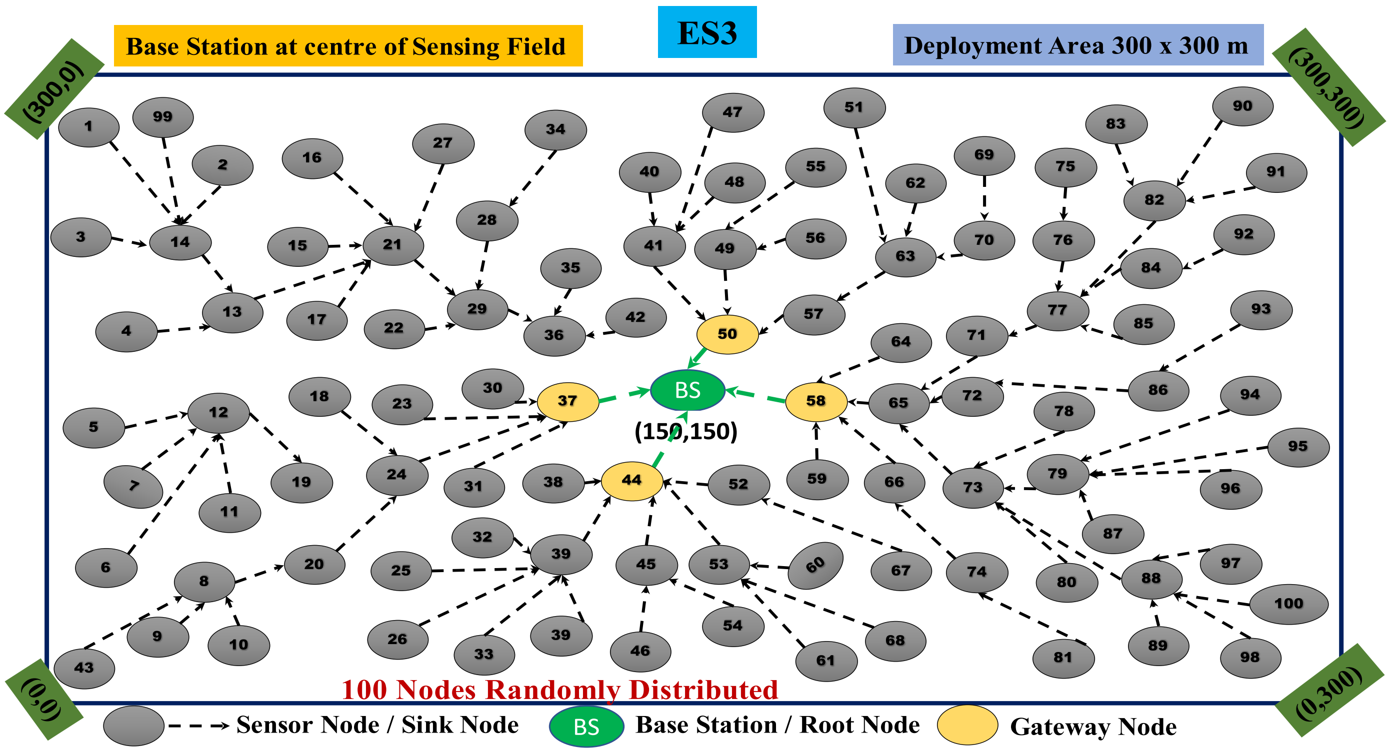


Figure 13: ES3 (Energy Efficient Sink Selection Scheme) protocol routing process when the BS is placed inside the sensing field

Drawback:

- It uses only *free-space model* (d^2 *power loss*) for propagation of sensor information.
- The energy consumption of tier heads located in tier T_3 which is far from BS will deplete its energy faster than normal and so it leads to reselection of tier heads frequently in the tiers far from BS.

4.1.4 ES3 protocol (Efficient Sink Selection Scheme)

Aim: To increase network performance and reduce average energy consumed by the node in large-scale sensor networks.

Protocol Description:

ES3 [29,30] uses *tree-based* routing. Here the nodes in ES3 network are classified into three types: Base station / Root Node, sink node/sensor node and Gateway node. Root node is the first node existed in the network with its hop zero. This node search for child nodes by sending broadcast packets. Here every sensor node can act as a sink node for relaying information of cluster of sensor nodes. Gateway nodes are utilized to handle the communication load of BS. The routing process in ES3 when BS is placed at center of sensing field is shown in Figure 13.

The energy model used in ES3 is same as in CEER network as explained from Eq. (2) and Eq. (3). The weight of sensor node when acting as a sink, differs from its calculation of CEER network as shown in Eq. (31) [31]

$$W_{S_n} = \frac{\rho}{dist_m} + \frac{\vartheta}{C_m+1} + \sigma R_m + \frac{\tau}{HoP_m+1} \quad (31)$$

Here, W_{S_n} is the weight of sink node S_n. C_m is the current count of child nodes connected to sink node S_n. R_m is sink node S_n left over energy. $dist_m$ is the distance between current node and Sink node S_n . We define the root node hop to be 0 and gate way nodes hop by default is set to 1. HoP_m is the hop of sink node S_n. ρ, ϑ, σ, and τ are standardized parameters of four variables $dist_m$, C_m , R_m and HoP_m and are set at values 15, 6, $^1/_{20}$, 11 respectively during the simulation.

Advantages:

- Suitable for Large-scale sensor networks
- Provides balance between various network performance metrics like Energy consumption, communication distance, hop count and child node count.

Table 2: Comparison between the PEGASIS, PRRP, ES3 and CEER protocols

	PEGASIS	PRRP	ES3	CEER
Scalability (Increasing number of nodes after network establishment)	very low	Low	Medium	High
Transmission delays	Very high, since it uses a simple control token passing approach	Medium, as only the tier heads of the extreme tiers have to use the intermediate tier heads to relay the information to BS	Quite high, as no proper clustering structing is maintained during the data transmission phase	Quite low, as the local quadrant information is routed directly to zone co-ordinators and one among them will take the responsibility to relay the information to basestation
Distribution of sensor Nodes in sensing field	Randomly Distributed	Uniform Random distribution	Randomly Distributed	Uniform Random Distribution
Control Message Overhead	Medium	Low as the network area is divided into tiers and overhead is added to data when the information is relayed through neighbour tier heads	Medium	Very low as the overhead is added to the data only when an entire zonal information has to be relayed through neighbour zonal heads in single or dual hop to reach the basestation
Uniform Distribution of Energy	High, when implemented will lesser node count as it has cain network architecture	Medium, due to its tier-based network architecture as the extreme right or extreme left tier heads faces few problems	Quite High, as it approach is tree based hierarchical approach	Quite High, as the network area is uniformly divided into equal zones and further into equal quadrants
Energy Efficiency	Low	Low as the sensor nodes in extreme tier heads has to dissipate more energy.	Very high, as it balances distance, hop, number of child nodes and residual energy	Very high, as it uses centroid clustering approach to find the best zonal head of a particular zone
Inter cluster/ quadrant structure	1-Hop	1-Hop	Multihop	1-Hop
Algorithm Complexity	High	Medium	High	Medium

Table 3: Simulation Parameter Table of ES3, PRRP and CEER networking protocols

Parameter	**Value**		
Protocol	**PRRP**	**ES3**	**CEER**
Network area (m)	100 × 100 m	100 × 100 m	100 × 100 m
Number of nodes	100	100	100
Sensor deployment	Uniform Random	Random	Uniform Random
Location of the sink	Centre of the Sensing Field	Centre of the Sensing Field	Centre of the Sensing Field
Channel Access Mechanism	MAC IEEE 802.11	MAC IEEE 802.11	MAC IEEE 802.11
Propagation Model	Free space Model	Free space/Multipath fading	Free space/Multipath fading
Total Simulation Time (sec)	200 Sec	200 sec	200 Sec
Initial Energy of each Node (J)	2 Joule	1.5 Joule	2 Joule
Packet Size (*l*)	64 Bytes	64 Bytes	64 Bytes
Transmitting/Reception Energy	400nJ/byte	*0.4μJ/byte*	50*nJ/bit*
ϵ_{fsm} (d^2 *power loss*) (Amplifier Energy) E_{amp}	10pJ/bit/m^2	$0.01 nJ/bit/m^2$	$10 pJ/bit/m^2$
ϵ_{mpf} (d^4 *power loss*)	NA	$1.3 nJ/bit/m^2$	$1.3 nJ/bit/m^2$
Energy Threshold (initial) E_{th}	1.5J	*0.75J*	1*J*

E_{th} (For one simulation Round)	0.4 mJ	0.4 mJ	0.4 mJ
Bandwidth	0.2 Mbps	0.2 Mbps	0.2 Mbps
Beam Forming Energy (*nJ/bit*)	5	5	5
Capture Threshold	10	10	10
Carrier Sense Threshold	3.652e-8 i.e., 250 m	22.427e-8 i.e., 50 m	3.652e-8 i.e., 250 m
Receiver Threshold	3.652e-8 i.e., 250 m	22.427e-8 i.e., 50 m	3.652e-8 i.e., 250 m
Interface Queue Length for Node	50 Packets	2 Packets	50 Packets
Receiving Power of Node	0.1 W	0.395 W	0.1 W
Transmitting Power of Node	0.2 W	0.66 W	0.2 W
Route Timeout	10 sec	10 sec	10 sec
Active Route Timeout	10 sec	10 sec	10 sec
Reverse Route Life	6 sec	6 sec	6 sec
Broadcast ID save	Up to 6 Times	Up to 6 Times	Up to 6 Times
Route Request Retries	3 Times	3 Times	3 Times
Maximum Route Request Timeout	10 Sec	10 Sec	10 Sec
Node Traversal Time	0.03 Sec	0.03 Sec	0.03 Sec
Local Repair Wait Time	0.15 Sec	0.15 Sec	0.15 Sec
Network Diameter	30 Hops	10 Hops	30 Hops
Route Reply Wait Time	1 Sec	1 Sec	1 Sec
Address Resolution Protocol Delay	0.01 Sec	0.01 Sec	0.01 Sec
Hello Interval	5 Sec	5 Sec	5 Sec
Allowed Hello Loss	3 Sec	3 Sec	3 Sec
Bad Link Lifetime	3 Sec	3 Sec	3 Sec
Max Hello Interval	(1.25 * Hello Interval)	(1.25 * Hello Interval)	(1.25 * Hello Interval)
Min Hello Interval	(0.75 * Hello Interval)	(0.75 * Hello Interval)	(0.75 * Hello Interval)
RTQ Maximum Length	64 Packets	64 Packets	64 Packets
RTQ Timeout	30 Sec	30 Sec	30 Sec
Max Errors Allowed	100	100	100

Drawback:

- Delay in data transmission due to more control overheads and normalized overheads added to sensed data transfer process.

4.2 Simulation performance metrics with results

The simulation parameters and their values used in NS-2 simulation of protocols PRRP, ES3 and CEER are displayed in Table 3 and their comparison results with conventional LEACH and traditional PEGASIS are shown in graphs under respective sections below.

4.2.1 Comparison in terms of Total Energy Consumption of Sensor Network

Augmenting the energy balance in sensor nodes minimizes the energy dissipation during network communications. Total energy consumption of sensor network is the energy consumed to perform transmission, reception and data aggregation.

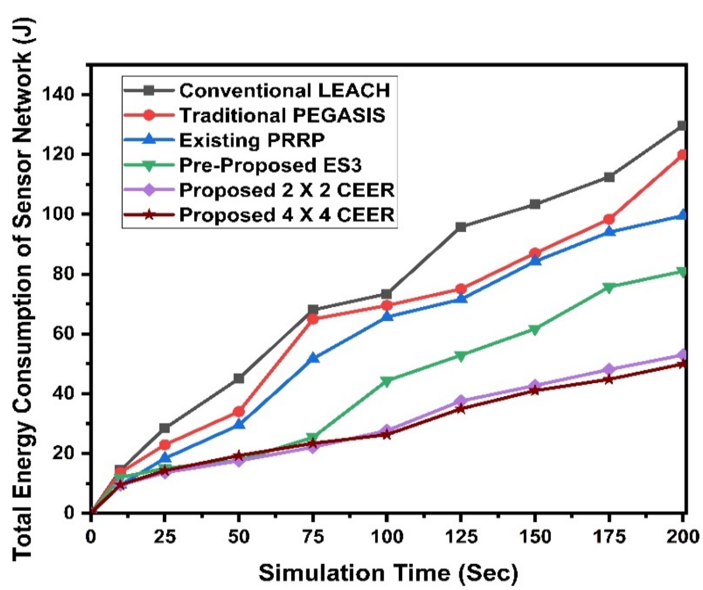


Figure 14: Total energy consumed by proposed and pre-existing sensor networks

in a simulation round of t = 200 Sec when BS is placed at centre of sensing area with co-ordinates (50,50)

When the protocols LEACH, PEGASIS, PRRP, ES3 and CEER are simulated under the same platform of NS-2 version 2.32 for a simulation round of t = 200 Sec, CEER network proved to consume less energy than the other network as shown in Figure 14. The energy saving table of CEER networks when compared to others is shown in Table 4.

Table 4: Outperformance by CEER network in terms of Energy at the end of simulation round

	Comparison with ES3	Comparison with PRRP	Comparison with PEGASIS	Comparison with LEACH
2 X 2 CEER	Saves 28 J	Saves 46.627 J	Saves 66.9947 J	Saves 76.80 J
4 X 4 CEER	Saves 31 J	Saves 49.632 J	Saves 70 J	Saves 79.80 J

4.2.2 Comparison in terms of Average Energy Consumption of Sensor Node in Sensing Field [32,33]

The performance metric "Average Energy consumption of sensor node" is vital in CEER as it determines its network lifetime. In any other protocol this metric has its importance as it shows how efficient is the proposed or implemented algorithm to compete with the existing works. In our scenario, when this metric result is compared with existing works as shown in Figure 15, CEER network saves energy consumed by a sensor node as its algorithm finds energy effective sink for every sensor node deployed in the network area and transmit the data through it to reach the basestation. Indirectly, the energy load is distributed equally and effectively among all the sensor nodes in the network.

From Figure 15, CEER network shows that a sensor node deployed inside its network can have longer lifespan i.e., around 62% than conventional LEACH, 56% than traditional PEGASIS, 38% than existing PRRP and 22% than our pre-proposed ES3.

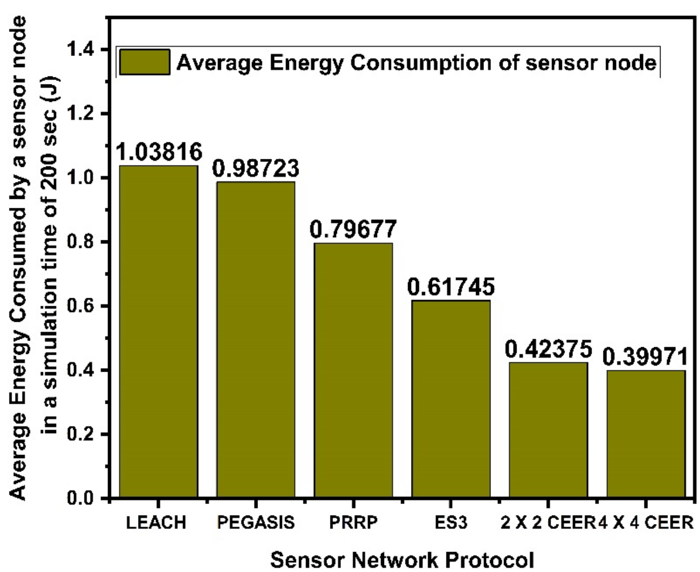


Figure 15: Average Energy consumed by a sensor node in simulated network protocols at the end of simulation time of t = 200 sec when the basestation is placed at co-ordinates (50,50) in the network area.

4.2.3 Comparison in terms of packet delivery ratio (PDR) of Sensor Network

Packet Delivery Ratio (PDR) is defined as the percentage ratio of packets received by the basestation and the generated packets by the source sensors. In *ns* terms it is defined as the ratio between the client of packages originated by the "application layer" CBR (Constant Bit Rate) sources and the package count received by CBR sink (BS) at the final destination.

$$Packet\ Delivery\ Ratio\ (PDR) = \frac{(\sum Total\ packets\ received\ by\ BS\ at\ time\ t) * 100}{\sum Total\ packets\ sensed\ by\ sensor\ nodes\ in\ network\ at\ time\ t} \quad (32)$$

PDR in the CEER network is high as this hierarchical protocol is constructed with stable routing operational procedure. As shown in Figure 16, at the beginning of the network initialization and initial data transfer phase, the PDR of CEER network might be low, but as the network enters the stable operation phase as mentioned in section 3, the PDR is quite good and more as compared to other simulated works under simulation time t and BS at the centre of the sensing field.

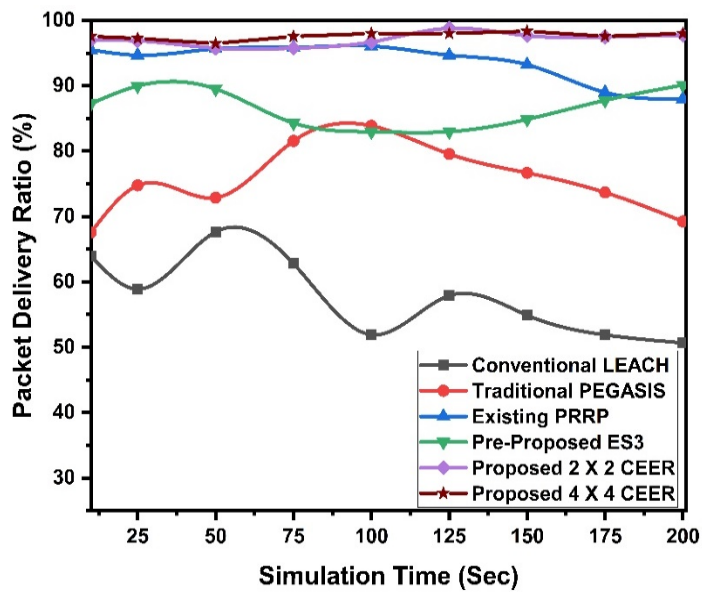


Figure 16: Packet delivery ratio (PDR) of simulated sensor networks in a simulation time t = 200 Sec, when the BS is placed at centre of sensing field at location co-ordinates (50,50)

4.2.4 Comparison in terms of packet loss percentage in Sensing Field

Packet loss rate is calculated in wireless sensor networks to check the reliability of a communication network path. This packet loss percentage is equal to the unreceived packets by the BS to the total packets sensed by sensor nodes in the network. Packet loss occurs due to failure of sensor node that carries data across the sensing area of the network.

$$Packet\ loss\ (P.L) = \begin{pmatrix} Total\ packets\ sensed\ by \\ sensor\ nodes\ in\ network \\ in\ simulation\ time\ t \end{pmatrix} - \begin{pmatrix} Total\ Packets\ received \\ by\ BS \\ in\ simulation\ time\ t \end{pmatrix} \quad (33)$$

Packet Loss Percentage =

$$\frac{(P.L*100)}{\text{Packets sensed by sensor nodes in network in simulation time } t} \quad (34)$$

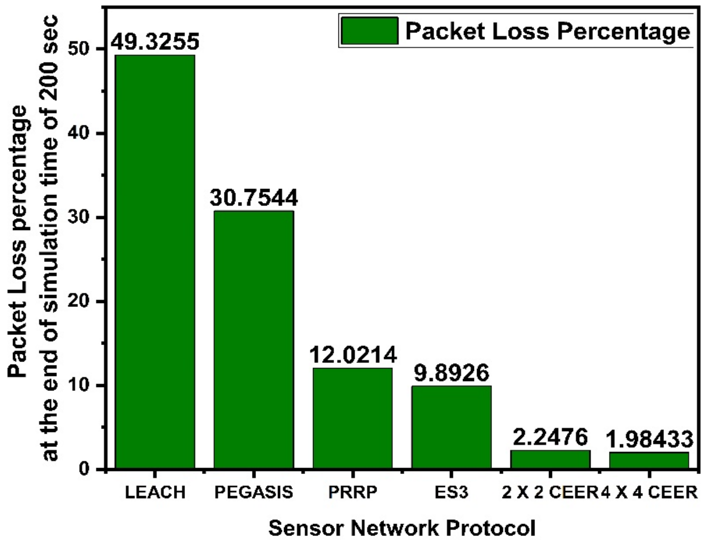


Figure 17: Packet Loss Percentage of sensor networks simulated in time $t = 200$ sec

Results from Figure 17 shows acceptable packet loss for CEER networks. Reliable protocols react to packet loss automatically and so as the CEER network as its nodes save the excess packets received in one transmission round in the queue memory to avoid further packet loss. Moreover, when compared to conventional LEACH and traditional PEGASIS protocols, CEER network packet loss is more than 47%, 28% lesser respectively. Similarly, when compared to existing PRRP protocol, it is around 10% lesser.

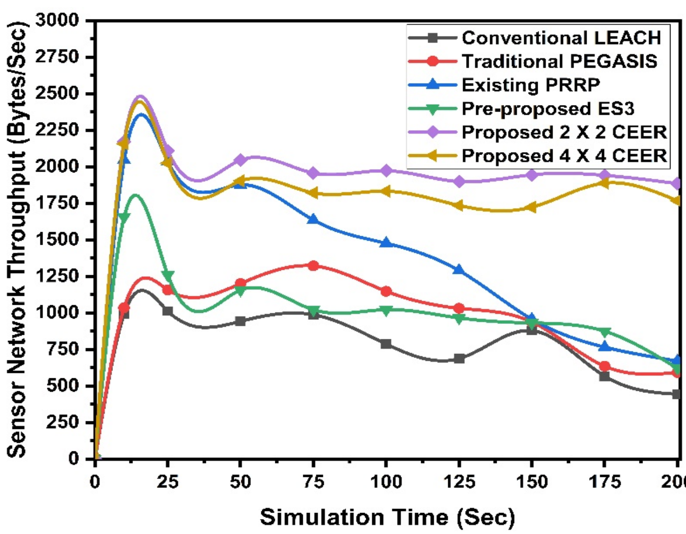


Figure 18: Sensor network throughput measured in bytes/sec of the protocols simulated in a simulation time of $t = 200$ sec when the BS is placed at locational co-ordinates (50,50)

4.2.5 Comparison in terms of Sensor Network Throughput

Throughput is the rate at which information is sent through the network. Network throughput in CEER networks is an amount of data transferred successfully from source sensors to base station at time *t* sec and its result graph is shown in Figure 18. High ratio of unsuccessful packet delivery leads to lower throughput and reduces the network lifespan. Throughput capacity is usually measured in *bits/sec*.

*Throughput (bits/sec) = (Packet count successfully received at BS * packet size)/Total time t spent in delivering those packets from source sensor to base station* (35)

Packet loss, latency and jitter are all related to reduce the sensor network throughput. All these three metrics are proved to be lesser in CEER network than the compared protocols when simulated based on the assumptions made in section 3.

5 Conclusions and Further proceedings

In this research work, we propose a clustering centroid-based energy efficient routing (CEER) protocol to solve the problem of forming zones and its clusters to transmit information energy efficiently to the basestation or sink node. An optimized path is achieved through CEER networks in the zones formed, by selecting energy efficient zone co-ordinators and zone heads. Energy centroid routing is proposed in this paper by considering the residual energy of sensor nodes and communication distance to the sink/basestation. The data path between zone heads and basestation is optimized at both one-hop level and two-hop levels by relaying the information to basestation via intermediate zone heads to avoid packet loss at the zone heads and sometimes directly to basestation itself if the transmission distance between the zone head and the basestation is less. From the simulation results, when the basestation is placed at center of the sensing field, CEER networks can transmit an ample information by consuming less amount of energy. Its simulation results proved that the average energy consumption of a sensor node inside CEER network consumes energy i.e., atleast 60% and 55% less than the conventional LEACH and traditional PEGASIS protocols respectively and atmost 35% than the existing PRRP protocol.

In future work, we have planned to optimize the path between the Zone Heads (ZH) and the Basestation (BS) by introducing multi-hop communication between them unlike the current dual hop and to implement it in IoTN scenario when the BS is placed outside the sensing field at the co-ordinates (50,125). In case if the basestation is still far reachable that the ZH has to dissipate most of its energy for data transfer, we planned to optimize the path using ANT routing strategy to avoid the packet loss.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] N. P. R. Kumar and G. J. Bala, "An Energy Efficient Quadrant Based Position Responsive Routing Protocol," 2019 2nd International Conference on Signal Processing and Communication (ICSPC), Coimbatore, India, 5-10, 2019. https://doi.org/10.1109/ICSPC46172.2019.8976781.

[2] H. Ali et al., "Clustering methods for cluster-based routing protocols in wireless sensor networks: Comparative study", Comparative study. International Journal of Applied Engineering Research, **12**(21), 11350-11360, 2017. http://www.ripublication.com

[3] W. Xinhua and W. Sheng, "Performance Comparison of LEACH and LEACH-C Protocols by NS2," 2010 Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science, Hong Kong, 254-258, 2010. https://doi.org/10.1109/DCABES.2010.58.

[4] N. Wang and H. Zhu, "An Energy Efficient Algorithm Based on LEACH Protocol," 2012 International Conference on Computer Science and Electronics Engineering, Hangzhou, 339-342, 2012. https://doi.org/10.1109/ICCSEE.2012.150.

[5] L. Li and C. Liu, "An Improved Algorithm of LEACH Routing Protocol in Wireless Sensor Networks," 2014 8th International Conference on Future Generation Communication and Networking, Haikou, 45-48, 2014. https://doi.org/10.1109/FGCN.2014.18.

[6] S. Gupta and N. Marriwala, "Improved distance energy based LEACH protocol for cluster head election in wireless sensor networks," 2017 4th International Conference on Signal Processing, Computing and Control (ISPCC), Solan, 91-96, 2017. https://doi.org/10.1109/ISPCC.2017.8269656.

[7] R. Dasondhi, M. Singh, D. Kulhare, "An Algorithm for Balanced Cost Cluster-Heads Selection for Wireless Sensor Network" International Journal of Engineering Research & Technology, **01**(10), 2012. IJERTV1IS10526

[8] N. Enami, R. A. Moghadam and K. D. Ahmadi, "A new neural network based energy efficient clustering protocol for Wireless Sensor Networks," 5th International Conference on Computer Sciences and Convergence Information Technology, Seoul, 40-45, 2010. https://doi.org/10.1109/ICCIT.2010.5711026.

[9] S. Mostafavi, V. Hakami, "A new rank-order clustering algorithm for prolonging the lifetime of wireless sensor networks" Int J Commun Syst., 33(e4313), 2020, https://doi.org/10.1002/dac.4313

[10] P. Sasikumar and S. Khara, "K-Means Clustering in Wireless Sensor Networks," *2012* Fourth International Conference on Computational Intelligence and Communication Networks, Mathura, 140-144, 2012. https://doi.org/10.1109/CICN.2012.136.

[11] W. Fakhet, S. E. Khediri, A. Dallali and A. Kachouri, "New K-means algorithm for clustering in wireless sensor networks," 2017 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC), Gafsa, 67-71, 2017. https://doi.org/10.1109/IINTEC.2017.8325915.

[12] Jorio A. Sanaa El Fkihi, Brahim Elbhiri, and Driss Aboutajdine "An Energy-Efficient Clustering Routing Algorithm Based on Geographic Position and Residual Energy for Wireless Sensor Network" Journal of Computer Networks and Communications, 2015, https://doi.org/11.10.1155/2015/170138.

[13] D. Takaishi, H. Nishiyama, N. Kato and R. Miura, "Toward Energy Efficient Big Data Gathering in Densely Distributed Sensor Networks" IEEE Transactions on Emerging Topics in Computing,, **2**(3), 388-397, 2014. https://doi.org/10.1109/TETC.2014.2318177.

[14] R. Vijayashree and C. Suresh Ghana Dhas., "Energy efficient data collection with multiple mobile sink using artificial bee colony algorithm in large-scale WSN", Automatika, **60**(5), 555-563, 2019, https://doi.org/10.1080/00051144.2019.1666548

[15] G. Kumar and J. Singh, "Energy efficient clustering scheme based on grid optimization using genetic algorithm for wireless sensor networks," 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, 1-5, 2013. https://doi.org/10.1109/ICCCNT.2013.6726634.

[16] N. Thakur and R. K. Chauhan, "Conservation of energy by using grid clustering in wireless sensor networks," 2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), Waknaghat, 591-596, 2016. https://doi.org/10.1109/PDGC.2016.7913192.

[17] A. Ray and D. De, "Energy efficient clustering protocol based on K-means (EECPK-means)-midpoint algorithm for enhanced network lifetime in wireless sensor network," IET Wireless Sensor Systems, **6**(6), 181-191, 2016. https://doi.org/10.1049/iet-wss.2015.0087.

[18] A. A. Shaikh and D. J. Pete, "Spatial Correlation and Centroid Based Clustering in Wireless Sensor Network," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 1-5, 2018. https://doi.org/10.1109/ICCUBEA.2018.8697416.

[19] R. Daniel and K. N. Rao, "EEC-FM: Energy Efficient Clustering based on Firefly and Midpoint Algorithms in Wireless Sensor Network," KSII Transactions on Internet and Information Systems, **12**(8), 3683-3703, 2018. https://doi.org/10.3837/tiis.2018.08.008

[20] Q. Kashif et al., "Optimized Cluster-Based Dynamic Energy-Aware Routing Protocol for Wireless Sensor Networks in Agriculture Precision", Journal of Sensors. 1-19, 2020. https://doi.org/10.1155/2020/9040395.

[21] S. Loganathan, J. Arumugam, "Energy centroid clustering algorithm to enhance the network lifetime of wireless sensor networks"., Multidim Syst Sign Process 2019. https://doi.org/10.1007/s11045-019-00687-y

[22] J. Shen, A. Wang, C. Wang, P. C. K. Hung and C. Lai, "An Efficient Centroid-Based Routing Protocol for Energy Management in WSN-Assisted IoT," IEEE Access, 5, 18469-18479, 2017. https://doi.org/10.1109/ACCESS.2017.2749606.

[23] H. T. Friis, A note on a simple transmission formula.Proc. IRE, 34, 1946.

[24] T. S. Rappaport, Wireless communications, principles and practice.Prentice Hall, 1996.

[25] R. Hetal, V. Sangeeta, A. Mohammad, "Comparative Study of PEGASIS Protocols in Wireless Sensor Network. IOSR Journal of Computer Engineering," 16, 25-30, 2014. https://doi.org/10.9790/0661-16512530

[26] R. K. Yadav and A. Singh, "Comparative study of PEGASIS based protocols in wireless sensor networks," 1st India International Conference on Information Processing (IICIP), Delhi, 1-5, 2016. https://doi.org/10.1109/IICIP.2016.7975320.

[27] N. Zaman, A.B. Abdullah, "Energy Optimization through Position Responsive Routing Protocol (PRRP) in Wireless Sensor Network," International Journal of Information and Electronics Engineering, **2**(5), 748-751, 2012. https://doi.org/10.7763/IJIEE.2012.V2.199

[28] N. Zaman and A. B. Abdullah, "Position Responsive Routing Protocol (PRRP)," 13th International Conference on Advanced Communication Technology (ICACT2011), Seoul, 644-648, 2011.

[29] T. Qiu, X. Liu, L. Feng, Y. Zhou and K. Zheng, "An Efficient Tree-Based Self-Organizing Protocol for Internet of Things," IEEE Access, 4, 3535-3546, 2016. https://doi.org/10.1109/ACCESS.2016.2578298.

[30] Nalluri, Raj Kumar & Bala, G. "An Efficient Energy Saving Sink Selection Scheme with the Best Base Station Placement Strategy Using Tree Based Self Organizing Protocol for IoT", Wireless Personal Communications **109**(2), 869-895, 2019. https://doi.org/10.1007/s11277-019-06595-5.

[31] H. Zhang, P. Chen and S. Gong, "Weighted spanning tree clustering routing algorithm based on LEACH," 2010 2nd International Conference on Future Computer and Communication, Wuha, 223-227, 2010. https://doi.org/10.1109/ICFCC.2010.5497381.

[32] S. Krit & L. Elmaimouni, "Energy consumption in wireless sensor network: simulation and comparative study of flat and hierarchical routing protocols"., IADIS International Journal on Computer Science and Information Systems. 12, 109-125, 2017.

[33] H. Oudani, J. Laassiri, S. Krit and L. El Maimouni, "Comparative study and simulation of flat and hierarchical routing protocols for wireless sensor network," 2016 International Conference on Engineering & MIS (ICEMIS), Agadir, 1-9, 2016. https://doi.org/10.1109/ICEMIS.2016.7745357.