



An energy and coverage sensitive approach to hierarchical data collection for mobile sink based wireless sensor networks

Saugata Roy¹ · Nabajyoti Mazumdar² · Rajendra Pamula¹

Received: 27 November 2019 / Accepted: 4 June 2020 / Published online: 22 June 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

In recent years, wireless sensor networks experience the energy hole problem as the most critical issue due to the heavy data forwarding load on the proximate sensor nodes to the sink. The best known solution found by the current state-of-the-art approaches for the energy hole problem is the Mobile Sink (MS) strategy. However, allowing the MS to visit every node for data collection incurs high data delivery latency, which may not be feasible in delay-sensitive applications. Thus, in this paper, restricted mobile sink motion is considered, where the MS halts at a limited number of locations stated as sojourn locations and all nodes disseminate their data to the nearby sojourn locations. The data dissemination to the sojourn location is achieved via a cluster-based routing protocol which aims to preserve the sensor nodes' energy to enhance the network lifetime. Furthermore, analogous to network lifetime, extending the coverage lifetime is of equal importance in many coverage sensitive applications of WSN. Thus, this article incorporates the coverage parameter to the proposed protocol in order to preserve the network coverage despite certain nodes die. Based on the sojourn locations, the proposed routing algorithm ensures that each cluster data is disseminated to the MS following the minimum hop path to limit the data delivery delay. Experimental results demonstrate the efficacy of the proposed protocol over several state-of-the-art protocols with respect to different metrics like network lifetime, coverage ratio, energy efficiency, packet delivery ratio, end-to-end delay, etc.

Keywords Wireless sensor network (WSN) · Energy hole problem · Mobile sink (MS) · Rendezvous node (RN) · Coverage significance

1 Introduction

With the advent of Internet of Things (IoT) (Glaroudis et al. 2020; Mabodi et al. 2020; Seyedi and Fotuhi 2020), objects, computing devices, machines in the contemporary world can stay coupled via a network where data transfer takes place in an unmanned way. The crucial component of IOT is Wireless Sensor Network (WSN), the basis of various large-scale IoT applications like smart cities, smart farming, smart grids, air pollution monitoring, etc (Rashid and Rehmani 2016; Deebak and Al-Turjman 2020; Sundhari and Jaikumar 2020). WSN is a large collection of small embedded systems known as wireless sensor nodes having sensing,

processing (a microcontroller), communication (wireless transceiver), and a power unit. These nodes are usually positioned over a target area to measure various parameters from the physical world and periodically report to the remote sink or Base Station (BS). The BS is connected to the centralized cloud server to convey the WSN data for further processing. However, such WSN based IoT framework may encounter a serious challenge in terms of energy preservation of the sensor nodes. As the sensor nodes are deployed in a hard-to-reach location, the replacement or recharging of the energy depleted nodes is not feasible (Rashid and Rehmani 2016). Therefore, the utmost goal is to make the WSN energy optimized in order to prolong its lifetime. It has been observed that maximum energy consumption in a WSN occurs due to data dissemination rather than processing or sensing (Sundhari and Jaikumar 2020; Heinzelman et al. 2000). The exploitation of multi-hop data transmission has been extensively used in the field of WSN to mitigate the foregoing issue (Mazumdar and Om 2018).

✉ Saugata Roy
saugataroy15@gmail.com

¹ Department of Computer Science and Engineering, Indian Institute of Technology (ISM), Dhanbad, India

² Department of Information Technology, Central Institute of Technology, Kokrajhar, India

In the recent past, the notion of cluster based data dissemination approach (Ullah 2020) has evolved as a fundamental methodology for prolonging the average lifetime of the WSNs. This methodology suggests a set of spatially closed sensors to be grouped to form a cluster and from each cluster, a leader node referred as Cluster Head (CH) is nominated. Incorporating multi-hop routing strategy to clustering (Bozorgi and Bidgoli 2019; Alaei and Yazdanpanah 2019) evenly distributes the network load and eliminates the data redundancy as data aggregation is performed at each cluster level. It also simplifies the data dissemination process as only the CH nodes act as the routing agents. However, cluster based multi-hop routing imposes additional relay load on CHs which are in the vicinity of the BS as these CHs have to relay data packets from distant CHs (inter cluster traffic). Such additional data forwarding load drains their energy much faster which leads them to premature death and makes the BS isolated from the rest of the network. This creates an *Energy Hole* in the network (Ren et al. 2016; Ramos et al. 2016). Unequal clustering strategy (Elkamel et al. 2019; Vijayalakshmi and Senthilkumar 2019) has been adopted by the researchers to alleviate this issue by producing smaller size clusters in the vicinity of the BS so that more CHs are available to share the inter cluster relay load to address the energy hole problem.

Nevertheless, in static sink scenario, the hotspots around the sink do not change and therefore the nodes in the proximity of sink suffer from a heavy concentration of data traffic. Hence, sink mobility has evolved as a better alternative to successfully resolve the energy hole problem (Wang et al. 2017; Krishnan et al. 2019; Abo-Zahhad et al. 2015). Unlike static sink, as the mobile sink (MS) moves the hotspots around the sink also change. As a result, the heavy data traffic nearby the sink gets distributed throughout the network which benefits attaining even energy depletion. From the previous research (Wang et al. 2017; Krishnan et al. 2019; Abo-Zahhad et al. 2015; Yarinezhad 2019; Christopher and Jasper 2020), MS trajectory based data gathering can be categorized into two groups. In the first one (Wang et al. 2017; Krishnan et al. 2019), MS reaches the individual CHs and collects data from them while the second one (Abo-Zahhad et al. 2015; Yarinezhad 2019; Christopher and Jasper 2020) allows the MS to stop only at some fixed locations to collect sensor data. The first approach minimizes the energy consumption of the sensors while experiencing a serious data delivery delay. On the contrary, the second approach has relatively higher energy consumption as the data from sensors reach via multi-hop fashion but achieves lower data gathering delay. Later, the concept of Rendezvous Node (RN) (Mottaghi and Zahabi 2015; Sharma et al. 2017; Mehto et al. 2020) has been introduced which buffers the incoming data from the sensor nodes and forwards them to MS in a single-hop fashion when MS comes within its proximity.

Although RN based data collection using MS has shown substantial energy preservation, it may lead to heavy packet exchange in the network if the data dissemination from sensor node to RN is not managed efficiently. Such unwanted packet exchange results in higher energy depletion of the WSN. Thus an efficient protocol should be designed for the data dissemination between sensor node and RN. Moreover, the WSN applications enable every individual sensor node to operate autonomously (Deebak and Al-Turjman 2020; Sundhari and Jaikumar 2020). Because for each decision making the frequent interaction between the centralized controller and the sensor nodes may lead to excessive data traffic in case of a large-scale WSN. Hence, the protocols designed for a large-scale WSN should be distributed in nature as the distributed one permits each sensor node to take decisions independently based on some available local information. This information are gathered by means of control packet exchanges which are confined only in the communication range of the sensors. Considering the challenges associated with sensor nodes, development of an efficient distributed algorithm is a major concern.

Since last decade, coverage (Soro and Heinzelman 2009; Gu et al. 2014; Mazumdar and Om 2017; Chen et al. 2019) of the target area has also become one of the primary concern along with energy awareness in the field of sensor networks. The term coverage refers to how much portion of the target area is successfully covered by the sensor nodes (Gu et al. 2014; Mazumdar and Om 2017). In WSN based coverage sensitive applications like battlefield surveillance, air pollution monitoring, defence technology etc., it is required to cover every point of the target area by at least k number of sensor nodes ($k \geq 1$) which is termed as area coverage or full coverage (Gu et al. 2014). But preserving a full coverage of the target area becomes a serious challenge when the nodes start dying particularly from the sparsely populated region.

1.1 Motivation and contribution

This subsection explores the major challenges emerged in the cluster based data dissemination with mobile sink. In addition, the gist of major contributions of this research are stated to address the respective challenges.

- *Energy* The protocols designed for energy optimized WSN should aim at maximizing the network lifetime regardless of their applications. In this context, MS based data dissemination has emerged as a paramount approach to conserve the sensor nodes' energy. However, challenges associated with mobile sink based routing can cause the overall energy consumption to increase if it is not managed efficiently. Thus, an efficient routing protocol should be designed to reduce the overhead of this operation in order to preserve the network energy.

To address the foregoing challenge this research mainly focuses to develop an energy optimized multi-objective routing algorithm by considering the energy sensitive parameter.

- **Reliability** In continuous monitoring applications, the WSN infrastructure is considered as reliable if the entire region of interest is covered by the sensor nodes. But rapid energy consumption of the sensors due to their heavy communication and computation duty reduces the average lifetime of the WSN which appears as a threat to the coverage sensitive network. Thus, the protocols designed for WSN should aim to preserve the network coverage despite the death of a certain number of nodes. Considering this issue, the proposed clustering and routing algorithms incorporate the coverage sensitive parameter that preserves the network coverage until a certain percentage of nodes die.
- **Scalability and performance** For large-scale WSN, the protocols designed should have the constant message and time complexity i.e., the protocols should be distributed in nature so that each node can execute them autonomously. Moreover, the delay sensitive applications of WSN demand the routing algorithms to ensure minimal data delivery latency. This article effectively resolves the aforementioned challenges by proposing a decentralized protocol that minimizes the message exchange overhead significantly. The proposed routing algorithm also guarantees the data dissemination from the cluster heads to their respective MS sojourn locations following a minimum hop path; accordingly, it reduces the data delivery latency of the network.

1.2 Organization

The rest part of this article is organized as follows. Section 2 discusses the existing clustering and routing approaches for the static and mobile sink. Section 3 describes the energy model, network model as well as lists the fundamental assumptions and terminologies. Section 4 elaborates operation of the proposed protocol in detail while Sect. 5 analyses the proposed protocol by means of dissipated energy calculation and corresponding lemmas. Section 6 demonstrates the effectiveness of the proposed protocol with the help of experimental evaluation which is followed by the conclusion in Sect. 7.

2 Related works

In the last two decades, WSN has evolved as a popular research domain among the research community particularly due to its large-scale applications (Rashid and Rehmani 2016). Different technologies like Internet of Things, cloud

computing have also integrated with WSN to enrich their application domain (Deebak and Al-Turjman 2020; Sundhari and Jaikumar 2020; Bhatia and Sood 2018). There are different research challenges for WSN based application protocols like, security (Jamali and Fotohi 2016, 2017; Fotohi and Bari 2020; Fotohi et al. 2020), localization (Gumaida and Luo 2019; Wang et al. 2018b), energy conservation (Mittal and Srivastava 2020; Anisi et al. 2008, 2012, 2013), etc. Considering the limited energy of the WSN nodes, energy conservation has evolved as a paramount need of such networks. There are different energy conservation techniques studied for decades (Abdul-Salaam et al. 2016), among which clustering has gained maximum attention (Ullah 2020). Low-Energy Adaptive Clustering Hierarchy (LEACH) (Heinzelman et al. 2000) proposed the first clustering technique to divide the whole network into several non-overlapping clusters where Cluster Head (CH) election is accomplished by a random probabilistic model. Later, different hierarchical clustering protocols were introduced where each CH delivers their data to the BS following an energy-aware multi-hop route (Bozorgi and Bidgoli 2019; Alaei and Yazdanpanah 2019). However, a severe difficulty arises from the clustering with multi-hop routing is Energy Hole problem (Ren et al. 2016; Ramos et al. 2016). This issue has been mitigated by several unequal clustering strategies (Elkamel et al. 2019; Vijayalakshmi and Senthilkumar 2019) where the cluster size reduces gradually as the CH approaches the BS.

However, the mentioned cluster based routing algorithms do not conserve full coverage of the network especially after the first node dies as they do not take into account any coverage aware metric. Two coverage aware clustering algorithms namely CPCP (Soro and Heinzelman 2009) and ECDC (Gu et al. 2014) enhance the coverage lifetime of the network by assuring that nodes from a densely populated area are chosen as better CH candidates than a sparsely populated area. CPCP selects the CHs by considering only coverage aware metrics whereas ECDC proposes an integrated protocol that employs both energy and coverage aware metrics for CH selection. But both of these algorithms suffer from the energy hole problem. Coverage aware unequal clustering algorithms proposed by Mazumdar and Om (2017) and Chen et al. (2019) diminish the energy hole problem as well as ensure a full coverage network over an adequate period of time.

To fix the energy hole issue in a more efficient way, several Mobile Sink (MS) based algorithms have been studied by Wang et al. (2017, 2018a), Krishnan et al. (2019), Yarinezhad (2019), Christopher and Jasper (2020), Sharma et al. (2017) and Mehto et al. (2020). Wang et al. (2017) and Krishnan et al. (2019) exploit the ant colony algorithm to establish an optimal MS route for cluster data gathering. But Krishnan et al. (2019) suggest a dynamic clustering approach to be performed in each round to attain better

energy balancing of the network. The studies presented by Abo-Zahhad et al. (2015), Yarinezhad and Hashemi (2018, 2019) and Christopher and Jasper (2020) reveal that MS needs not to reach at individual CHs for data collection, rather it stops at certain sojourn locations to gather the cluster data. An MS-based Adaptive Immune Energy Efficient clustering Protocol (MSIEEP) is constructed by Abo-Zahhad et al. (2015) to mitigate the energy hole problem by minimizing the dissipated energy and overhead of control packets throughout the network. To reduce the number of hop counts the proposed protocol partitions the target area into R number of equal-sized regions, each of which leads to one sojourn location of the MS. Wang et al. (2018a) propose an Enhanced PEGASIS (EPEGASIS) algorithm where the sensor nodes adjust their communication range dynamically to conserve energy based on their distance from the MS. The nodes closest to the MS will, therefore, have shorter communication range in order to conserve their energy and prolong their lifetime. In EPEGASIS, distant nodes use longer communication range during the routing process which enhances packet failure probability. Yarinezhad (2019) develops an MS based routing algorithm that utilizes a virtual nested ring architecture with router nodes to store and update the current MS location. This kind of infrastructure prevents the network from unnecessary flooding and increasing collisions but the failure of the router node may affect the overall network performance. Apart from Abo-Zahhad et al. (2015), Wang et al. (2018a) and Yarinezhad (2019), several significant MS based data collection approaches have been studied by Yarinezhad and Hashemi (2018, 2019) and Christopher and Jasper (2020) which support the clustering technique in the form of virtual grids or cells. Yarinezhad and Hashemi (2018, 2019) present two virtual cellular structure based routing protocols RCC and RBGM that address efficient data dissemination in WSN by defining several routing rules. Both the protocols, instead of advertising the latest sink position over the network, update the routing between the cell headers with minimal energy consumption and delay when the MS makes a move into a new cell. A Dynamic Hexagonal Grid Routing Protocol (DHGRP) presented by Christopher and Jasper (2020) performs a dynamic routing to share the updated sink position with minimal cost to the necessary Grid Heads (GH). However, in literature (Yarinezhad and Hashemi 2019, 2018; Christopher and Jasper 2020), the GH in the active grid (the current grid of MS) acts as the data collection point for the entire network. This demands a large buffer size of the sensor nodes otherwise, an uncontrollable packet drop may result.

Various research works on rendezvous based data gathering in mobile sink scenario have been explored by Mottaghi and Zahabi (2015), Sharma et al. (2017) and Mehto et al. (2020) which utilize the notion of rendezvous nodes (RN). Mottaghi and Zahabi (2015) combine the concept of

LEACH, MS and RN to present an optimized clustering algorithm that exhibits better performance than traditional LEACH. But Mottaghi and Zahabi (2015) encounter longer data transmission distance due to its single-hop routing strategy; consequently it suffers from a high packet drop ratio. A rendezvous based routing protocol is designed by Sharma et al. (2017) where a virtual cross area is defined as a rendezvous region in the middle of the sensing area by following two data transmission modes. The first mode permits the source node to forward its data to the nearest backbone node which in turn forwards them to MS. In the second mode, each sensor gets the location information about the MS from the closest backbone node and then makes direct communication with the MS based on that information. The main drawback of Sharma et al. (2017) is that the backup strategy may fail during the death of a single node in the routing chain. Mehto et al. (2020) present a virtual grid based rendezvous and sojourn point nomination procedure that significantly reduces the RP reconstruction latency. But the foregoing strategy would be incapable of collecting sensor data when all the candidate RPs are dead inside the energy efficient search region of a grid.

Table 1 depicts the summary and comparison of the related existing protocols. It is obvious from the earlier literature that efficient data delivery with coverage sensitivity plays a key role in energy and coverage aware applications of WSN. In light of these facts, a novel energy and coverage aware cluster based routing approach using mobile sink is proposed in this paper.

3 System model

Before intensively going through the proposed work, we discuss the fundamental models, assumptions and terminologies used throughout this article. The system model comprises of *Energy Model*, *Network Model and Assumptions* followed by *Relevant Terminologies* which are briefly elaborated in the later subsections.

3.1 Energy model

We use the basic radio model shown in Heinzelman et al. (2000) for energy dissipation of the sensors. The radio model says that to transmit a data packet of size b bits to a distance δ the radio expands energy:

$$E_T(b, \delta) = b \times \epsilon_{elec} + b \times \epsilon_{amp} \times \delta^e, \quad (1)$$

where ϵ_{elec} is the energy dissipation per bit in the electronic circuit, ϵ_{amp} is the energy dissipation per bit in the amplifier, and e is the path loss exponent whose value depends on transmission distance δ . If distance δ is less than threshold

Table 1 Summary and comparison of related works

Protocol	Coverage awareness	Sink trajectory	Structure type	Multiple sinks	Delay	Hot spot probability
Chen et al. (2019)	Yes	NA	Clustering	No	High	High
Krishnan et al. (2019)	No	Predetermined	Clustering	No	Low	Medium
Wang et al. (2018a)	No	Random	Tree	No	Medium	Medium
Mottaghi and Zahabi (2015)	No	Predetermined	Clustering	No	Medium	Medium
Sharma et al. (2017)	No	Random	Chain	No	High	Medium
Abo-Zahhad et al. (2015)	No	Predetermined	Clustering	No	Low	Medium
Yarinezhad (2019)	No	Random	Nested rings	Yes	Medium	Low
Christopher and Jasper (2020)	No	Random	Hierarchical hexagonal	No	Medium	Low
Yarinezhad and Hashemi (2018)	No	Random	Hierarchical grid	No	Low	Low
Yarinezhad and Hashemi (2019)	No	Predetermined	Hierarchical grid	No	Low	Low
Proposed	Yes	Predetermined	Clustering with RNs	Yes	Low	Low

distance δ_{th} then according to free space (*fs*) model e is set to 2 and $\epsilon_{amp} = \epsilon_{fs}$ otherwise, according to multipath (*mp*) model e is set to 4 and $\epsilon_{amp} = \epsilon_{mp}$. Threshold distance δ_{th} is derived as

$$\delta_{th} = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}$$

Hence, Eq. 1 can be expressed as

$$E_T(b, \delta) = \begin{cases} b \times \epsilon_{elec} + b \times \epsilon_{fs} \times \delta^2 & \text{if } \delta < \delta_{th} \\ b \times \epsilon_{elec} + b \times \epsilon_{mp} \times \delta^4 & \text{if } \delta \geq \delta_{th} \end{cases} \quad (2)$$

Similarly, to receive a b bits of data packet radio expands energy:

$$E_R(b) = b \times \epsilon_{elec} \quad (3)$$

For b bits data aggregation the radio expands energy as:

$$E_{agg} = b \times \epsilon_{da} \quad (4)$$

where ϵ_{da} is the energy dissipation factor for one bit of data aggregation.

3.2 Network model and assumptions

This section introduces the network configuration and introductory assumptions used throughout the proposed work and experimental evaluation. Suppose, the network is made up of a Set $S = \{s_1, s_2, s_3, \dots, s_n\}$ of n number of sensor nodes and a Mobile Sink (MS). Each sensor s_i is deployed at location (x_i, y_i) over a Region of Interest (RoI) of size $M \times N \text{ unit}^2$. The MS moves throughout RoI along a certain trajectory and stops at some specific sojourn points for sensor data collection. Following rudimentary assumptions are made to construct the proposed protocol:

- After deployment, all nodes remain stationary throughout its lifetime.
- Each node is concerned about their geographical position with the help of well-known localization algorithms (Gumaida and Luo 2019; Wang et al. 2018b).
- All sensors are homogeneous in nature i.e., all are having equal battery power, communication range, and sensing capacity.
- The intra cluster data are highly associated, therefore can be aggregated as a whole whereas the inter cluster data are not correlated and cannot be aggregated.
- Transmission power of each node is balanced according to the propagation distance by some power control techniques.
- All the sensors are well synchronized with respect to their timer values (Benzaid et al. 2017).
- In order to guarantee connectivity, the radius for inter cluster communication range (R_{com}) is at least two times larger than intra cluster communication radius (R_c) i.e., $R_{com} \geq R_c$

3.3 Relevant terminologies

The basic definitions used in the proposed protocol are listed below.

1. *Neighbour Set* A node s_j can be considered as a neighbour of node s_i if s_j lies in its intra cluster range $R_c(i)$. Therefore the neighbour set of s_i is given as

$$Neigh(i) = \{s_j \mid s_j \in (S - s_i) \wedge \delta(s_i, s_j) < R_c(i)\},$$

where function $\delta()$ represents the Euclidean distance between two nodes in $2 - D$ space.

$$\delta(s_i, s_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

2. **Neighbour Centrality ($Neigh_cen(i)$)** This parameter measures the spatial density of a node s_i i.e., how much it is dense with respect to the average distance of its neighbours. Thus, centrality of a node s_i can be calculated as,

$$Neigh_cen(i) = \frac{\sum \delta(s_i, s_j)}{|Neigh(i)| \times R_max}, \quad (5)$$

where $s_j \in Neigh(i)$ and $Neigh_cen(i) \in [0, 1]$.

3. **Coverage Significance (CS)** This parameter measures how much sensing area of a node is overlapped with its neighbours (Mazumdar and Om 2017). For a node s_i it can be calculated as

$$CS(i) = 1 - \left(\frac{\text{overlapped sensing area of } s_i (A_{ov}^i)}{\text{sensing area of } s_i (A_{sen}^i)} \right), \quad (6)$$

where $CS(i) \in [0, 1] \forall i$. It is clearly understood that lower CS value indicates higher sensing area overlapping. It should be noted that this parameter is employed in the proposed protocol to estimate the overlapped sensing area of the sensor nodes. Thus, it aids in coverage sensitive decision making.

4. **Coverage Ratio (CR)** It measures what fraction of the RoI ($M \times M$ unit²) is covered by the sensing ranges of the alive nodes. Coverage ratio in r th round can be calculated as:

$$CR^r = \frac{\bigcup_{i=1}^{n_{alive}^r} A_{sen}^i}{M \times N}, \quad (7)$$

where, $n_{alive}^r \leq n$ denotes the number of alive nodes in r th round. The initial coverage ratio (at 1st round) can be expressed as:

$$CR^1 = \frac{\bigcup_{i=1}^n A_{sen}^i}{M \times N}. \quad (8)$$

CR is a quality measure parameter which reflects the coverage proportion of the RoI in every network operation round of the proposed protocol.

5. **Coverage lifetime (CL)** It is the time elapsed (in rounds) from the beginning of the network operation to the first drop down of the initial coverage ratio CR^1 (Eq. 7). Coverage lifetime of a WSN can be mathematically expressed in terms of round r as

$$\begin{aligned} CL &= r \\ &\text{subject to } CR^r \leq CR^1 \\ &\text{where } CR^1 = CR^2 = \dots = CR^{r-1} \end{aligned} \quad (9)$$

6. **Min-Max Normalization** It is a scaling technique which ensures a given data to be confined to the new range from its existing range as follows

$$val_{norm} = \frac{val - min}{max - min} (max_{new} - min_{new}) + min_{new}, \quad (10)$$

where, val is the given value, max and min are the upper and lower bound of the existing range respectively, $[min_{new}, max_{new}]$ is the new range, and val_{norm} is the normalized value of val . In the proposed protocol, the input parameters of the multi-criteria function like cost function (in Eq. 17) may produce different impacts as they are measured on different scales. This issue can be resolved by adopting the min-max normalization technique which scales every parameter in the range $[min_{new}, max_{new}]$. According to the proposed model, $[min_{new}, max_{new}]$ is taken as $[0, 1]$ and thus Eq. 10 is reduced to

$$val_{norm} = \frac{val - min}{max - min} \quad (11)$$

5. **Weighted Product Model (WPM)** This well-known model is employed to solve a multi-criteria decision-making problem. It first constructs a weighted normalized decision matrix of m alternatives and n criteria. For each alternative, WPM determines the Preference Score (PS) by multiplying each criterion raised to the power of its relative weight.

$$PS_{pi} = \prod_{j=1}^n C_{ij}^{W_j}, \text{ for } i = 1, 2, \dots, m \quad (12)$$

Here, a higher preference value indicates a better alternative. WPM sharply changes the PS value of the alternatives (candidate nodes in the proposed protocol) such that a small change in the input parameters causes a large change in function output. It is worth mentioning that the input parameters such as residual energy, neighbour centrality, coverage significance etc. marginally vary from one node to another. Hence, it is important to have a function for which small variation in the input parameters results a significant variation in the output; accordingly it helps in making decision for multi-criteria problem.

4 Proposed work

In this section, an *energy and coverage aware hierarchical data collection protocol with mobile sink* is proposed which balances the energy consumption as well as prolongs the coverage lifetime of the network. The operation of the proposed protocol is divided into four elementary phases namely (i) Prepare phase, (ii) Set Up phase, (iii) Routing phase, and (iv) Steady phase. Figure 1 shows the working of different phases of this protocol. Later subsections explore all these phases in detail.

4.1 Prepare phase

At the beginning of the network operation, all sensor nodes set their ID as $s_i.ID = i$ and status as $s_i.status = 'ON'$ (acronym for Ordinary Nodes) considering that they have not yet taken part in any operation. Mobile Sink (MS), on the other hand, defines the number of sojourn locations where it stops to collect the sensor data. In order to obtain n_{sp} number of uniformly distributed sojourn locations of the MS, the Base Station (BS) calculates the centre of the RoI (c_x, c_y) as follows

$$c_x = \frac{M}{2}, \quad c_y = \frac{N}{2}$$

BS then constructs a virtual circular shape centred at (c_x, c_y) whose circumference is employed to determine the foregoing sojourn locations. Therefore, for an arbitrary radius R the coordinates (loc_x^i, loc_y^i) of i th sojourn location are calculated as:

$$loc_x^i = R \cos\left(\frac{2\pi}{n_{sp}}(i + 1)\right) + c_x, \quad loc_y^i = R \sin\left(\frac{2\pi}{n_{sp}}(i + 1)\right) + c_y \tag{13}$$

where n_{sp} is the number of sojourn locations and $i = 1, 2, \dots, n_{sp}$. Figure 2 shows the MS trajectory for 6 sojourn locations. The rest part of the prepare phase can be divided into following sub phases:

4.1.1 Rendezvous node (RN) selection

Initially, the MS moves to each sojourn location computed using Eq. 13 and broadcasts a beacon message (containing its current sojourn loc info) within the communication range of sensor nodes. Thus, all nodes (ON) in the vicinity of MS's sojourn location will receive this message. Upon receiving it, a node sets its status as 'RN' and stores the current MS's sojourn location. Similarly, a set of RNs is selected around

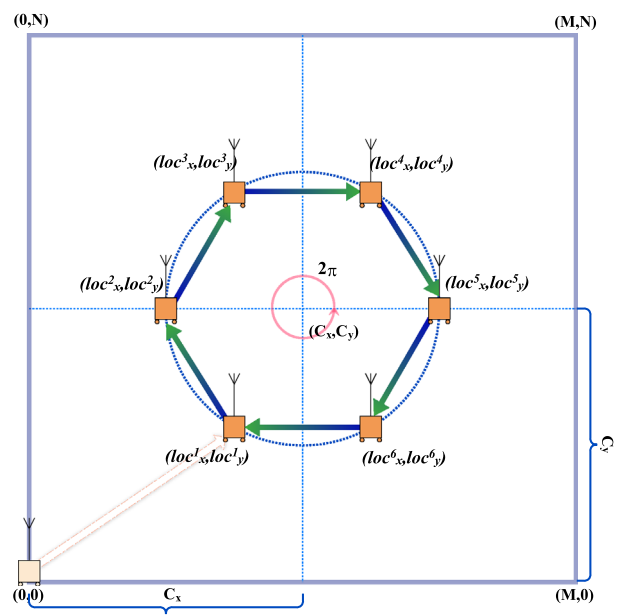


Fig. 2 MS sojourn locations for $n_{sp} = 6$

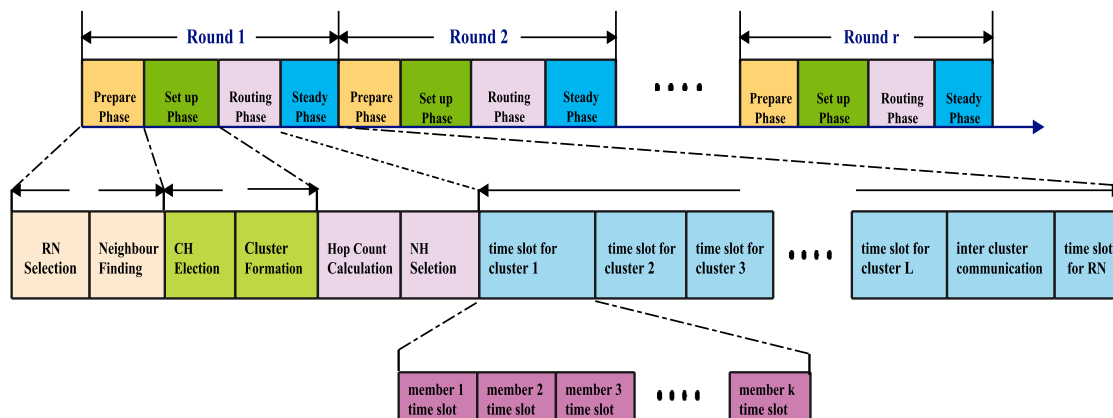


Fig. 1 Phases of the proposed protocol

each sojourn location of the MS. These RNs act as a buffer to cache the data from CHs and deliver to the MS when it stops at their respective sojourn location. Thus, RNs are responsible for delivering the network data to MS.

4.1.2 Neighbour finding

In this subphase, every node broadcasts a control packet (a WELCOME message containing node_ID, node_location, res_Energy) to announce its presence in the network. The message broadcasting is done in the MAC sublayer by means of CSMA/CA protocol to avoid the collision. After receiving the WELCOME packets, each node will be aware of their neighbours along with corresponding information.

4.2 Setup phase

This phase uniformly divides the whole network (except the RNs) into a number of clusters by adopting a distributed clustering technique. In clustering, a member node is permitted to forward their sensed data only to corresponding Cluster Head (CH). CH, in turn, aggregates the data from their members to minimize the data redundancy in WSN and then sends the aggregated data with its own data to MS by means of multi-hop data transmission.

4.2.1 CH election

The vital part of the clustering process is to select the appropriate CHs for the energy constrained network. All the nodes in RoI compete for CH election procedure which begins with their corresponding timer function ($timer_val(i)$) calculation. Utilizing this value, CHs will be nominated among the ONs. Initially, each node voluntarily starts their stopwatch. As the countdown begins, the corresponding timer function value for each node decays. Therefore, it is obvious that shorter the timer function value higher the probability of being CH. Each node s_i waits until its $timer_val(i)$ value expires. As soon as the value expires, it declares itself as a CH by setting its status $s_i.Status$ as 'CH' and broadcasts an advertise message AD_{CH_MSG} in the radius R_C of its neighbourhood area.

AD_{CH_MSG} contains node ID and residual energy information. Upon receiving AD_{CH_MSG} from s_i , each neighbouring node s_j manually stops their stopwatch and identifies itself as a cluster member by setting $s_j.Status$ as 'CM'. Then it updates their candidate CH list $Can_list(j)$ by adding s_i as its probable CHs.

$$Can_list(j) = \{i \mid j \in Neigh(i) \wedge s_i.Status = \}CH\}.$$

It may happen that a non-CH node s_j may receive more than one CH advertisement messages if s_j falls within the cluster range of multiple CHs.

4.2.2 Cluster formation

After the CH election, each member node s_j needs to be associated with the final CH which is selected from their $Can_list(j)$. Selection of final CH is done by considering two parameters: residual energy and distance with the member node. In order to balance the network load, candidate CH s_k with high residual energy $E_{res}(k)$ and shorter distance $\delta(j, k)$ is preferable. To make a compromise between these two conflicting parameters WPM (Eq. 12) is employed

$$select(E_{res}(k), \delta(j, k)) = [E_{res}(k)]^{W_1} \times [1 - \delta(j, k)]^{W_2}, \quad (14)$$

where $select(E_{res}(k), \delta(j, k))$ function for node s_j computes the selection value for its candidate CH s_k , $k \in Can_list(j)$ and W_1, W_2 are the weight factors which compromises between energy and distance. Since the network is uniformly clustered and our key concern is to make the WSN well energy distributed, we consider more weight on energy. Normally

$$W_1 > W_2 \text{ and } W_1 + W_2 = 1.$$

Candidate CH s_k with highest $select()$ value, will be considered as the final CH of s_j . After the final CH selection, the non-CH node s_j sends a $JOIN_{REQ}$ message to its selected CH in order to join the cluster. The proposed clustering procedure is concisely described in Algorithm 1.

Algorithm 1: CH election & cluster formation

Data: A large number of wireless sensor nodes deployed in the RoI
Result: A well distributed set of clusters

```

1 begin
2   for each node  $s_i \in S$  do
3     1. Collect the neighbourhood information to determine  $CS(i)$ ,
        $Neigh\_cen(i)$  and  $E_{res}(i)$ 
4     2. Calculate the timer function value  $timer\_val(i)$  using Eq. 15
5   end
6   for each node  $s_i \in S$  do
7     if ( $timer\_val(i) == 0$ ) then
8       Node  $s_i$  announces itself as CH by broadcasting  $AD_{CH\_MSG}$  in the
9       neighbour range  $R_C$  and set  $s_i.status == 'CH'$ 
10    if  $s_j$  receives  $ADD_{CH\_MSG}$  from node  $s_i$  then
11      1.  $s_j$  stops the competition and set  $s_j.status == 'CM'$ 
12      2.  $s_j$  updates its candidate CH list by  $Can\_list(j) = Can\_list(j) \cup \{i\}$ 
13    end
14    for each node  $s_j \in S \wedge s_j.status == 'CM'$  do
15      1. Calculate the selection function  $select(E_{res}(k), \delta(j, k))$  for each
16      candidate CH  $s_k$  in the
17       $Can\_list(j)$  using Eq. 14.
18      2. Choose the final CH of  $s_j$  for which
19       $select(E_{res}(CH), \delta(j, CH)) =$ 
20       $max \{select(E_{res}(k), \delta(j, k)) \mid \forall k \in Can\_list(j)\}$ 
21      3. Node  $s_j$  replies the final CH with a  $JOIN_{REQ}$  message to be associated
22      as a member.
23    end
24  end
25 end

```

4.2.3 Timer function ($timer_val(i)$) analysis

The aforementioned timer function (in 4.2.1) considers (i) residual energy (E_{res}), (ii) neighbour centrality ($Neigh_{cen}$), and (iii) coverage significance (CS) as the three distinct input parameters to successfully nominate the suitable CHs. The first parameter E_{res} of the timer function favours to choose nodes having higher residual energy as CH considering the additional responsibility of the leader, accordingly it is inversely proportional to the timer value. On the other hand, from Eq. 5, it can be observed that a node with lower $Neigh_{cen}$ value infers better CH candidate as such CH will result in optimal intra cluster communication cost. Equation 6 asserts that nodes having lower CS value (i.e., higher sensing area overlapping) is always suitable for good CH candidate as its death leaves the least amount of area uncovered. Therefore, $Neigh_{cen}$ and CS value are directly proportional to the timer computation.

$$timer_val(i) = \epsilon_i \times \left\{ \frac{Neigh_cen(i) \times CS(i)}{E_{res}(i)} \right\}, \quad (15)$$

where ϵ_i is a randomly distributed value in the range [0.9,1] used to reduce the probability of two nodes having the same timer value.

4.3 Routing tree construction phase

In this phase, an optimal loop-free routing path is constructed through which cluster data are relayed to the MS.

4.3.1 Hop count calculation

Initially, all RNs set their hop count to 1 and the CHs set their hop count to ∞ . To initiate the hop count computation process each RN and CH waits for a certain period of time before broadcasting its current hop count information within its inter cluster communication range (R_{com}). The waiting time for node s_i (RN or CH) is calculated as follows:

$$waiting(i) = \left\lceil \frac{\delta(i, MS)}{X} \right\rceil \times t_{hop}, \quad (16)$$

where constant X is large enough to scale $\frac{\delta(i, MS)}{X} \in [0, 1]$ and t_{hop} is the maximum allowed waiting time for the hop count calculation phase. The waiting time is formulated in such a way that RN's timer will expire before all the CHs' timer. Once the timer of a node s_i (RN or CH) expires, it broadcasts a $HOP_ADV(i)$ message containing its current hop count ($hop(i)$) within R_{com} . Upon receiving this message, another CH s_j may update its hop count ($hop(j)$) and candidate next hop list ($NH_list(j)$) considering the following conditions illustrated in the Lines 16–24 of Algorithm 2.

Algorithm 2: Finding Hop Count & candidate NH list of each CH

```

1 begin
2   for each node  $s_i \in S$  do
3     if ( $s_i.status == 'RN'$ ) then
4       | Set  $hop(i) = 1$ 
5     else if ( $s_i.status == 'CH'$ ) then
6       | Set  $hop(i) = \infty$ 
7     end
8   for each node  $s_i \in S$  do
9     | if ( $s_i.status == 'RN'$  or  $'CH'$ ) then
10      | Calculate  $waiting_i$  using Eq.16
11    end
12  for each node  $s_i \in S$  do
13    | if ( $s_i.status == 'RN'$  or  $'CH'$ ) then
14      | if ( $waiting(i) == 0$ ) then
15        |  $s_i$  broadcasts  $HOP\_ADV(i)$  in  $R_{com}$ 
16        for each CH  $s_j$  receiving  $HOP\_ADV(i)$  do
17          | if ( $hop(j) > hop(i) + 1$ ) then
18            | 1. Reset  $hop(j) = hop(i) + 1$ 
19            | 2. Reset  $NH\_list(j) = \{i\}$ 
20          else if ( $hop(j) == hop(i) + 1$ ) then
21            | Update  $NH\_list(j) = NH\_list(j) \cup \{i\}$ 
22          else
23            |  $s_j$  ignores the  $HOP\_ADV(i)$  from  $s_i$ 
24          end
25        end
26      end
27    end
28  end
29 end

```

It is to mention that the Line 17 of Algorithm 2 ensures that each CH s_j will have a minimum hop count to relay its data to the MS. Line 17 ensures that if a CH s_j has multiple next hop candidates to communicate with the Mobile Sink (MS) following the same minimum hop count, then all such candidates are considered as a member of $NH_list(j)$ set. On the other hand (Line 21), if CH s_j receives a $HOP_ADV(i)$ message from another CH s_i which will lead to a higher hop count than its current one, then s_j simply neglects such message. Thus, the candidate next hop list $NH_list(j)$ of CH s_j will consist a set CHs with a lower hop count value within its R_{com} . Unlike CHs, the RNs will have $NH_list = \emptyset$ as they directly communicate to the MS. It is worth mentioning that each CH will choose the most suitable CH as the next hop from its candidate NH list which is elaborated in the next section.

4.3.2 Next Hop (NH) selection

In order to relay its aggregated data, each CH s_i selects the NH from $NH_list(i)$ by means of a cost function. For each candidate next hop $s_j \in NH_list(i)$, the cost function $cost_{NH}(i, j)$ returns a value $\in [0, 1]$ where higher value denotes better NH candidate. The proposed routing algorithm designs the cost function in terms of four parameters namely

- (i) residual energy of the candidate NH s_j ($E_{res}(j)$).

- (ii) dissipated energy for data transmission between s_i and s_j ($E_T(i, j)$).
- (iii) next hop degree of candidate NH i.e., $|NH_list(j)|$.
- (iv) coverage significance of candidate NH ($CS(j)$).

$$i.e., cost_{NH}(i, j) = f(E_{res}(j), E_T(i, j), |NH_list(j)|, CS(j)) \quad (17)$$

Cost function analysis To understand the basics of the cost function we explored the fundamental objectives of each parameter used as input in this function.

Objective 1: Maximize $E_{res}(j)$ The purpose of this parameter is to make the routing process energy aware. Hence, the CH prefers to choose a next hop node among its candidates having higher energy.

Objective 2: Minimize $E_T(i, j)$ The second parameter, $E_T(i, j)$ is considered to make the best choice from the sender's point of view. Hence, the CH estimates the energy required to communicate with the candidate NH and prefers the one that leads to minimal consumption of transmission energy.

Objective 3: Maximize $|NH_list(j)|$ The third parameter denotes the next hop degree of the candidate NH s_j . The objective of this parameter is to select a reliable NH for a CH. A node is considered to be more reliable NH if it has a higher number of next hop candidates i.e., multiple routing choices. Thus a reliable CH can select the best relay node among the different alternatives. On the other hand, a candidate NH with limited next hop degree is forced to relay its data following the same path repeatedly which drastically

diminishes the residual energy of the routing path. Thus it is more logical for a CH to choose the next hop with higher reliability.

Objective 4: Minimize $CS(j)$ A lower value of coverage significance indicates a higher degree of sensing area overlapping. Therefore, to prolong the coverage lifetime it is advantageous to select the NH from the dense area since the death of sensors from this region affects network coverage lifetime less significantly. Hence, minimal value for this parameter is employed to construct a better cost function.

Considering the importance of the mentioned objectives we have constructed the above cost function by combining these four parameters.

But each parameter produces a different impact on the cost function as they are measured on a different scale. We solve this issue by adopting min–max normalization technique by Eq. 11 which scales every parameter values in the range [0, 1]. Thus the normalized value of $Eres(j)$ is evaluated as

$$Eres_{norm}(j) = \frac{Eres(j) - Eres_{min}(j)}{Eres_{max}(j) - Eres_{min}(j)}$$

where $Eres_{max}(j) = \max\{Eres(j) \mid \forall j \in NH_list(i)\}$ and $Eres_{min}(j) = \min\{Eres(j) \mid \forall j \in NH_list(i)\}$.

Likewise, we can define the normalized values of other parameters as $ET_{norm}(i, j)$, $|NH_list_j|_{norm}$, and $CS(j)_{norm}$. As mentioned above, the parameters are conflicting in nature because the cost function will be optimized only when the first and third parameters are maximized while the second and fourth parameters are minimized. The trade-off between these parameters can be resolved with the help of WPM (Eq. 12) as:

$$Cost(i, j) = Eres_{norm}(j)^{A_1} \times (1 - ET_{norm}(i, j))^{A_2} \times |NH_list(j)|_{norm}^{A_3} \times (1 - CS(j)_{norm})^{A_4} \quad (18)$$

The formulated cost function values ranges in [0, 1] where higher value indicates better next hop candidate.

4.4 Steady phase

The steady phase is responsible for data collection throughout the network and comprises of two sub-phases as (a) intra cluster transmission (b) inter cluster transmission depending on the perspective of routing strategy.

4.4.1 Intra cluster transmission

For each cluster, CMs transmit their sensed data to the corresponding CH obeying the Time Division Multiple Access (TDMA) schedule wherein each CM has a predefined time slot. CMs are allowed to send data only in this fixed slot to

avoid the collision among the data packets. CH, on the other hand, performs aggregation on the gathered data from all of its members. Data aggregation employs signal processing method which compresses the information into a single packet in order to minimize the number of data packets in the network.

4.4.2 Inter cluster transmission

After the routing tree construction phase, all the CHs are aware of the hop count required to deliver data to the Mobile Sink (MS). Based on this hop count value, each CH computes a routing delay as follows

$$delay(i) = e^{\frac{1}{hop(i)}} \quad (19)$$

It is to be noted that CHs with higher hop count are assigned to a smaller delay which assures that the routing process is initiated with CHs having maximum hop count. Once the delay timer of a CH expires, it transmits the aggregated data to the selected Next Hop (NH) node according to the previously constructed route. The same process continues till all the CH data are available to the RNs.

To collect the network data from the RNs, the mobile sink moves to one of the sojourn locations and wakes up the surrounding RNs via broadcasting a wake-up message. Upon receiving the wake-up message, RNs transmit the aggregated data to the MS. After waiting for a certain time called sojourn time which is large enough to accumulate all the data from its adjacent RNs, MS switches to the next sojourn location. This process continues until MS visits all the sojourn locations to ensure the whole data collection. Since CH has to carry out more burden than the CMs, their energy dissipation is accelerated faster in every round. To make the WSN energy balanced, the proposed protocol periodically alternates the role of CHs if and only if any one of CH's remaining energy goes down a predefined threshold value. On the other hand, RNs secure the connectivity of the MS sojourn locations with the rest of the network. Hence, extending the RNs lifetime is another major concern in MS based data collection. In view of RNs' lifetime, the energy parameter of NH selection cost function (in Eq. 18) will force the CHs to consider such RNs having higher residual energy for data forwarding.

4.5 Multiple sinks support

For a large-scale network data collection from a limited number of sojourn locations by a solo Mobile Sink (MS) may not be tolerated under the delay sensitive applications of WSN. This motivates us to exploit the notion of multiple mobile sinks in order to alleviate the average data delivery delay. The proposed protocol suggests N_{sp} number of sojourn points in the Region of Interest (RoI) for n_{ms} number of mobile sinks

where each sojourn point will be visited by a single MS for data collection. In other words, every MS will traverse a set of N_{sp}/n_{ms} number of sojourn points. The assigning of sojourn points to each MS will be carried out by the base station. It is noteworthy that the primary objective of the proposed protocol is to disseminate the cluster data to the Rendezvous Points (RN) which are predefined. Thus in the presence of multiple mobile sinks, once the corresponding sojourn locations of the mobile sinks are declared, the proposed protocol can be executed to disseminate the network data to the predefined sojourn locations as discussed in the protocol.

5 Analysis of the proposed protocol

5.1 Calculation of dissipated energy

For each node, dissipated energy refers to its total energy consumption due to data and control packet exchanges throughout the network operation in one round. To calculate dissipated energy we evaluate the energy consumption in each phase of the proposed protocol (Sect. 4) according to the energy dissipation radio model (Sect. 3.1). Afterward, we determine the dissipated energy for each node s_i in each round based on its status (whether it is RN, CH or CM).

Energy consumption in prepare phase Dissipated energy for each RN due to receiving the beacon message from the MS is calculated as follows

$$E_{bea}(b_{cp}) = b_{cp} \times \epsilon_{elec}, \quad (20)$$

where b_{cp} is the size of a control packet. During neighbour finding, dissipated energy for each node s_i is calculated as

$$E_{nei}(i, b_{cp}, R_{com}) = \underbrace{b_{cp} \times \epsilon_{elec} + b_{cp} \times \epsilon_{amp} \times R_{com}^e}_{\text{broadcasting WELCOME packets in } R_{com}} + \underbrace{\eta_w^i \times b_{cp} \times \epsilon_{elec}}_{\text{receiving WELCOME packets}}, \quad (21)$$

where η_w^i is the number of WELCOME messages received by each node s_i

Energy consumption in clustering phase During CH advertisement, dissipated energy for each sensor s_i (except RNs) can be calculated as

$$E_{adv}(i, b_{cp}, R_C) = \underbrace{b_{cp} \times \epsilon_{elec} + b_{cp} \times \epsilon_{amp} \times R_C^e}_{\text{broadcasting } ADD_{CH} \text{ message}} + \underbrace{\eta_i \times b_{cp} \times \epsilon_{elec}}_{\text{receiving } ADD_{CH} \text{ message}}, \quad (22)$$

where η_i is the number of ADD_{CH} messages received by each node s_i from their neighbours. During cluster formation, dissipated energy for each CH s_i and its corresponding CM s_j are calculated respectively as:

$$E_{CH_C}(i, b_{cp}, R_C) = \underbrace{m_i \times b_{cp} \times \epsilon_{elec}}_{\text{receiving } JOIN_{REQ} \text{ messages}} + \underbrace{b_{cp} \times \epsilon_{elec} + b_{cp} \times \epsilon_{amp} \times R_C^e}_{\text{broadcasting TDMA message}}, \quad (23)$$

$$E_{CM_C}(j, b_{cp}, \delta) = \underbrace{b_{cp} \times \epsilon_{elec} + b_{cp} \times \epsilon_{amp} \times \delta(i, j)^e}_{\text{sending } JOIN_{REQ} \text{ message}} + \underbrace{b_{cp} \times \epsilon_{elec}}_{\text{receiving TDMA message}}, \quad (24)$$

where m_i is the number of $JOIN_{REQ}$ messages received by each cluster head s_i from non-CH nodes and $\delta(i, j)$ is the distance between CH s_i and its corresponding member s_j .

Energy consumption in routing and data transfer phase During routing tree construction, dissipated energy for each RN and CH can be respectively calculated as:

$$E_{RN_R}(i, b_{cp}, R_{com}) = \underbrace{b_{cp} \times \epsilon_{elec} + b_{cp} \times \epsilon_{amp} \times R_{com}^e}_{\text{broadcasting } HOP_ADV \text{ message}}, \quad (25)$$

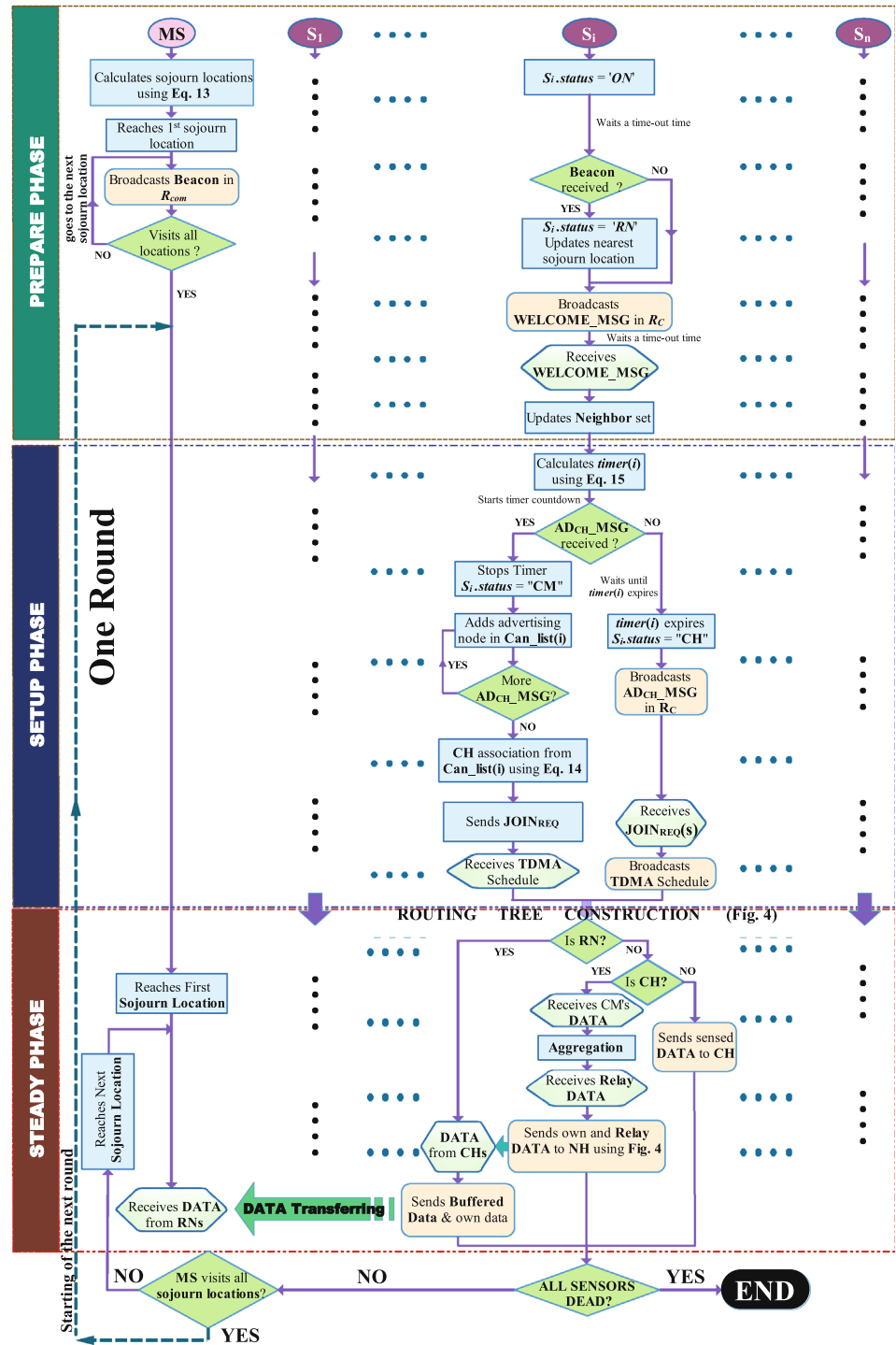
$$E_{CH_R}(i, b_{cp}, R_{com}) = \underbrace{\mu_i \times b_{cp} \times \epsilon_{elec}}_{\text{receiving } HOP_ADV \text{ messages}} + \underbrace{b_{cp} \times \epsilon_{elec} + b_{cp} \times \epsilon_{amp} \times R_{com}^e}_{\text{broadcasting } HOP_ADV \text{ message}}, \quad (26)$$

where μ_i is the number of HOP_ADV messages received by CH s_i from RN or other CHs having lower hop count.

During data transfer, dissipated energy for each CM, CH, and RN can be respectively calculated as:

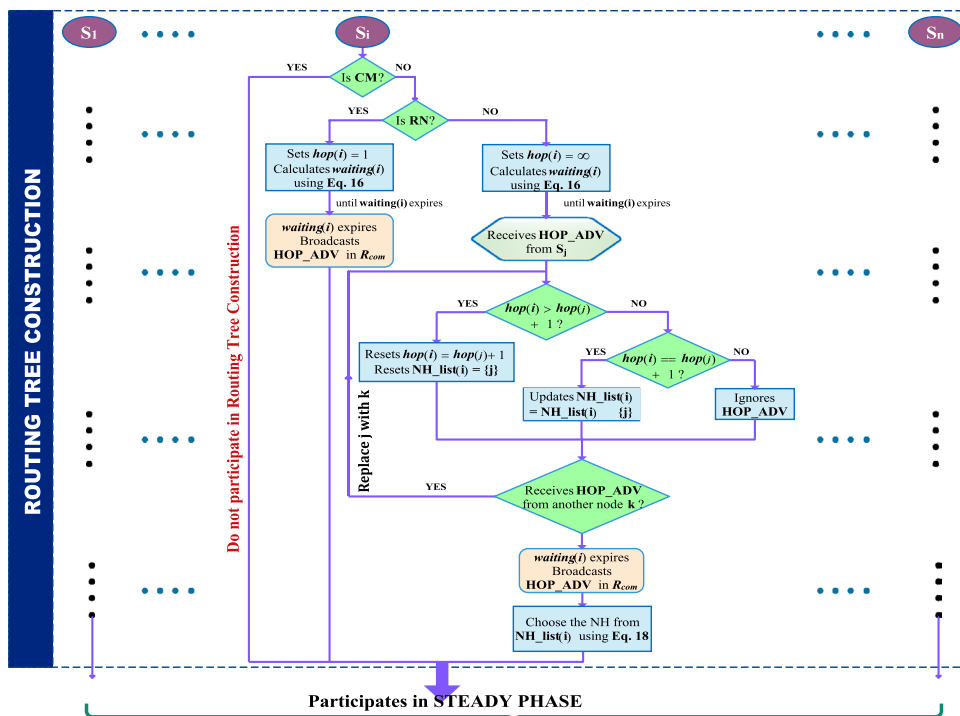
$$E_{CM_D}(j, b, \delta) = \underbrace{b \times \epsilon_{sen}}_{\text{sensing of data packet of } b \text{ bits}} + \underbrace{b \times \epsilon_{amp} \times \delta(i, j)^e}_{\text{transmitting of data packet to corresponding CH}}, \quad (27)$$

Fig. 3 Flowchart of the proposed protocol



$$\begin{aligned}
 E_{CH,D}(i, b, \delta) = & \underbrace{b \times \epsilon_{sen}}_{\text{sensing of data packet of } b \text{ bits}} + \underbrace{m_i \times b \times \epsilon_{elec}}_{\text{receiving of data packets from members}} + \underbrace{(m_i + 1) \times b \times \epsilon_{da}}_{\text{aggregating of received data packets along with own packet}} \\
 & + \underbrace{rel_i \times b \times \epsilon_{elec}}_{\text{receiving of data packets from other CHs}} + \underbrace{(rel_i + 1)(b \times \epsilon_{elec} + b \times \epsilon_{amp} \times \delta(i, NH_i)^e)}_{\text{transmitting of aggregated packet with relay packets to NH}}
 \end{aligned} \tag{28}$$

Fig. 4 Routing tree construction



$$\begin{aligned}
 E_{RN_D}(k, b, \delta) &= \underbrace{b \times \epsilon_{sen}}_{\text{sensing of data packet of } b \text{ bits}} + \underbrace{rel_k \times b \times \epsilon_{elec}}_{\text{receiving of data packets from CHs}} \\
 &+ \underbrace{(rel_k + 1)(b \times \epsilon_{elec} + b \times \epsilon_{amp} \times \delta(i, MS)^e)}_{\text{transmitting of relay packets with own packet to MS}},
 \end{aligned}
 \tag{29}$$

where b is the size of a data packet, m_i is the number of member nodes of CH s_i and $relay_i$ is the number of relay packets received by CH s_i from the distant cluster heads.

5.2 Theorems

Theorem 1 The running time complexity of the clustering algorithm is linear i.e., $\mathcal{O}(n)$, where n is the total number of sensors in the WSN.

Proof The proposed clustering algorithm can be divided into CH election and cluster formation phases. In CH election, every node (except the RNs) calculates its timer value and starts a stopwatch for timer countdown. As soon as the timer expires, it announces itself as a CH by broadcasting the ADD_{CH} message in its intra cluster communication range. Nodes for which a ADD_{CH} message (from its neighbouring node) comes before the timer expires, stop waiting for timer and turns into a non-CH node. It is clearly understood that the whole CH election process (calculating timer value,

broadcasting ADD_{CH} message etc.) is performed concurrently in each node. Hence, the running time complexity of the CH selection process is constant i.e., $\mathcal{O}(1)$. During cluster formation, every non-CH node is associated with its final CH from its candidate CH list by means of select() function. In worst case scenario, the number of candidate CHs is $(n - 1)$ i.e., $\mathcal{O}(n)$ where n is the total number of nodes. Hence the running time complexity of the proposed clustering algorithm can be calculated as:

$$T(n) = \mathcal{O}(1) + \mathcal{O}(n) = \mathcal{O}(n)$$

□

Theorem 2 The control message complexity for clustering is $\mathcal{O}(1)$ per node and $\mathcal{O}(n)$ for the entire network.

Proof As stated by the proposed clustering algorithm, as soon as timer value expires respective nodes announce their leadership by broadcasting ADD_{CH} control messages in their intra cluster range and then each non-CH node is associated with its final CH by sending the $JOIN_{REQ}$ control message to the final CH. Hence, clustering costs one control message per CH and CM i.e., $\mathcal{O}(1)$. On the other hand, since RNs do not participate in clustering and if there are k number of RNs in the network ($k \ll n$) then clustering causes $n - k$ number of control message flow in the entire network which is a linear function of n i.e., $\mathcal{O}(n)$ (Figs. 3, 4). □

Table 2 Experimental setup

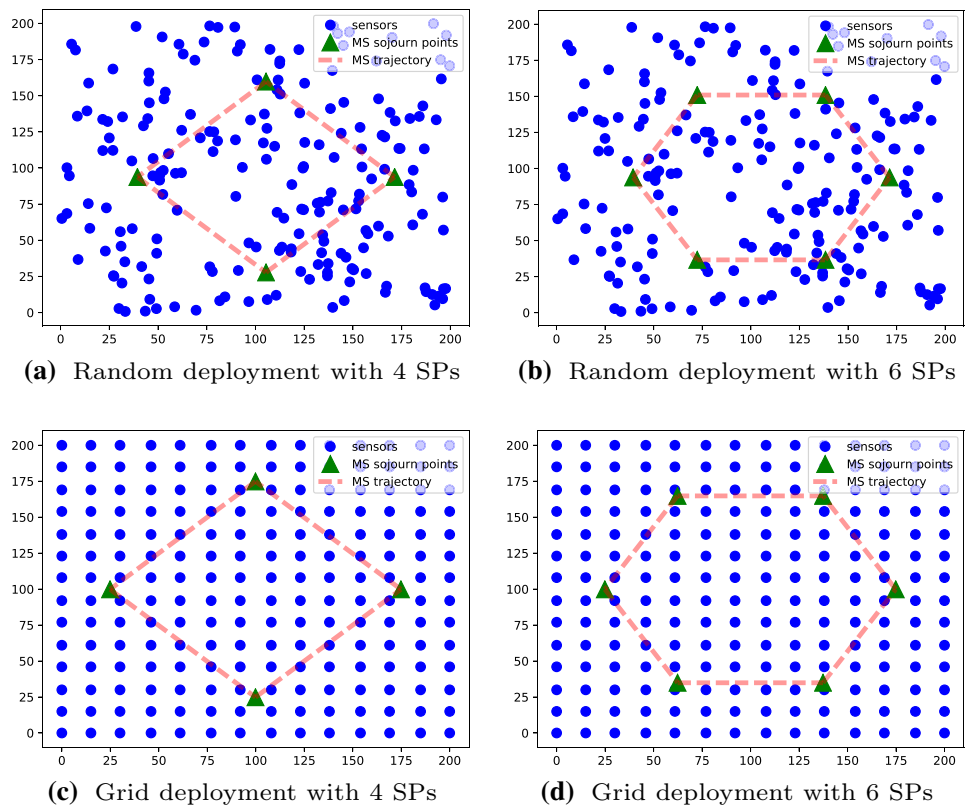
Region of Interest (RoI)	200 × 200 m ²
Total no. of sensors (N)	200–600
Initial energy of each sensor (E_{init})	0.5 J
Intra cluster radius (R_c)	20 m
Inter cluster radius (R_{com})	40 m
Data packet size (b)	4000 bits
Control packet size (b_{cp})	200 bits
Energy dissipation per bit in the electronic circuit (ϵ_{elec})	50 nJ
Free space energy dissipation per bit in the amplifier (ϵ_{fs})	10 pJ/bit/m ²
Multi path energy dissipation per bit in the amplifier (ϵ_{mp})	0.0013 pJ/bit/m ⁴
Path loss exponent (e)	2
Energy dissipation factor for one bit data aggregation (ϵ_{da})	5 nJ/bit/signal

Theorem 3 *The proposed clustering and routing algorithms are completely distributed in nature.*

Proof An algorithm in WSN is considered as distributed if each sensor node in the network can take a decision (for example CH election, cluster formation, next hop selection, etc.) based on the local information only i.e., the whole algorithm can be executed on every single node without any command of the MS. According to the proposed protocol, necessary parameters for CH election and cluster formation phases are residual energy, neighbour centrality, coverage

significance, and so on which are all locally available to the sensors. In the routing phase, the cost function for Next Hop (NH) selection requires parameters like residual energy of candidate NH, next hop degree of candidate NH, coverage significance of candidate NH etc., which can be evaluated by the local information available in the inter cluster range only. Thus, the proposed protocol performs its whole operation by allowing every node to share and communicate their information only within the inter cluster range without requiring any kind interference of the MS which proves the distributed nature of the protocol.

Fig. 5 Initial deployment scenarios



Theorem 4 *The proposed routing algorithm ensures the delivery of cluster data to its respective MS sojourn location in a loop-free manner.*

Proof We prove the above theorem by assuming that looping exists in the constructed routing tree (Proof by contradiction). A loop will occur around sensor s_i if the data packet delivered by s_i to s_j is again redelivered to s_i via some other sensor s_k i.e., the routing path will be like $s_i \rightarrow s_j \rightarrow \dots s_k \rightarrow s_i$. But in accordance with proposed routing algorithm, a CH s_i selects its next hop s_j from the $NH_list(i)$ consisting a list of CHs at lower hop within s_i 's communication range. Thus, for the selected next hop node of s_i $hop(i) > hop(j)$. It means $hop(i) > hop(j) > \dots hop(k) > hop(i)$. Hence, by transitivity $hop(i) > hop(i)$ which is logically incorrect. Thus the proposed routing algorithm assures a loop free routing path from each CH to MS sojourn location.

6 Experimental analysis

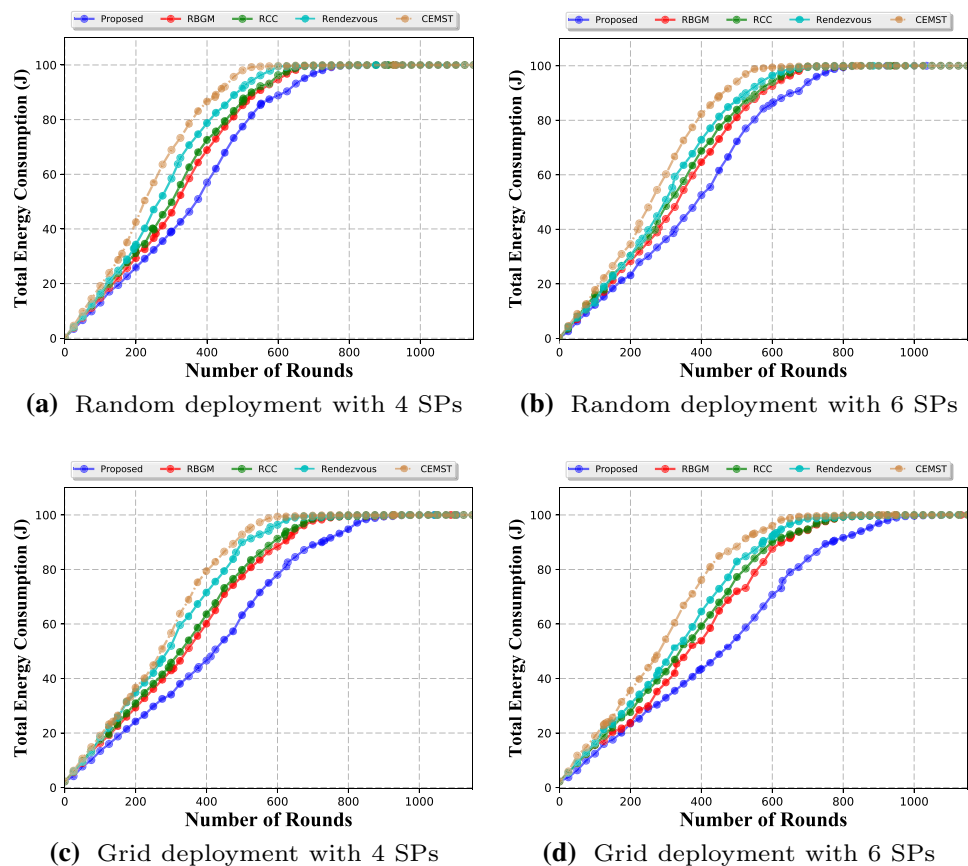
Implementing the WSN clustering and routing protocols in a real-life scenario is quite laborious for large-scale networks. Hence, the necessity of simulation tools has emerged to analyze and validate the performance of the WSN protocols. This section conducts an extensive simulation analysis to

demonstrate the improvements of the proposed protocol over the similar existing protocols namely CEMST (Chen et al. 2019), Rendezvous (Sharma et al. 2017), RCC (Yarinezhad and Hashemi 2018), and, RBGM (Yarinezhad and Hashemi 2019). The whole set of simulations has been executed using Python programming language (version 2.7) on the development environment Spyder 3.1.2. Every experiment is executed in an octa-core Intel(R) Xeon(R) processor server running the Windows Server 2012R2 operating system. Furthermore, NetworkX Python library is employed to design, manipulate, and examine the hierarchical routing structure of the five studied protocols. A tabular representation of the experimental setup based on (Heinzelman et al. 2000) is given below (Table 2).

6.1 Simulation environment

In order to verify the flexibility of the proposed protocol under various WSN applications, this simulated analysis takes into account both in-door and out-door deployments. In-door applications uses grid-based node deployment technique mostly in small area applications like smart building (Farsi et al. 2019). The random deployment technique is used for out-door applications with larger area like environment monitoring (Sharma et al. 2016). Figure 5a, b depict randomly distributed WSNs for 200 number

Fig. 6 Total Energy Consumption where $n = 200$



of nodes with a mobile sink having 4 and 6 Sojourn Points (SP) respectively. Likewise, Fig. 5c, d do the same for WSNs following grid deployment. Furthermore, it has been observed from the empirical studies that varying the density of the deployed sensors might have a crucial impact on the network performance metrics (Yue and He 2018). Therefore, this paper demonstrates the scalability of the proposed protocol by accomplishing the simulation analysis for the aforementioned scenarios with different number of nodes. It is worth mentioning that each scenario is simulated 50 times to get a stable output and the average of these is presented herein as a final result.

6.2 Performance evaluation

Based on the following metrics, the performance of the proposed protocol is evaluated:

1. *Energy consumption (in Joule)* It denotes the total energy consumption of the network due to the data and control packets transmission, reception, and, aggregation during clustering, routing and, data gathering phases. In the first round, is can be expressed as:

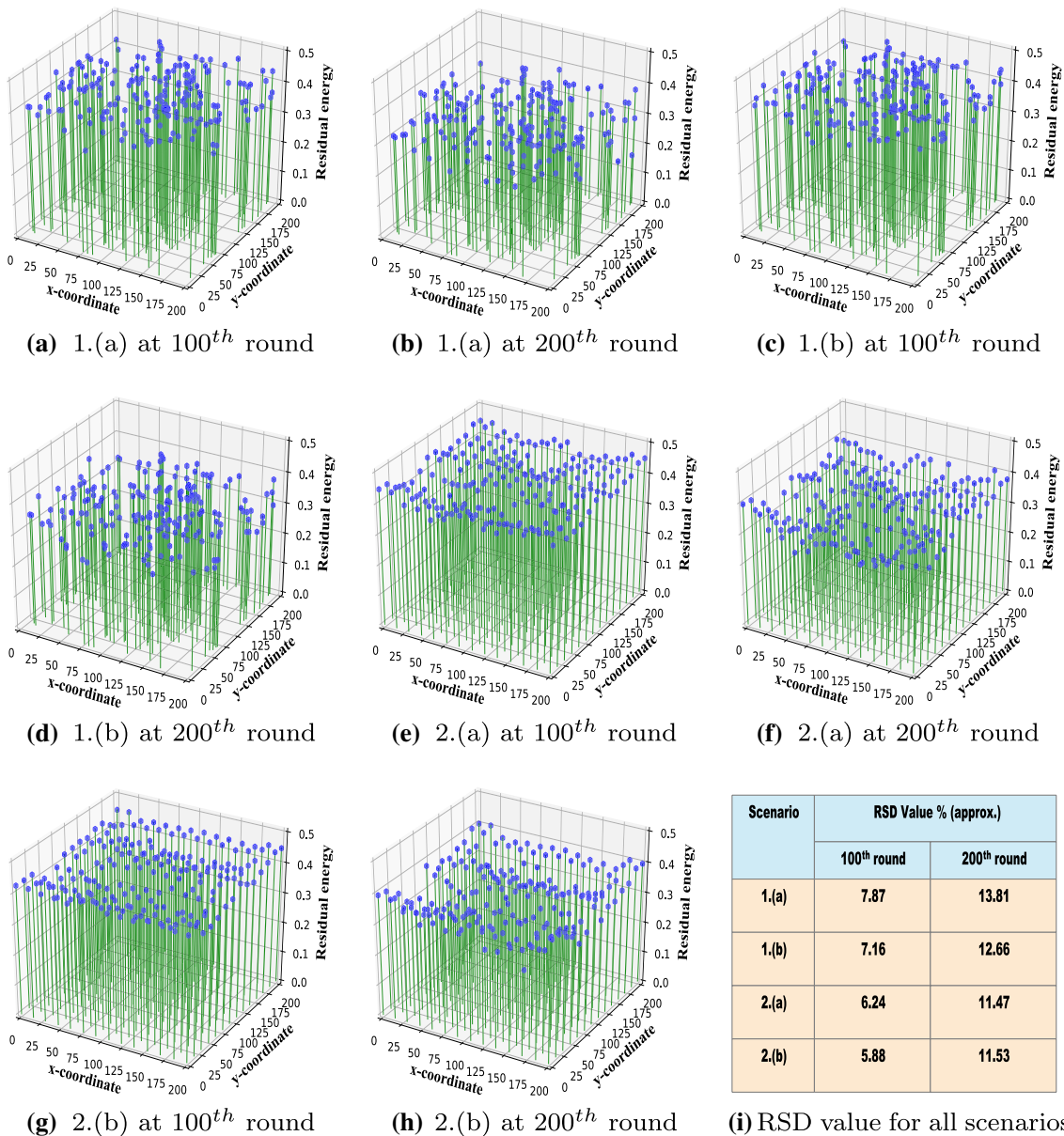


Fig. 7 Residual energy population and dispersion of the proposed protocol

$$E_{con}^1 = \sum_{i=1}^n e_1^i,$$

where e_1^i is the energy consumption of i th node in 1st round and n is the total number of nodes. Later rounds calculate this metric as a cumulative sum of the current and previous rounds' energy consumption e.g. energy consumption in r th can be evaluated as:

$$E_{con}^r = \sum_{i=1}^n e_1^i + \sum_{i=1}^n e_2^i + \dots \text{ } r \text{ times} \tag{30}$$

2. *Number of alive nodes* This metric denotes the total number of nodes whose residual energy is still greater than the energy threshold value ϵ (minimum energy required to perform the network operation). Hence, in a round R , a node i is declared as alive if

$$\left(E_{init} - \sum_{r=1}^{R-1} e_r^i \right) > \epsilon \text{ i.e., } E_{res}^i > \epsilon \tag{31}$$

where E_{init} is the initial energy of each node and e_r^i is the energy consumption of i th node in r th round.

3. *First Node Die (FND)* It is referred to the number of rounds after which the first alive node runs out its energy i.e., its residual energy drops below the energy threshold value ϵ . FND can be mathematically expressed as follows:

$$FND = \min \{ R^i \mid i = 1, 2, \dots, n \}$$

$$\text{subject to } \left(E_{init} - \sum_{r=1}^{R^i} e_r^i \right) \leq \epsilon \tag{32}$$

where R^i denotes the lifetime of i th node.

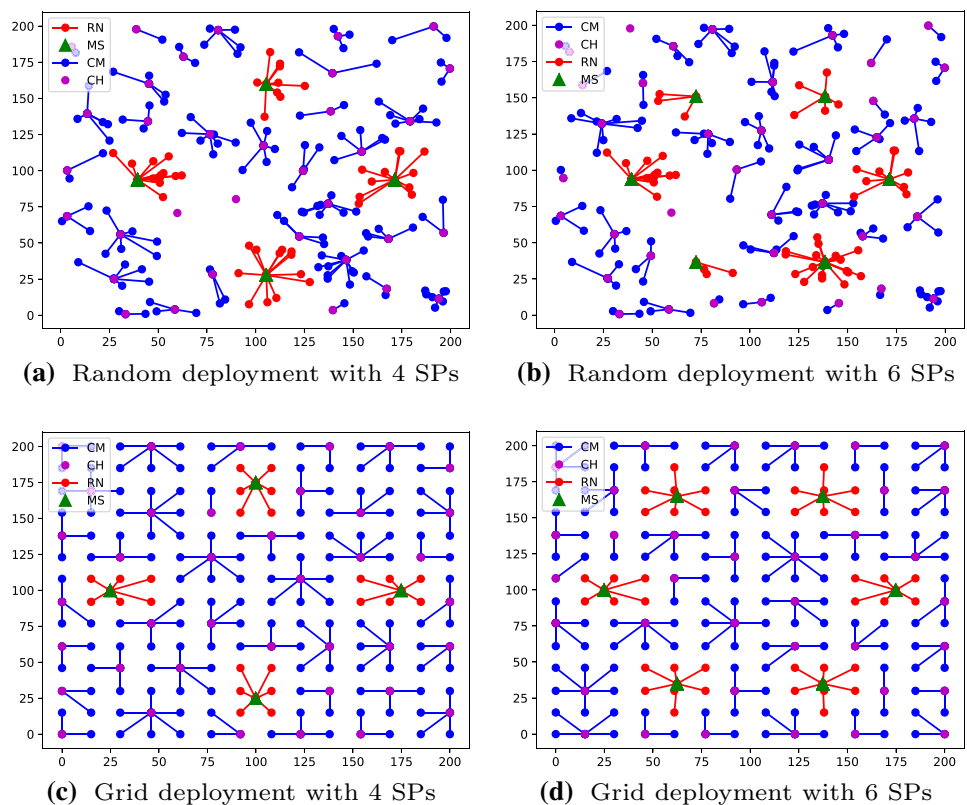
4. *Packet Delivery Ratio (PDR)* This metric generally defines the reliability property of the network. It denotes the rate of data packets received at the mobile sink i.e.,

$$PDR = \frac{\sum \text{No. of packets received by MS}}{\sum \text{No. of packets transmitted}}. \tag{33}$$

In unattended WSNs where the sink nature is sporadic and the deployment environment is hostile, packets reached to sink are always lesser in number than packets transmitted. As stated by random uniformed model (Abo-Zahhad et al. 2015), the probability of data packet loss dynamically increases with the increase in the distance between source i and destination j and is defined by

$$P_{loss}(i,j) = \begin{cases} 0 & \delta(i,j) \in [0, 50) \\ 0.01 * (\delta(i,j) - 50) & \delta(i,j) \in [50, 100] \\ 1 & \delta(i,j) \in (100, \infty) \end{cases} \tag{34}$$

Fig. 8 Cluster formation at 200th round where $n = 200$



If the link probability between the source and the next hop is greater than $P_{loss}(i, j)$ then the packet is assumed to be successfully delivered otherwise, dropped.

5. **Coverage ratio** This metric (Eq. 7) helps to determine that how long a WSN is capable to retain the full coverage criteria.
6. **End-to-End Delay (E2Ed)** It is the time (in ms) needed for a packet generated by some sensor node to be successfully received by the sink. It involves transmission, propagation, queuing and processing delays. The average delay of all the nodes in the WSN can be evaluated as:

$$E2Ed = \frac{\sum_{i=1}^n P_i \times (t_a - t_g)^i}{\sum_{i=1}^n P_i}, \tag{35}$$

where t_g is single packet generation time at the sensor node s_i , t_a is packet arrival time at the sink, P_i is the number of packets generated at sensor node s_i and successfully received by the sink, and, n is the total number of sensor nodes.

6.2.1 Energy efficiency

The key concern of any energy aware WSN is to preserve the residual energy of the sensor nodes which reflects the longevity of the network. In this regard, we have presented a comparison graph (Fig. 6) among the proposed and related existing protocols considering the total energy consumed in the network. It is obvious that the proposed one conserves more energy as compared to the others in all the considered scenarios (random and grid). In CEMST (Chen et al. 2019) despite using efficient clustering and routing strategy, the mobile sink strategy is not involved, thus the data routing load is not uniformly distributed. On the contrary, Rendezvous (Sharma et al. 2017) suggests a rendezvous based data collection strategy using sink mobilization but may suffer from long chain multi-hop routing. In comparison with

CEMST and Rendezvous, RRC (Yarinezhad and Hashemi 2018) and RBGM (Yarinezhad and Hashemi 2019) attain better energy distribution throughout the network with the help of virtual grid based hierarchical routing protocols. However, these protocols impose an additional overhead on the active grid headers as they have to relay the whole network data to the mobile sink (MS). This inflicts an uneven energy dissipation of the sensor nodes. Herein, an efficient hierarchical data dissemination protocol is designed to evenly diffuse the clustering and routing load across the network. Figure 7 shows the residual energy population of all nodes in the network at 100th and 200th round which reveals the uniform energy dissipation of the nodes in the proposed protocol. Thus, it validates the successful elimination of the energy hole problem. In order to better understand the uniformity of energy dissipation, Relative Standard Deviation (RSD) of residual energy population is presented in this section. It measures the dispersion of all nodes' residual energy around its mean value and can be calculated as:

$$RSD = \frac{\text{standard deviation of residual energy population}(\sigma)}{\text{mean of residual energy population}(\mu)} \times 100,$$

where lower RSD value corresponds to the low variation of residual energy distribution i.e., achieves better uniformity in terms of energy dissipation. Figure 7(i) tabulates the RSD values of Fig. 7a–h. The numerator σ and denominator μ are evaluated as:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (E_{res}(i) - \mu)^2}{n}} \quad \text{and} \quad \mu = \frac{\sum_{i=1}^n E_{res}(i)}{n}.$$

Finally, Fig. 8 shows a network instance of the clustered WSN at 200th round under different network scenarios to reveal in the uniform formation of clusters throughout the network.

Fig. 9 Network lifetime for different node densities ($n = 200, 400,$ and, 600)

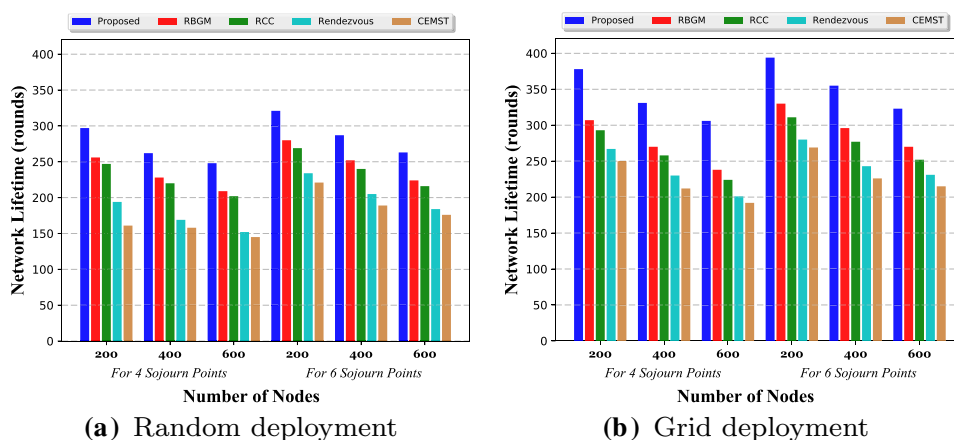


Fig. 10 Number of alive nodes

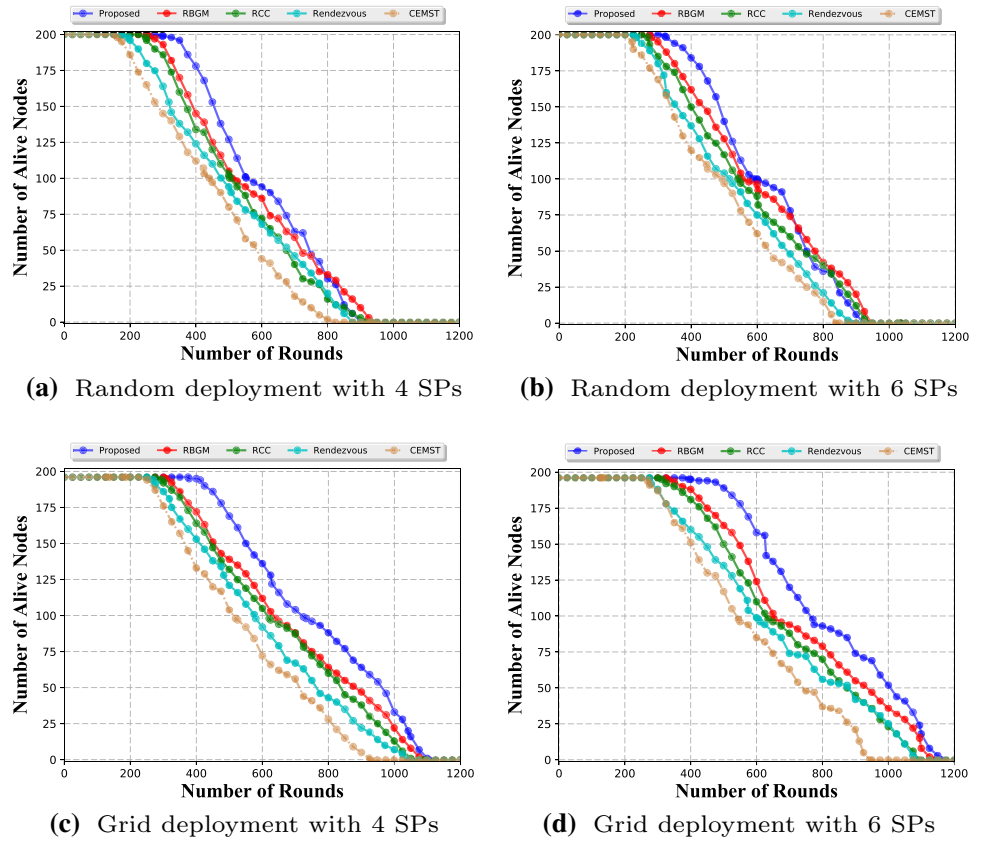


Fig. 11 Coverage ratio where $n = 200$

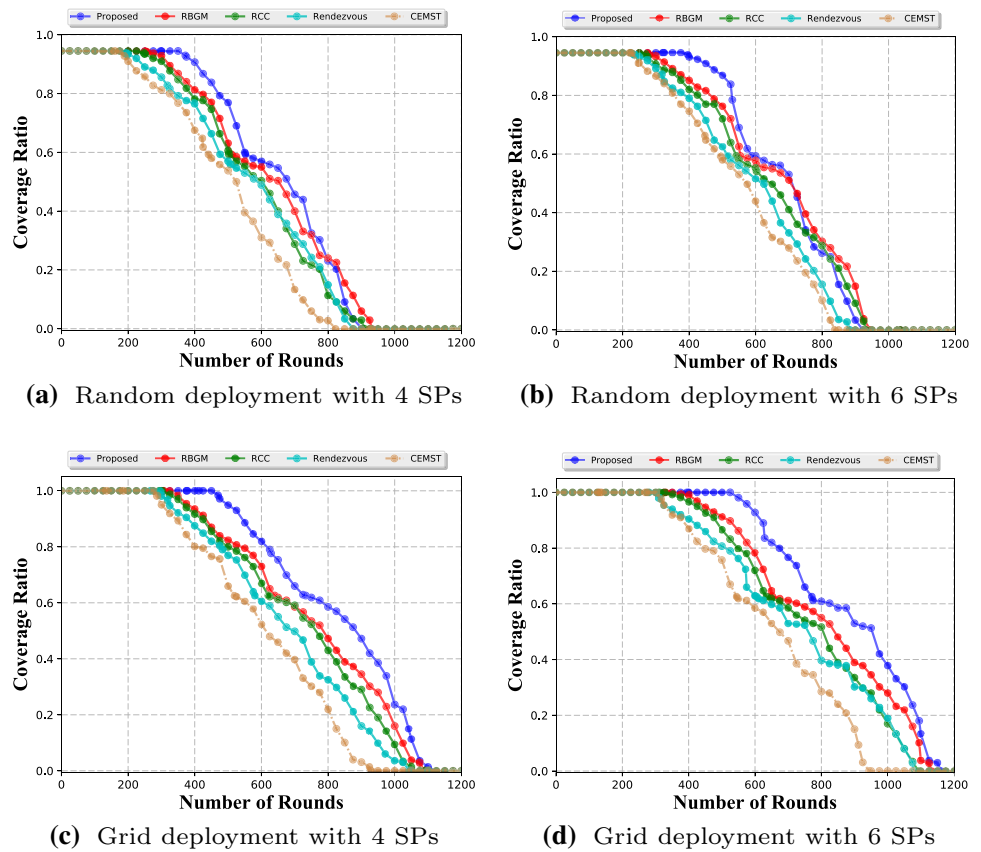
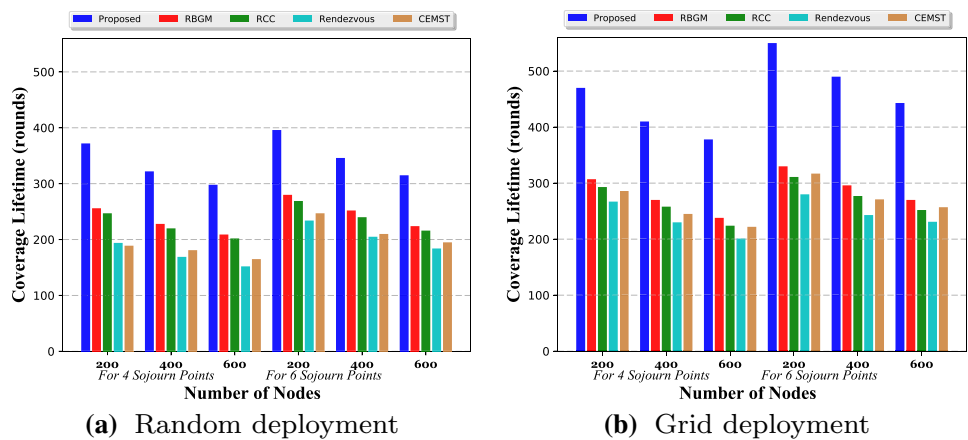


Fig. 12 Coverage lifetime for different node densities ($n = 200, 400, \text{ and } 600$)

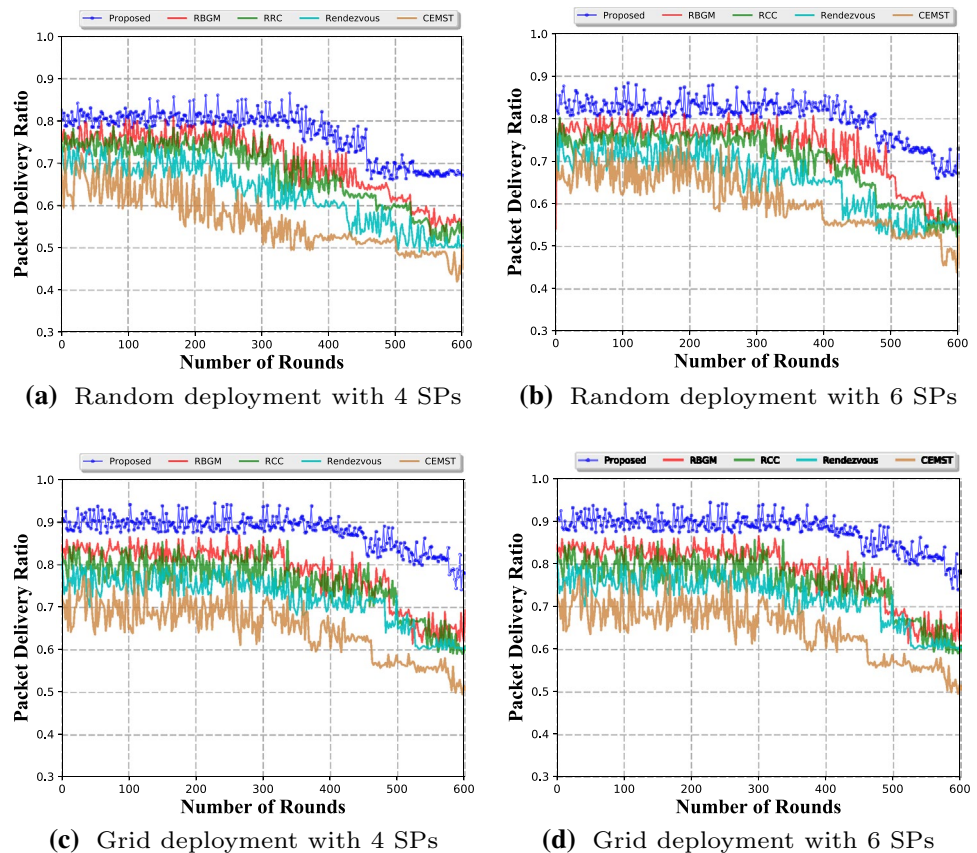


6.2.2 Network lifetime

The terminology for network lifetime has been expressed in several ways in the existing literature (Mazumdar and Om 2017, 2018). Most of the articles set a benchmark for network lifetime as First Node Die (FND) (Eq. 32) while some of the articles measure network lifetime by the time duration of Half of the Nodes Alive (HNA) in the network. This article considers the network lifetime as FND. Figure 9 provides insights into network lifetime in the aforementioned protocols by presenting bar graphs for different node

densities ($n = 200, 400, \text{ and } 600$) under different network scenarios. From Fig. 9a, it is observed that the proposed protocol enhances the network lifetime by 13–19% \times than RBGM (Yarinezhad and Hashemi 2019), 19–23% \times than RCC (Yarinezhad and Hashemi 2018), 40–63% \times than Rendezvous (Sharma et al. 2017), and, 45–84% \times than CEMST protocol (Chen et al. 2019). From Fig. 9b, we can notice that proposed method improves the network lifetime by 19–23% \times than RBGM, 27–37% \times than RCC, 41–52% \times than Rendezvous, and, 46–59% \times than CEMST protocol. Thus, the proposed one authenticates its scalability by attaining a

Fig. 13 Packet delivery ratio where $n = 200$



consistent performance over others for different node densities. Enhancement of network lifetime is obtained due to minimal energy dissipation per round caused by efficient clustering and routing procedure. Yet network lifetime with respect to FND is a very strict measure for the network performance evaluation since the network may deliver a considerable performance even after the death of few nodes. As a result, a line graph with respect to the number of alive nodes in each round is presented to demonstrate the betterment of the proposed protocol over related existing ones till HNA (Fig. 10).

6.2.3 Coverage preservation

One of the major concerns in coverage sensitive applications of WSN is full coverage preservation of the monitored area. This article deeply considers the full coverage of the target area. At the time of initial deployment, all the protocols satisfy the full coverage criteria but as the network operation goes on for a number of rounds, the energy of the sensor nodes starts decreasing and may result in the death of some sensors. Under such circumstances, it becomes a challenge to sustain the initial coverage ratio (Eq. 8) of the network. Moreover, in random sensor deployment, most state-of-the-art protocols fail to retain their initial coverage ratio as soon as the first node dies (FND) as they do not take into account any coverage sensitive metric during clustering as well as routing phase. Hence, the death of sensors from the dense region is not guaranteed and likewise, Rendezvous (Sharma et al. 2017), RRC (Yarinezhad and Hashemi 2018), and RBGM (Yarinezhad and Hashemi 2019) barely extend their coverage lifetime (Eq. 9) after FND. On the other hand, CEMST (Chen et al. 2019) and the proposed protocol incorporate the coverage significance (Eq. 6) as an additional parameter during Cluster Head (CH) and Next Hop (NH) selection phase which allows the CHs and NHs to be chosen from the dense region. This leads to the death of a CH or NH from the dense region which may leave the

smaller portion of the target area uncovered than the sparse region, and consequently maintains a higher coverage ratio (Eq. 7) after FND. Fig. 11 reveals that the proposed protocol surpasses the compared protocols in terms of coverage ratio (Eq. 7) for 200 number of nodes. To verify the competence of the proposed protocol under the impact of different node densities, we have further presented a bar graph (as shown in Fig. 12) of coverage lifetime for 200, 400, and 600 number of nodes. From Fig. 12a, it is noticed that the proposed protocol extends the coverage lifetime by 37–45% × than RBGM (Yarinezhad and Hashemi 2019), 44–50% × than RCC (Yarinezhad and Hashemi 2018), 68–96% × than Rendezvous (Sharma et al. 2017), and 61–97% × than CEMST protocol (Chen et al. 2019). From Fig. 12b, we can observe that the proposed method improves the coverage lifetime by 36–65% × than RBGM, 59–77% × than RCC, 76–1.01 × than Rendezvous, and 64–81% × than CEMST protocol.

6.2.4 Packet delivery

Successful delivery of data packets to the mobile sink implies the network reliability level i.e., higher the Packet Delivery Ratio (PDR) (Eq. 33) better the link quality and, consequently improves the network Quality of Service (QoS). Figure 13 manifests the efficacy of the proposed protocol over the similar ones with respect to the metric PDR for 200 number of nodes. CEMST (Chen et al. 2019) suffers from a lower PDR as it maintains a higher network hop count due to the presence of static sink. Rendezvous (Sharma et al. 2017) on the other hand, obeys a chain based routing where the death of a single rendezvous node breaks the chain topology and results in an increasing packet drop. Methods suggested in RCC (Yarinezhad and Hashemi 2018) and RBGM (Yarinezhad and Hashemi 2019) experience comparatively lower network hop count but impose a heavy burden on the active cell header which may lead to the buffer overflow, thus enhances the probability of packet drop. In contrast, the proposed protocol distributes the data

Fig. 14 Average PDR till FND for different node densities ($n = 200, 400,$ and, 600)

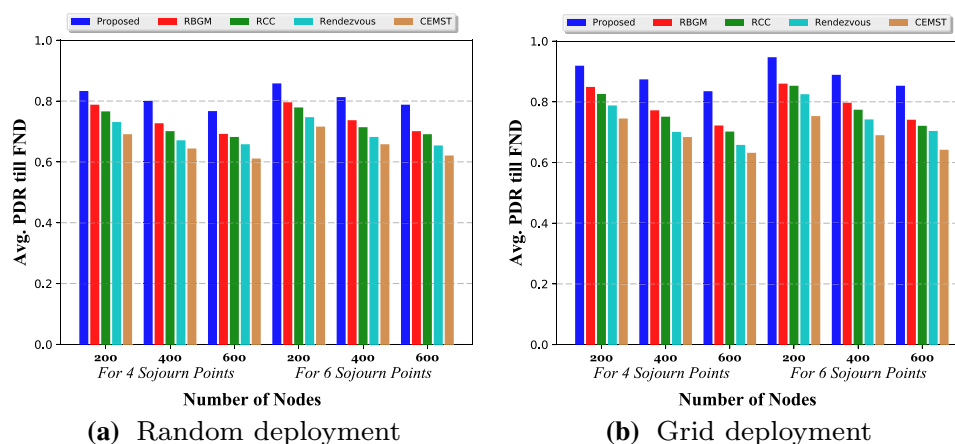


Fig. 15 End to end delay where $n = 200$

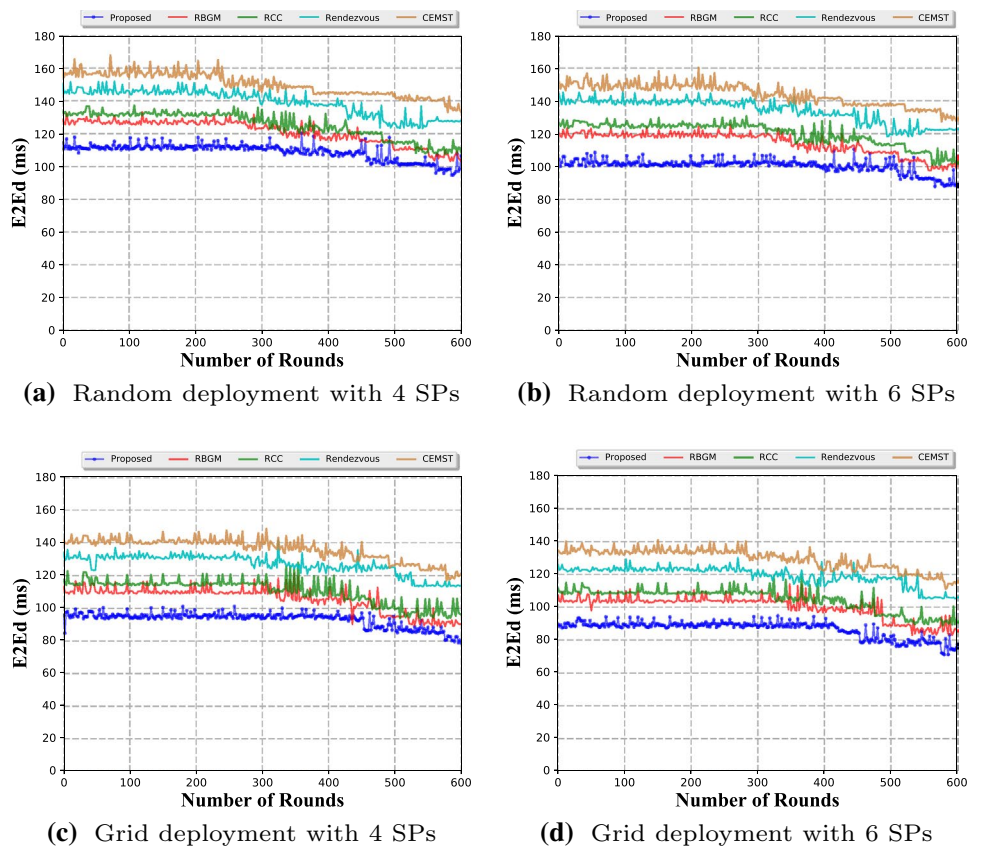
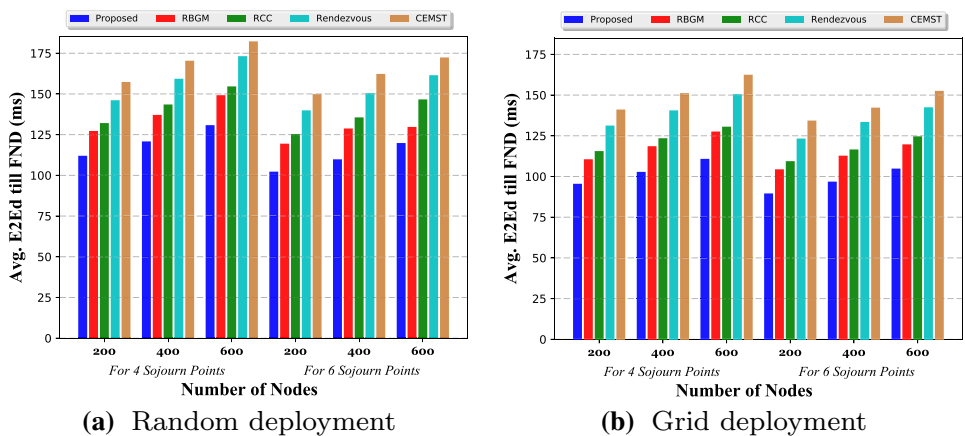


Fig. 16 Average E2Ed till FND for different node densities ($n = 200, 400, \text{ and } 600$)



forwarding load among the multiple RNs along with better hop count minimization; accordingly it achieves a higher PDR. Furthermore, Fig. 14 shows a bar representation of the foregoing protocols with respect to average PDR till FND for 200, 400, and 600 number of nodes. It is observed from Fig. 14a that the proposed protocol improves the average PDR value till FND by 6–12% × than RBGM (Yarinezhad and Hashemi 2019), 9–14% × than RCC (Yarinezhad and Hashemi 2018), 14–20% × than Rendezvous (Sharma et al. 2017), and 20–27% × than CEMST protocol (Chen et al. 2019). Likewise, Fig. 14b exhibits that the proposed method

improves the average PDR value till FND by 8–16% × than RBGM, 11–19% × than RCC, 15–27% × than Rendezvous, and 23–33% × than CEMST protocol.

6.2.5 End-to-end delay

A lower value of average end-to-end delay ($E2Ed$) (Eq. 35) makes the network feasible for delay-sensitive applications of WSN and therefore this article aims to achieve minimal data delivery delay to the Mobile Sink (MS). Figure 15 presents a comparison graph for 200 number of nodes to

demonstrate the dominance of the proposed protocol over the related existing ones in terms of average end-to-end delay of the network (Eq. 35). It is worth mentioning that both the CEMST (Chen et al. 2019) and Rendezvous (Sharma et al. 2017) encounter higher $E2Ed$ value due to the presence of long-chain multi-hop routing. Despite minimizing the network hop count, RCC (Yarinezhad and Hashemi 2018) and RBGM (Yarinezhad and Hashemi 2019) assign extra load on the active cell header which inflicts the buffer overflow, and thus leading to dropped packets and increasing network delay. On the other hand, the proposed routing protocol guarantees a minimum hop data delivery path to the MS and obtains a lower $E2Ed$ value. In addition, a bar graph of the foregoing protocols with respect to average $E2Ed$ till FND is presented in Fig. 16 for 200, 400, and 600 number of nodes. It is observed from Fig. 16a that the proposed protocol reduces the average $E2Ed$ till FND by 8–17% \times than RBGM (Yarinezhad and Hashemi 2019), 18–23% \times than RCC (Yarinezhad and Hashemi 2018), 30–37% \times than Rendezvous (Sharma et al. 2017), and 39–47% \times than CEMST protocol (Chen et al. 2019). Similarly, Fig. 12b exhibits that the average $E2Ed$ value of the proposed protocol till FND is improved by 14–17% \times than RBGM, 18–22% \times than RCC, 36–38% \times than Rendezvous, and 45–50% \times than CEMST protocol. It is to be noted that with the increasing number of nodes the network congestion will be high; accordingly, the average $E2Ed$ value increases.

7 Conclusion and future work

This article proposes an energy and coverage aware hierarchical routing protocol which alleviates the energy hole problem by means of restricted mobile sink motion. The proposed protocol is completely distributed in nature which results in reduced communication and computation overhead. Inclusion of the coverage parameter in the proposed protocol ensures the enhancement of the coverage lifetime till a certain number of nodes die. Moreover, the routing strategy guarantees a minimal hop path for each CH to disseminate their cluster data to the appropriate MS sojourn location. To justify the effectiveness of the proposed protocol over the compared ones we have performed the experimental analysis and comparisons in terms of different metrics like network lifetime, total energy consumption, coverage ratio, packet delivery ratio, etc.

However, this work considers a set of predetermined MS sojourn locations for data collection. A possible future direction of this work may consider finding the optimal sojourn locations particularly with multiple mobile sinks support. In addition, we have assumed that the MS has no power and buffer constraints to collect data from the sensor network. Thus, further research may include energy and buffer

constrained MS which is an evolving challenge in applications like the Internet of Things (IoT), smart environment, Internet of Nano Things (IoNT), and others.

References

- Abdul-Salaam G, Abdullah AH, Anisi MH, Gani A, Alelaiwi A (2016) A comparative analysis of energy conservation approaches in hybrid wireless sensor networks data collection protocols. *Telecommun Syst* 61(1):159–179
- Abo-Zahhad M, Ahmed SM, Sabor N, Sasaki S (2015) Mobile sink-based adaptive immune energy-efficient clustering protocol for improving the lifetime and stability period of wireless sensor networks. *IEEE Sens J* 15(8):4576–4586
- Alaei M, Yazdanpanah F (2019) Eelcm: an energy efficient load-based clustering method for wireless mobile sensor networks. *Mob Netw Appl* 24(5):1486–1498
- Anisi MH, Rezazadeh J, Dehghan M (2008) Feda: fault-tolerant energy-efficient data aggregation in wireless sensor networks. In: 2008 16th international conference on software, telecommunications and computer networks, IEEE, pp 188–192
- Anisi MH, Abdullah AH, Razak SA (2012) Efficient data gathering in mobile wireless sensor networks. *Life Sci J* 9(4):2152–2157
- Anisi MH, Abdullah AH, Razak SA (2013) Energy-efficient and reliable data delivery in wireless sensor networks. *Wirel Netw* 19(4):495–505
- Benzaïd C, Bagaa M, Younis M (2017) Efficient clock synchronization for clustered wireless sensor networks. *Ad Hoc Netw* 56:13–27
- Bhatia M, Sood SK (2018) Internet of things based activity surveillance of defence personnel. *J Ambient Intell Humaniz Comput* 9(6):2061–2076
- Bozorgi SM, Bidgoli AM (2019) Heec: a hybrid unequal energy efficient clustering for wireless sensor networks. *Wirel Netw* 25(8):4751–4772
- Chen DR, Chen LC, Chen MY, Hsu MY (2019) A coverage-aware and energy-efficient protocol for the distributed wireless sensor networks. *Comput Commun* 137:15–31
- Christopher VB, Jasper J (2020) Dhgrp: dynamic hexagonal grid routing protocol with mobile sink for congestion control in wireless sensor networks. *Wirel Pers Commun* 112:1–20
- Deebak B, Al-Turjman F (2020) A hybrid secure routing and monitoring mechanism in IOT-based wireless sensor networks. *Ad Hoc Netw* 97:102022
- Elkamel R, Messouadi A, Cherif A (2019) Extending the lifetime of wireless sensor networks through mitigating the hot spot problem. *J Parallel Distrib Comput* 133:159–169
- Farsi M, Elhosseini MA, Badawy M, Ali HA, Eldin HZ (2019) Deployment techniques in wireless sensor networks, coverage and connectivity: a survey. *IEEE Access* 7:28940–28954
- Fotohi R, Bari SF (2020) A novel countermeasure technique to protect wsn against denial-of-sleep attacks using firefly and hopfield neural network (HNN) algorithms. *J Supercomput*. <https://doi.org/10.1007/s11227-019-03131-x>
- Fotohi R, Firoozi Bari S, Yusefi M (2020) Securing wireless sensor networks against denial-of-sleep attacks using RSA cryptography algorithm and interlock protocol. *Int J Commun Syst* 33(4):e4234
- Glaroudis D, Iossifides A, Chatzimisios P (2020) Survey, comparison and research challenges of iot application protocols for smart farming. *Comput Netw* 168:107037
- Gu X, Yu J, Yu D, Wang G, Lv Y (2014) Ecdc: an energy and coverage-aware distributed clustering protocol for wireless sensor networks. *Comput Electr Eng* 40(2):384–398

- Gumaida BF, Luo J (2019) Novel localization algorithm for wireless sensor network based on intelligent water drops. *Wirel Netw* 25(2):597–609
- Heinzelman WR, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd annual Hawaii international conference on system sciences, IEEE, p 10
- Jamali S, Fotohi R (2016) Defending against wormhole attack in manet using an artificial immune system. *New Rev Inf Netw* 21(2):79–100
- Jamali S, Fotohi R (2017) Dawa: defending against wormhole attack in manets by using fuzzy logic and artificial immune system. *J Supercomput* 73(12):5173–5196
- Krishnan M, Yun S, Jung YM (2019) Dynamic clustering approach with ACO-based mobile sink for data collection in WSNS. *Wirel Netw* 25(8):4859–4871
- Mabodi K, Yusefi M, Zandiyan S, Irankhah L, Fotohi R (2020) Multi-level trust-based intelligence schema for securing of internet of things (IOT) against security threats using cryptographic authentication. *J Supercomput*. <https://doi.org/10.1007/s11227-019-03137-5>
- Mazumdar N, Om H (2017) Distributed fuzzy logic based energy-aware and coverage preserving unequal clustering algorithm for wireless sensor networks. *Int J Commun Syst* 30(13):e3283
- Mazumdar N, Om H (2018) Distributed fuzzy approach to unequal clustering and routing algorithm for wireless sensor networks. *Int J Commun Syst* 31(12):e3709
- Mehto A, Tapaswi S, Pattanaik K (2020) Virtual grid-based rendezvous point and sojourn location selection for energy and delay efficient data acquisition in wireless sensor networks with mobile sink. *Wirel Netw* 26:1–17
- Mittal N, Srivastava R (2020) An energy efficient clustered routing protocols for wireless sensor networks. In: Recent trends and advances in artificial intelligence and internet of things. Springer, New York, pp 581–596
- Mottaghi S, Zahabi MR (2015) Optimizing leach clustering algorithm with mobile sink and rendezvous nodes. *AEU Int J Electron Commun* 69(2):507–514
- Ramos HS, Boukerche A, Oliveira AL, Frery AC, Oliveira EM, Loureiro AA (2016) On the deployment of large-scale wireless sensor networks considering the energy hole problem. *Comput Netw* 110:154–167
- Rashid B, Rehmani MH (2016) Applications of wireless sensor networks for urban areas: a survey. *J Netw Comput Appl* 60:192–219
- Ren J, Zhang Y, Zhang K, Liu A, Chen J, Shen XS (2016) Lifetime and energy hole evolution analysis in data-gathering wireless sensor networks. *IEEE Trans Ind Inform* 12(2):788–800
- Seyedi B, Fotohi R (2020) Niashpt: a novel intelligent agent-based strategy using hello packet table (HPT) function for trust internet of things. *J Supercomput*. <https://doi.org/10.1007/s11227-019-03143-7>
- Sharma S, Puthal D, Jena SK, Zomaya AY, Ranjan R (2017) Rendezvous based routing protocol for wireless sensor networks with mobile sink. *J Supercomput* 73(3):1168–1188
- Sharma V, Patel R, Bhadauria H, Prasad D (2016) Deployment schemes in wireless sensor network to achieve blanket coverage in large-scale open area: a review. *Egypt Inform J* 17(1):45–56
- Soro S, Heinzelman WB (2009) Cluster head election techniques for coverage preservation in wireless sensor networks. *Ad Hoc Netw* 7(5):955–972
- Sundhari RM, Jaikumar K (2020) Iot assisted hierarchical computation strategic making (HCSM) and dynamic stochastic optimization technique (DSOT) for energy optimization in wireless sensor networks for smart city monitoring. *Comput Commun* 150:226–234
- Ullah Z (2020) A survey on hybrid, energy efficient and distributed (heed) based energy efficient clustering protocols for wireless sensor networks. *Wirel Pers Commun* 112:1–29
- Vijayalakshmi V, Senthilkumar A (2019) Uscdrp: unequal secure cluster-based distributed routing protocol for wireless sensor networks. *J Supercomput* 76:1–16
- Wang J, Cao J, Sherratt RS, Park JH (2017) An improved ant colony optimization-based approach with mobile sink for wireless sensor networks. *J Supercomput* 74:1–13
- Wang J, Gao Y, Yin X, Li F, Kim HJ (2018a) An enhanced pegasis algorithm with mobile sink support for wireless sensor networks. *Wirel Commun Mob Comput*. [10.1155/2018/9472075](https://doi.org/10.1155/2018/9472075)
- Wang Z, Zhang B, Wang X, Jin X, Bai Y (2018b) Improvements of multihop localization algorithm for wireless sensor networks. *IEEE Syst J* 99:1–12
- Yarinezhad R (2019) Reducing delay and prolonging the lifetime of wireless sensor network using efficient routing protocol based on mobile sink and virtual infrastructure. *Ad Hoc Netw* 84:42–55
- Yarinezhad R, Hashemi SN (2018) A cellular data dissemination model for wireless sensor networks. *Pervasive Mob Comput* 48:118–136
- Yarinezhad R, Hashemi SN (2019) An efficient data dissemination model for wireless sensor networks. *Wirel Netw* 25(6):3419–3439
- Yue YG, He P (2018) A comprehensive survey on the reliability of mobile wireless sensor networks: taxonomy, challenges, and future directions. *Inf Fusion* 44:188–204

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.