A multi-tier based clustering framework for scalable and energy efficient WSN-assisted IoT network

Anurag Shukla[1] · Sarsij Tripathi[1]

Abstract

The effectiveness of wireless sensor network (WSN) in Internet of Thing (IoT) based large scale application depends on the deployment method along with the routing protocol. The sensor nodes are an important component of WSN-assisted IoT network running on limited and non-rechargeable energy resource. The performance of WSN-assisted IoT is decreased, when network is deployed at large area. So, developing robust and energy-efficient routing protocol is a challenging task to prolong the network lifetime. In contrast to the state-of-the-art techniques this paper introduces Scalable and energy efficient routing protocol (SEEP). SEEP leverages the multi-hop hierarchical routing scheme to minimize the energy consumption. To achieve scalable and energy efficient network, SEEP employs a multi-tier based clustering framework. The network area in SEEP is divided into various zones with the help of proposed subarea division algorithm. The number of zones in the network are increased as the network size increases to avoid long-distance communication. Every zone is divided into certain number of clusters (sub-zones) and the number of clusters are increased towards the base station, whereas the zone width is decreased. In every cluster, some of the optimal nodes are promoted as a Relay Node (RN) and Cluster Head (CH). Normal nodes send their sensed data to the base station via local RN and CH in a multi-hop way. Furthermore, propose protocol provides a trade-off between distance and energy to prolong the network lifetime. In the proposed framework, static and mobile scenarios have been considered by applying Random walk and Random waypoint model for node mobility in simulation to make it more realistic as the various application of WSN-assisted IoT. The effectiveness of SEEP is examined against LEACH, M-LEACH, EA-CRP, TDEEC, DEEC, SEP, and MIEEPB, and result indicates that SEEP performs better for different network metrics: network lifetime, scalability, and energy efficiency.

Keywords Internet of Things · Hierarchical framework · Mobility · Random waypoint model · Random walk · Energy depletion · Network lifetime · Scalability

1 Introduction

In recent times, the area of microelectronics has become more advanced and lead the research and development on low-cost wireless sensor nodes, resource constraint devices, and tiny nodes. Wireless sensor network (WSN) plays a vital role in many WSN-assisted IoT applications [1]. WSN-assisted IoT has a broad range of applications that includes a smart parking system, industrial wireless network, healthcare monitoring system, border security monitoring, and animal monitoring system [2, 3]. The sensor nodes, actuation, radio frequency identification (RFID), and many more devices are combined and form a WSN-assisted IoT network to obtain some common objective. The participate nodes in WSN-assisted IoT make the physical object conscious about the various real attributes in the deployed network, such as hearing, monitoring, feeling, and trigger an event with the coordination of other devices [4]. The generated sense data in a network are transmitted to the Base station (BS) by the routing protocol. Most of the participating devices in WSN-assisted IoT networks are run on limited non-rechargeable energy sources such as a battery. Generally, WSN-assisted IoT

✉ Anurag Shukla
anurag672@yahoo.com

Sarsij Tripathi
stripathi.cs@nitrr.ac.in

[1] Department of Computer Science, NIT Raipur, GE Road, Raipur 492010, India

applications deployed in the rugged and dense environment, and it is complicated to replace or add the energy sources in the sensor. Therefore, efficient energy management is a critical issue in WSN-assisted IoT [5, 6].

The WSN-assisted IoT has brought many applications and makes daily human life more convenient. The deployed network in some of the WSN-assisted IoT applications may remain static or dynamic, depending on the application requirements. For example, border monitoring and environment monitoring supports static network whereas, smart healthcare monitoring system and animal monitoring system support dynamic network. In a dynamic network, nodes move freely from one location to another. The movement of the nodes varies with respect to the time, along with their velocity, pause time, and acceleration. Therefore, in the past various mobility models were proposed to mimic the user mobility patterns, and their results can vary with respect to each other. By review of earlier reported work, we found that Random walk and Random waypoint is the most popular mobility model for simulating the mobile network of WSN-assisted IoT. In the Random walk mobility model, the node selects their destination with random speed and direction with zero pause interval. It is a memoryless model, and information of previous states does not affect the future state. The node in the Random waypoint mobility model is very similar to Random walk with a small variation. The nodes in Random waypoint can choose their destination with a pause interval under some defined range, and future state depends on the previous state [7, 8].

The WSN-assisted IoT networks are deployed at a vast area as compared to WSN; also, numbers of nodes (mostly resource constraint) are large as compared to the WSN network. Hence, traditional WSN methods may not be effective in scalable WSN-assisted IoT network. To overcome the scalability and network lifetime issue, many researchers adopted the cluster-based hierarchical framework [9, 10]. The prime focus in cluster-based routing protocols is on the efficient Cluster Head (CH) selection and organizing of remaining nodes in the clusters. The nodes within the cluster send their data to the CH. The CH has accountability for the data aggregation of local cluster member and conveys the aggregated data to the base station or adjacent CHs toward the Base Station (BS). This results as extra load on CH which leads to more energy depletion as compare to normal nodes [11, 12]. The CH near to BS depletes their energy faster as compare to other nodes because it also acts as a Relay Node (RN) for aggregated data sent by distant CHs. To address all these issues, in this paper, a multi-tier based clustering framework and subdivision technique are proposed for network deployment along with efficient routing protocol that aims

to achieve load balancing in the network and transform the long-distance communication into the shorter multi-hop distance communications so that network lifetime can be increased. The major contribution of this paper can be summarized as follow:

(A) A multi-tier hierarchical framework based on subdivision technique for node deployment that aims to reduce long distance communication. The proposed multi-tier framework can support load balancing, and scalability.

(B) An effective RN and CH selection method for the proposed framework has been introduced. The proposed approach supports load balancing in the network.

(C) In this work three node deployment scenarios have been considered to make multi-tier hierarchical framework more realistic as various applications of WSN-assisted IoT. The considered scenarios are: (a) Sensor nodes will be static in the network (b) Senor nodes are allowed to move within their local cluster boundary only (c) Sensor node are able to move within the entire network area.

(D) The proposed work in this paper has resulted in improved network lifetime along with lesser energy consumption when compared with LEACH, M-LEACH, EA-CRP, TDEEC, DEEC, SEP, and MIEEPB.

The structure of a paper is as follows: literature review of static and dynamic WSN routing protocols are given in Sect. 2. The system framework, mobility scenarios, subdivision algorithm and energy model are formulated in Sect. 3. The Scalable and Energy-Efficient routing protocol (SEEP) for the proposed framework is presented in Sect. 4. Section 5 will discuss the simulation environment. Section 6 discusses about the simulation results of the proposed framework and their comparative analysis for static and mobile node scenarios under various mobility models. Later, the conclusion has been discussed.

2 Related work

The challenge in WSN-assisted IoT is slightly different from the traditional wireless and wired network due to vast and scalable networks. The classical routing protocol is based on IP based protocol and does not match the requirements of WSN-assisted IoT due to less support of scalability. Hence, many effective routing methods have been introduced to cope with the demands of WSN-assisted IoT. Most of the routing protocols can be classified into three categories: hierarchical or location-based routing and

data-centric. In hierarchical routing, nodes are organized into the clusters, and some of the optimal nodes in the cluster are promoted as a Cluster Head (CH). The local member of clusters sends the data to their CH. CH gathers the data into one packet through the data aggregation scheme and relays the data to the base station (BS). The main job of normal nodes is to sense the network area and transmit the sensed data to their CH.

One of the most inspiring and popular hierarchical routing protocols is known as Low Energy Adaptive Cluster Hierarchy (LEACH). LEACH) is a multi-hop routing protocol, where the node in the network is grouped into clusters. The network lifetime of LEACH is divided into rounds, where each round is subsequently divided into the set-up and steady-state phase. In the set-up phase, the network is organized into clusters. Some of the nodes in the cluster are promoted as CH by the probability function (based on random function, which varies between 0 to 1). In a steady-state phase, CH aggregates the sensed data into one packet and consequently conveys to BS directly [13]. LEACH is not suitable for the WSN-assisted IoT network due to some drawbacks. LEACH does not include any relaying method between CH and BS. When the network is deployed in a large area, the distance between CH and BS is increased, and it leads to massive energy depletion. Secondly, the CH selection method in LEACH is randomized and does not assure the best node selection as CH. Mhatre and his colleague come with the variant form of LEACH, which is known as a multi-hop form of LEACH (M-LEACH). In the steady-state phase of M-LEACH, CH sends the data to the BS via other intermediate CHs in a multi-hop way instead of direct transmission as LEACH. M-LEACH gains a better network lifetime and supports uniform energy depletion [14]. Loscri and his colleagues introduced the two-level hierarchy routing protocol (TL-LEACH). TL-LEACH based on the random rotation of primary and secondary cluster heads and made a two-level hierarchy. TL-LEACH distributes the energy load in a better way, specifically when the density of the network is large [15]. To overcome from the LEACH drawbacks, centralized from of LEACH (LEACH-C) is introduced. In LEACH-C, BS has the accountability to discover the optimal node as CHs, along with the best node to handle this selection process. The CH selection method of LEACH-C is slightly better as compared to LEACH random selection method. Before forming the cluster in LEACH-C, BS requests to all nodes in the network to send their location and current remaining energy in the set-up phase of every round. The direct communication between sensor nodes and BS puts a heavy burden on the nodes [16]. The authors in [17] came with a better solution and proposed a Fixed Low Energy Adaptive Clustering Hierarchy Protocol (LEACH-F). In LEACH-F, the set-up

overhead phase is eliminated, and BS elects CHs and their local member one time until the network becomes dead. CH role rotates with other cluster members to provide uniform energy consumption, which is resulted in a better network lifetime. EESAA protocol [18] added residual energy as one more parameter for interchange between sleep and active mode to conserve energy. The Hybrid Energy-Efficient Distributed Clustering protocol (HEED) introduced a method for CH selection, which selects according to the hybrid of node residual energy and the node degree [19]. After going through the above-reported protocols, we identified many issues for Cluster Head (CH) selection and their arrangement, which are not adequate as per the basic requirements of the WSN-assisted IoT network. Robustness is an essential feature to maintain network connectivity and data transmission. But due to the early death of nodes near to BS, the protocols reported in this section may not maintain network connectivity. These routing protocols work on the common objective as energy optimization. But they do not account for other WSN-assisted IoT features like scalability, load balancing, etc. Therefore, they are not preferred for scalable and real-time WSN-assisted IoT application.

Huang et al. proposed a novel deployment scheme for green IoT network [20]. The proposed scheme is based on a hierarchical framework, along with communication constraints. The sensor nodes are not allowed to communicate with each other directly, and they can communicate via RN. The sensor sends the sensed data to the RN, and RN transmits the sensor data to the base station. When the network area is vast, direct communication between RN to base station cause the massive energy depletion and decreases the network lifetime.

Rani et al. introduced Minimum Energy Consumption Chain-based Cluster Coordinator Protocol (ME-CBCCP) [21] protocol for WSN-assisted IoT network. The MB-CBCCP consider hierarchical cluster framework for node placement along with some communication constraints to prolong the network life. This technique selects the RNs randomly among the shortest distance nodes and does not consider any distance parameter for the next RN selection. So, some of RNs can be overlapped in each other sensing area. It causes additional RN participation to cover the respective cluster. Hence, this cause additional system cost, and resource utilization is inefficient. The MB-CBCCP pre-decides the possible number of nodes as RNs randomly, which is not preferable. The number of RN selection can vary from cluster to cluster, and it should depend on the node density of the clusters.

The authors in [22] proposed an improved version of LEACH and called Power-Efficient Gathering in Sensor Information Systems (PEGASIS). PEGASIS makes the chain of node and sensor node transmit and receive data

from their nearest neighbours. In all chains, one node in the network is selected randomly as a leader node. All nodes in the network communicate to each other in such a way that data reaches the base station by leader node. PEGASIS provides a better lifetime as compared to LEACH due to equal distribution of power consumption, which results in less energy depletion per round in network. Jafri and his colleagues modified the PEGASIS protocol and included the multi-head chain along with the sink mobility concept. The sink is allowed to move in a fixed trajectory from one region to the other. It waits for a sojourn time at the sojourn location (the location where the sink temporarily stays for data collection). Sojourn location is predefined before the network deployment. Sojourn time can be defined as the time interval for which sink stays at a specific decided position and collects data from the chain leaders. This approach increases the network lifetime as compare to PEGASIS [23].

To prolong the network lifetime, many researchers introduced the heterogeneity-aware protocol by using a hierarchical clustering scheme. SEP, DEEC, and TDEEC are one of them. SEP is an extension of LEACH, and it works on two levels of heterogeneity, and the level of heterogeneity cannot be increased by more than two. SEP increases the stability period (time interval before the first node dies) due to more powerful heterogeneous nodes in terms of energy [24]. Distributed Energy-Efficient clustering (DEEC) is an improved protocol over SEP, where it can take two or multi-level heterogeneity. The Cluster Heads (CHs) in DEEC are selected based on dynamic probability. It can be calculated by the ratio of residual energy of every node and the average energy of the entire nodes in the network. Node with the highest energy becomes a CH [25]. Threshold Distributed Energy-Efficient Clustering (TDEEC) is an enhanced form of DEEC. The CH selection in TDEEC is based on the modified threshold value, and it is calculated based on the ratio of node's residual energy and average energy of network with concern to the optimal number of CH [26].

Darabkh and his colleagues proposed an energy-efficient EA-CRP routing protocol. It is based on the layering and clustering framework. EA-CRP performs well for the scalable network, but when the network size is increased, it forms less number of layers in the network and which is resulted in long-distance communication between the layers. Therefore, it provides less network lifetime for large network area [11]. Priyan et al. proposed a framework for the Internet of Vehicle (IoV) based on the Random waypoint mobility model. This healthcare monitoring system considered the simulation scenario, which includes mobile ambulances, mobile doctors, and patience. The Performance Rank (PR) is used for the selection of best ambulance (mobile node) among all the available ambulances.

PR is calculated for every ambulance by these parameters: Euclidean distance from the patient to the mobile ambulance, medical capacity, and the number of patients currently using this ambulance. The ambulance with minimum PR is selected for a needy patient [27]. The prime focus of the above-discussed protocols is to achieve energy-efficient communication to increase the network lifetime for the static network. These routing protocols are not suited for WSN-assisted IoT applications due to complex nature and less support for scalability. In this paper, Scalable and Energy-Efficient routing protocol (SEEP) has proposed along with multi-tier based clustering framework. In this work, a strategy for the zone and cluster formation is introduced, which is varied along with the size of the network area to meet the requirement of an energy-efficient, scalable network. Furthermore, proposed framework considered static and mobile scenarios by applying Random walk and Random waypoint model for node mobility while evaluations to make it more realistic as the various application of WSN-assisted IoT.

3 System model

In this section, an overview of the deployment framework with various scenarios is presented. Subsequently, the sub-area division algorithm will be discussed to organize the network into zones and clusters. Later, various assumptions related to communication constraints and energy model will be covered.

3.1 Proposed framework

Figure 1 represents the multi-tier based clustering framework. The nodes in this framework either can be static or mobile, which depends on the application requirement. The node mobility concept brings by applying Random walk (RW) and Random waypoint (RWP) mobility model. The proposed framework can be used for those WSN-assisted IoT applications, where participating nodes in the network are static and dynamic, for example, Border security force [21], Animal Tracking and Monitoring System [28], Smart healthcare system [27]. In the multi-tier framework, the rectangular shape network area is considered. The top layer is configured as a base station layer, and the network area is divided into various zones. Every zone is divided into subzones, and it is denoted as a cluster. The areas of zones are reduced toward the BS, and the numbers of clusters are increased toward the BS to balance the network load uniformly. The Lower zone (zone 1) has the maximum area with the minimum number of clusters, and the upper zone (zone 3) has a minimum area with the maximum number of clusters. The reason can be listed here that the upper zone
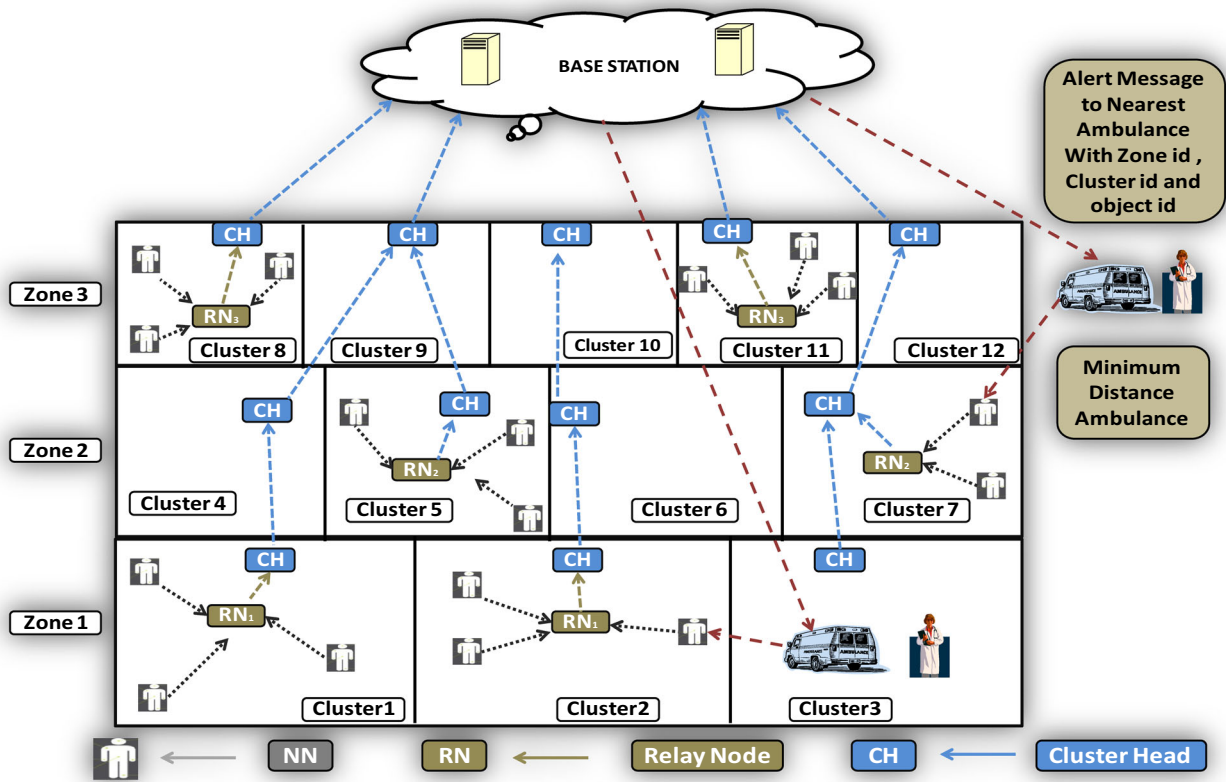
Fig. 1 A multi-tier based clustering framework

not only sends their aggregated data towards the BS, but it also relaying traffic of lower zones. In other words, the intracluster communication distance is reduced as moving towards the BS by increasing the number of clusters with decreasing zone area. The clusters have the same size within the local zone.

In this work, a smart health care application is considered for the multi-tier based clustering framework (Fig. 1). In a smart healthcare system, persons are equipped with IoT wearable devices to collect their clinical data continuously. The clinical data is transmitted to the base station (BS) via a Relay Node (RN) and Cluster Heads (CHs). The clinical data is sent to a local cluster RN, where the RN forwards this data to their CH, and CH sends the data to the upper zone CH. This process is repeated until the data is received by the BS from the topmost layer CH. Inter-Zone communication is based on some constraints, which will be discussed in the upcoming section. When the patient's health is critical, the alert message will be forwarded to the Base Station (BS), and BS broadcast this message with all mobile ambulance along with patient id, zone id, and their current cluster id so that the ambulance can reach within the minimum time. The wearable or implanted IoT devices in the patient body cannot be recharged or changed on a

daily basis, so there is a need of efficient communication method to avoid inconvenience.

3.2 Different scenario of deployed network

The proposed a framework covers various WSN-assisted IoT applications. It is extension of work carried out by Darabkh et al. [11] for static IoT network. We modified and enhanced their work for mobile WSN-assisted IoT application also. Three different practical network deployment scenarios are considered, and they are discussed below.

Scenario 1 In this deployment, nodes remain static after network deployment and their pseudo code is given in Algorithm 2. This type of topology used in many WSN-assisted IoT applications such as environment monitoring [21].

Scenario 2 In this scenario we considered the case where nodes are free to move after network deployments, but node movements are constrained within their respective zone boundary. The pseudo code for this scenario and their graphical representation are given in Algorithm 3 and Fig. 2 respectively. This type of network topology is used in animal monitoring application of WSN-assisted IoT, where the animal can roam in their respective

Fig. 2 Node movements within their respective zone (Scenario 2)
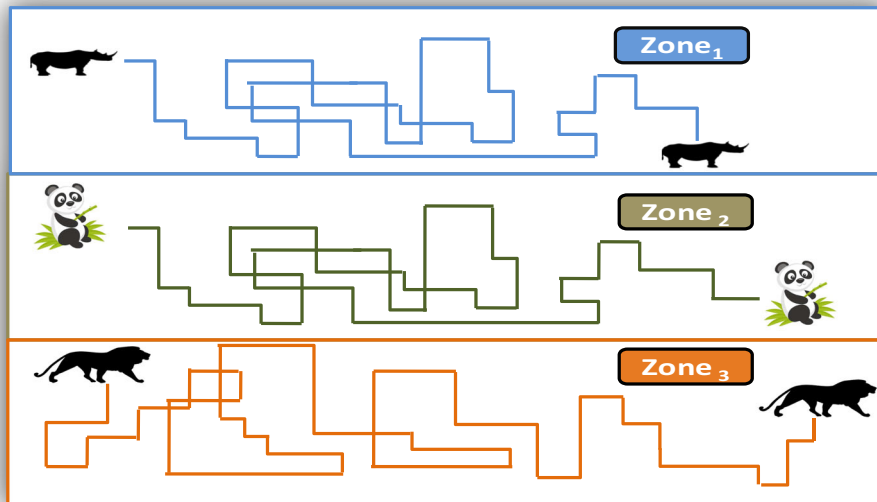
terrain [29]. We assumed that animals are attached with some wearable IoT device for monitoring purpose and it is represented in Fig. 2.

Scenario 3 In this deployment, nodes are free to move from one zone to another zone across entire network area and their pseudo code to deploy this network is given in Algorithm 4. This type of topology used in Smart healthcare monitoring system based on WSN-assisted IoT, where the users can move in entire city [30]. We have taken an assumption that humans are attached with medical sensors and represented in Fig. 3.

3.3 Network area subdivision algorithm

After many numbers of simulations, the proposed framework division techniques have been derived to increase the network lifetime. Moreover, the proposed framework is generic and simple so that it can be applicable for many WSN-assisted IoT applications.

Let R is the sensor radius, D_A is the diameter of RN sensing area, D_G, and S is the diagonal and side length of the maximum size rectangle within the RN sensing area. Here, width represents the network area width.

The number of zones will be formed in the network as follows:

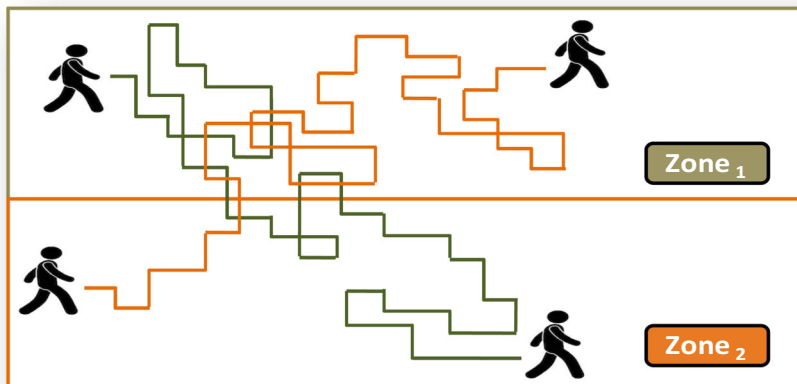$$N_Z = \frac{Width}{R} + 1 \tag{1}$$

We know that

$$D_A = 2R$$

$$D_G = S \times \sqrt{2}$$

We can see form the Fig. 4 that diagonal and diameter have same length. Hence,

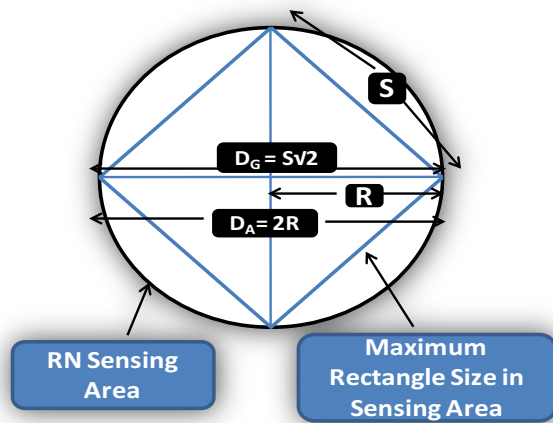**Fig. 3** Nodes movement in the entire network (Scenario 3)

Fig. 4 Maximum size rectangle in RN sensing area

$$D_A = D_G$$
$$2R = S \times \sqrt{2}$$
$$S = 2R/\sqrt{2}$$

The area of rectangle is:

$$S \times S$$
$$(2R/\sqrt{2}) \times (2R/\sqrt{2})$$
$$2R^2$$

Let assume X and Y are the length and width of the last zone respectively.

The area of zone is $X \times Y$.

The maximum numbers of clusters (N_{C_1}) will be formed to cover the bottom most zone as follows:

$$N_{C_1} = \left\lceil \frac{X \times Y}{2R^2} \right\rceil \tag{2}$$

Bearing in the mind that the numbering of zones starting from 1 at the bottom most layer. Number of clusters created in the next following layers is defined as follows:

$$N_{C_i} = N_{C_{i-1}} + 1, \quad 1 < i \leq N_Z \tag{3}$$

The width of zones is formulated as following:

$$Width_{Z_L} = \left\lfloor \frac{R}{\sqrt{3}} \right\rfloor \tag{4}$$

$$Width_{Z_M} = \left\lceil \frac{Totalwidth - Width_{Z_L}}{N_Z - 1} \right\rceil \tag{5}$$

$$Width_{Z_K} = Width_{Z_K - 1} + \left\lfloor \frac{Width_{Z_m} - Width_{Z_L}}{M - 1} \right\rfloor \tag{6}$$

$$Width_{Z_1} = Totalwidth - \sum_{i=2}^{N_Z} Width_{Z_i} \tag{7}$$

here Z_L represents the upper zone (nearest to BS), Z_1 represents the lower zone, Z_M represents the next zone to the middle one (M = Round ($N_Z/2$) + 1), Z_K represents the zone apart from define above (1 < K < M or M < K < Z_L).

It can be seen from Eq. 4, width of Z_L zone is proportional to $\frac{R}{\sqrt{3}}$. After many numbers of experiments for different network area and different sensing radius, we got the better network lifetime as compare to other value. Therefore, in this work, we considered the $\sqrt{3}$ in denominator instead of other value.

3.4 System constraints

In this framework, all nodes can perform the data reception, data aggregation, and transmission. But the nodes, which are Normal Nodes (NNs) can transmit the data only; whereas other types serve nodes (RNs and CHs) can perform the data reception, aggregation, and transmission to reduce the energy consumption. Communication between NNs can take place through local cluster RNs only. RN forwards the aggregated data to the local CH. However, cluster heads of local zone (CH$_{LZ}$) can communicate only two nearest neighboring cluster head of upper zone (CH$_{UZ}$), where zone id of CH$_{UZ}$ is exactly one greater than zone id of CH$_{LZ}$. CH transmits the aggregated data to the BS via intermediate CH$_{UZ}$. Communication constraints of SEEP is summarized in the following points:

(a) Normal Node (NN) to Normal Node (NN)

   NNs are not allowed to communicate directly in SEEP framework, even if NNs belong to same cluster or zone.

(b) Normal Node (NN) to Relay Node (RN)

   NNs can send the data to their local cluster RN. But, NN are not allowed to communicate to other clusters RN.

(c) Relay Node (RN) to Cluster Head (CH)

   RN can send their aggregated data to the local cluster CH and RNs are not allowed to communicate to other cluster CH or upper zones CH.

(d) Cluster head of upper zone (CH$_{UZ}$) to Cluster head to local zone (CH$_{LZ}$)

   CH of lower zone can send their data to upper zone neighbor CH, which is discussed in detail in Sect. 4.3. If the upper zone CH become dead due to low energy, then lower zone CH can send the packet directly to BS.

Let p and q two nodes; they will be called neighbor nodes if they can communicate with each other via above discussed communication constraints. Let N(p) represents the set of p neighbors node, and A represents the adjacency

Fig. 5 a Data communication without data aggregation and b with data aggregation
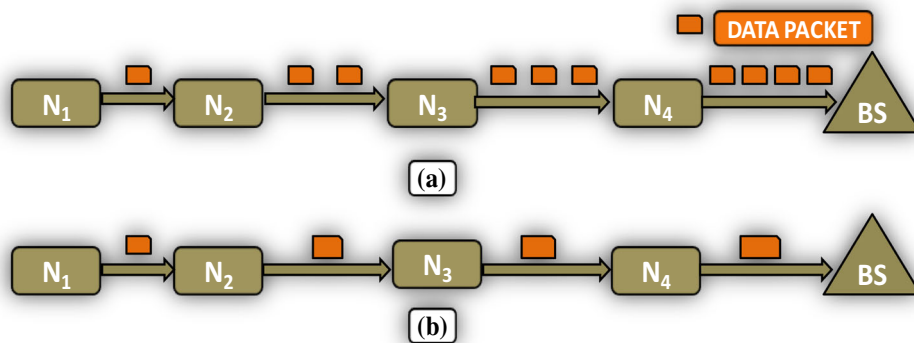
matrix. The adjacency matrix is used in energy depletion equation.

$$A = \begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1|N|} \\ A_{21} & A_{22} & \ldots & A_{2|N|} \\ \vdots & \vdots & \ddots & \vdots \\ A_{|N|1} & A_{|N|2} & \ldots & A_{|N||N|} \end{bmatrix} \qquad (8)$$

where $A_{pq} = 1$ if $q \epsilon N(p)$, otherwise $A_{pq} = 0$.

3.5 Energy model

The main source of energy depletion in IoT network is data transmission and reception. Sensor nodes deplete the energy in sensing and processing is very less. In this work, the first order radio model is followed [31] for energy depletion in data communication, which is given below:

$$E_s = L \times (E_{elec} + \in_X \times d^2) \qquad (9)$$

$$E_r = L \times E_{elec} \qquad (10)$$

$$E_l = L \times (E_{elec} + \in_X \times d^4) \qquad (11)$$

The energy depletion in short distance and long-distance communication is identified by Eqs. (9) and (11) respectively. Data reception at RN and CH nodes and their utilized energy in this task is computed by Eq. (10). Energy depletion per unit time for each node is calculated by:

$$E_a = \sum_{b \in RN} A_{ab}.F_{ab}.(E_{elec}^{NN} + \in_1 .d_{ab}^2) \quad \forall a \in NN \qquad (12)$$

$$E_b = \sum_{a \in NN} A_{ab}.F_{ab}.E_{elec}^{RN} + \sum_{c \in CH_{LZ}} A_{bc}.F_{bc}.(E_{elec}^{RN} + \in_2 .d_{bc}^2)$$
$$\forall b \in RN \qquad (13)$$

$$E_c = \sum_{b \in RN} A_{bc}.F_{bc}.E_{elec}^{CH_{LZ}}$$
$$+ \sum_{d \in CH_{UZ}} A_{cd}.F_{cd}.(E_{elec}^{CH_{LZ}} + \in_3 .d_{cd}^2)$$
$$\forall c \in CH_{LZ} \qquad (14)$$

$$E_d = \sum_{c \in CH_{LZ}} A_{cd}.F_{cd}.E_{elec}^{CH_{UZ}}$$
$$+ \sum_{e \in CH_{UZ}UBS} A_{de}.F_{de}.(E_{elec}^{CH_{UZ}} + \in_4 .d_{de}^2) \qquad (15)$$
$$\forall d \in CH_{UZ}$$

$$E_e = \sum_{d \in CH_{LZ}UCH_{UZ}} A_{de}.F_{de}.E_{elec}^{BS} \quad \forall e \in BS \qquad (16)$$

E_a, E_b, E_c, E_d, E_e denote the energy depleted by NN, RN, CH$_{LZ}$, CH$_{UZ}$, and BS respectively in data transmission and reception. The symbols E_{elec}^{NN}, E_{elec}^{RN}, $E_{elec}^{CH_{LZ}}$, $E_{elec}^{CH_{UZ}}$, E_{elec}^{BS} indicate the energy consumption in the radio electronics of NN, RN, CH, CH and BS. $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$ are used as node amplifier for NN, RN, CH$_{LZ}$ and, CH$_{UZ}$ respectively. F_{ab} represents the data rate between node a and b. Equation (12) denotes the energy consumed in data transmission processes between NN to RN within the intra cluster. The energy consumed in data receiving processes from NN to RN and sending the data from RN to CH$_{LZ}$ (cluster head of local cluster) by RN is computed by Eq. (13). The energy consumed by CH$_{LZ}$ for the data reception from RN and transmission to the CH$_{UZ}$ (upper zone nearest neighbor's CH) is represented by Eq. (14). Equation (15) expresses the energy consumed by CH$_{UZ}$ for the data transmission process to upper zone CH or the BS and the data receiving processes by the CH$_{LZ}$. Equation (16) shows the consume energy at the BS layer. All of the above equations exclude the energy depletion in signaling data by the BS to nodes, because it is negligible as compared to data transmission and receiving [20, 21].

3.6 Aggregation model

In this framework, all nodes (except NN) can perform data aggregation, which significantly reduces the energy depletion of the sensor nodes. Figure 5 represents the simple scenario of data aggregation, where nodes are located in a network. Each Node sends the data packet to the BS. In the first scenario, node start the communication without data aggregation, and total identified number of

packets is n*(n + 1)/2. In the second scenario, nodes send the packets with data aggregation and the total observed packet is n. Each node aggregates its data with the pre-

ceding node's received packet. So, the total number of reduced packets would be (n + 1)/2 with data aggregation in the network [31].

Algorithm 1. Scalable Energy Efficient Routing Protocol (SEEP)

1: **Procedure** Routing Phase (Number of Nodes N, Network Length L, Network Width W)
2: Nodes deployments in the network area   ▷ N sensor nodes are randomly distributed
3: BS computes the various divisions such as number of zones along with their width, and number of clusters in every zone by Equation 1 to 7
4: **declare** zone_id = {1, 2, 3... J}   ▷ Id for every zone 1 to J
5: **declare** clust_id = {1, 2, 3... K}   ▷ Id for every cluster 1 to K
6: **declare** NN = {1, 2, 3... N}   ▷ All NN ∈ N
7: **declare** RN = { }   ▷ Set containing node id's of relay nodes
8: **declare** CH = { }   ▷ Set containing node id's of cluster heads
9: **declare** local_zone = {}   ▷ Current zone
10: **declare** local_cluster = {}   ▷ Current cluster
                    % Node behaviour either static or mobile based on discussed scenarios %
11: Assign the ids to every zone and cluster
12: Algo = Any scenario (Scenario 1 / Scenario 2 / Scenario 3) can be selected based on application
            requirements and their corresponding algorithm are Algo 2, Algo 3, Algo4
13: [CH, RN, Threshold$_{CH}$, Threshold$_{RN}$] = Algo (local_zone, local_cluster, RN, CH, rn, ch, p, J)
14:                       **% Data Routing Phase %**
15: **for** every zone, Current$_{zone\_id}$ ≤ J **do**
16: │ local_zone = Current$_{zone\_id}$
17: │ **for** every cluster ∈ local_zone **do**
18: │ │ local_cluster = cluster
19: │ │ **Send** the packet from u (u ∈ local_cluster$_{NN}$) to v (v ∈ local_cluster$_{RN}$)
20: │ │ **Send** the packet from v (v ∈ local_cluster$_{RN}$) to w (w ∈ local_cluster$_{CH}$)
21: │ **end for**
22: │ ch1 = nearest neighbor cluster CH, where ch1.zone_id = local_zone.zone_id + 1
23: │ ch2 = nearest neighbor cluster CH, where ch2.zone_id = local_zone.zone_id + 1
24: │ distance$_1$ = $\sqrt{(w_x - ch1_x)^2 + (w_y - ch1_y)^2}$
25: │ distance$_2$ = $\sqrt{(w_x - ch2_x)^2 + (w_y - ch2_y)^2}$
26: │ **if** $\frac{ch1.energy}{distance_1}$ > $\frac{ch2.energy}{distance_2}$
27: │ │ x = ch1
28: │ **else**
29: │ │ x = ch2
30: │ **end if**
31: │ **Send** packets in order from w → x → y (w ∈ local_clust$_{CH}$, x ∈ upper_zone_ neighbor_clust$_{CH}$, y ∈ BS)
32: **end for**
33: **for** every zone, Current$_{zone\_id}$ ≤ J **do**
34: │ local_zone = Current$_{zone\_id}$
35: │ **for** every cluster ∈ local_zone
36: │ │ local_cluster = cluster and CH, RN ∈ local_cluster
37: │ │ **if** CH.energy ≤ Threshold$_{CH}$ **then**
38: │ │ │ CH Selection Phase   ▷ Algorithm 5
39: │ │ **end if**
40: │ │ **if** RN.energy ≤ Threshold$_{RN}$ **then**
41: │ │ │ RN Selection Phase   ▷ Algorithm 6
42: │ │ **else**
43: │ │ │ Data Routing Phase   ▷ Line 14 - 32
44: │ │ **end if**
45: │ **end for**
46: **end for**
47:**end Procedure**

---

4 Proposed scalable and energy-efficient routing protocol (SEEP)

This section introduces the proposed routing protocol. SEEP is a self- sustained protocol with multi-tier based clustering framework that aims to increase the network lifetime by utilizing load balancing and uniform energy depletion among the nodes in the scalable WSN-assisted IoT network. The description of SEEP includes various phases. The first phase is network set up phase followed by network scenario 1, 2 and 3. Further RN selection is explained and finally CH selection is discussed.

Algorithm 2. Scenario 1(local_zone, local_cluster, RN, CH, rn, ch, p, M)
1: **procedure** Nodes remain static in network
2: **for** every zone, Current$_{zone\_id}$ ≤ M **do**
3: │   local_zone = Current$_{zone\_id}$
4: │   **for** every cluster ∈ local_zone **do**
5: │   │   local_cluster = cluster
6: │   │   P = Number of nodes in local_cluster
7: │   │   [ch, Threshold$_{CH}$] = **ClusterHead_Selection** (local_zone, local_cluster, P)
8: │   │   [rn, Threshold$_{RN}$] = **RelayNode_Selection** (local_zone, local_cluster, P, ch)
9: │   │   CH ← CH ∪ ch
10: │   │   RN ← RN ∪ rn
11: │   **end for**
12: **end for**
13: **end procedure**

Algorithm 3. Scenario 2(local_zone, local_cluster, RN, CH, rn, ch, p, M)
1: **procedure declare** T = {1 t}   ▷ Travelling time of mobile node
2: **declare** θ = {0 180}   ▷ Travelling direction of mobile node
3: **declare** V = {v$_{lower}$ v$_{upper}$}   ▷ Travelling velocity of mobile node
4: **declare** x$_1$ = { }   ▷ Starting x coordinate of local_zone length
5: **declare** x$_2$ = { }   ▷ Ending x coordinate of local_zone length
6: **declare** y$_1$ = { }   ▷ Starting y coordinate of local_zone width
7: **declare** y$_2$ = { }   ▷ Ending x coordinate of local_zone width
% Mobility with constraints: Node movement in between their respective zones area **%**
8: **for** every zone, zone_id ≤ M **do**
9: │   local_zone = zone
10: │   P = Number of nodes in local_zone
11: │   x$_1$= Starting x coordinate of local_zone length
12: │   x$_2$= Ending x coordinate of local_zone length
13: │   y$_1$= Starting x coordinate of local_zone width
14: │   y$_1$= Ending x coordinate of local_zone width
15: │   **for** every nodes p ∈ P **do**
16: │   │   T = rand(1,t), θ = rand(0,180), V = rand(v$_{lower}$, v$_{upper}$)
17: │   │   p$_x$ = p$_{x-1}$ + V × T × cosθ   ▷ p$_x$ and p$_{x-1}$ are new and previous x coordinate of node p
18: │   │   p$_y$ = p$_{y-1}$ + V × T × sinθ   ▷ p$_y$ and p$_{y-1}$ are new and previous y coordinate of node p
│   │   where x$_1$ ≤ p$_x$ ≤ x$_2$ and y$_1$ ≤ p$_y$ ≤ y$_2$
19: │   **end for**
20: **end for**
21: **for** every zone, Current$_{zone\_id}$ ≤ M **do**
22: │   local_zone = Current$_{zone\_id}$
23: │   **for** every cluster ∈ local_zone **do**
24: │   │   local_cluster = cluster
25: │   │   P = Number of nodes in local_cluster
26: │   │   [ch, Threshold$_{CH}$] = **ClusterHead_Selection** (local_zone, local_cluster, P)
27: │   │   [rn, Threshold$_{RN}$] = **RelayNode_Selection** (local_zone, local_cluster, P, ch)
28: │   │   CH ← CH ∪ ch
29: │   │   RN ← RN ∪ rn
30: │   **end for**
31: **end for**
22: Cluster information update based on nodes new location in network
33: **end procedure**

Algorithm 4. Scenario 3(local_zone, local_cluster, RN, CH, rn, ch, p, M)

1: **procedure declare** N   ▷ Total nodes in network
2: **declare** T = {1 t}   ▷ Travelling time of mobile node
3: **declare** θ = {0 180}   ▷ Travelling direction of mobile node
4: **declare** V = { v_{lower} v_{upper} }   ▷ Travelling velocity of mobile node
5: **declare** x_1 = Starting x coordinate of network length
6: **declare** x_2 = Ending x coordinate of network length
7: **declare** y_1 = Starting y coordinate of network width
8: **declare** y_2 = Ending x coordinate of network width
% Mobility with no constraints: Node movement in between entire network area **%**
9: **for** every nodes n ∈ N **do**
10:    T = rand(1,t), θ = rand(0,180), V = rand(v_{lower}, v_{upper})
11:    n_x = n_{x-1} + V × T × cosθ   ▷ n_x and n_{x-1} are new and previous x coordinate of node n
12:    n_y = n_{y-1} + V × T × sinθ   ▷ n_y and n_{y-1} are new and previous y coordinate of node n
     where $x_1 \leq n_x \leq x_2$ and $y_1 \leq n_y \leq y_2$
13: **end for**
14: **for** every zone, $Current_{zone\_id} \leq M$ **do**
15:    local_zone = $Current_{zone\_id}$
16:    **for** every cluster ∈ local_zone **do**
17:      local_cluster = cluster
18:      P = Number of nodes in local_cluster
19:      [ch, $Threshold_{CH}$] = **ClusterHead_Selection** (local_zone, local_cluster, P)
20:      [rn, $Threshold_{RN}$] = **RelayNode_Selection** (local_zone, local_cluster, P, ch)
21:      CH ← CH ∪ ch
22:      RN ← RN ∪ rn
23:    **end for**
24: **end for**
25: Cluster and zone information update based on nodes new location in network
26: **end procedure**

4.1 SEEP routing protocol

In SEEP, each node in the network sends their data to the BS in a multi-hop way. In algorithm 1 from line 1-11, BS computes the various divisions related to network. After that SEEP will deploy the network for the selected scenario (Scenario will be selected by the user according to application requirements). Furthermore, CH and RN selection will be done via BS with algorithm 5 and 6 for the selected scenario. After the high responsibility node (RN and CH) selection, sensor nodes transmit their sensed data to the local cluster's RN. RN aggregate the neighbor data into its own message and a single message forwards to local CH. CH relays the aggregated data to preferable neighbor CH towards the BS. Preferable CH is selected based on the ratio of residual energy and their distance to lower zone CH. CH with more ratio will be selected as a neighbor CH for relaying the data to the upper zone (line 12-32). To make the system effective, high responsibility node's energy is compared with the energy threshold and if it is found to be less than the new eligible node will be selected as high responsibility nodes (line 33–46). The whole processes will be iterated repeatedly until all nodes in the network become dead. CH and RN selection will be done by base station, whenever nodes have residual energy less than the threshold energy. The control flow of SEEP represents in Fig. 6.

4.2 Set up phase

In the first phase of SEEP, sensor nodes are deployed in a rectangular shape area. It is good to keep in mind that BS is aware of network attributes such as length and width of a network, and node position (by using global position system) [11, 21]. Furthermore, BS performs the computation at once related to various divisions such as the number of zones, cluster per zone, and width of every zone in the network by Eqs. 1 to 7. The zone and cluster id are assigned to every zone and cluster. Subsequently, BS sends the detailed message to every node in network related to their zone id, cluster id, and neighbors list. Nodes reside in the same cluster are neighbors. Nodes behaviour remains either static or mobile according to the nature of various WSN-assisted IoT applications.

Algorithm 5. Cluster Head (CH) Selection Phase

1: **Procedure** ClusterHead_Selection(local_zone, local_cluster, Array of nodes p)

2: **declare** bs_location = {x y}   ▷ Base station location in network

3: Find the Euclidean distance between all sensor nodes in the clusters and base station

4: **for** every nodes p ∈ P **do**

5:      $distance = \sqrt{(p_x - bs_x)^2 + (p_y - bs_y)^2}$

6:      $Node_{distance} = Node_{distance} + distance$

7: **end for**

8: $Avg\_distance = \frac{Node_{distance}}{|P|}$

9: **Threshold_{CH}** = Compute the energy consumption for sending the data at Avg_distance by Equation 9
            and 11

10: **for** every nodes p ∈ P **do**

11:      **if** p.energy > Threshold_{CH}

12:          **calculate** chance of election **CE**

13:          $distance = \sqrt{(p_x - bs_x)^2 + (p_y - bs_y)^2}$

14:          $CE = \left(\frac{p.\text{Residual Energy}}{\text{Initial Energy}}\right) \times \left(\frac{1}{distance}\right)$

15:      **end if**

16: **end for**

17: **Sort** nodes P in the descending order based on their CE values and store in sort_P

18: ch = first node of sort_P

19: **end procedure**

4.3 CH selection

After the network set up phase, CH selection will be invoked by algorithm 5 for the selected scenario. The CH with more remaining energy can support the uniform energy depletion in the network. On the other end, CH near to BS can decrease the communication distance in between and prolong the network lifetime. In algorithm 5 from line 3 to 8, the Euclidean distance and average distance are calculated between nodes and BS. Subsequently, the value of an energy threshold is calculated to check the eligibility of nodes to become a CH. The energy consumed for sending the data for an average distance will be Threshold, and it can be calculated by Eqs. 9 and 11. In the proposed
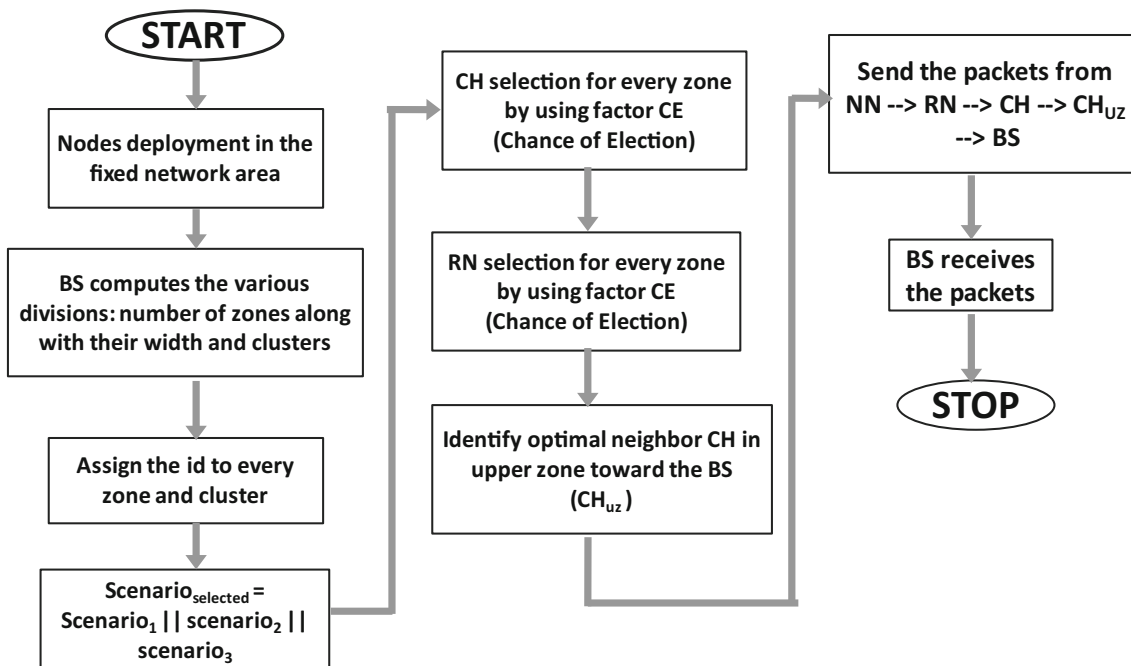

Fig. 6 Flow chart depicting the working of SEEP

framework, lower zone cluster's CH send the data to upper zone nearest neighbor CH. But in the worst case, if the entire intermediate cluster CHs runs out of energy, then lower zone CH can communicate directly to BS. In order to increase the effectiveness of the network, the average distance between node and BS is considered for threshold calculation (Algorithm 5, line 8). After that, the CE (chance of election) is computed for every node in the cluster whose residual energy is higher than the calculated Threshold_{CH}. In CE, the ratio of nodes residual energy and initial energy along with their distance to BS is considered (Algorithm 5, line 10-16). The node with more residual energy and less distance from BS will have more chance to become CH. In line 17–18, the nodes are sorted in descending order based on their CE factor and node with highest CE promoted as CH.

4.4 RN selection

After the CH selection in every zone, RN selection will take place. It can be noticed from one radio energy model [32]; energy expenditure is proportional to the communication distance. Therefore, in this phase, three parameters are considered for RN selection: (1) node with maximum remaining energy; (2) node with the shortest distance to other neighbor nodes (3) node with the minimum distance to CH. The pseudo code of RN selection is given in algorithm 6. Firstly, in algorithm 6 from line 3 to 9, the average distance between every node in a cluster and their CH are calculated. Subsequently, the energy threshold for RN is calculated. The node with more energy than the threshold will qualify for tentative RN. The Threshold_{RN} for RN is computed based on RN functionality. The job of an RN is to perform data reception from all their neighboring nodes. After the data reception, RN aggregates the data along with neighbors data into a single fixed message and forwards it to the CH. Therefore, we considered the required energy consumption to accomplish these tasks for threshold evaluation and formula is given from line 10–11. To make RN selection effective, the CE (chance of election) is evaluated for every node. CE is a combination of the node's residual energy, the average distance to neighbor nodes, and distance to CH. Nodes with maximum CE will be promoted as RN (line 12–25).

Algorithm 6. Relay Node (RN) Selection Phase

1: **Procedure** RelayNode_Selection(local_zone, local_cluster, Array of nodes p, ch)
2: **declare** ch_location = {x y}   ▷ Local cluster CH location in network
3: Find the Euclidean distance between all sensor nodes in the cluster and local cluster CH
4: **for** every nodes p ∈ P **do**
5:     $distance = \sqrt{(p_x - ch_x)^2 + (p_y - ch_y)^2}$
6:     $distance_{p\,to\,ch} \leftarrow distance_{p\,to\,ch} \cup distance$
7:     $sum\_distance = sum\_distance + distance$
8: **end for**
9: $Avg\_distance = \frac{sum\_distance}{|P|}$
10: **Threshold$_{RN}$** = Compute the energy consumption for data receiving by the (|P| -1) neighbors, energy
                cost for data aggregation of (|P|-1) neighbours into tentative selected RN data, energy
                cost for data transmission at Avg_distance from tentative RN to ch
11: **Threshold$_{RN}$** = (|P|-1)×Energy$_{Receiving}$ + (|P|) × Energy$_{Aggregation}$ + Energy$_{Transmission}$
12: **for** every node p ∈ P **do**
13:     **if** p.energy > Threshold$_{RN}$
14:        **for** every node q ∈ P **do** (except p)
15:           $distance_{pq} = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}$
16:           $Node_{distance} = Node_{distance} + distance_{PQ}$
17:        **end for**
18:        **calculate** chance of election **CE**
19:        $Node_{avg\_distance} = \frac{Node_{distance}}{|P|}$
20:        $CE_p = \left(\frac{p.Residual\ Energy}{Initial\ Energy} \right) \times \left(\frac{1}{Node_{avg\_distance}} \right) \times \left(\frac{1}{distance_{p\,to\,ch}} \right)$
21:        $CE \leftarrow CE \cup CE_p$
22:     **end if**
23: **end for**
24: **Sort** nodes P in the descending order based on their CE values and store in sort$_P$
25: rn = first node of sort$_P$
26: **end procedure**

---

5 Experimental set-up

In this work, we followed the network parameters as given in Table 1. To validate the performance of the proposed protocol in terms of network lifetime, we considered four metrics for network lifetime comparison, which includes First Node Dead Statistics (FND_Statistics), Half Node Dead Statistics (HND_Statistics), Last Node Dead Statistics (LND_Statistics) and the number of dead nodes after each round. The number of packets received by BS through hierarchical topology (NN → RN → CH → BS) is denoted as one round. FND_Statistics, HND_Statistics, and LND_Statistics is the duration between the rounds, when the communication begins and the rounds where the first node, half number of nodes and last node become dead in the network [33, 34].

The simulation of a multi-tier based clustering framework is represented in Fig. 7 and simulation of the proposed framework along with SEEP has been performed in MATLAB. The network area in figure is divided into 4 zones. It can be noticed from the figure that Zone 1 has one cluster and as we move close to BS, the width of every zone is decreasing, and number of clusters are increasing. The reason is that lower zone CH sends the aggregated data to upper zone neighbor CH only. However, upper zone CH sends the aggregated data of local cluster and also relaying the lower zone data towards the BS. The relaying traffic becomes higher near the BS zones. In this figure, communication between NN to RN represented in red dotted line. RN forwards the aggregated data into one packet to their local zone CH (Blue line). CH relays the data to selected optimal neighbor CH towards the BS (Green Line).

6 Result analysis and discussion

In this section, the effectiveness of our proposed approach is analysed via results obtained and compared the performance with existing protocols in literature as SEP [24], DEEC [25], LEACH [13], M-LEACH [14], TDEEC [26], MIEEPB [23], ME-CBCCP [21] and EACRP [11].

6.1 Network lifetime

In this section results are discussed for the various network lifetime metrics (FND, HND, LND and total number of dead nodes after each rounds) to analyse the significance of the proposed approach.

6.1.1 Network lifetime comparison for static techniques

In this section, the performance of proposed work is analysed for static routing protocols under increasing
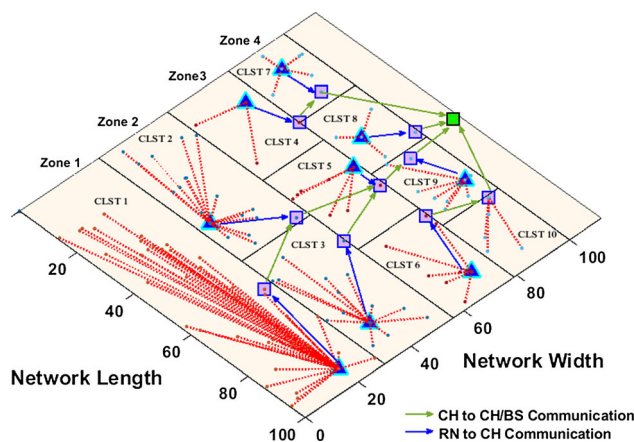

Fig. 7 Simulation of multi-tier based clustering framework

| Table 1 Simulation parameters	Parameters	Values
	E_0 (Initial energy of deployed node)	0.5 J
	E_{elec} for all the nodes	50 nJ/bit
	E_l represent the energy depletion for long distance communication	0.0013 pJ/bit/m^4
	E_s represent the energy depletion for short distance communication	10 pJ/bit/m^2
	Energy used in beam forming E_{bf}	5 nJ/bit
	L (number of bits in a packet)	6400 bits
	Control packet zine	200 bits
	RN Communication Range	30 m
	Speed Interval	[2.2–4.2] m/s
	Walk Interval	[0–1] m/s
	Pause Interval	[4–6] s
	Θ Direction Interval	[− 180, 180]
	Topology	[Static, mobile]
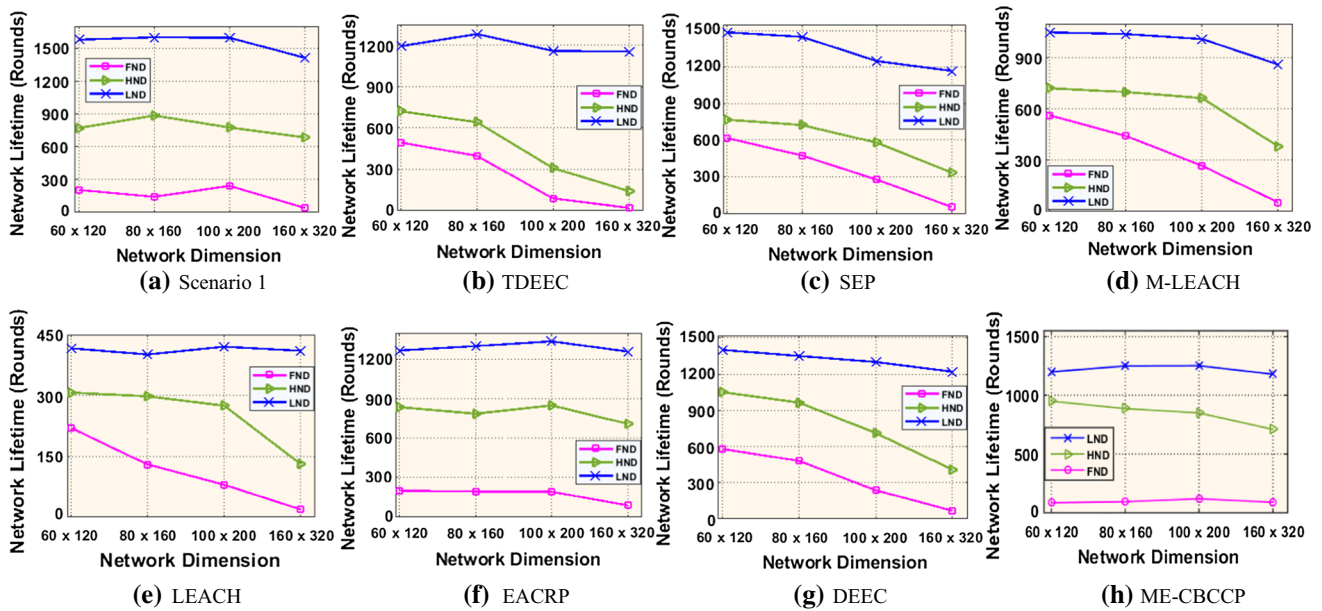
Fig. 8 Network lifetime Comparison for the static routing protocols under various network topologies: Area = 60 × 120 m², 80 × 160 m², 100 × 200 m², 160 × 320 m², Nodes = 200

network size have been increased and observed that the proposed approach prolongs the network lifetime and performs well for scalable network. In this work, FND, HND, and LND metrics are considered to examine the performance for scalable network.

It can be noticed from Fig. 8a that static protocol supports the better network lifetime, and the last node died in the network approximately after 1500 rounds for different network dimensions. When network size was increased from 60 × 120 m² to 100 × 200 m², the performance of static SEEP almost remains stable due to more number of zones and cluster formation against increased network size. The large communication distance is divided into small multi-hop communication, and additional number of CH distributes the load of the network. The network lifetime of SEEP is slightly decreased when network size reaches 160 × 320 m², whereas the performance of SEP, M-LEACH, DEEC are decreasing continuously due to large distance communication between nodes for increased network size, which cause higher energy consumptions. LEACH, ME-CBCCP, TDEEC, and EA-CRP remains almost stable for increased network size but supports less network lifetime as compared to SEEP, for LND metric. In the compared protocols, SEP performs well and gained network lifetime up to 1500 rounds for network size of 60 × 120 m², but as soon as the network area increases, its performance decreases.

In this work, the EA-CRP framework is modified to meet the requirement of a scalable WSN-assisted IoT

network. The number of zone formation should be increased with the increasing network area so that short distance communication, as well as load balancing, can be achieved to enhance network lifetime. When the network area extends from 60 × 120 m² (7200 m²) to 60 × 120 m² (51,200 m²), EA-CRP divides the network into 4 to 5 zones, which is very less as compared to the ratio of the extended network area. Therefore, EA-CRP performance is less effective for increased network size. It can be seen from Table 2 that the number of zone is increasing in SEEP with the increasing network size as compared to EA-CRP, which results in a better network lifetime.

6.1.2 Network lifetime comparison for dynamic techniques

In this section, the impact of mobility is analysed on the performance of the proposed framework with increased network size. The behaviour of node movement for Random waypoint and Random walk is depicted in Figs. 9 and 10. In figures, the pattern of the mobility using the Random waypoint mobility model as per the parameters given in Table 1.

It can be observed from Figs. 9 and 10 that the node starts at position 1 and ends at position 50, followed by multiple steps. The node stays at every step for some amount of time between 4 and 6 s (pause interval). This can be seen from Fig. 9 that some positions are represented by two numbers. The reason is that the node stays at these

Table 2 Zone formation for increasing network area

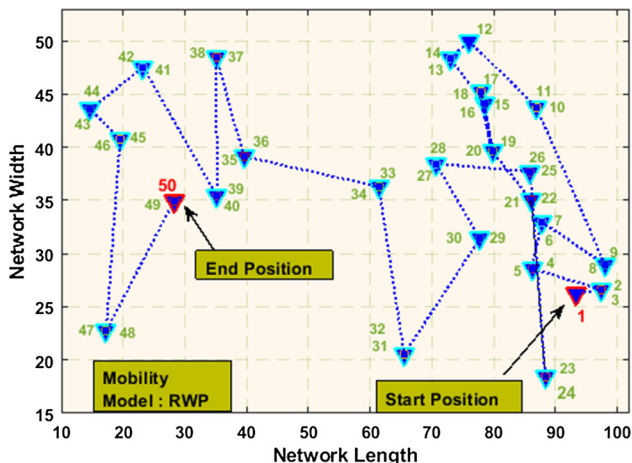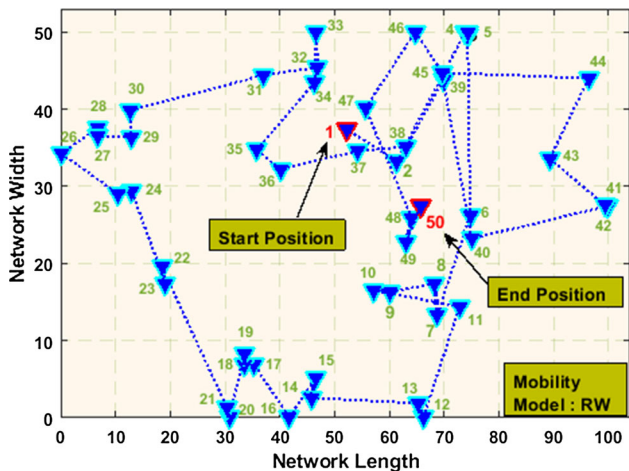| Protocol	Number of zone formation for various network dimension			
	60×120 m^2	80×160 m^2	100×200 m^2	160×320 m^2
SEEP	8	10	12	18
EACRP	4	4	5	6

Fig. 9 Mobility model: random waypoint

Fig. 10 Mobility model: Random walk

positions for some amount of time, which depends on the pause interval. Figure 10 shows the traveling pattern of the mobile node using the Random walk mobility model. The distance covered by a mobile node in Fig. 10 is more as compared to Fig. 9. It is due to the pause interval is zero in Random walk, and nodes remain moving until the simulation times expire.

Figure 11 shows the network lifetime comparison for scenario 2, 3, and MIEEPB for four network sizes: 60×120 m^2, 80×160 m^2, 100×200 m^2, and 160×320 m^2. It can be seen from Fig. 11a, b, d, e that the

proposed framework supports better network lifetime with FND and HND metrics as compare to MIEEPB. The nodes in scenario 2 and 3 are mobile in the network and may change their position after every round. Therefore, after each round new CH and RN selection may be done with respect to changed network as per new nodes position, and it leads to uniform energy depletion. One interesting point we observe, if the mobile nodes cover a large distance in the network, then it supports a better network lifetime.

As we mentioned earlier, while discussing the deployment scenarios, nodes in scenario 2 can move within their zones only, whereas nodes in scenario 3 can move across the network. So, nodes in scenario 3 will travel more distance. Furthermore, as we saw in Figs. 9 and 10 that, nodes with the Random walk model travelled more distance as compared to the Random waypoint mobility model. If the covered distance by the mobile nodes is large, it will change the network as per the new position of mobile nodes. The nodes as high responsibility node (RN or CH) in the previous round (before the mobility) will have less probability to be reselected as high responsibility node in the new round. Hence, in every round new node (most of the time) will be selected as RN and CH by the CE (as discussed in Algorithm 5 and 6). Based on the above discussion, it can conclude that energy depletion in the entire network will be uniformed. Also, it can be noticed from Fig. 11 that scenario 3 performs better as compared to scenario 2. Further, scenarios with Random walk supports better network lifetime as compared to Random waypoint.

6.1.3 Number of alive nodes versus round number

The dead node trend of compared protocols represented in Fig. 12. It can be observed from Fig. 12a that static SEEP (scenario 1) performs better and supports maximum lifetime with respect to other static protocols.

To make it more clear, network lifetime is also compared with FND, HND, and LND statistics in Fig. 13. In Figs. 12b and 13b, a comparison has been made between dynamic protocol MIEEPB, scenario 2, and 3 with Random walk and Random waypoint model. We observed that SEEP with scenario 3 under RWP gains a better network lifetime. SEEP also analysed for all considered scenarios in Figs. 12c and 13c. The plots illustrate that Scenario 2 and 3 perform better as compared to scenario 1 with FND and HND statistics because of the reselection of nodes as CH
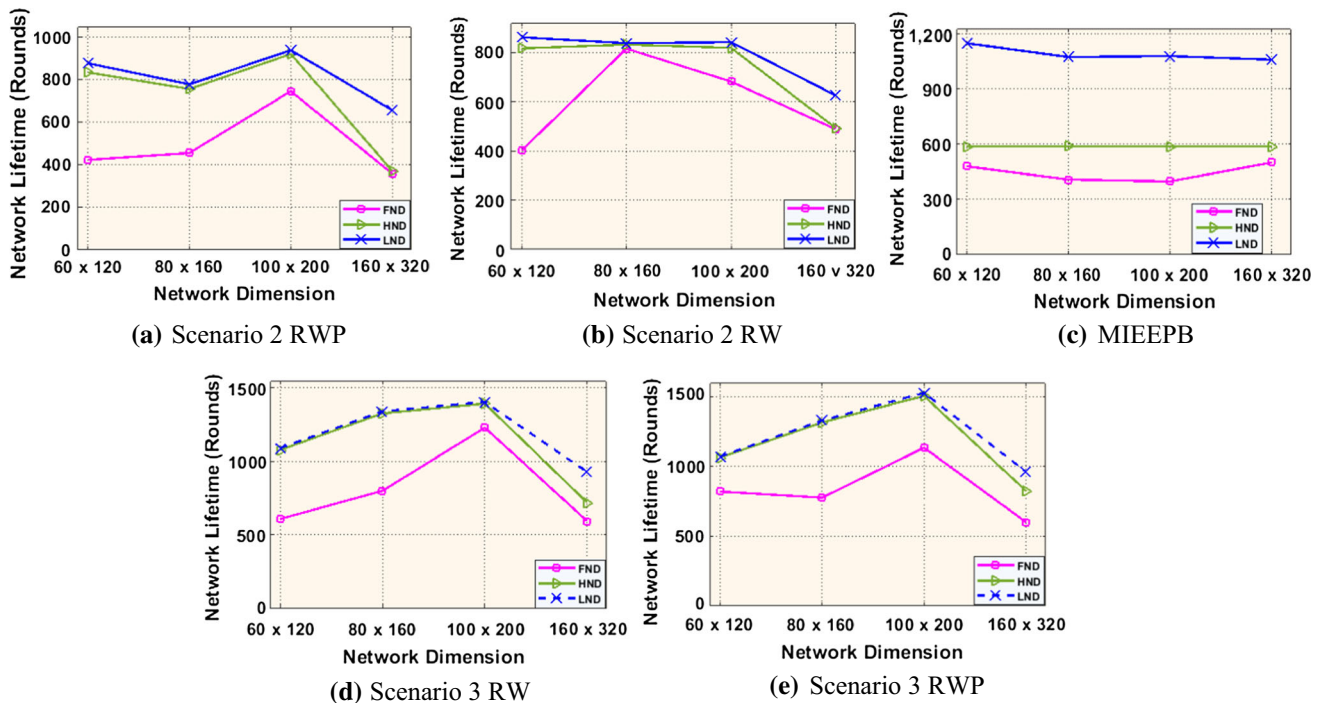
Fig. 11 Network lifetime Comparison for the dynamic routing protocols under various network topologies: Area = 60×120 m^2, 80×16 m^2, 100×200 m^2, 160×320 m^2, Nodes = 100

and RN for new node structure in every round due to mobility. As we discussed earlier, high responsibility nodes (CH and RN) perform data reception, aggregation, and transmission, whereas normal nodes only send the data to RN; therefore, RN and CH deplete more energy. Further, due to mobility, node location changes after every round. The result supports the fact stated earlier in Sect. 6.1.2 that the mobility of nodes results in uniform energy consumptions and leading to better network lifetime. It can be noticed from Figs. 12c and 13c that as the number of the round is increased, scenario 1 performs better in terms of network lifetime for LND statistics.

In Tables 3 and 4, we have summarized the network lifetime gain of SEEP compared to static and dynamic protocols under FND, HND, and LND statistics, and it is evaluated by Eq. 17. It can be noticed from Table 3 that SEEP (Scenario 1) performs remarkably well as compared to standard static routing protocols. DEEC performs slightly better with FND statistics due to heterogeneity (nodes in the network is heterogeneous in terms of energy), and the total initial energy of the network is more as compared to SEEP, which will be discussed in the next section. In Table 4, it can be noticed that SEEP (Scenario 3, RWP) performs better as compared to other dynamic protocol with FND, HND, and LND statistics.

$$\text{Performance gain} = \frac{SEEP - Other\,Protocol}{Other\,Protocol} \times 100\% \tag{17}$$

6.2 Energy consumptions

In this section, we analyse the energy consumption of SEEP for static and dynamic technique for both mobility model against standard routing protocol as reported in literature.

6.2.1 Total energy consumptions

It is described as the total energy depletion by NNs, RNs, and CHs in each round. The Total Energy Depletion is defined as:

$$TED = \sum_{ROUNDS} \left[E_{Tx}^{NN} + E_{Tx,Rx}^{RN} + E_{Tx,Rx}^{CH} \right] \tag{18}$$

Figure 14a–c represents the total energy depletion of the compared protocols for static, dynamic, and all scenarios of SEEP from the first round to last round. It can be observed from Fig. 14a that DEEC and SEP have more energy in the initial round due to the heterogeneous nature that means initial energy of node are different. The plot illustrates that the proposed protocol has a uniform energy depletion curve as compared to other routing protocols till 1500 rounds.
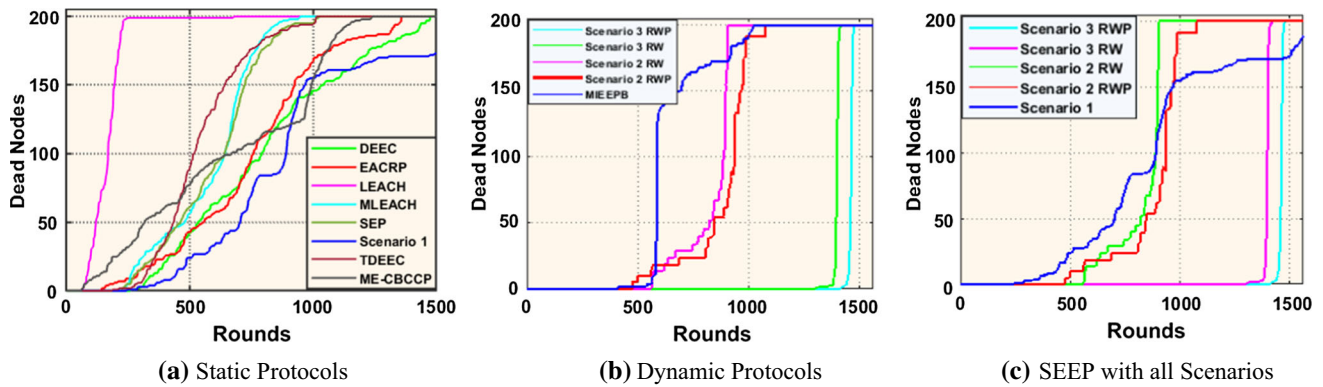
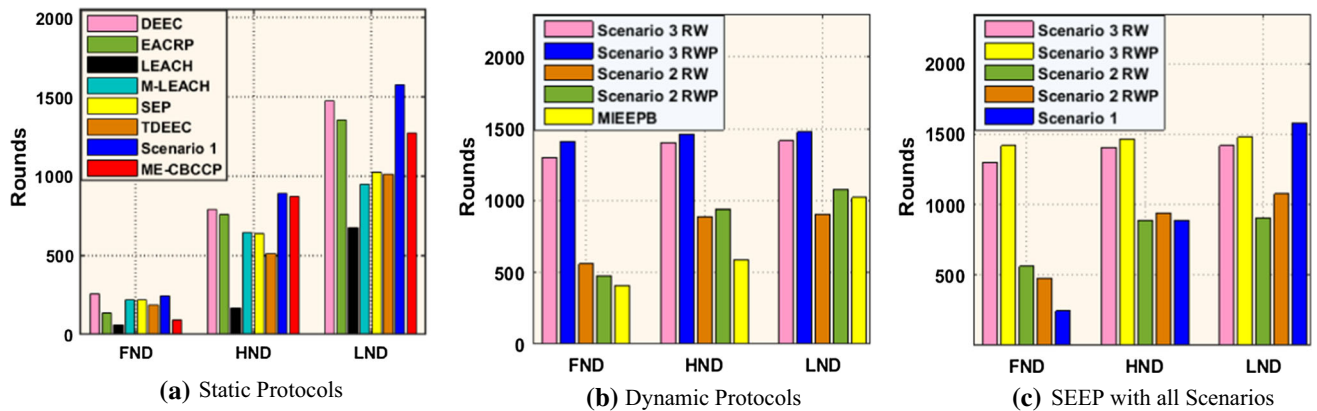**Fig. 12** Dead node statistics for network topology: Area = 200 × 200, m², Nodes = 200

Fig. 13 FND, HND, and LND statistics for network topology: Area = 200 × 200, m², Nodes = 200

Table 3 Performance gain of SEEP (Scenario 1) over static protocols

Protocols	FND (%)	HND (%)	LND (%)
DEEC	− 5.4	12.74	7.19
EACRP	77.53	17.70	16.43
LEACH	308	427.21	133.38
M-LEACH	11.36	38.35	66.84
SEP	10.85	39	54.29
TDEEC	31.72	73.34	56.12
ME-CBCCP	104.16	4.82	24.42

Table 4 Performance gain of SEEP (Scenario 3, RWP) over dynamic protocols

Protocols	FND (%)	HND (%)	LND (%)
SEEP: Scenario 2 (RW)	152.40	64.67	63.06
SEEP: Scenario 2 (RWP)	197.47	56.24	37.32
SEEP: Scenario 3 (RW)	8.83	4.57	4.32
MIEEPB	244.52	148.97	44.29
SEEP: Scenario 2 (RW)	152.40	64.67	63.06

The static SEEP has a better lifetime (1530 rounds) because of the energy consumption per round is minimal as compared to other static and dynamic protocols which have maximum network lifetime (1440-1490 rounds). In dynamic protocols, SEEP (scenario 3) with RWP gained a better network lifetime (1490 rounds). It can be noticed from Fig. 14b that SEEP (scenario 3) with RWP also depletes its energy uniformly (straight curve) till the last round.

6.2.2 Energy balancing

To analyse energy balancing, 200 nodes were deployed randomly in 200 × 200 m² network area. The residual energy of sensor nodes after 200 rounds is represented in Fig. 15a–g for static routing protocols. When the nodes are deployed in a large network area, then the communication distance between nodes is increased. In a scalable WSN-assisted IoT network, energy balancing [35] plays an essential role in increasing the network lifetime. It can be observed in Fig. 15a that SEEP (scenario 1) supports energy balancing efficiently. In LEACH, SEP, ME-CBCCP and EA-CRP, some of the nodes have depleted their energy
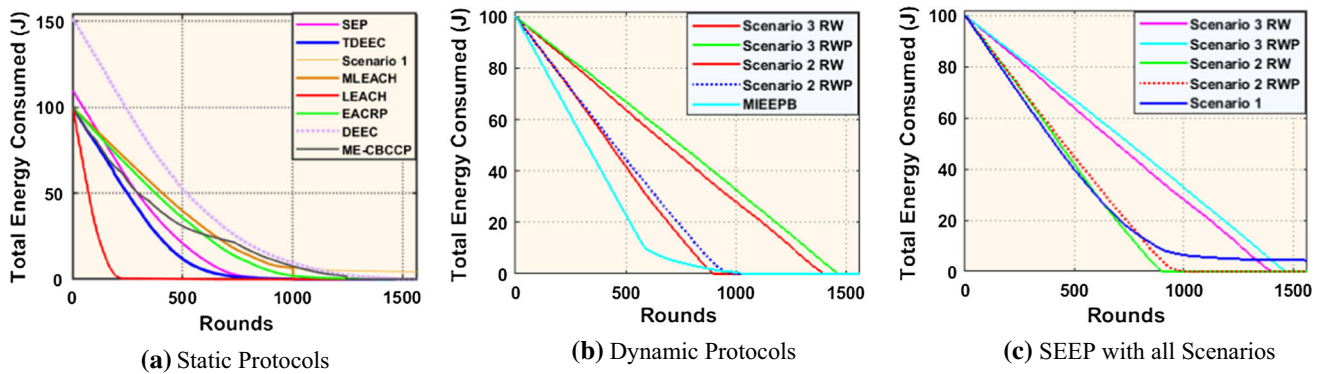
(a) Static Protocols　　　　**(b)** Dynamic Protocols　　　　**(c)** SEEP with all Scenarios

Fig. 14 Total Energy consumption statistics for network topology: Area = 200 × 200, m², Nodes = 200

(a) Scenario 1　　　**(b)** TDEEC　　　**(c)** SEP　　　**(d)** M-LEACH

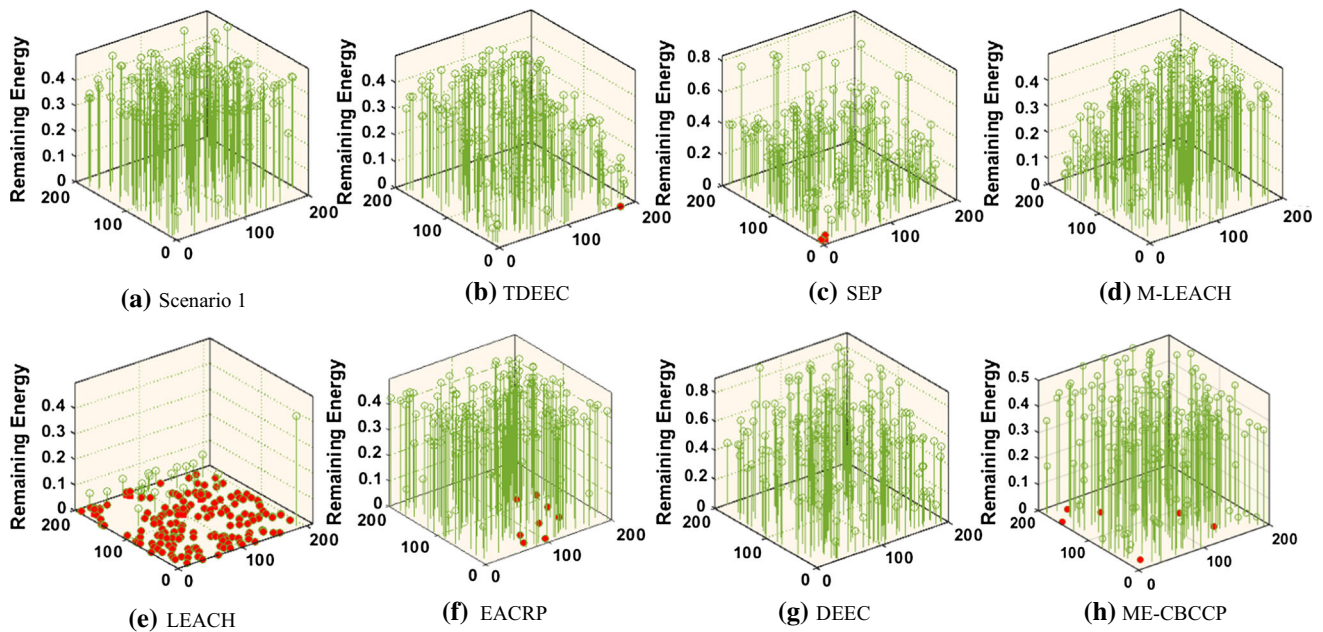**(e)** LEACH　　　**(f)** EACRP　　　**(g)** DEEC　　　**(h)** ME-CBCCP

Fig. 15 Remaining energy of sensors after 200 rounds for static routing protocols, Network Area = 200 × 200 m², Nodes = 200

completely, which is depicted in red colour. Among all compared protocol, LEACH supports minimum energy balancing (many of the nodes become dead). It can be noticed in Fig. 15d, g that M-LEACH and DEEC got less energy efficiency as compared to SEEP.

The residual energy of sensor nodes for dynamic routing protocols has been represented in Fig. 16a–e. It can be observed from Fig. 16 that SEEP with scenarios 2 and 3 (with both mobility models) got better energy efficiency as compared to MIEEPB.

Tables 5 and 6 bring the insight view of results and proof that SEEP for scenario 1 and SEEP with RWP for scenario 3 perform better as compared to static and dynamic protocols, respectively. We split the energy into 10 levels with the 0.05 J step. It can be noticed from Table 5 that 99 nodes are recorded after 200 rounds in the

second level of energy followed by EA-CRP (98 nodes), ME-CBCCP (68) and DEEC (99 nodes). But in EA-CRP 10 nodes belong to the last level, and DEEC and SEP have more initial energy due to heterogeneity (as discussed in Sect. 6.2.1). Among them, LEACH performance is at least and 190 nodes in the last level.

It can be observed from Table 6 that Scenario 3 performed better with both mobility models as compared to scenarios 1, 2, and MIEEPB. 199 and 193 nodes come in the second level of SEEP (scenario 3) with RW and RWP mobility model. These results assure that SEEP with scenario 1 and SEEP with scenario 3 have better network lifetime as compared to static and dynamic protocols, respectively.
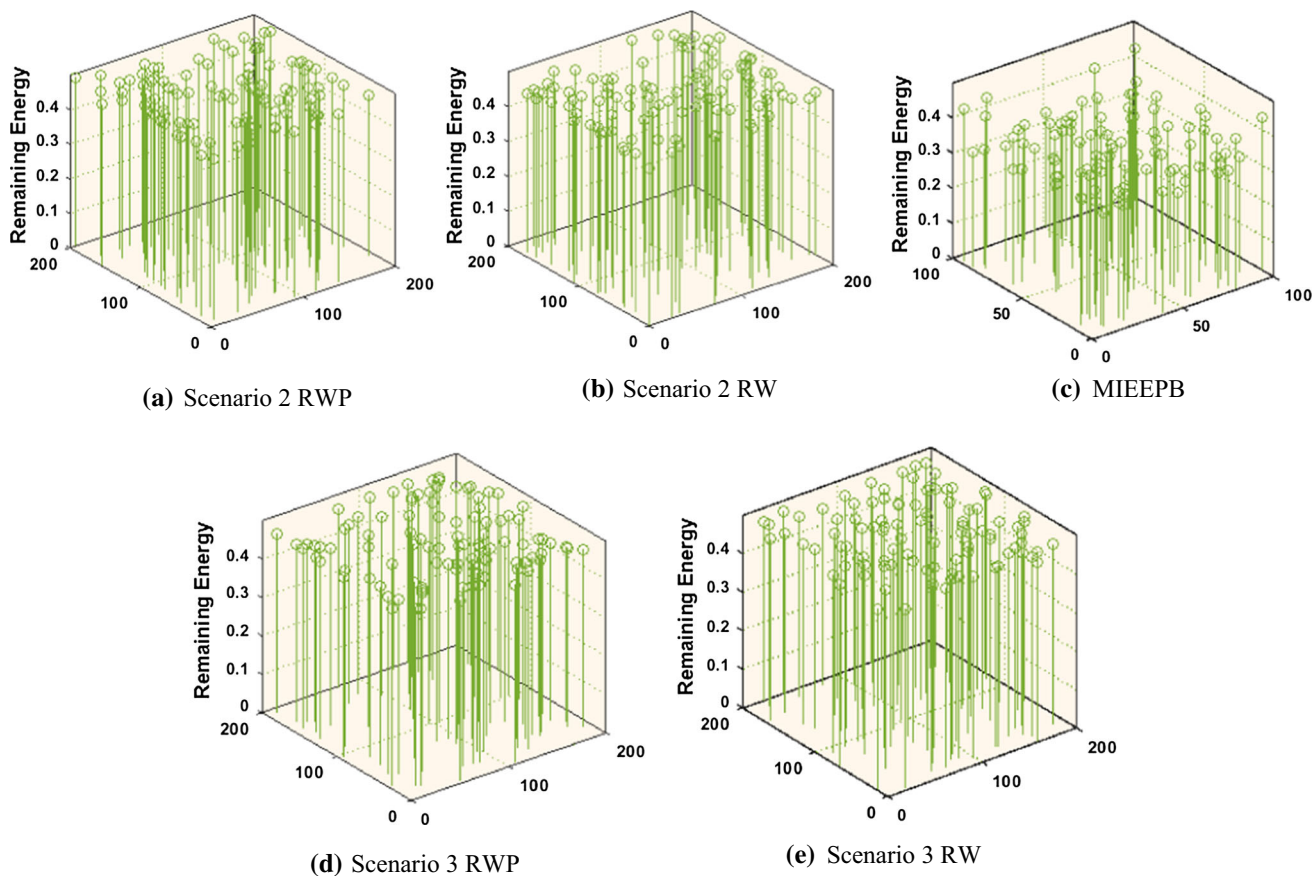
(a) Scenario 2 RWP   **(b)** Scenario 2 RW   **(c)** MIEEPB

(d) Scenario 3 RWP   **(e)** Scenario 3 RW

Fig. 16 Remaining energy of sensor after 200 rounds for dynamic routing protocols, Network Area = 200 × 200 m², Nodes = 200

Table 5 Energy distribution level after 200 rounds for static routing protocols

Levels	Energy range (J)	Node distributions according to their energy level							
		SEEP (Scenario 1)	TDEEC	SEP	M-LEACH	LEACH	EACRP	DEEC	ME-CBCCP
1	[0.45–0.50]	0	0	16	0	1	0	0	26
2	[0.40–0.45]	99	3	2	2	0	98	99	68
3	[0.35–0.40]	45	36	67	70	0	32	22	29
4	[0.30–0.35]	24	59	55	58	0	23	53	13
5	[0.25–0.30]	13	49	15	29	0	16	12	11
6	[0.20–0.25]	13	27	15	13	0	7	3	8
7	[0.15–0.20]	5	16	17	15	0	2	9	4
8	[0.01–0.15]	1	8	10	13	32	7	0	10
9	[0.05–0.01]	0	0	0	0	0	5	2	0
10	[0.0–0.05]	0	2	3	0	190	10	0	31

6.2.3 Stability factor

Section 6.2.2 talked about energy balancing specifically up to some particular round (200). To evaluate the variability of energy depletion in every round, we compute the summation of standard deviation for every node up to 600 rounds and term it as stability factor (SF): [35]

$$\text{Stability factor (SF)} = \sqrt{\frac{\sum_{i=1}^{N}(E_{avg} - E(i)_{Currentround})^2}{N}}$$

(19)

Where N is the total number of nodes, E_{avg} is the average energy depletion of the nodes in the current round,

Table 6 Energy distribution level after 200 rounds for dynamic routing protocols

Levels	Energy range (J)	Node distributions according to their energy level					
		SEEP: Scenario 1	SEEP: Scenario 2		SEEP: Scenario 3		MIEEPB
			RW	RWP	RW	RWP	
1	[0.45–0.50]	0	0	0	0	0	19
2	[0.40–0.45]	99	0	0	193	199	3
3	[0.35–0.40]	45	155	153	6	1	94
4	[0.30–0.35]	24	33	39	1	0	25
5	[0.25–0.30]	13	6	8	0	0	32
6	[0.20–0.25]	13	6	0	0	0	20
7	[0.15–0.20]	5	0	0	0	0	7
8	[0.01–0.15]	1	0	0	0	0	0
9	[0.05–0.01]	0	0	0	0	0	0
10	[0.0–0.05]	0	0	0	0	0	0

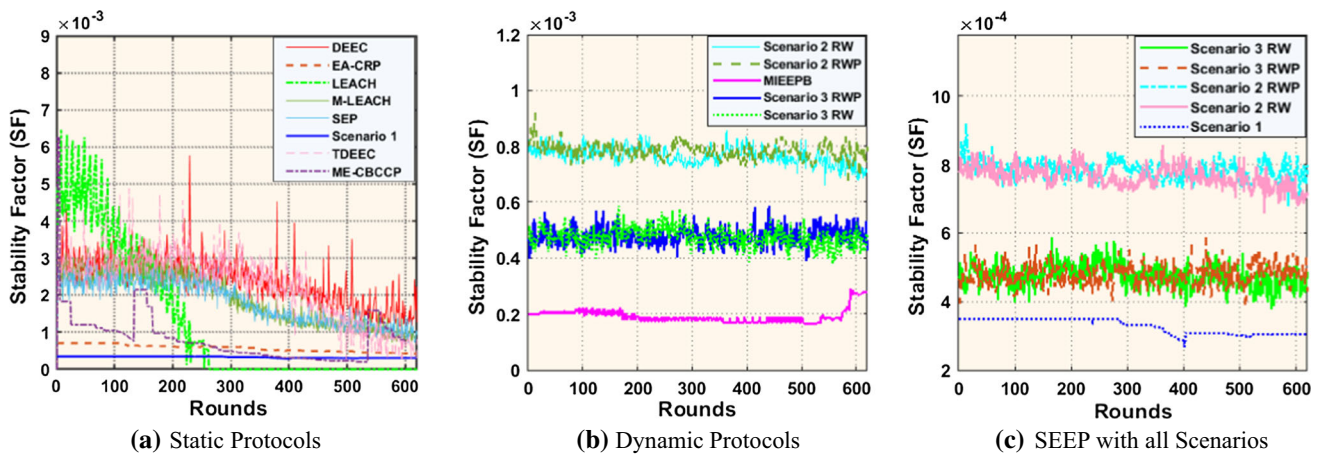**Fig. 17** Stability factor for network topology: Area = 200×200, m^2, Nodes = 200

and E(i) is the consumed energy by the ith node in the current round.

We have calculated the SF for static, dynamic, and all scenarios of SEEP and which are represented in Fig. 17a–c. It can be observed from the figure that SEEP (scenario 1) performs better as compared to other static protocols due to more zone and cluster formation for increasing network size. As we discussed in Sect. 3.1, SEEP transforms the long-distance communication into small multi-hop distance communication and support load balancing, which is resulted in less and consistent Stability Factor (SF).

In LEACH and M-LEACH, CH sends the packet directly to BS. The CHs, which are located at a large distance from the BS, run out of energy in the earlier round due to large distance communication. In ME-CBCCP protocol, the high responsibility nodes (RN, CH and Cluster Coordinator) selection are random, which cause large distance communication distance between nodes, and resulted in less network lifetime. SEP, DEEC, and TDEEC work on the heterogeneity (some nodes have more initial

energy) concept in order to increase network lifetime. But in these techniques, CH attempts packet transmission to BS via nearest neighbor, but if they are unable to find a desire neighbor, then the direct communication takes place from CH to BS. Therefore, the discussed protocol suffers from the SF problem. EA-CRP performs better as compared to the discussed protocol (LEACH, M-LEACH, ME-CBCCP, SEP, DEEC, TDEEC, MIEEPB) and use multi-hopping communication between node and BS. But it does not perform well for increased network size.

It can be noticed from Fig. 17b that MIEEPB SF obtained a low curve as compared to mobile scenarios of SEEP. It is due to the base station mobility concept used in MIEEPB for data collection, as we discussed in Sect. 2. In MIEEPB nodes remain static, and BS moves in the entire network, which minimize the distance between sensor nodes and base station. Therefore, MIEEPB obtained less SF. Generally, in IoT applications, this kind of scenario is not considerable, and BS remains stable at some specific area.

7 Conclusion

Sensor nodes are equipped with small batteries and obtain energy from these batteries is very limited. Energy conservation is one of the most critical matters in the development of WSN-assisted IoT network at large scale. In this paper, a novel energy-efficient routing protocol based on the subdivision algorithm along with a multi-tier based clustering framework has proposed. We also introduced a strategy for zone and cluster formation, which is varied along with the size of the network area to meet the requirement of energy-efficient, scalable network. The significant advantage of our proposed protocol is summarized as communication between inter clusters at a small distance in a multi-hopping way for large area network. SEEP has been analysed for different deployment scenarios, mobility models, and performance metrics, which mainly includes network lifetime and energy depletion related statistics. Furthermore, it has been compared with existing standard protocols, and experimental results validate the significance of SEEP in terms of mentioned performance metrics. The performance gain of SEEP over the traditional method summarize below.

1. The proposed SEEP for static network receives network lifetime better by 7.19% than DEEC, 16.43% than EACRP, 133.8% than LEACH, 66.84% than MOD-LEACH, 56.12% than TDEEC, 24.42% than ME-CBCCP and 54.29% than SEP protocols.
2. The proposed SEEP (for scenario 3, RWP) receives network lifetime better by 63.06% than scenario 2 with RW, 37.32% than scenario 2 with RW, 4.32% than scenario 3 with RW, and 44.29% than MIEEPB.
3. We analysed the proposed framework for all considered scenarios; among them, the static approach gains maximum network lifetime.
4. The energy consumption is uniform in nature in our proposed approach that leads to increase network lifetime. Almost 84% of nodes are having residual energy in the range of [0.30–0.45 J] after completing 200 rounds (as given in Table 5).
5. We also analysed both the mobility model for our framework and from simulation results, it can be concluded that Random waypoint performs slightly better as compare to Random walk.

References

1. Jing, Q., Vasilakos, A. V., Wan, J., Lu, J., & Qiu, D. (2014). Security of the Internet of Things: perspectives and challenges. *Wireless Networks, 20*(8), 2481–2501.

2. Tsai, C. W., Lai, C. F., & Vasilakos, A. V. (2014). Future Internet of Things: open issues and challenges. *Wireless Networks, 20*(8), 2201–2217.

3. Oliveira, L. M., & Rodrigues, J. J. (2011). Wireless sensor networks: A survey on environmental monitoring. *JCM, 6*(2), 143–151.

4. Guleria, K., & Verma, A. K. (2019). Comprehensive review for energy efficient hierarchical routing protocols on wireless sensor networks. *Wireless Networks, 25*(3), 1159–1183.

5. Zhao, N., Yu, F. R., & Sun, H. (2015). Adaptive energy-efficient power allocation in green interference-alignment-based wireless networks. *IEEE Transactions on Vehicular Technology, 64*(9), 4268–4281.

6. Tang, J., So, D. K., Zhao, N., Shojaeifard, A., & Wong, K. K. (2018). Energy efficiency optimization with SWIPT in MIMO broadcast channels for Internet of Things. *IEEE Internet of Things Journal, 5*(4), 2605–2619.

7. Bettstetter, C., Hartenstein, H., & Pérez-Costa, X. (2004). Stochastic properties of the random waypoint mobility model. *Wireless Networks, 10*(5), 555–567.

8. Cai, Y., Wang, X., Li, Z., & Fang, Y. (2014). Delay and capacity in MANETs under random walk mobility model. *Wireless Networks, 20*(3), 525–536.

9. Hisham, M., Elmogy, A., Sarhan, A., & Sallm, A. (2019). Energy efficient scheduling in local area networks. *Wireless Networks, 20*(3), 1–14.

10. Yadav, R. N., Misra, R., & Saini, D. (2018). nergy aware cluster based routing protocol over distributed cognitive radio sensor network. *Computer Communications, 129,* 54–66.

11. Darabkh, K. A., Al-Maaitah, N. J., Jafar, I. F., & Ala'F, K. (2018). EA-CRP: A novel energy-aware clustering and routing protocol in wireless sensor networks. *Computers & Electrical Engineering, 72,* 702–718.

12. Darabkh, K. A., Hawa, M., Saifan, R., & Ala'F K. (2017). A novel clustering protocol for wireless sensor networks. In *2017 international conference on wireless communications, signal processing and networking (WiSPNET)* (pp. 435–438), Chennai, India. IEEE.

13. Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks. In *Proceedings of the 33rd annual Hawaii international conference on system sciences, 2000* (pp. 10-pp). IEEE.

14. Mhatre, V., & Rosenberg, C. (2004). Homogeneous vs heterogeneous clustered sensor networks: a comparative study. In *ICC* (pp. 3646–3651).

15. Loscri, V., Morabito, G., & Marano, S. (2005). A two-levels hierarchy for low-energy adaptive clustering hierarchy (TLLEACH). In *Vehicular Technology Conference, 2005. VTC-2005-Fall.* 2005 IEEE 62nd (Vol. 3, pp. 1809–1813). IEEE.

16. Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications, 1*(4), 660–670.

17. Heinzelman, W.B. (2000). *Application-specific protocol architectures for wire-less networks*. Ph.D. dissertation, Massachusetts Institute of Technology, USA.

18. Shah, T., Javaid, N., & Qureshi, T. N. (2012). Energy efficient sleep awake aware (EESAA) intelligent sensor network routing protocol. In *2012 15th international multitopic conference (INMIC)* (pp. 317–322). IEEE.

19. Younis, O., & Fahmy, S. (2004). HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing, 3*(4), 366–379.

20. Huang, J., Meng, Y., Gong, X., Liu, Y., & Duan, Q. (2014). A novel deployment scheme for green internet of things. *IEEE Internet of Things Journal, 1*(2), 196–205.

21. Rani, S., Talwar, R., Malhotra, J., Ahmed, S. H., Sarkar, M., & Song, H. (2015). A novel scheme for an energy efficient Internet of Things based on wireless sensor networks. *Sensors, 15*(11), 28603–28626.

22. Lindsey, S., & Raghavendra, C. S. (2002) PEGASIS: Power-efficient gathering in sensor information systems. In *Aerospace conference proceedings, 2002*. IEEE (Vol. 3, pp. 3–3). IEEE.

23. Jafri, M. R., Javaid, N., Javaid, A., & Khan, Z. A. (2013). Maximizing the lifetime of multi-chain pegasis using sink mobility. arXiv preprint arXiv:1303.4347.

24. Smaragdakis, G., Matta, I., & Bestavros, A. (2004). *SEP: A stable election protocol for clustered heterogeneous wireless sensor networks*. Boston: Boston University Computer Science Department.

25. Qing, L., Zhu, Q., & Wang, M. (2006). Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks. *Computer Communications, 29*(12), 2230–2237.

26. Saini, P., & Sharma, A. K. (2010). Energy efficient scheme for clustering protocol prolonging the lifetime of heterogeneous wireless sensor networks. *International Journal of computer applications, 6*(2), 30–36.

27. Priyan, M. K., & Devi, G. U. (2017). Energy efficient node selection algorithm based on node performance index and random waypoint mobility model in internet of vehicles. *Cluster Computing, 21*(1), 213–227.

28. Al-Turjman, F. M., Hassanein, H. S., & Ibnkahla, M. A. (2013). Efficient deployment of wireless sensor networks targeting environment monitoring applications. *Computer Communications, 36*(2), 135–148.

29. Keertana, P., Vanathi, B., & Shanmugam, K. (2017). A survey on various animal health monitoring and tracking techniques. *International Research Journal of Engineering and Technology, 4*(2), 533–536.

30. Baskar, S., Shakeel, P. M., Kumar, R., Burhanuddin, M. A., & Sampath, R. (2020). A dynamic and interoperable communication framework for controlling the operations of wearable sensors in smart healthcare applications. *Computer Communications, 149,* 17–26.

31. Sajwan, M., Gosain, D., & Sharma, A. K. (2019). CAMP: cluster aided multi-path routing protocol for wireless sensor networks. *Wireless Networks, 25*(5), 2603–2620.

32. Darabkh, K. A., El-Yabroudi, M. Z., & El-Mousa, A. H. (2019). BPA-CRP: A balanced power-aware clustering and routing protocol for wireless sensor networks. *Ad Hoc Networks, 82,* 155–171.

33. Chang, H., & Tassiulas, L. (2000). Energy conserving routing in wireless ad-hoc networks. In *Proceedings IEEE INFOCOM 2000. conference on computer communications. Nineteenth annual joint conference of the IEEE computer and communications societies* (pp. 22–31). IEEE.

34. Han, Z., Wu, J., Zhang, J., Liu, L., & Tian, K. (2014). A general selforganized tree-based energy-balance routing protocol for wireless sensor network. *IEEE Transactions on Nuclear Science, 61*(2), 732–740.

35. Sajwan, M., Gosain, D., & Sharma, A. K. (2018). Hybrid energy-efficient multi-path routing for wireless sensor networks. *Computers & Electrical Engineering, 67,* 96–113.

Anurag Shukla is a Ph.D. research scholar in National Institute of Technology, Raipur, India. He did Masters of Technology in Computer Science and Engineering Department with specialization in Information Technology at Birla Institute of Technology Mesra, Ranchi, India in 2015. He did Bachelor of Engineering in Computer Science and Engineering from R.G.P.V. Technical University in 2011. His research area is WSN, IoT and symbolic execution. He has published more than 12 research papers at international journals and conference proceedings.

Sarsij Tripathi received his Ph.D. and Master of Technology (M.Tech.) in Computer Science and Engineering from Motilal Nehru National Institute of Technology Allahabad, India in year 2013 and 2008 respectively. He received his B.Tech. (CSE) from Rohailkhand University, Bareilly, India in 2002. Currently he is working as an Assistant Professor at the Department of Computer Science and Engineering, Motilal Nehru National Institute of Technology, Allahabad. His research interests are related to real time distributed systems, adhoc networks and evolutionary computing. He has published more than 15 research papers at international journals and conference proceedings.