Peer-to-Peer Networking and Applications
https://doi.org/10.1007/s12083-019-00777-5

# A novelistic approach for energy efficient routing using single and multiple data sinks in heterogeneous wireless sensor network

Sandeep Verma[1] · Neetu Sood[1] · Ajay Kumar Sharma[2]

Received: 29 June 2018 / Accepted: 19 June 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

**Abstract**
In this paper, Improved Dual Hop Routing protocol (IDHR) and Multiple data sink-based Energy Efficient Cluster-based routing protocol (MEEC) are proposed. The Cluster Head (CH) selection in IDHR and MEEC is done by incorporating node density parameter along with other parameters, namely energy and distance between the node and the sink. In MEEC, multiple data sinks are employed to pact with the burden on the relaying nodes involved in data forwarding. The node density factor proves to be adherent for energy preservation of nodes by abating the average communication distance between the nodes and respective CH. The employment of multiple data sinks in MEEC avoids any dual hop communication between CHs and sink, which in turn alleviates hot-spot problem and ameliorates network longevity. As evident from the simulation results, IDHR and MEEC individually outperform competitive protocols, namely SEECP, DRESEP and TEDRP with respect to various performance metrics. Furthermore, the performance comparison of MEEC and IDHR is investigated to limelight the essence of employing multiple data sinks for a network.

**Keywords** Multiple data sinks · MEEC · IDHR · Clustering · Heterogeneous routing protocol · Hot-spot problem

## 1 Introduction

Recent advancements in the area of wireless communications and miniature electronics have fostered the development of tiny sized, low-cost, multifunctional sensor nodes that communicate in the wireless medium [1]. These sensor nodes have processing and communicating capabilities that facilitate them to perform networking in any attended or unattended (hostile/remote) areas. The sensor nodes form a network to collect data from their proximate surroundings and thereafter coordinate among themselves for execution of certain tasks. Later, these nodes send orchestrated data to the data sink or the base station. The network formed by these nodes is termed as Wireless Sensor Network (WSN) [2]. The list of applications in which sensor network has a significant role to play, are only limited by the human imagination. A few of its applications include remote healthcare [3], disaster management, environment and industrial monitoring, reconnaissance and targeting system, battlefield surveillance, air pollution and agricultural monitoring [4, 5].

The battery consumption of a sensor node is directly dependent on the way the nodes communicate with each other or with the sink. The grouping of sensor nodes, i.e., clustering not only enhances the scalability of the network but also helps in acquiring energy balancing among nodes. Cluster Head (CH) disseminates the data from the cluster member nodes and forward it to the next CH or to the sink depending upon the single or dual hop communication [6]. There are various parameters, namely residual energy, initial energy, distance factor, etc. that are adherent for selecting any node as CH. However, node density factor is yet not included in various routing strategies aiming for network longevity. When the network area is large, the nodes get deployed at higher average distance from each other. Selecting a particular node which is surrounded by higher number of nodes as a CH, will make it easier for that node (CH) to disseminate data from cluster member nodes at a reduced distance. It will further result in energy preservation of cluster nodes in that particular cluster

✉ Sandeep Verma
sandi287@gmail.com

Neetu Sood
soodn@nitj.ac.in

Ajay Kumar Sharma
sharmaajayk@rediffmail.com

[1] Department of ECE, Dr. B. R, Ambedkar National Institute of Technology, Jalandhar 144011, India

[2] I. K. Gujral Punjab Technical University, Kapurthala 110040, India

Published online: 06 July 2019

[7]. CH also performs data aggregation other than the task of data collection, consequently, it devours energy much faster than the other nodes of the cluster. The heterogeneous wireless sensor network (HWSN) has been promising to balance the energy consumption of nodes declared as CHs [8]. It is noted that energy heterogeneity can be introduced in the nodes at any level of energy which basically comprises of nodes equipped with extra energy resources [9]. Therefore, in this work, our one of the concerns is to improve the method of CH selection in HWSN vis-à-vis state-of-the-art routing protocols.

The placement of sink is mostly dependent upon the topology of the network. In some unattended applications, where the targeted area is quite large, the sink is supposed to be placed outside that area and eventually the nodes communicate through multi-hop communication [10]. While doing so, hot-spot problem or energy hole occurs due to the excessive relaying load on the nodes nearby to the sink as shown in Fig. 1 (a) [11]. With the single sink in operation, the nodes undergo multi hop communication encountering the following problems [12].

a) Procrastination in data delivery is caused due to the congestion around the single sink.
b) Scalability gets affected due to the expansion of network that physically ends up in enhancing the communicating distance between nodes and the sink.
c) Single sink dependence of the network is highly vulnerable to the 'No Communication' condition whenever sink failure occurs due to any particular reason.
d) Frequent selection or frequent rotation of relay node leads to the enhancement in number of overheads.

There has been a great magnitude of research which deals with hot-spot problem. Many of the researchers have given significant attention to relieve the burden on the relaying nodes located near to the data collecting sink [12, 13]. The hot-spot problem is taken into consideration in the various corona-based models constructed to acquire uniform energy consumption in the network [11]. The techniques incorporated for corona-based models include addition of relay nodes, sink mobility, ammendments in the number and width of corona, etc. However, despite of many such endeavours made in the direction of alleviation of hot-spot problem [14], it still comes out to be a NP (Non Polynomial)-hard problem. To deal with such scenario, we employ multiple data sinks outside the network as shown in Fig. 1 (b). It tends to reduce the energy consumption of nodes by reducing effective communication distance.

The sensor nodes have finite energy resources (e.g., 1.2 V, <0.5 AH batteries) [15, 16], therefore, the energy of these nodes should be preserved to acquire higher network lifetime. Whenever the sensor nodes are operated in unattended areas, it becomes cumbersome to replace or charge their batteries. It is noted that once the battery of a node is exhausted, it is said to be dead and if it happens with all other nodes, then the whole network becomes dead. Therefore, our prominent focus is on enhancing the network lifetime and the stability period (number of rounds covered till first node is dead) of the network.

### 1.1 Major contributions

In order to address the above discussed essential issues related to the 'CH selection in HWSN' and 'to pact with hot-spot problem', the following contributions are reported.

a) Two protocols, namely IDHR (using single sink) and MEEC (using multiple sinks) are proposed that involve dual and single hop communication, respectively. The CH selection is done by incorporating node density factor along with the residual energy of a node. In addition, the distance of a node from the only sink (in IDHR) and from the nearest sink (in MEEC) is also considered.
b) To improve inter cluster communication in IDHR, a new dual hop decisive strategy is implemented. It defines the

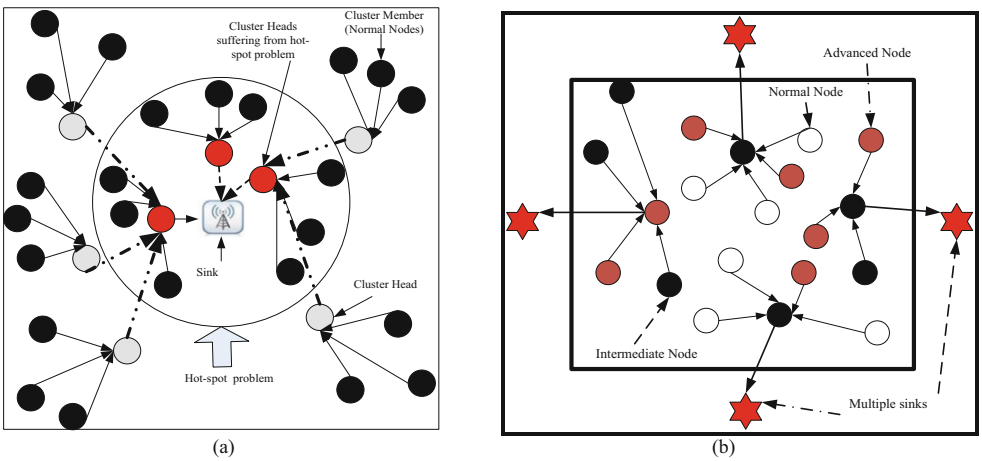



**Fig. 1** (a) Single Sink causing hot-spot problem (b) Network with four sinks

radius of the inner circle by an average distance of all the nodes from the sink. The CH selection is favored inside the circle by incorporating the distance factor in its probability expression to reduce the distance between the nodes and the sink.

c) The hot-spot problem is mitigated by the MEEC protocol as it involves single hop communication from each CH to the nearest sink in the network. It preserves a significant amount of energy of each CH node which is selected by following the same approach as in IDHR.

d) The state-of-the-art heterogeneous protocols, namely DRESEP [17], SEECP [18] and TEDRP [19] are taken into consideration for the performance evaluation of the proposed protocols on the benchmark of different performance metrics.

The aforementioned contributions and corresponding discussions are comprehensively discussed in Section 3.3 where the proposed protocols are also distinguished from the state-of-the-art protocols.

The rest of the manuscript is organized as follows. Section 2 explains the related work done in the direction of achieving enhanced stability period in the HWSN. Some of the research work employing multiple data sinks is also discussed at the end of this section. Section 3 discusses the state-of-the-art heterogeneous protocols; DRESEP [17] and SEECP [18] with their shortcomings. The network model and proposed methodology of IDHR and MEEC are discussed in Section 4. The results and discussions for IDHR and MEEC are comprehensively presented in Section 5. The conclusion of the entire contribution made to this paper is reported in Section 6. A little direction to the future work is also given along with conclusion. Thereafter, the references incorporated in the proposed work are listed.

## 2 Related work

A plethora of research work has been reported that contributed in the enhancement of network stability and longevity of HWSN [9]. The cluster-based routing has played a crucial role in reducing the number of transmissions of sensor nodes to a sink. Clustering renders scalability to the network that makes it possible to extend the physical boundaries of the network with respect to the significant number of nodes deployed in the network. The literature work reported in this manuscript reviews the research contributions made in the direction of three aspects. These aspects include CH selection for HWSN to implement energy efficient routing, various routing strategies encountered to alleviate hot-spot problem and lastly the routing techniques based on multiple data sinks in the network.

### 2.1 CH selection for HWSN

To investigate the first aspect out of the aforementioned ones, it is observed that the potential researchers have focused on making the clustering energy efficient by introducing various methods for the selection of CHs [20] . The CH firstly, orchestrates the data from the cluster members by collecting and aggregating it and later forwards that data to the sink [9]. The concept of joint clustering and routing rendered energy balancing to the network [21]. The clustering is done in two phases; set up and steady-state phase. In set up phase, the selection of CH is performed and 'data transmission' is done in the steady-state phase. The methods of CH selection by various state-of-the-art protocols, are discussed in Table 1. The parameters, namely residual energy, distance within the nodes or between nodes and sink, initial energy, and node density are investigated for various heterogeneous routing protocols.

The research work in HWSN started with energy heterogeneity at two levels in SEP [22]. The main concern in the heterogeneous protocol has been the advancements in Probability Based Threshold Formula (PBTF) for the cluster head selection. The subsequent contributions by the heterogeneous routing protocols namely, DEEC [23], EEHC [24], DDEEC [25], EDDEEC [26], BEENISH [27], TSEP [28] and P-SEP [29], TEDRP [19] are discussed in Table 1. The corresponding research gaps of the aforementioned protocols are also reported. Fersi et al. in [30] proposed CLEVER (Cluster-based Energy-aware Virtual Ring Routing) protocol that exploited virtual identity-based routing to perform intra and inter cluster routing. DRESEP [17] and SEECP [18] protocols used dual hop communication for the nodes to send their data to the sink. However, due to the dual hop communication, it resulted in heavy burdening of relay nodes as they are supposed to transfer data from the farthest nodes. A moment comes during data transmission when the energy of relaying nodes gets fully drained. As a result, the data is not forwarded to the sink and a hot-spot is created and this scenario is termed as a hot-spot problem. Some of the research works have followed the frequent rotation of relay nodes for their dual hop or multi-hop communication. However, it is observed that the systematic rotation of relay nodes still causes increase in overheads that consequently depletes their energy.

### 2.2 Routing strategies encountered to alleviate hot-spot problem

Multitudinous routing strategies have been proposed to alleviate energy-hole problem [31–37]. The tabular analysis of such various state-of-the-art techniques is presented in Table 2. It covers several aspects, namely energy-hole mitigation approaches, key findings and corresponding research gaps. It is observed that the pattern in which data dissemination is done from all the nodes by a sink, corresponds to many-

**Table 1** Comparative study of state-of-the-art heterogeneous routing protocols

| Study reference | Name of Protocols | Heterogeneity level | No. of sink | Reactive/ Proactive | Mode of Communication | Hot-spot Problem | No. of CH fixed | CH selection based on | | | | | | Research Gap |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Initial Energy | Residual Energy | Total Energy | Average Energy | Distance | Node Density | |
| Smaragdakis et al. (2004) [22] | SEP | 2 | 1 | Pro-active | Single-hop | No | × | ✓ | × | × | × | × | × | Not suitable for multi-level |
| Qing et al. (2006) [23] | DEEC | 2 | 1 | Pro-active | Single-hop | No | × | × | ✓ | × | ✓ | × | × | Penalization of high energy nodes |
| Kumar et al. (2009) [24] | EEHC | 3 | 1 | Pro-active | Single-hop | No | × | × | × | × | × | × | × | Energy of nodes is not considered for CH selection |
| Elbhiri et al. (2010) [25] | DDEEC | 2 | 1 | Pro-active | Single-hop | No | × | × | ✓ | × | ✓ | × | × | Not suitable for multi-level heterogeneity |
| Javaid et al. (2013) [26] | EDDEEC | 3 | 1 | Pro-active | Single-hop | No | × | × | ✓ | × | ✓ | × | × | Not suitable for multi-level heterogeneity |
| Qureshi et al. (2013) [27] | BEENISH | 4 | 1 | Reactive | Single-hop | No | × | × | ✓ | × | ✓ | × | × | Penalization of high energy nodes |
| Kashaf et al. (2012) [28] | TSEP | 3 | 1 | Reactive | Single-hop | No | × | × | × | × | × | × | × | Energy factor not included for CH selection |
| Kumar et al. (2015) [17] | DRESEP | 3 | 1 | Reactive | Dual hop | Yes | × | × | ✓ | × | × | ✓ | × | Pre-fixed circular radius in random deployment scenario |
| Kumar et al. (2016) [18] | SEECP | 3 | 1 | Reactive | Dual hop | Yes | ✓ | × | ✓ | × | ✓ | × | × | In-efficient selection of circular radius for dual hop comm. |
| Paola et al. (2017) [29] | P-SEP | 2 | 1 | Pro-active | Single-hop | No | × | × | × | × | × | × | × | Selects CH randomly, No preference to advanced node |
| Kumar et al. (2017) [19] | TEDRP | 3 | 1 | Reactive | Dual hop | Yes | × | × | ✓ | × | × | × | × | Hot-Spot Problem exists |
| Proposed Work | IDHR | 3 | 1 | Reactive | Dual Hope | Yes | × | ✓ | ✓ | × | ✓ | ✓ | ✓ | Hot spot problem prevails marginally |
| | MEEC | 3 | 4 | Reactive | Single-hop | No | × | ✓ | ✓ | × | ✓ | ✓ | ✓ | – |

**Table 2** Evolutionary advancements in energy hole mitigation techniques

| Study reference | Name of protocols/ | Methods considered for energy hole mitigation | Key Findings | Pitfalls |
|---|---|---|---|---|
| Li and Mohapatra [11] (2007) | Analytical Model | – Investigated different methods for factors, namely traffic compression, node density, hierarchical deployment, source bit rate | – Compression and aggregation technique helps in reducing the burden on the ring closest to the sink | – An insight to network lifetime and throughput is missing.<br>– Energy hole is not mitigated completely |
| Wu et al. [31] (2008) | Theoretical Model | – Used non-uniform nodes distribution strategy<br>– Proposed an energy efficient routing among different coronas | – The proposed technique saved energy of a node by 90%, | – Increased the cost of the network by adding more nodes<br>– The routing technique is not explained comprehensively |
| Liu et al. [32] (2010) | Unequal clustering mechanism | – The size of cluster is made smaller near to sink and size increases as the farthest cluster is reached.<br>– Cluster radius rotation policy is proposed. | – Network lifetime is enhanced comprehensively | – The number of overheads will increase with cluster radius rotations and with the varying size of CHs<br>– Energy hole still exists |
| Liu et al. [33] (2011) | Theoretical Analysis of energy hole | – A theoretical analysis is given to acquire the timing, size and the location of energy hole and it also estimated network lifetime theoretically. | – The first ever attempt to present such theoretical model for energy hole.<br>– Also provides the optimal network parameters for acquiring enhanced network lifetime. | – The proposed research didn't cater for the space and time elements of energy hole.<br>– The analysis of energy hole is done on very traditional protocols. |
| Liu [13] (2013) | ODDRC | – Balances the energy consumption within the gradient along with the outside the gradient sinking sensor network.<br>– Mixed routing strategy is adopted to acquire non uniform energy distribution | – 1st attempt to target the energy consumption within the same gradient (corona)<br>– Optimal data distribution ratio is computed | – The comparative performance of Network lifetime against rounds is not given. The proposed scheme is not evaluated for sufficient number of performance metrics.<br>– Energy hole is still not fully avoided |
| Ren et al. [34] (2015) | Theoretical Model | – Traffic load, energy consumption and network lifetime is estimated<br>– After analysis of aforementioned parameters, routing scheme is proposed. | – Network lifetime is estimated and boundary of energy hole is determined<br>– A theoretical analysis is done to investigate the temporal and spatial evolution of energy hole | – The sensor nodes located in the same region are considered to have similar energy which is not possible due to difference in geographical positions.<br>– Dividing the network into different smaller regions makes the network complex. |
| Gupta and Pandey [35] (2016) | Improved EUDAC | – CH selection is done on the basis of residual energy and distance to the sink and the competitive radius is determined on the basis of distance, residual energy and node density.<br>– The relay metric is used for relaying the data to the neighboring nodes | – Improved EADUC outperforms EADUC and HUCL protocol<br>– A control over clustering overhead helps in lessening the energy consumption | – CH is selected only on the basis of energy and distance.<br>– The energy hole still exists as the proposed methods is deprived of acquiring the energy balancing. |
| Naureen et al. [36] (2017) | Aggregation-based Topology learning (ATL) | – Three methods for topology data coding are proposed that are Full, Partial and Revised Partial Mapping<br>– Data payload length and number of messages that are transmitted and received by a node are considered for evaluation of ATL protocol | – Identifies energy holes by receiving and aggregating the topology data, RPM performs better than other two methods in carrying topology data<br>– ATL helps in identifying energy holes at a minimum cost of energy | – Data payload length has to be optimized for acquiring optimal network performance<br>– No consideration is reported to simplify the proposed complex algorithm<br>– The estimation of energy hole doesn't consider the temporal aspect for identifying the energy hole problem |
| Mohemed et al. [14] (2017) | OHCR and OHA | – OHCR control the overheads due to its local nature<br>– OHA renders high adaptability to the network | – Enhances stability period along with network lifetime<br>– Limits the tradeoff between network overhead and connectivity | – The algorithm is quite complex.<br>– Energy-hole avoidance is not convincing as the main focus is towards enhancing the network lifetime. |
| Akila and | | – Geo-CHs are located along with the Primary CH so as to deal with hot spot problem | – Load is effectively balanced between nodes that results in enhanced network lifetime | – Incorporation of additional nodes increases cost of the network furthermore, the declaration messages about the |

**Table 2** (continued)

| Study reference | Name of protocols/ | Methods considered for energy hole mitigation | Key Findings | Pitfalls |
|---|---|---|---|---|
| Venkatesan [37] (2018) | Zonal Clustering Algorithm (ZCA) | – It ensures the load of single cluster on CH nearest to sink. | | Geo-CHs will increase the overheads and will lead to energy consumption. |

to-one network as also analyzed by several researchers through corona-based model [38]. The energy-hole problem was firstly addressed by the Li et al. in [11] by proposing a mathematical model for energy hole mitigation pertaining to the corona-based model. Liu et al. in [39] proposed GlobalSame and RingSame algorithms in the direction of mitigating hot-spot problem in adaptive MPSK-based WSN. Abdulla et al. in [40] proposed a hybrid approach that combined two routing strategies; flat and hierarchical routing for reducing the energy consumption and exploring the data compression, respectively. Several other attempts contributed in mitigating the energy-hole as discussed in Table 2 [31–37]. However, most of them were acquainted with the complex algorithms and many of them did not really focus on the network lifetime enhancement. It is comprehended that despite of numerous schemes reported for designing the corona-based model, the hot-spot problem still sustained [32, 35].

It is worth noting that in case of single hop communication from CH to sink, there is no hot-spot problem. In order to exploit single hop communication in large area network, the only possible appropriate solution that assures energy balancing is employing multiple data sinks. It is contemplated from the above study of the heterogeneous protocols that they have employed single sink in their approach towards acquiring network longevity. In some applications, where the network area is hostile and placement of sinks is supposed to be done outside the network, the scenario of multiple data sink becomes significant [41].

## 2.3 Routing techniques based on multiple data sinks in the network

The network topology that employs multiple data sinks, play an important role to relax the constraints of limited energy resources in large area applications. Azad and Chockalingam [41] focused on the limits on network lifetime by using single and joint optimization employing multiple data sinks. It aimed to achieve the optimality of sink location to acquire energy efficient network. Q-learning based Adaptive Zone Partition (QAZP) scheme helped in avoiding the binding of sensor nodes to nearest sink to avoid their exhaustion [42]. It is comprehended that the network employing multiple data sinks, takes less time to deliver data packets to sink in comparison to the network having single sink [43]. Some of the research works also utilized the optimization approach for binding the multiple data sinks with the deployed nodes. Safa et al. [44] used Genetic Algorithm (GA) for route selection of a sensor node for data transmission to the corresponding sink. Apart from binding the nodes with multiple sinks, there had been some routing algorithms which worked for enhancing the network longevity. Energy-Optimized Routing Algorithm (EORA) was proposed for multi-sink WSNs [45]. It introduced avoid and load balancing strategy

for low energy nodes and multiple sinks, respectively. Deng et al. in [46] addressed the problem of data dissemination using multiple sinks by designing a suboptimal online algorithm that require very little knowledge in advance. Masdari and Naghiloo [47] targeted the congestion problem prevailing in the network with a single sink in operation. The authors applied the fuzzy logic-based sink selection for the nodes to mitigate congestion around any sink. The QoS and security concerns are of paramount importance in some essential applications. The Quality of Service (QoS) aware routing technique significantly improves energy saving and reduces the delay [48]. To fulfill the security related requirements, the routing among nodes can be secured by implementing various security related mechanisms [49].

After the extensive insight into the above discussed three aspects aiming to foster the network performance, some crucial concerns are discovered. Firstly, there is an absence of a heterogeneous routing technique that not only focus on the energy efficient CH selection but also alleviates the energy hole in the network. Secondly, the studies exploiting the multiple data sinks in the network highlighted their significant role in bringing an energy balance to the network. Moreover, it is observed that state-of-the-art protocols, namely DRESEP, SEECP outperformed the other reactive protocols in terms of network stability and longevity. In the subsequent sections, the aforementioned protocols have been discussed in detail to highlight the pitfalls in their routing frameworks.

## 3 System framework of DRESEP and SEECP

In this section, the system frameworks which define the set up and steady state phase of protocols namely DRESEP and SEECP, are discussed. These protocols have similar inter-cluster communication strategy with respect to each other. However, the criteria for defining the 'circle radius' inside the network is different. These protocols emphasize on the CH selection and propose different decisive strategy for undergoing inter-cluster communication. The clustering process followed in DRESEP protocol is explained below.

### 3.1 DRESEP

The two phases i.e., set up and steady state phase that are involved in the operation of DRESEP, are explained as follows.

#### 3.1.1 Set up phase of DRESEP

The CH selection in DRESEP followed the same criteria as in NEAP [50] but it added distance parameter to its threshold formula for CH selection. The values of the probabilities for each type of nodes in DRESEP are computed as follow.

$$\mathbf{P_N} = \frac{\mathbf{P}}{(\mathbf{1} + \mathbf{m}\alpha + \mathbf{m_0}\beta)} \tag{1}$$

$$\mathbf{P_{IN}} = \frac{\mathbf{P} \times (\mathbf{1} + \beta)}{(\mathbf{1} + \mathbf{m}\alpha + \mathbf{m_0}\beta)} \tag{2}$$

$$\mathbf{P_{AN}} = \frac{\mathbf{P} \times (\mathbf{1} + \alpha)}{(\mathbf{1} + \mathbf{m}\alpha + \mathbf{m_0}\beta)} \tag{3}$$

The terms or notations used in above eq. (1–3) are defined as follows. $P_N$, $P_{IN}$ and $P_{AN}$ represent the probabilities for normal, intermediate, and advanced nodes, respectively. The notations $\alpha$ and $\beta$ are the energy fractions of advanced and intermediate nodes, respectively, whereas $m$ and $m_0$ are fractions of the number of advanced and intermediate nodes, respectively. $P$ is the probability that defines the number of CHs in the network. The threshold values for normal, intermediate and advanced nodes are computed as follows.

$$\mathbf{T(N_N)} = \left\{ \frac{\mathbf{P_N}}{\mathbf{1-P_N}\left(\mathbf{r} \times \mathbf{mod}\left(\frac{\mathbf{1}}{\mathbf{P_N}}\right)\right)} \times \frac{\mathbf{D_{avg}}}{\mathbf{D_{BS}}} \times \left[\frac{\mathbf{E_{CNT}}}{\mathbf{E_{MAX}}} + \left(\mathbf{r_s div}\left(\frac{\mathbf{1}}{\mathbf{P_N}}\right)\right)\left(\mathbf{1}-\frac{\mathbf{E_{CNT}}}{\mathbf{E_{MAX}}}\right)\right] \right\} \tag{4}$$

$$\mathbf{T(I_N)} = \left\{ \frac{\mathbf{P_{IN}}}{\mathbf{1-P_{IN}}}\left(\mathbf{r} \times \mathbf{mod}\left(\frac{\mathbf{1}}{\mathbf{P_{IN}}}\right)\right) \times \frac{\mathbf{D_{avg}}}{\mathbf{D_{BS}}} \times \left[\frac{\mathbf{E_{CNT}}}{\mathbf{E_{MAX}}} + \left(\mathbf{r_s div}\left(\frac{\mathbf{1}}{\mathbf{P_{IN}}}\right)\right)\left(\mathbf{1}-\frac{\mathbf{E_{CNT}}}{\mathbf{E_{MAX}}}\right)\right] \right\} \tag{5}$$

$$\mathbf{T(A_N)} = \left\{ \frac{\mathbf{P_{AN}}}{\mathbf{1-P_{AN}}}\left(\mathbf{r} \times \mathbf{mod}\left(\frac{\mathbf{1}}{\mathbf{P_{AN}}}\right)\right) \times \frac{\mathbf{D_{avg}}}{\mathbf{D_{BS}}} \times \left[\frac{\mathbf{E_{CNT}}}{\mathbf{E_{MAX}}} + \left(\mathbf{r_s div}\left(\frac{\mathbf{1}}{\mathbf{P_{AN}}}\right)\right)\left(\mathbf{1}-\frac{\mathbf{E_{CNT}}}{\mathbf{E_{MAX}}}\right)\right] \right\} \tag{6}$$

In eq. (4–6), the terms $T(N_N)$, $T(I_N)$, and $T(A_N)$ denote the threshold values for normal, intermediate and advanced nodes, respectively. $E_{CNT}$ and $E_{MAX}$ are the current and maximum initial energy of corresponding node under threshold computation for CH selection. $D_{avg}$ is the average distance between all nodes and sink, and $D_{BS}$ is the distance between current node and sink. The term $r_s$defines the number of consecutive rounds in which a node has not been CH and *div* is used to compute the divergence of the probability of corresponding node. A random number is generated for each type of node and if the random value is greater than the threshold value generated for that type { $T(N_N)$, $T(I_N)$, and $T(A_N)$}, a node is said to be CH else it will operate as a cluster member.

**Pitfalls for using deterministic method in DRESEP** The CH selection with above discussed deterministic method has some serious shortcomings that are explained as following.

a) Initially, as soon as the network comes into operation, the chances are there that any node can elect itself as CH. In that case, it is very rare that $r_s$ will reach to the value $(\frac{1}{P_X})$ and thereby making $(r_s div \frac{1}{P_X})$ equals to 0, where $P_X$ is $P_N$, $P_{IN}$, and $P_{AN}$ in case of normal, intermediate and

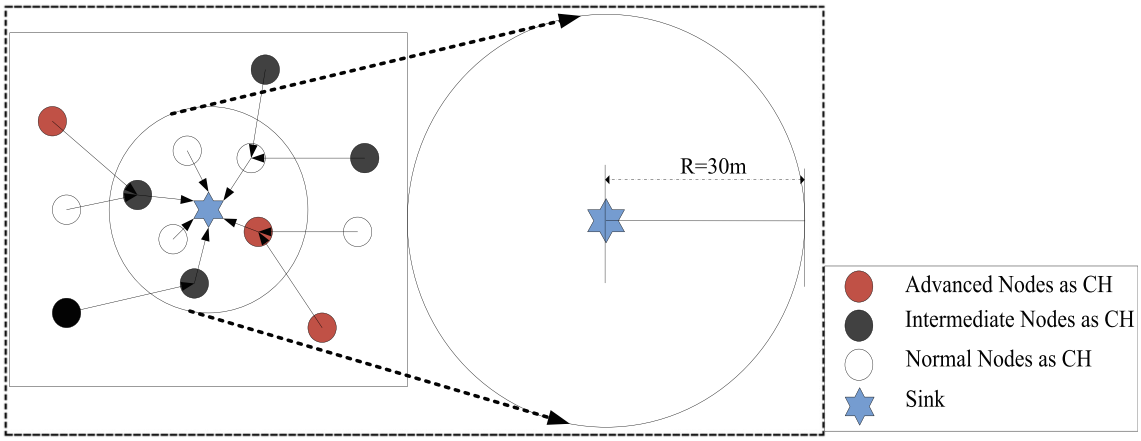



**Fig. 2** Data transmission in DRESEP

advanced nodes, respectively. Therefore, the significance of $r_s$ in the eq. (4–6) becomes very small.

b) As the data transmission proceeds, the energy ratio i.e., $\left(\frac{E_{CNT}}{E_{MAX}}\right)$ decreases. In that case, the factor '$\left(r_s div \frac{1}{P_X}\right)\left(1-\frac{E_{CNT}}{E_{MAX}}\right)$' has to be adjusted to avoid low threshold value for each type of nodes and it is also observed that this factor has significance only when $r_s$ reaches to value $(\frac{1}{P_X})$. It indicates that the node had not selected itself to be CH for the last $(\frac{1}{P_X})$ rounds. Consequently, the number of elected CHs in those rounds becomes very small. In contrast, if a node is being selected as CH in the start of rounds, then its chances of getting elected as CH gets lower irrespective of its residual energy. Consequently, it creates unbalance in the number of elected CHs.

#### 3.1.2 Steady state phase of DRESEP

The dual hop communication among CH nodes is fixed on the circular radius which is taken as 30 m as shown in Fig. 2. It is a major stumbling block for acquiring energy efficient communication among nodes undergoing inter-cluster communication. It is due to the fact that when nodes are randomly deployed in the network, the inner area of circular region with pre-fixed radius might get deployed with a smaller number of nodes vis-à-vis to the region outside the circular area. As a result, it leads to the burdening those few nodes (located inside the circular area) for relaying the data coming from the outer circular area. Hence, energy imbalance is created in the network.

### 3.2 SEECP

SEECP is an extension to DRESEP where former focuses on the enhancement of stability period and latter worked towards network longevity. The set up phase and steady state phase of SEECP are discussed as follows.

#### 3.2.1 Set up phase

In this protocol, a fixed number of CHs are selected in a deterministic manner that consider the sole factor of residual energy of nodes. This method of electing CHs is quite inefficient as the deployment of nodes is random, so ignoring distance factor will exert heavy cost on the nodes in terms of energy consumption. It is due to the reason that in case when nodes located at the farther distance from the sink are selected as CH, the energy consumption in the network is proliferated. Consequently, it will reduce the network lifetime.

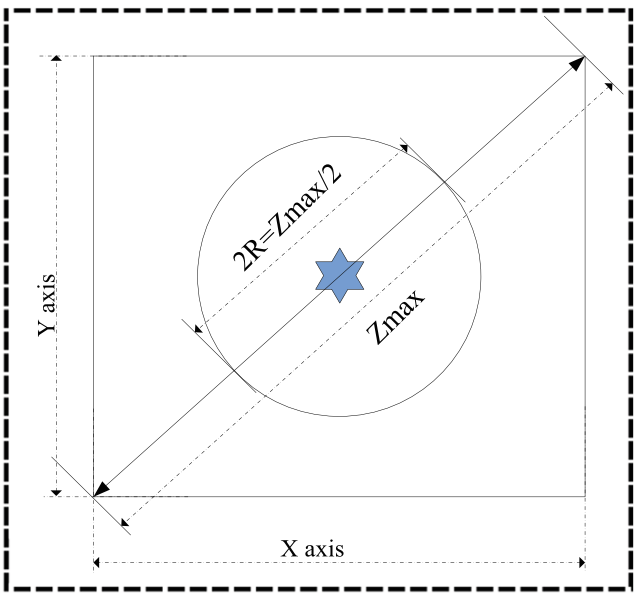



**Fig. 3** Data transmission in SEECP

#### 3.2.2 Steady-state phase

This phase follows the same procedure as of DRESEP but in SEECP, the different decisive strategy is employed to decide for a node to pursue dual hop communication. The inter-cluster communication is decided by defining a circle around the sink whose radius is computed through physical geometry in the network as shown in Fig. 3. This physical geometrical dependence of circular region makes the inter-cluster communication inefficient which is attributed to the random deployment of nodes. It not only adds the complexity to the network but also deprives the network of its scalability characteristics. This approach for defining the circular region around the sink limits the network for its exploration to numerous large area applications.

### 3.3 Novelty and significant contributions of proposed work as compared to DRESEP, SEECP and TEDRP protocols

It is encapsulated from the extensive review reported in Subsection 2.1, 2.2 and 2.3, that there is an absence of a heterogeneous routing protocol that not only focuses on CH selection but also aims to alleviate hot-spot problem. The contributions and applications of proposed work is discussed below.

A. Contributions and application of IDHR

a) The proposed algorithm IDHR incorporates the energy density factor for the CH selection. While doing so, it will foster the selection of a node as a CH which has the highest number of nodes in its vicinity. Eventually, it will help in reducing the effective distance of cluster member nodes from the CH and thereby lowering the energy consumption of nodes. However, the aforementioned competitive protocols only considered the energy and distance parameters that is insufficient to acquire better network performance.
b) Furthermore, the deterministic expression employed for the CH selection in DRESEP and SEECP had serious drawbacks, which are comprehensively discussed in Subsection 3.1.1 (A). Therefore, the probabilistic method of CH selection is employed for IDHR.
c) Besides the energy efficient CH selection in IDHR, the inter-cluster communication is also improved. Unlike the competitive protocols, it computes the average distance of the nodes from the sink for deciding about the single or dual hop communication. It is quite essential as the nodes are randomly deployed and adopting this way of inter-cluster communication helps in energy balancing in the network.

IDHR is basically developed for large area applications where sink can be placed inside the network. Basically, it is suitable for attended applications like agricultural operations.

B. Contributions and application of MEEC

a) The hot-spot problem occurs whenever the dual or multi hop communication is involved in the network. In order to extenuate this concern completely, the only solution is to introduce single hop communication in the network. However, for large area applications in which sink is placed outside the network, employing single hop communication will result in long haul communication. Consequently, the network performance will degrade. To pact with such scenarios, MEEC helps in alleviating the hot-spot problem by involving single hop communication as it considers the multiple sinks placed outside each periphery of the network.
b) It is the first ever attempt reported in the literature so far that not only improves CH selection as done by IDHR but also alleviates hot-spot problem altogether. Unlike MEEC, the protocols DRESEP, SEECP, TEDRP and IDHR are operated with single sink, but still encounters the hot-spot problem and are not suitable for unattended applications.

The one of the striking contributions of MEEC is its applicability for unattended applications for example forest fire detection, volcanic eruption, etc. As the human intervention is not possible in such hostile regions, therefore multiple sinks are placed outside the network area, hence MEEC is favorite for such applications.

## 4 Proposed system framework for IDHR and MEEC

The protocols IDHR and MEEC have some network assumptions, and these protocols follow same radio energy model which are discussed as follow. Hence, the operations of protocols IDHR and MEEC are individually discussed.

### 4.1 Network assumptions

There are following network assumptions that are adopted in proposed routing protocols operating in HWSN.

a) The sensor nodes are deployed randomly in a rectangular shaped region of area *A*. All nodes are static in nature, and each one of them owns its unique identification number.
b) The sensor nodes have energy heterogeneity up to three levels and comprise of normal, intermediate, and advanced nodes. The energy of advanced nodes is highest

among three genres and normal nodes have the least energy as compared to the others.

c) These sensor nodes are location unaware as they don't have any GPS-enable antenna on and off their circuitry.
d) Once the nodes deployment is done, battery recharging of sensor nodes is not possible due to unavailability of any recharging resource. It necessitates the energy preservation for network longevity.
e) Whenever the sensor nodes are deployed randomly in network field where the data sinks are located around each periphery, each sink broadcasts *HELLO* message which is acknowledged by each node. Further, *HELLO_neighbor* message is broadcasted by each node in its neighborhood, i.e., up to its communication range. With the help of these messages, distance from the nearest sink and neighboring nodes is computed.
f) The sensor nodes can fuse their data after receiving from the neighbor node and finally one fused data packet is generated.
g) In case of MEEC protocol, four data sinks are located systematically around each periphery of the network. There is the persistent and adequate power supply which releases all energy constraints on these data sinks. Moreover, these are identical to each other, and they are enriched with memory, energy, and computation.
h) The sensor nodes consume their energy by following a fundamental radio energy model (discussed below), and these sensor nodes have symmetric radio link in between them.
i) The wireless communication among nodes is considered to be a secured one. The security issues related to proposed protocol are beyond the scope of this paper.

## 4.2 Radio energy consumption model

In this paper, the fundamental radio energy consumption model [17] is used for protocols IDHR and MEEC. It decides for the amount of depletion of energy for all the nodes that communicate in the network. The nodes are involved in processes namely, data transmission, reception and aggregation that devours their energies. Furthermore, the energy is also consumed while switching on the circuitry of these sensor nodes. The protocols IDHR and MEEC are discussed as follows.

## 4.3 IDHR

In this section, the set up and steady state phase for the proposed IDHR protocol are discussed. The essential contributions reported in IDHR are the energy efficient 'CH selection (in set-up phase) and the inter-cluster communication (in steady state phase)'. The set up phase is discussed as follows.

### 4.3.1 Set up phase

As soon as the random deployment of nodes is done in the targeted area, the nodes get connected to each other with the exchange of messages. The structure of the network employing heterogeneous nodes is discussed as follow.

A. Network Model

The total number of normal, advanced and intermediate nodes are given by a set of eqs. (7–9). In these eqs., the proportions of a number of advanced and intermediate nodes are represented by $m$ and $\boldsymbol{m_0}$ fractions. The total number of advanced, intermediate, and normal nodes are represented by $\boldsymbol{N_{ADV}}$, $\boldsymbol{N_{INT}}$ and $\boldsymbol{N_{NRM}}$, respectively.

$$\mathbf{N_{ADV}} = \mathbf{n} \times \mathbf{m} \tag{7}$$

$$\mathbf{N_{INT}} = \mathbf{n} \times \mathbf{m_0} \tag{8}$$

$$\mathbf{N_{NRM}} = \mathbf{n} \times (\mathbf{1}-\mathbf{m}-\mathbf{m_0}) \tag{9}$$

The advanced and intermediate nodes are $\alpha$ and $\beta$ times higher in energy as compared to normal nodes, respectively as shown in the eq. (10–15). $\boldsymbol{E_{ADV}}$, $\boldsymbol{E_{INT}}$, and $\boldsymbol{E_{NRM}}$ represent the energy of all advanced, intermediate, and normal nodes, respectively. The total energy of the network is given by $\boldsymbol{E_T}$ which is devised in eq. (15). The notations $\alpha$ and $\beta$ are the energy fractions of advanced and intermediate nodes, respectively. The energy of normal node is given by $\boldsymbol{E_O}$. The computation of energy of the network is done as follows.

$$\mathbf{E_{ADV}} = \mathbf{E_O} \times (\mathbf{1} + \boldsymbol{\alpha}) \times \mathbf{n} \times \mathbf{m} \tag{10}$$

$$\mathbf{E_{INT}} = \mathbf{E_O} \times (\mathbf{1} + \boldsymbol{\beta}) \times \mathbf{n} \times \mathbf{m_0} \tag{11}$$

$$\mathbf{E_{NRM}} = \mathbf{E_O} \times (\mathbf{1}-\mathbf{m}-\mathbf{m_0}) \times \mathbf{n} \tag{12}$$

$$\mathbf{E_T} = \mathbf{E_{ADV}} + \mathbf{E_{INT}} + \mathbf{E_{NRM}} \tag{13}$$

$$\mathbf{E_T} = \mathbf{E_O} \times (\mathbf{1} + \boldsymbol{\alpha}) \times \mathbf{n} \times \mathbf{m} + \mathbf{E_O} \times (\mathbf{1} + \boldsymbol{\beta}) \times \mathbf{n} \times \mathbf{m_0} + \mathbf{E_O} \times (\mathbf{1}-\mathbf{m}-\mathbf{m_0}) \times \mathbf{n} \tag{14}$$

$$\mathbf{E_T} = \mathbf{n} \times \mathbf{E_O} \times (\mathbf{1} + \boldsymbol{\beta} \times \mathbf{m_0} + \mathbf{m} \times \boldsymbol{\alpha}) \tag{15}$$

These equations give the comprehensive description of the heterogeneous structure of proposed protocols. The CH selection for IDHR is discussed as follows.

B. CH selection in IDHR

When the network is established, it is the cluster-based routing that owes the responsibility of data dissemination from the network. As CH performs data collection, an appropriate selection of CH is of paramount importance for any network in operation. IDHR follows the probability-based threshold formula for CH selection unlike the deterministic method in SEECP protocol. The CH selection in IDHR is made on the following parameters:

a) The ratio of residual energy ($E_{res}$) to the average energy i.e., $E_{avg}$ (computed by eq. (20)) of nodes, is given by factor i.e., $\left(\frac{E_{res}}{E_{avg}}\right)$ for a particular node in the network.

b) The ratio of distance between '$i^{th}$ node (where $i$ ranges from 1 to total number of nodes) and sink' i.e., $D(i)$ and an average distance between 'all nodes and sink' denoted by $D_{avg}$, is given by the factor i.e., $\left(\frac{D(i)}{D_{avg}}\right)$. This factor is incorporated to foster the selection of CH within the circular region around the sink.

c) The node density factor i.e. $N_{DNSTY}$ is used to reduce the effective distance of CH from its respective cluster member nodes.

d) An energy threshold factor i.e., $E_{THS}$ is employed to avoid the condition of frequent selection of high energy nodes as CHs. This condition is termed as 'penalization of high energy nodes'.

After incorporating aforementioned factors, the probability $P(i)$ for each type of node to become CH is constructed by using following eq. (16).

$$\mathrm{P(i)} = \begin{cases} \dfrac{\mathrm{P_{opt}} \times \mathrm{E_{res}} \times \mathrm{D(i)} \times \mathrm{N_{DNSTY}}}{\left((1 + \mathrm{m} \times \alpha + \mathrm{m_0} \times \beta) \times \mathrm{D_{avg}} \times \mathrm{E_{avg}}\right)} & \text{for normal nodes if } \mathrm{E_{res}} \geq \mathrm{E_{THS}} \\ \dfrac{\mathrm{P_{opt}} \times \mathrm{E_{res}} \times (1 + \mathrm{a}) \times \mathrm{D(i)} \times \mathrm{N_{DNSTY}}}{\left((1 + \mathrm{m} \times \alpha + \mathrm{m_0} \times \beta) \times \mathrm{D_{avg}} \times \mathrm{E_{avg}}\right)} & \text{for advanced nodes if } \mathrm{E_{res}} \geq \mathrm{E_{THS}} \\ \dfrac{\mathrm{P_{opt}} \times \mathrm{E_{res}} \times (1 + \beta) \times \mathrm{D(i)} \times \mathrm{N_{DNSTY}}}{\left((1 + \mathrm{m} \times \alpha + \mathrm{m_0} \times \beta) \times \mathrm{D_{avg}} \times \mathrm{E_{avg}}\right)} & \text{for intermediate nodes if } \mathrm{E_{res}} \geq \mathrm{E_{THS}} \\ \mathrm{k} \times \dfrac{\mathrm{P_{opt}} \times \mathrm{E_{res}} \times (1 + \beta) \times \mathrm{D(i)} \times \mathrm{N_{DNSTY}}}{\left((1 + \mathrm{m} \times \alpha + \mathrm{m_0} \times \beta) \times \mathrm{D_{avg}} \times \mathrm{E_{avg}}\right)} & \text{for all nodes if } \mathrm{E_{res}} < \mathrm{E_{THS}} \end{cases} \tag{16}$$

The notations used in above eq. (16) are defined as follows. $P_{opt}$ is the probability to define the number of CHs in the network. The term $k$ is the optimal constant value which comes into operation when residual energy of nodes becomes less than the threshold energy. The node density factor used in the probability expression is computed in eq. (17).

$$\mathbf{N_{DNSTY}(i)} = \mathbf{1} - \frac{\left(\frac{\mathbf{1}}{\mathbf{Nc}} \sum_{\mathbf{i=1}}^{\mathbf{Nc}} \mathbf{D_{i-j}}\right)}{\mathbf{100}} \tag{17}$$

$D_{i-j}$ is the distance of each $i^{th}$ node from the $j^{th}$ node, where $i$ and $j$ are equal to total nodes in the network. $D_{i-j}$ is computed as follows in eq. (18).

$$\mathbf{D_{i-j}} = \sqrt{(\mathbf{node(i).x} - \mathbf{node(j).x})^2 + (\mathbf{node(i).y} - \mathbf{node(j).y})^2} \tag{18}$$

The terms $node(i).x$ and $node(i).y$ are the cartesian coordinate for $i^{th}$ node. Also, $node(j).x$ and $node(j).y$ are the cartesian coordinates for $j^{th}$ node. The distance between nodes and sink i.e., $D(i)$ is computed as follows in eq. (19) where $sink.x$ and $sink.y$ are the cartesian coordinates for $j^{th}$ node.

$$\mathbf{D(i)} = \sqrt{(\mathbf{node(i).x} - \mathbf{sink.x})^2 + (\mathbf{node(i).y} - \mathbf{sink.y})^2} \tag{19}$$

$E_{avg}$ is the average energy of all the nodes in the network, computed by the eq. (20) where $E_T$ is the total energy of whole network (computed by eq. (15)) and $n$ being the total number of nodes in the network.

$$\mathbf{E_{avg}} = \frac{\mathbf{1}}{\mathbf{n}}(\mathbf{E_T}) \tag{20}$$

The value of $k$ ranges from 0 to 1, and its optimal value is computed after acquiring successful simulations with maximum stability period for the network. The threshold formula for the IDHR follows the eq. (21).

The probability value $P(i)$ computed in the eq. (21) is used further to compute the threshold $T(i)$ for each $i^{th}$ node given as follows.

$$\mathbf{T(i)} = \begin{cases} \dfrac{\mathbf{P(i)}}{\left(\mathbf{1} - \mathbf{P(i)}\left(\mathbf{r} \bmod \left(\dfrac{\mathbf{1}}{\mathbf{P(i)}}\right)\right)\right.} & \text{if } \mathbf{S(i)} \in \mathbf{G} \\ \mathbf{0} \end{cases} \tag{21}$$

In above eq. (21), $r$ is the current round, $S(i)$ is the current $i^{th}$ node (where $i$ ranges from 1 to $n$) under threshold computation and $G$ is the group of nodes which are yet to become CH. It is noted that $G$ is the set of $G$' and $G$" which comprise of group of intermediate and advanced nodes, respectively that are yet to become CH. The random number i.e., $R_N$ is generated for each type of node and the corresponding 'adaptive random number' denoted by $ARN_o$ is computed as given below in eq. (22).

$$\mathbf{ARN_o} = \frac{\mathbf{N} - \mathbf{D_N}}{\mathbf{N}} \times \mathbf{R_N} \tag{22}$$

In above eq. (22), $N$ and $\boldsymbol{D_N}$ denote the total nodes and the number of dead nodes, respectively whilst the current round $\boldsymbol{R_N}$ is in progress. The reason for incorporating $\boldsymbol{ARN_o}$ is to make it sure that there are uniform number of CHs during network operation therefore, the dependence on $\boldsymbol{R_N}$ has to be put under control. As the number of dead nodes increases in the network consequently, the number of CHs also decreases. Therefore, in IDHR, the computation of $\boldsymbol{R_N}$ is modified and it is made adaptive to $\boldsymbol{D_N}$ as given by the eq. (22). It is noted that $\boldsymbol{D_N}$ increases with the passage of rounds due to continuous energy expenditure. As a result, $\boldsymbol{ARN_o}$ is also decreased which ultimately favors the CH selection in the later stages of data transmission. Furthermore, $\boldsymbol{ARN_o}$ is compared with $\boldsymbol{T(i)}$ for each node. If $\boldsymbol{ARN_o}$ for $i^{th}$ node is less than the threshold value $\boldsymbol{T(i)}$ (computed in eq. (21)), then that node is selected as CH otherwise the node will stay as cluster member node.

#### 4.3.2 Steady state phase

This phase pacts with the data transmission in the network when the set up phase is already established. It covers the intra-cluster and inter-cluster communication which are explained below.

**Algorithm 1 IDHR Algorithm**
**Input:** *n, $r_{max}$, $ARN_o$, single_sink=(50,50);*
**Output:** **A=**{CH_S, dead nodes, alive nodes};

```
Random deployment of nodes and placement of sink (see Section 2.2)
Calculate E_T by using eq. (7-15);
CH_S=0;
for r=1 to r_max do
    alive=n;
    dead=0;
    for i=1:n do
        if E(i)≤ 0 then
            dead=dead+1;
            if (dead==n) then
                all_dead=r;
            end if
            alive=alive-dead;
        end if
    end for
    for i=1:n do
        if E_(i) > 0 then
            Calculate D_i−j by using eq. (17);
            Calculate D_(i) by using eq. (18);
            Calculate N_DNSTY(i) by using eq. (19);
            Calculate E_avg by using eq. (20);
            Calculate P(i) by using eq. (16);
            Calculate T(i) by using eq. (21);
            Calculate ARN_o by using eq. (22);
            if T(i)> ARN_o then
              CH_S=CH_S+1;
              data transmission follows as explained in Section 2;
            else
              i^th node ←TDMA slot for data tx to CH;
              Update E_(i) by using radio energy model;
            end if
        end if
    end for
    if (dead==n)
    break;
    end if
end for
return A
```

a) Intra-cluster communication: The selected CHs collect data from the nodes by assigning a TDMA slot to each cluster member node, after that it aggregates the data and send it to the sink.
b) Inter-cluster communication: The inter-cluster communication is decided upon the circular region that is defined around the sink as shown in Fig. 4. In IDHR, the radius of the circle is the average distance of all the nodes from the sink unlike the DRESEP and SEEP. It favors more CH selection within the region and thereby reducing the effective communicating distance of CHs from the sink. However, if the distance of any CH from the sink is more than the $D_{avg}$ then it passes its data to the nearest CH that lies inside the circular region. Otherwise, CH sends data directly to the sink. As the nodes in the network are randomly scattered, this approach helps in the reducing the energy consumption by minimizing the distance between the CHs and sink.

#### 4.3.3 Explicative remarks for algorithm 1

The Algorithm 1 highlights the steps taken into consideration for IDHR protocol. The input to the algorithm is acquainted with $n$ (i.e., the total number of nodes deployed randomly), $r_{max}$ (i.e., the maximum number of rounds for the network operation), $ARN_o$ (Adaptive Random Number) and sink placement at the middle of the network. The algorithm aims to render energy efficient CH selection and increased number of alive nodes to the network. The status of dead nodes over the passage of rounds is also obtained. Line 1 states about the random deployment of heterogeneous nodes and placement of sink in the middle of the network. Line 2 calculates the value of total energy i.e., $E_T$ by using eq. (7–15) which is then incorporated in the construction of probabilistic formula for CH selection. Line 3 initiates the *CH_S* variable which is to store number of CHs framed in the network. Line 4 starts the loop that run for each round ranging from 1 to $r_{max}$ and it covers different steps of operations defined within the Line 4 and Line 37. Line 5 and Line 6 initiates the 'alive' and 'dead' status of nodes just before the energy consumption which gets updated after every round. Line 7-Line 15 define a *for* loop that checks for the dead nodes status with respect to their current energy. Line 16-Line 33 runs the *for* loop to execute CH selection task for each node with a condition mentioned in Line 17 for each node with $\mathbf{E}_{(i)} > 0$. Line 25 checks for the threshold value and if it is greater than $ARN_o$ then CH is selected and *CH_S* is updated. Otherwise, node is declared as cluster member node and it transmits data to the CH according to the TDMA slot. Line 34-Line 36 covers a termination condition to check if all nodes are dead at any particular round. Line 38 returns *A* which is a set of *CH_S*, dead nodes and alive nodes.

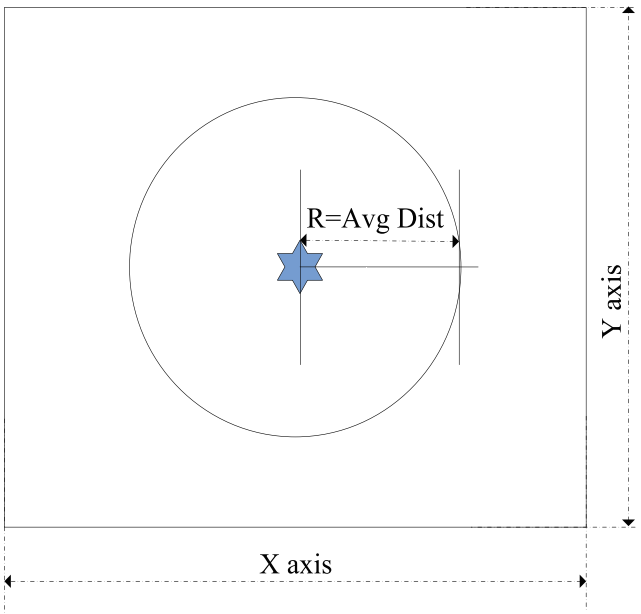



**Fig. 4** Data transmission in IDHR

### 4.4 MEEC

MEEC is designed for a network having multiple data sinks in the network. Therefore, along with the CH selection in the set up phase, the most prominent contribution by the MEEC protocol is the mitigation of hot-spot problem in the steady state phase. The clustering process for the MEEC is discussed as follows.

#### 4.4.1 Set up phase

MEEC considers the process similar to IDHR for set up phase for its network formation. The deployment of heterogeneous nodes is random but uniform, however, the placement of sink differs to that of IDHR. Four data sinks are placed systematically in the middle of each periphery outside the network. They are placed at 10-m distance from the boundary of each side of the square shaped region. The rest of the network structure stays same as that of IDHR.

A. CH selection

The selection of CH is made in the same way as adopted in IDHR. The probabilistic formula for CH selection is quite similar to that of IDHR. The proposed probability $P(i)$ for each type of node to become CH, is constructed by using following eq. (23).

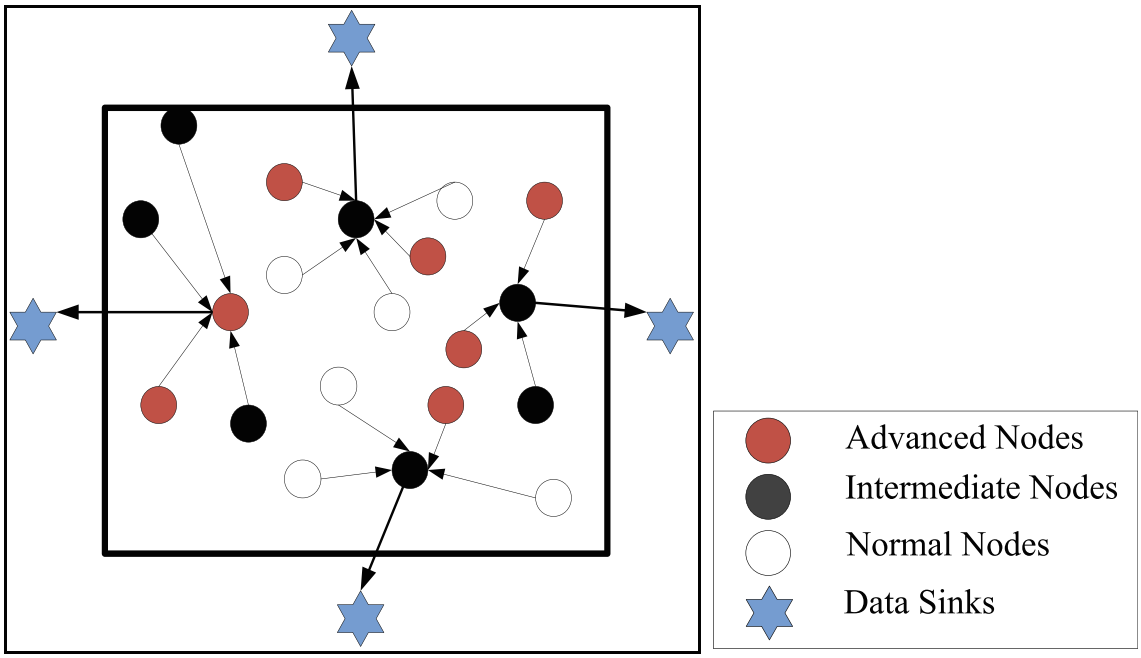



**Fig. 5** Data transmission scenario in MEEC

$$P(i) = \begin{cases} \dfrac{P_{opt} \times E_{res} \times \mathbf{D_N(i)} \times N_{DNSTY}}{\left((1 + m \times \alpha + m_0 \times \beta) \times \mathbf{D_{avg_NS}} \times E_{avg}\right)} & \text{for normal nodes if } E_{res} \geq E_{THS} \\ \dfrac{P_{opt} \times E_{res} \times (1 + a) \times \mathbf{D_N(i)} \times N_{DNSTY}}{\left((1 + m \times \alpha + m_0 \times \beta) \times \mathbf{D_{avg_NS}} \times E_{avg}\right)} & \text{for advanced nodes if } E_{res} \geq E_{THS} \\ \dfrac{P_{opt} \times E_{res} \times (1 + \beta) \times \mathbf{D_N(i)} \times N_{DNSTY}}{\left((1 + m \times \alpha + m_0 \times \beta) \times \mathbf{D_{avg_NS}} \times E_{avg}\right)} & \text{for intermediate nodes if } E_{res} \geq E_{THS} \\ k \times \dfrac{P_{opt} \times E_{res} \times (1 + \beta) \times \mathbf{D_N(i)} \times N_{DNSTY}}{\left((1 + m \times \alpha + m_0 \times \beta) \times D_{\mathbf{avg_NS}} \times E_{avg}\right)} & \text{for all nodes if } E_{res} < E_{THS} \end{cases} \quad (23)$$

In above eq. (23), the terms $\boldsymbol{D_N(i)}$ and $\boldsymbol{D_{avg_NS}}$ are the distance factors that are incorporated in MEEC unlike to $\boldsymbol{D(i)}$ and $\boldsymbol{D_{avg}}$ which were included in IDHR. $\boldsymbol{D_N(i)}$ represents the distance of $i^{th}$ node ($i$ ranges from 1 to total number of nodes i.e.,100) from *sink(p)* located at the minimum distance. In eq. (24), the terms $\boldsymbol{sink(p).x}$ and $\boldsymbol{sink(p).y}$ denote the Cartesian coordinates of $\boldsymbol{sink(p)}$ where $p = \{1,2,3,4\}$ which corresponds to the four sinks employed outside the network. $\boldsymbol{D_{avg_NS}}$ denotes the average distance of all nodes (lying in the vicinity of nearest sink) from the nearest CH in a cluster as given in eq. (25). The number of nodes in a particular cluster is defined by $\boldsymbol{C_N}$. The terms $\boldsymbol{node_{CH}.x}$ and $\boldsymbol{node_{CH}.y}$ denote the Cartesian coordinates of CH node.

$$\mathbf{D_N(i)} = \mathbf{argmin_{sink}(p)} \sqrt{(\mathbf{node(i).x} - \mathbf{sink(p).x})^2 + (\mathbf{node(i).y} - \mathbf{sink(p).y})^2} \quad (24)$$

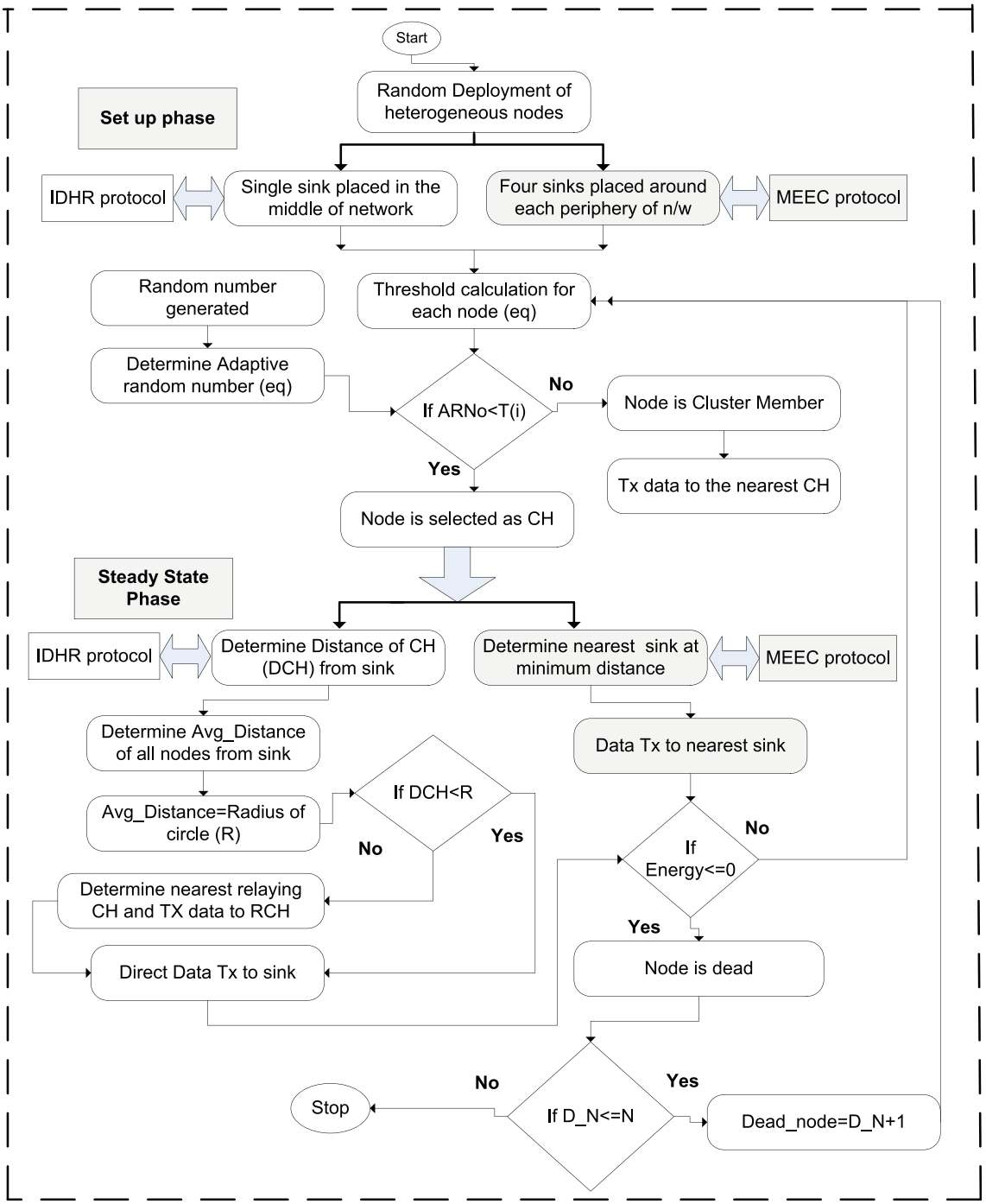



**Fig. 6** Flowchart for IDHR and MEEC protocols

$$\mathbf{D}_{\mathbf{avg}_{\mathbf{NS}}} = \frac{\mathbf{1}}{\mathbf{C}_{\mathbf{N}}} \sum \left( \sqrt{(\mathbf{node}_{\mathbf{CH}}.\mathbf{x} - \mathbf{node}(\mathbf{i}).\mathbf{x})^2 + (\mathbf{node}_{\mathbf{CH}}.\mathbf{y} - \mathbf{node}(\mathbf{i}).\mathbf{y})^2} \right) \tag{25}$$

The rest of the parameters in eq. (23) have same definitions as explained for eq. (16) in IDHR. After probability computation for IDHR, the similar steps of $\boldsymbol{ARN_o}$ generation and threshold computation are performed for MEEC. Ultimately, the node with high threshold value than $\boldsymbol{ARN_o}$, is declared as CH and rest of the nodes remain cluster member nodes.

#### 4.4.2 Steady state phase

The steady state phase of MEEC is significantly better than IDHR in the context of avoiding hot-spot problem. The intra-cluster and inter-cluster communication are discussed as follow.

**Intra-cluster communication** It follows the methodology for intra-cluster communication similar to that of IDHR. The cluster members send their data to CH in their defined TDMA slots.

**Inter-cluster communication** The one of the most promising solution that extenuate the hot-spot problem completely, is finding a way to implement routing without any relaying nodes. As a result, the effective distance of communication among nodes and sink will be decreased. It is worth noting that there is no inter-cluster communication in MEEC, due to the fact that the communication between the sink and the CH is made a single hop. To implement the aforementioned solution, four sinks are employed outside the periphery of the network as shown in Fig. 5. CH sends their data to the nearest sink on a single hop based on RSSI (Received Signal Strength Indicator). It is done to abate the energy expenditure evolved due to the dual hop communication among CHs.

**Algorithm 2 MEEC Algorithm**
**Input:** *n, rmax, A.RN, sink1=(50, 110); sink2=(50, -10); sink3=(110,50); sink4=(-10, 50);*
**Output: A=** {CH_S, dead nodes, alive nodes};

```
Random deployment of nodes and placement of multiple data sinks (see Section 2.2)
Calculate E_T by using eq. (7-15);
CH=0;
for r=1 to r_max do
    alive=n;
    dead=0;
    for i=1:n do
      if E(i)≤ 0 then
      dead=dead+1;
        if (dead==n) then
          all_dead=r;
        end if
        alive=alive-dead;
      end if
    end for
    for i=1:n do
      if E_(i) ≥ 0 then
        Calculate D_N(i) by using eq. (24);
        Calculate D_avg_NS by using eq. (25);
        Calculate N_DNSTY(i) by using eq. (19);
        Calculate E_avg by using eq. (20);
        Calculate P(i) by using eq. (23);
        Calculate T(i) by using eq. (21);
        Calculate ARN_o by using eq. (22);
        if T(i)> ARN_o then
          CH_S=CH_S+1;
          data transmission follows as explained in Section 2;
        else
          i^th node ←TDMA slot for data tx to CH;
          Update E_(i) by using radio energy model;
        end if
      end if
    end for
    if (dead==n)
      break;
    end if
end for
return A
```

#### 4.4.3 Explicative remarks for algorithm 2

The Algorithm 2 works in the quite similar direction as that of Algorithm 1. However, it differs in some significant aspects which are highlighted in this discussion about Algorithm 2. The first crucial difference is the input parameters considered for Algorithm 2 that incorporates the employment of four data collecting sinks located at 10 m distance from each periphery of the network. The rest of the parameters stay same as that of Algorithm 1. The steps of the algorithm 2 are as follows. Line 1 to Line 15 are same as discussed in Algorithm 1. Line 16 and Line 17 starts the *for* loop for each node and *if* loop that checks for energy of each node whether it is greater than zero, respectively. Line 18-Line 33 runs the for loop to execute CH selection task for each node. It must be noted that Algorithm 2 differs from Algorithm 1 in terms of parameters used for calculation of *P(i)* computed in eq. (23). Line 18 and Line 19 calculates $\boldsymbol{D_N(i)}$ and $\boldsymbol{D_{avg_NS}}$, respectively which are distance factors that are incorporated in MEEC unlike $\boldsymbol{D(i)}$ and $\boldsymbol{D_{avg}}$ which were included in IDHR. $\boldsymbol{D_N(i)}$ represents the distance of $i^{th}$ node ($i$ ranges from 1 to total number of nodes i.e.,100) from nearest sink. $\boldsymbol{D_{avg_NS}}$ denotes the average distance of all nodes (lying in the vicinity of nearest sink) from the nearest CH a cluster as given in eq. (25). The rest of the steps of the Algorithm 2 are same as that of Algorithm 1.

e) After the CH selection, the network enters in the steady state phase. For IDHR, the distance of CH from the sink denoted by *DCH* is computed. Then the average distance of all nodes from the sink is computed. The radius of circular region with sink at the center point, is given by the value of average distance of all nodes from the sink.
f) Further it is checked, if *DCH* is less than *R* (circular radius) then data is directly transmitted to the sink otherwise, the data is forwarded to nearest relaying CH located inside the circular region.
g) In MEEC, the steady state phase differs to that of IDHR due to employment of multiple data sinks. When the CH is selected it discovers the nearest sink at the minimum distance.
h) The energy of each node is checked for both protocols (IDHR and MEEC) that operates in steady state phase. If the energy of a node is less than or equal to the zero, a node is said to be dead. Furthermore, if the number of dead nodes *(D_N)* become equal to the total nodes in the network then the network stops functioning. Otherwise, the counter to the number of dead nodes is incremented by 1 and similar process is repeated for the threshold computations.

### 4.5 Proposed methodology for IDHR and MEEC

The detailed description of the methodology adopted for IDHR and MEEC protocols demonstrated in Fig. 6, is explained as follows.

a) The whole network operates under two phases; set up and steady state phase. In the set up phase, the network operation starts with the random, but uniform deployment of heterogeneous nodes in the network area of pre-defined dimension.
b) IDHR employs single sink that is located in in the middle of the network but in MEEC, four sinks are deployed systematically around each periphery of the network and placed at a regular distance to each other.
c) The CH selection stays quite same for both the protocols as explained in Subsections 4.3.1 and 4.4.1 The eq. (21–22) are used to compute probabilities. In MEEC, the distance factor is modified due to the multiple data sinks (as discussed in Subsection 4.4).
d) The threshold calculation is performed for each node for both protocols. After generating random number, it is modified by incorporating dead nodes (given by eq. (22)). So, *ARNo* is generated for both protocols to ensure CH selection in each round. If *ARNo* is less than the threshold value, the node is said to be CH else it's a cluster member node which transmit data to the nearest CH.

### 4.6 Computational complexity analysis of algorithm 1 (IDHR) and algorithm 2 (MEEC)

The complexity analysis decides the applicability potential of any algorithm. The computational complexity of both algorithms i.e., IDHR and MEEC is dependent on some significant factors, namely $\boldsymbol{r_{max}}$ (maximum number of rounds iterated) and $n$ (total number of nodes in the network). The following lemma is devised to have a deep insight into the complexity analysis of these algorithms.

> Lemma 1 (:). IDHR and MEEC terminates in fixed iterations $\boldsymbol{r_{max} = O(1)}$ and has computational complexity of $\boldsymbol{O(r_{max} \times n)}$

As soon as the data transmission starts among the nodes, the nodes consume their energy in the subsequent rounds. The iterations keep on occurring until the total number of nodes are dead in the network. It is comprehended that the $\boldsymbol{r_{max}}$ iterations are inversely proportional to the number of dead nodes. As the total number of nodes are fixed in number and all nodes experience energy consumption therefore, at a particular value of $\boldsymbol{r_{max}}$, all nodes become dead. Hence, IDHR and MEEC terminate eventually.

The overall computational complexity of IDHR and MEEC is $\boldsymbol{O(r_{max} \times n)}$ due to the following reasons.

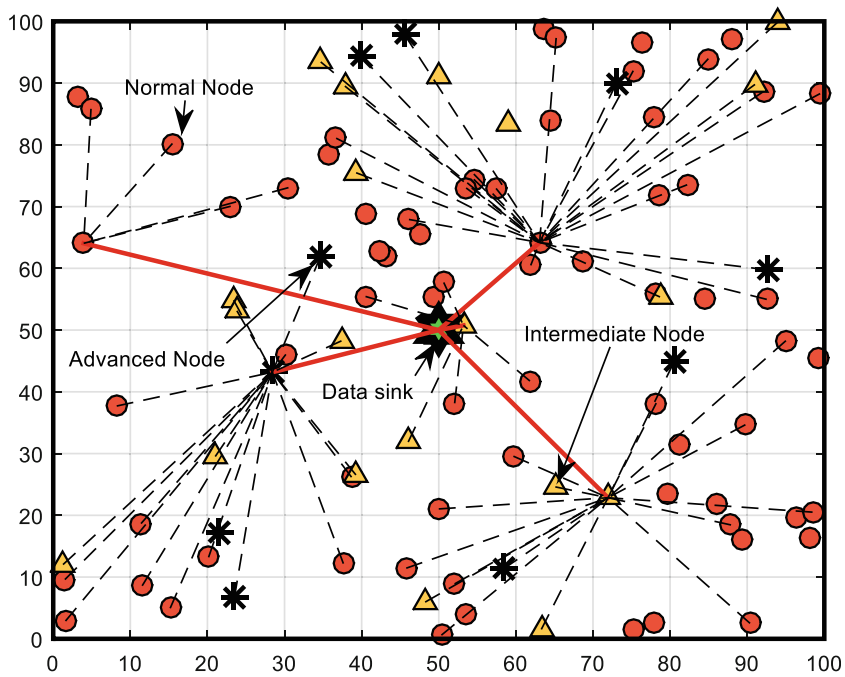
**Table 3** Simulation parameters

| Network Model and GA Parameters | Values |
|---|---|
| Network Area Size | $100 \times 100$ m$^2$ |
| Number of Nodes ($N$) | 100 |
| Number of data sinks for IDHR and MEEC | 1 and 4 |
| Initial energy of nodes (in Joules) ($E_o$) | 1 |
| Energy heterogeneity Node Type | 3-level; normal, intermediate and advanced nodes |
| Energy fraction of intermediate nodes ($\beta$) and advanced nodes ($\alpha$) | $\beta = 1$, $\alpha = 2$ |
| Number of intermediate nodes ($m$) and advanced nodes fraction ($m_o$) | $m = 0.2$, $m_o = 0.1$ |
| Energy required for running transmitter and receiver $\mathbf{E_{elc}}$ | 50nJ/bit |
| Threshold distance ($d_o$) | 87 m |
| Amplification energy required for smaller distance $d \leq d_o$ ($E_{efs}$) | 10pJ/bit/m$^2$ |
| Amplification energy required for larger distance $d > d_o$ ($E_{mp}$) | 0.0013pJ/bit/m$^4$ |
| Energy consumption incurred while data aggregation ($E_{da}$) | 5nJ/bit/signal |
| Data packet size | 2000bits |
| Number of Simulation run | 20 |
| Confidence interval | 95% |

a) Initially, the conditions to check the number of dead nodes is processed for each value of n for every value of $r_{max}$. Therefore, the computational complexity for this computation is $O(r_{max} \times n)$.
b) Thereafter, the process of CH selection is performed for $n$ values for every value of $r_{max}$. The eq. (16–22) in case of IDHR and eq. (17–28) in case of MEEC are computed for CH selection. However, the computational complexity contributed by aforementioned equations is negligible. Therefore, the computational complexity at this stage of CH selection is obtained as $O(r_{max} \times n)$.

By summarizing the complexities computed in points a) and b), the total computational complexity becomes **2** ×

**Fig. 7** Data transmission scenario of IDHR



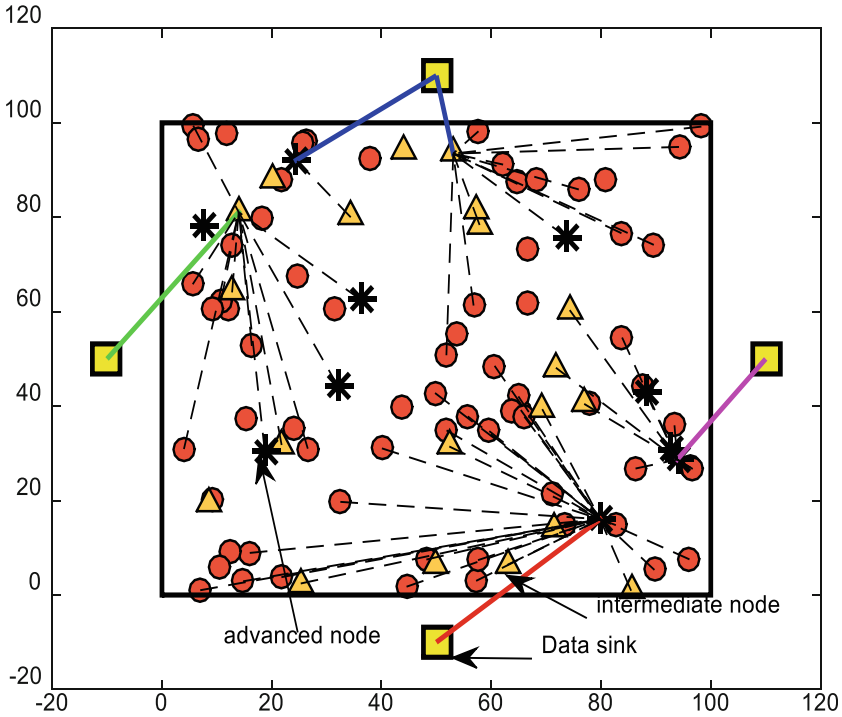



**Fig. 8** Data transmission scenario of MEEC

$O(r_{max} \times n)$, which is significantly equivalent to $O(r_{max} \times n)$. It is observed that the state-of-the-art protocols taken into consideration for performance evaluation of IDHR experiences similar computational complexity. It is due to the reason that the algorithm IDHR focuses on improving the CH selection and inter-cluster communication among the CHs. In case of MEEC, the algorithm focuses on finding the nearest sink for a CH node, but still added computational complexity is negligible. Consequently, the whole structure of iterations remains same as that of other competitive protocols. Therefore, the proposed algorithms are optimal in the context of computational complexity.

## 5 Results and discussions

The performance evaluation of IDHR is done against the state-of-the-art protocols TEDRP, SEECP and DRESEP, all working in the energy heterogeneous environment of WSN. These protocols aim for network longevity and enhanced stability

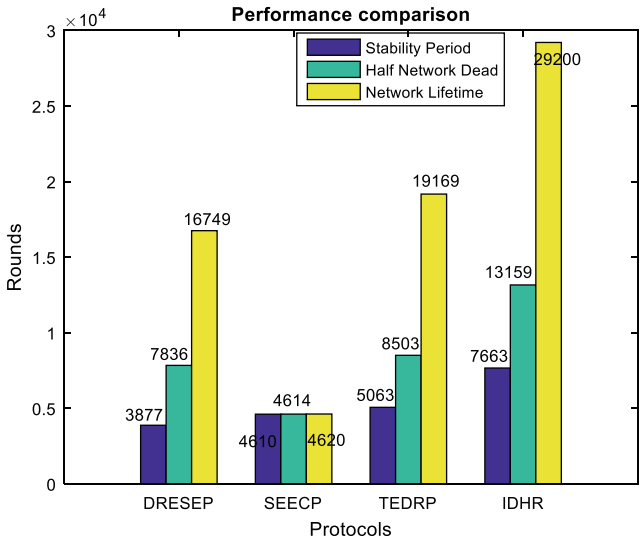



**Fig. 9** Stability Period, HND and Network lifetime comparison of IDHR with other protocols

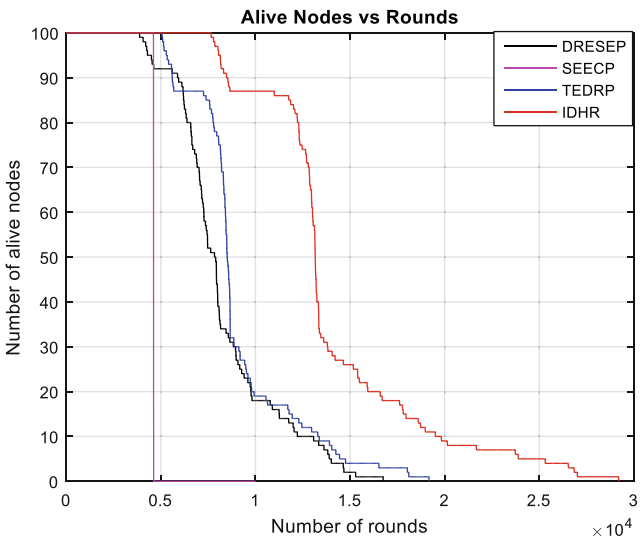



**Fig. 10** Comparison of alive nodes vs rounds of IDHR with other protocols

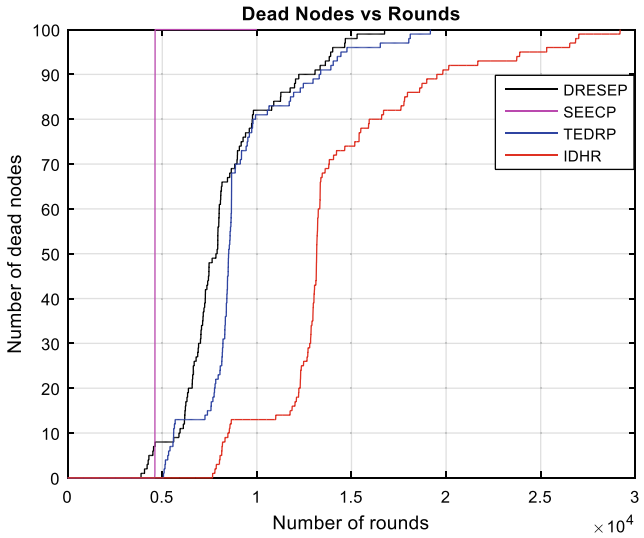



**Fig. 11** Comparison of dead nodes vs rounds of IDHR with other protocols

period for the given simulation environment. The essence of employing multiple data sinks is highlighted by the performance comparison of MEEC against the IDHR protocol. The performance of these protocols is evaluated on the benchmark of performance metrics that are defined as follow.

## 5.1 Performance metrics

There are some essential performance metrics which are utilized to investigate the performances of IDHR and MEEC against the aforementioned protocols. These metrics are discussed as follow.

a) Stability Period

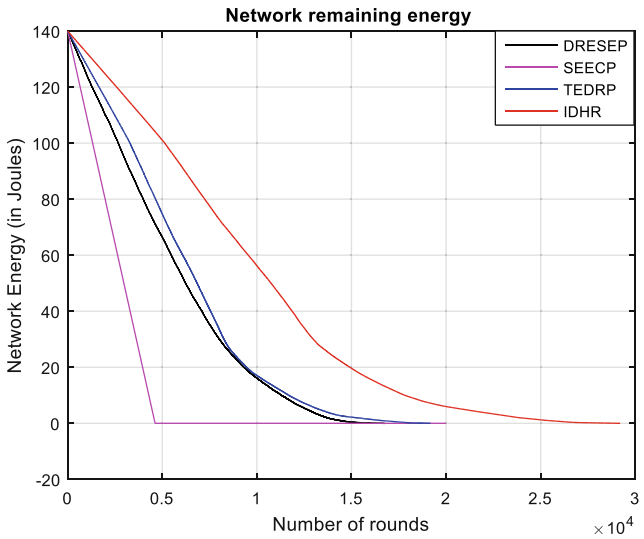



**Fig. 12** Comparison of Network's remaining energy of IDHR with other protocols

The reliable data dissemination is assured by this performance metric. It is defined as the number of rounds covered till the first node of any type i.e., normal, intermediate and advanced nodes, runs out of the energy completely. This factor becomes prominent when the loss of even little bit of information can cause heavy cost to some applications. The higher stability period assures the higher reliability of proposed routing protocols operating for any network.

b) Network Longevity

Network longevity can be defined as the number of rounds covered till all nodes are exhausted of energy while pursuing data transmission. When the network is for attended applications, the network longevity plays an important role for continuous data delivery till all nodes are dead.

c) Number of dead nodes against rounds

The network performance is also evaluated on this factor as it gives the status of a number of dead nodes with the passage of rounds. When a more number of rounds are covered against the number of alive nodes, the network performance is said to be upgraded.

d) Throughput

The number of data packets successfully transmitted to sink per unit time is termed as throughput. This is one of the recurring parameters for fostering the QoS (Quality of Service) so as to assure the reliability of the network. It is noteworthy to mention that the sole consideration of network longevity doesn't help in acquiring the best of network performance. Therefore, upgradation in QoS uplifts the network performance and enhances the credibility of proposed routing strategy.

e) Network's remaining energy

As the data transmission proceeds, the total energy of nodes i.e., the network's energy is consumed gradually while nodes communicate with the other nodes or with the sink. This metric helps in exposing the status of total energy of nodes after each round. When the routing strategy is an efficient one, the network preserves its energy and covers more number of rounds thereby resulting in acquiring network longevity.

## 5.2 Simulation settings

This section presents the simulation environment utilized for simulating the IDHR and MEEC protocols. The simulation software MATLAB version 2016 is installed on Windows 10 operating system. The system has further configuration as follows, Intel Core i3–540 processor operating at 3.07GHz

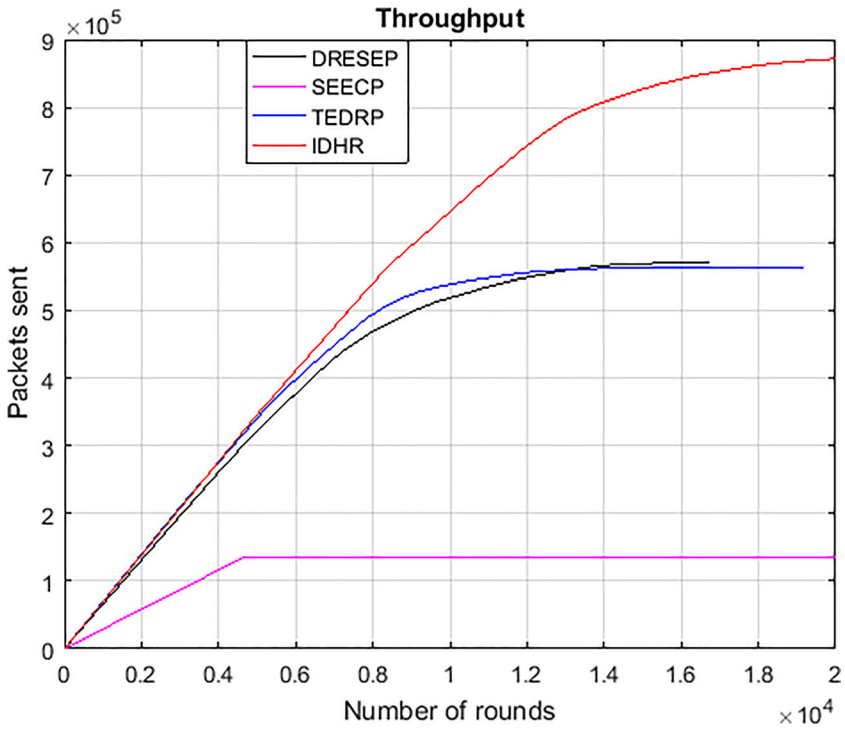



**Fig. 13** Comparison of throughput of IDHR with other protocols

and RAM used is of size 4 GB. The network area of 100 × 100 $m^2$ is simulated and 100 energy heterogeneous nodes are randomly deployed in it. The sink is placed in the middle of the network in the case of IDHR whereas four sinks are placed outside the network in case of MEEC. These four sinks are placed at (50, −10), (−110, 50), (50, 110) and (−10, 50) coordinates of the network area. The simulation parameters considered for IDHR and MEEC are listed in Table 3. The initial energy of normal nodes is 1 J and intermediate and advanced nodes are two and three fold to that of normal nodes, respectively. The number of intermediate and advanced nodes are in 20% and 10% of the total nodes in the network. The radio energy parameters are kept same as that being used in other protocols taken for comparative analysis. The Confidence Interval (CI) for the simulation run is 95%. The data transmission scenario for IDHR and MEEC are illustrated in Fig. 7 and

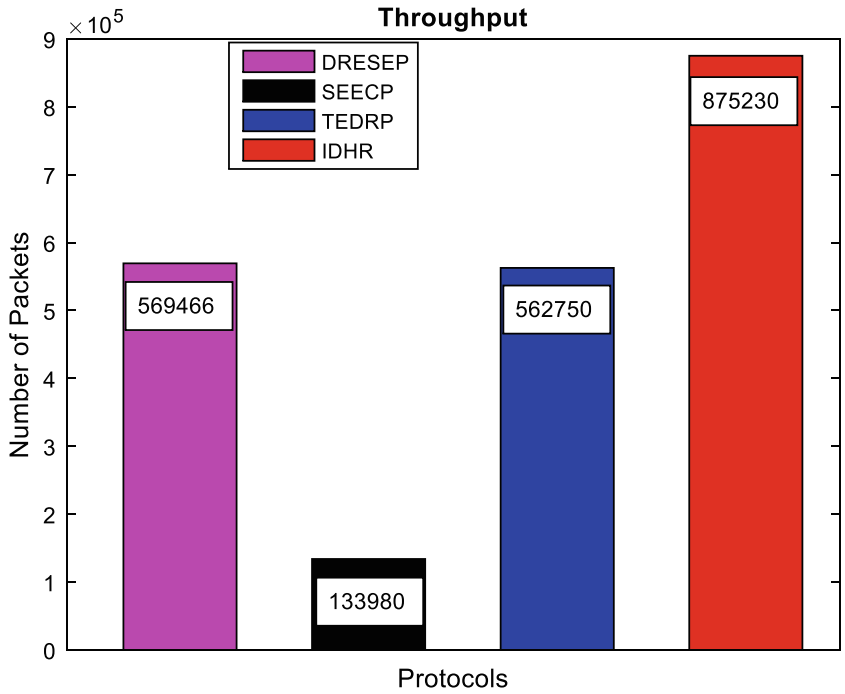



**Fig. 14** Comparison of number of data packets transmitted by each protocol

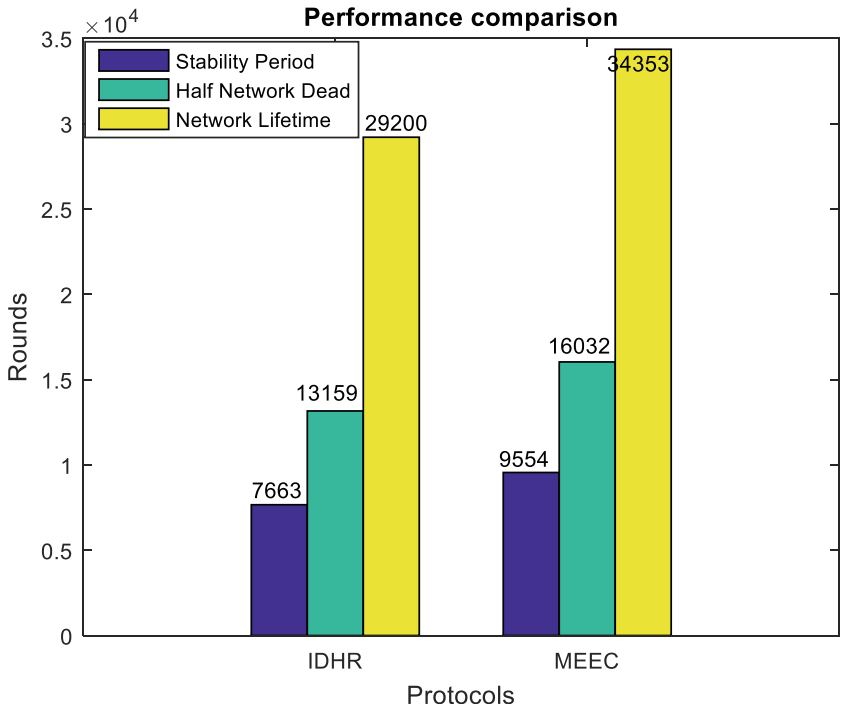



**Fig. 15** Stability Period, HND and Network lifetime comparison of MEEC with IDHR

Fig. 8, respectively. The value of $\alpha$ and $\beta$ are kept fixed and same for both IDHR and MEEC protocols.

## 5.3 Simulation analysis

The simulation analysis of IDHR and MEEC, is done on the grounds of performance metrics discussed in Section 5.1. The working process of IDHR is illustrated in the flowchart shown in Fig. 6. The simulation results of IDHR and MEEC are comprehensively discussed below.

### 5.3.1 Simulation results of IDHR protocol

The comparative analysis of the performance metrics for IDHR against the other protocols, is discussed as follows.

a) Stability Period

In IDHR, the first node depletes its energy after 7663 rounds but 5063, 4610 and 3877 rounds in the case of TEDRP, SEECP and DRESEP, respectively as illustrated in Fig. 9. It is evident that IDHR contributes to 51.35%, 66.22%

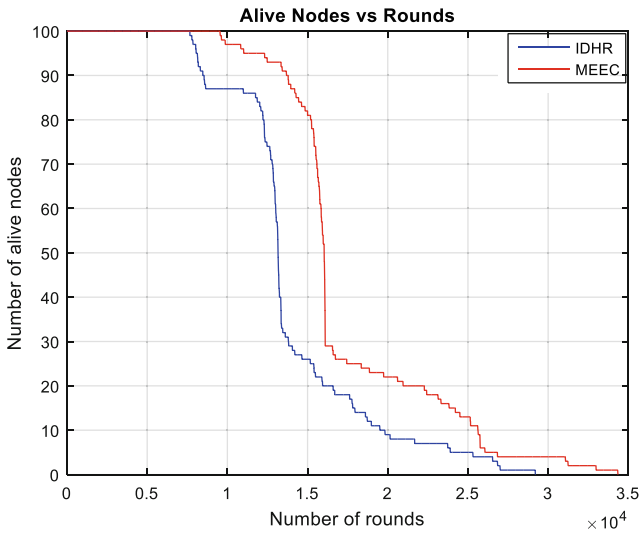



**Fig. 16** Comparison of alive nodes vs rounds of MEEC with IDHR

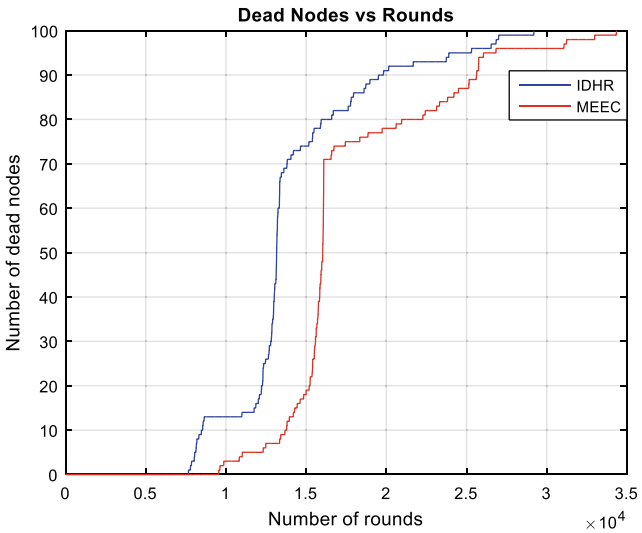



**Fig. 17** Comparison of dead nodes vs rounds of MEEC with IDHR

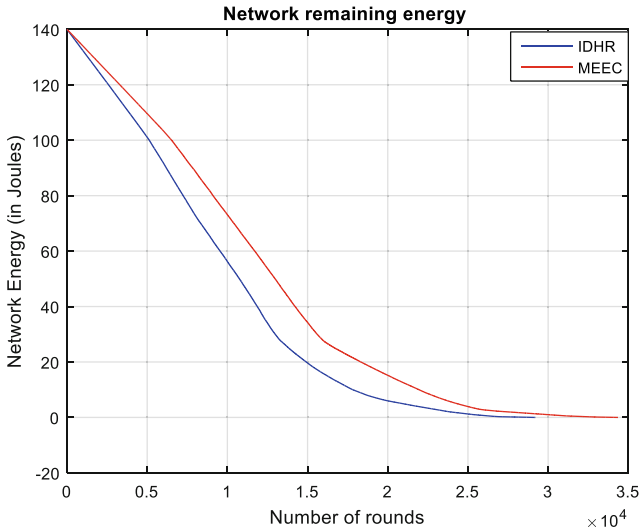



**Fig. 18** Comparison of Network's remaining energy of MEEC with IDHR

and 97.65% gigantic amelioration in stability period in comparison to the protocols TEDRP, SEECP and DRESEP, respectively.

The prominent reason behind such amelioration in stability period is the energy efficient CH selection that not only includes the distance and energy factors but also the node density factor. It makes it possible for every cluster to select a node as a CH which is surrounded by the maximum number of nodes along with the consideration of energy and distance factors. The reduced distance between nodes and its CH, helps in the energy preservation of normal nodes which normally tend to die first due to their minimum initial energy. Unlike the protocols TEDRP, SEECP and DRESEP that are taken into consideration for comparative analysis, IDHR assures the energy efficient CH selection that enhances stability period.

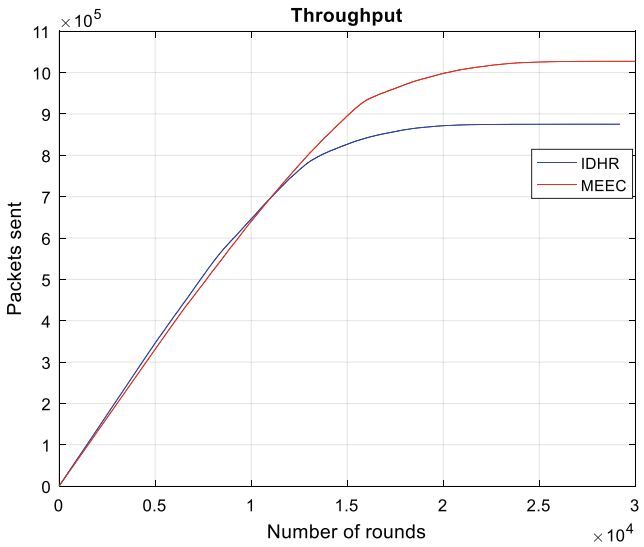



**Fig. 19** Comparison of throughput of MEEC with IDHR

b) Network Lifetime

It is observed from the Fig. 9 and Fig. 10 that IDHR covers 29,200 rounds whereas TEDRP, SEECP and DRESEP cover 19,169, 4620, and 16,749 rounds, respectively prior to energy exhaustion of all the nodes. IDHR accounts to 10,031 more rounds as compared to the TEDRP protocol, hence contributing in 52.31% enhancement in the network lifetime. Moreover, IDHR ameliorates network lifetime by 532% and 74.33% in comparison to SEECP and DRESEP protocols, respectively.

Such amelioration is reported not only due to the proposed CH selection but also by considering the average distance of all nodes from the sink as a radius of the circle. It helps in favoring the CH selection inside the circle so as to reduce the average distance of CHs from the sink. Furthermore, the node density factor reduces the communication cost for the sensor nodes within the cluster. It is due to the reason that node density factor favors the CH selection of a node which is surrounded by more neighboring nodes. Consequently, overall network energy is preserved leading to improved network lifetime.

c) Number of dead nodes against the rounds

The performance comparison of IDHR is illustrated in Fig. 11, which investigates the different percentage of dead nodes of the network with respect to the number of rounds. It is evident from Fig. 11 that First Node Dead (FND) for IDHR is 7663 rounds, which is 5063, 4610 and 3877 rounds in case of TEDRP, SEECP and DRESEP, respectively. Half Nodes Dead (HND) for IDHR is 13,159 rounds whereas, it is just 8503, 4614 and 7836 rounds in case of TEDRP, SEECP and DRESEP protocols, respectively. The improvement in Last node dead (LND) is also observed in IDHR as it covers 29,200 rounds whereas TEDRP, SEECP and DRESEP cover 19,169, 4620, and 16,749 rounds, respectively. It can be contemplated from the above analysis that IDHR covers more number of rounds at different stages of dead nodes. It is attributed to the reduced energy expenditure due to the energy efficient CH selection.

d) Network's remaining energy:

The rate of network's energy expenditure is measured by this metric. When the data transmission proceeds, the network's energy reduces. As shown in Fig. 12, IDHR performs better as compared to TEDRP, SEECP and DRESEP protocols, respectively in the context of covering a greater number of rounds while data transmission is in progress. IDHR covers maximum number of rounds but SEECP covers least number of rounds during the network operation. It is due to the reason that energy efficient selected CH consumes a very little stock of energy among all nodes. Moreover, the intra cluster

**Table 4** Comparative analysis of IDHR and MEEC with other protocols for different metrics

**Value of advanced fractions and quantity fractions of node** $m=0.1, m_o=0.2, \beta=1, \alpha=2$

| Protocols | No. of data sinks | Total Energy of Network (Joules) | Stability Period (rounds) | Half Node Dead (rounds) | Network Lifetime (rounds) | Throughput (packets) |
|---|---|---|---|---|---|---|
| DRESEP | 1 | 140 | 3877 | 7836 | 16,749 | 569,466 |
| SEECP | 1 | 140 | 4610 | 4614 | 4620 | 133,980 |
| TEDRP | 1 | 140 | 5063 | 8503 | 19,169 | 562,750 |
| **IDHR** | 1 | 140 | 7663 | 13,159 | 29,200 | 875,230 |
| **MEEC** | 4 | 140 | 9554 | 16,032 | 34,353 | 1,027,384 |

communication also preserves energy of nodes in the most efficient way. Therefore, when the energy of whole network is preserved, the nodes cover more number of rounds with the available stock of the remaining energy.

e) Throughput/Number of data packets sent to sink

In case of IDHR, as illustrated in Fig. 13 and Fig. 14, the throughput is enhanced comprehensively as it successfully transmits 875,230 data packets whereas TEDRP, SEECP and DRESEP transmit 562,750, 133,980 and 569,466 data packets, respectively. It is evident from the throughput comparative analysis, IDHR improves throughput by 55.52%, 553.25% and 53.69% as compared to TEDRP, SEECP and DRESEP protocols, respectively. Such improvement is reported due to energy efficient intra-cluster and inter-cluster communication. Nodes transmit data packets for a longer duration due to the network longevity acquired by the proposed protocol. The following subsection covers the simulation analysis of MEEC protocol.

### 5.3.2 Simulation analysis of MEEC protocol

The simulation settings to design MEEC is kept same as that of used in IDHR. The only difference in the designing MEEC, is the employment of four sinks outside the network. Performance investigation of MEEC is carried out against the other protocols by considering the performance metrics discussed in subsection 5.1.

a) Stability Period

It is illustrated from Fig. 15 that MEEC has a stability period of 9554 rounds whereas IDHR, TEDRP, SEECP and DRESEP have a stability period of 7663, 5063, 4610, and 3877 rounds, respectively. It is observed that MEEC contributes to 88.7%, 107.25% and 146.43% gigantic amelioration in stability period in comparison to the protocols TEDRP, SEECP and DRESEP, respectively. In case of MEEC, a greater number of rounds are covered with respect to the different number of alive nodes during the network run.

Moreover, as illustrated in Fig. 15, MEEC improves stability period of IDHR by 24.68% which is accounted to the employment of multiple data sinks due to which the geometric distance between the sink and CH is reduced comprehensively. The comparison of the number of alive nodes over the rounds is also done for MEEC against IDHR as shown in Fig. 16. MEEC covers more number of rounds for different number of alive nodes left in the network.

b) Network lifetime

It is reported that network lifetime of MEEC is improved by 17.65%, 79.21%, 643.57%, and 105.1% as compared to the IDHR, TEDRP, SEECP and DRESEP protocols, respectively. The total number of rounds covered by MEEC is 34,353 till all nodes are completely exhausted in their energy resources. Whereas, there are 29,200, 19,169, 4620 and 16,749 rounds in case of IDHR, TEDRP, SEECP and DRESEP protocols, respectively as shown in Fig. 15. The enhancement in the network lifetime is attributed to the fact that dual hop communication is transformed into single hop communication that leads to the energy saving of relaying nodes. The overall distance between node and sink is shortened. Furthermore, the comparative analysis of network lifetime of MEEC is done with IDHR. It is evident from the Fig. 16 that MEEC covers 5153 more rounds against the 29,200 rounds of IDHR. The reason behind amelioration in the network lifetime is the reduction in the energy consumption due to closely located data sinks. Due to the single hop and reduced distance communication, hot-spot problem is mitigated that helps in the energy balancing in the network thereby enhancing network longevity.

c) Number of dead nodes against rounds

The performance comparison of MEEC is illustrated in Fig. 17 in which MEEC covers a greater number of

**Table 5** Comparative percentage improvement by IDHR and MEEC

| Percentage (%) Improvement by IDHR and MEEC protocols | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Protocols | Stability Period | | Half Node Dead | | Network Lifetime | | Throughput | |
| | IDHR | MEEC | IDHR | MEEC | IDHR | MEEC | IDHR | MEEC |
| IDHR | – | 24.67 | – | 21.83 | – | 17.64 | – | 17.38 |
| DRESEP | 97.65 | 146.43 | 67.93 | 104.59 | 74.33 | 105.1 | 53.69 | 80.41 |
| SEECP | 66.22 | 107.25 | 185.19 | 247.46 | 532 | 643.57 | 553.25 | 666.82 |
| TEDRP | 51.35 | 88.7 | 54.75 | 88.55 | 52.32 | 79.21 | 55.52 | 82.56 |

rounds at different stages of dead nodes. It is evident from the Fig. 17 that First Node Dead (FND) for MEEC is 9554 rounds which is 7663, 5063, 4610 and 3877 rounds in case of IDHR, TEDRP, SEECP and DRESEP protocols, respectively. Half Nodes Dead (HND) for MEEC is 16,032 rounds whereas, it is just 13,159, 8503, 4614 and 7836 rounds in case of IDHR, TEDRP, SEECP and DRESEP protocols, respectively. HND for MEEC is improved by 21.83%, 88.55%, 247.46%, and 104.59% as compared to IDHR, TEDRP, SEECP and DRESEP protocols, respectively. The improvement in Last node dead (LND) is also observed in MEEC as it covers 34,353 rounds whereas IDHR, TEDRP, SEECP and DRESEP cover 29,200, 19,169, 4620, and 16,749 rounds, respectively. It can be contemplated from the Fig. 17 and above analysis that MEEC covers a greater number of rounds at different stages of dead nodes as compared to IDHR. It is due to the reduced energy expenditure as a result of energy preservation strategy applied in the CH selection as well as in the inter-cluster communication. It is found that MEEC improves the network lifetime by the significant improvement of 17.65%. It is again due to the reduced energy expenditure incurred due to employment of multiple data sinks outside the network.

d) Network's remaining energy

As the network operates in MEEC, it covers more number of rounds as discussed above. As illustrated in Fig. 18, the network's remaining energy of MEEC covers a greater number of rounds as compared to IDHR. It is due to the employment of multiple data sinks which helps in abating the network's remaining energy at a very low rate as compared to the other protocols. Furthermore, the residual energy is saved by the much larger amount due to the reduction in the number of communication hops. Consequently, it leads to the reduction in the effective communication distance from the CH to the respective sink.

e) Throughput/Number of data packets sent to sink

The total number of packets transmitted to the respective sink in case of MEEC is 1,027,384, whereas it is 875,230, 562,750, 133,980 and 569,466 data packets in case of IDHR, TEDRP, SEECP and DRESEP protocols, respectively. The improvement in throughput is obtained due to the network longevity achieved by MEEC which helps in dissemination and hence transmission of more number of packets in the network. In addition to this, MEEC also outperforms IDHR in throughput metric as illustrated in Fig. 19. It is observed that MEEC transmits 152,154 data packets more than that sent by IDHR i.e. improving throughput by 17.38%. It is due to the reduction in the average communicating distance of CH nodes from the sink that helps in fostering the data delivery at the sink.

### 5.4 Summarized analysis of IDHR and MEEC protocols

It can be encapsulated from the results obtained by simulation analysis of IDHR and MEEC that by incorporating multiple data sinks in the network; it not only enhances the reliability but also stability period in the network. The comparative analysis of proposed protocol is reported in Table 4. The energy fractions, $\alpha$ and $\beta$ are kept fixed and they have a huge role to play in deciding about the initial energy stock of the nodes. It is discerned from [51] that with the increase in the value of the energy fractions, the total energy of the network is enhanced. Consequently, the network performance is upgraded. To justify the comparison among the competitive protocols, the simulation platform is kept same for all of them as mentioned in Table 4. Furthermore, the Table 5 demonstrates the percentage improvement by IDHR and MEEC to the protocols (DRESEP, TEDRP and TEDRP) and IDHR, respectively. It is noted that the performance improvement for the MEEC is observed due to the proposed CH selection not solely due to the multiple data sinks.

Load balancing is also achieved with MEEC as compared to the other protocols in the single sink as well as in multiple data sinks scenarios. Hence, making it favorite for some apposite time critical unattended applications.

## 6 Conclusion

To pact with the limited battery resources of sensor nodes, two routing protocols, namely IDHR and MEEC are proposed that employ single and multiple data sinks, respectively to perform the data dissemination. The CH selection is incorporated with node density factor along with energy and distance to consume energy of nodes in well orchestral manner. IDHR employs dual hop communication by defining circular radius inside the network which is based on the average distance between the nodes and sink. Whereas, predefined fixed value of radius is incorporated in DRESEP and SEECP protocols. MEEC uses single hop communication among CHs and data sinks, consequently, removes hot-spot problem from the network. It is evident from the simulation analysis of IDHR that stability period is ameliorated by 51.35%, 66.22% and 97.65% as compared to TEDRP, SEECP and DRESEP protocols, respectively. In addition to this, IDHR outperformed aforementioned protocols for different performance metrics, namely network lifetime and throughput. MEEC reports the significant improvement vis-à-vis TEDRP, SEECP and DRESEP in stability period, network longevity and throughput metrics by 24.67%, 17.64% and 17.39%, respectively. In IDHR, such performance enhancement is achieved due to the proposed CH selection and energy efficient inter-cluster communication. Whereas, in MEEC, it is due to the reduced communication distance by using multiple data sinks. The gigantic performance of these protocols makes them favorite for various large area applications. In future, the placement of multiple data sinks can be optimized before their actual placement outside the network which will further improve network performance.

## References


1. Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E (2002) A survey on sensor networks. IEEE Commun Mag 40:102–114
2. Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E (2002) Wireless sensor networks: a survey. Comput Netw 38:393–422
3. Ortmann S, Langendoerfer P, Brzozowski M, Piotrowski K (2014) Wireless sensor networks: a key enabling Technology for Remote Healthcare. In: Systems Design for Remote Healthcare. Springer: 201–227
4. Piotrowski K, Sojka A, Langendoerfer P (2010) Wireless sensor networks can save lives-benefits and open issues. Sensoren Messsyst 2010:24
5. Ammari HM, Gomes N, Grosky WI, et al (2012) Review of applications of wireless sensor networks. Wirel Sens Netw Curr Status Future Trends 1
6. Abbasi AA, Younis M (2007) A survey on clustering algorithms for wireless sensor networks. Comput Commun 30:2826–2841
7. Soro S, Heinzelman WB (2009) Cluster head election techniques for coverage preservation in wireless sensor networks. Ad Hoc Netw 7:955–972
8. Yarvis M, Kushalnagar N, Singh H, et al (2005) Exploiting heterogeneity in sensor networks. Proceedings of the 24th annual joint conference of the IEEE computer and communications societies, IEEE: 878–890
9. Tanwar S, Kumar N, Rodrigues JJ (2015) A systematic review on heterogeneous routing protocols for wireless sensor network. J Netw Comput Appl 53:39–56
10. Tao Y, Zhang Y, Ji Y (2013) Flow-balanced routing for multi-hop clustered wireless sensor networks. Ad Hoc Netw 11:541–554
11. Li J, Mohapatra P (2007) Analytical modeling and mitigation techniques for the energy hole problem in sensor networks. Pervasive Mob Comput 3:233–254
12. Mottola L, Picco GP (2011) MUSTER: adaptive energy-aware multisink routing in wireless sensor networks. IEEE Trans Mob Comput 10:1694–1709
13. Liu T (2013) Avoiding energy holes to maximize network lifetime in gradient sinking sensor networks. Wirel Pers Commun 70:581–600
14. Mohemed RE, Saleh AI, Abdelrazzak M, Samra AS (2017) Energy-efficient routing protocols for solving energy hole problem in wireless sensor networks. Comput Netw 114:51–66
15. Healy M, Newe T, Lewis E (2008) Wireless sensor node hardware: a review. Sensors, IEEE: 621–624
16. Vieira MAM, Coelho CN, Da Silva DC, da Mata JM (2003) Survey on wireless sensor network devices. Proceedings of Conference on Emerging Technologies and Factory Automation, IEEE: 537–544
17. Mittal N, Singh U (2015) Distance-based residual energy-efficient stable election protocol for WSNs. Arab J Sci Eng 40:1637–1646
18. Mittal N, Singh U, Sohi BS (2017) A stable energy efficient clustering protocol for wireless sensor networks. Wirel Netw 23:1809–1821
19. Mittal N, Singh U, Sohi BS (2017) A novel energy efficient stable clustering approach for wireless sensor networks. Wirel Pers Commun 95:2947–2971
20. Liu X (2012) A survey on clustering routing protocols in wireless sensor networks. Sensors 12:11113–11153
21. Xu Z, Chen L, Chen C, Guan X (2016) Joint clustering and routing design for reliable and efficient data collection in large-scale wireless sensor networks. IEEE Internet Things J 3:520–532
22. Smaragdakis G, Matta I, Bestavros A (2004) SEP: a stable election protocol for clustered heterogeneous wireless sensor networks. Boston University Computer Science Department
23. Qing L, Zhu Q, Wang M (2006) Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks. Comput Commun 29:2230–2237
24. Kumar D, Aseri TC, Patel RB (2009) EEHC: energy efficient heterogeneous clustered scheme for wireless sensor networks. Comput Commun 32:662–667
25. Elbhiri B, Saadane R, Aboutajdine D, others (2010) Developed distributed energy-efficient clustering (DDEEC) for heterogeneous wireless sensor networks. Proceedings of 5th international symposium on I/V communications and Mobile network (ISVC), IEEE: 1–4
26. Javaid N, Qureshi TN, Khan AH et al (2013) EDDEEC: enhanced developed distributed energy-efficient clustering for heterogeneous wireless sensor networks. Procedia Comput Sci 19:914–919
27. Qureshi TN, Javaid N, Khan AH et al (2013) BEENISH: balanced energy efficient network integrated super heterogeneous protocol for wireless sensor networks. Procedia Comput Sci 19:920–925
28. Kashaf A, Javaid N, Khan ZA, Khan IA (2012) TSEP: threshold-sensitive stable election protocol for WSNs. Proceedings of 10th International Conference on Frontiers of Information Technology (FIT), IEEE: 164–168
29. Naranjo PGV, Shojafar M, Mostafaei H et al (2017) P-SEP: a prolong stable election routing algorithm for energy-limited heterogeneous fog-supported wireless sensor networks. J Supercomput 73: 733–755
30. Fersi G, Louati W, Jemaa MB (2016) CLEVER: cluster-based energy-aware virtual ring routing in randomly deployed wireless sensor networks. Peer–Peer Netw Appl 9:640–655
31. Wu X, Chen G, Das SK (2008) Avoiding energy holes in wireless sensor networks with nonuniform node distribution. IEEE Trans Parallel Distrib Syst 19:710–720

32. Liu A-F, Wu X-Y, Chen Z-G, Gui W-H (2010) Research on the energy hole problem based on unequal cluster-radius for wireless sensor networks. Comput Commun 33:302–321
33. Liu A-F, Zhang P-H, Chen Z-G (2011) Theoretical analysis of the lifetime and energy hole in cluster based wireless sensor networks. J Parallel Distrib Comput 71:1327–1355
34. Ren J, Zhang Y, Zhang K et al (2016) Lifetime and energy hole evolution analysis in data-gathering wireless sensor networks. IEEE Trans Ind Inform 12:788–800
35. Gupta V, Pandey R (2016) An improved energy aware distributed unequal clustering protocol for heterogeneous wireless sensor networks. Eng Sci Technol Int J 19:1050–1058
36. Naureen A, Zhang N, Furber S (2017) Identifying energy holes in randomly deployed hierarchical wireless sensor networks. IEEE Access 5:21395–21418
37. Akila IS, Venkatesan R (2018) An energy balanced geo-cluster head set based multi-hop routing for wireless sensor networks. Clust Comput 1–10
38. Asharioun H, Asadollahi H, Wan T-C, Gharaei N (2015) A survey on analytical modeling and mitigation techniques for the energy hole problem in corona-based wireless sensor network. Wirel Pers Commun 81:161–187
39. Liu A, Zhang D, Zhang P et al (2014) On mitigating hotspots to maximize network lifetime in multi-hop wireless sensor network with guaranteed transport delay and reliability. Peer–Peer Netw Appl 7:255–273
40. Abdulla AE, Nishiyama H, Kato N (2012) Extending the lifetime of wireless sensor networks: a hybrid routing algorithm. Comput Commun 35:1056–1063
41. Azad AP, Chockalingam A (2011) Enhancing lifetime of wireless sensor networks using multiple data sinks. Int J Sens Netw 9:139–157
42. Cheng S-T, Chang T-Y (2012) An adaptive learning scheme for load balancing with zone partition in multi-sink wireless sensor network. Expert Syst Appl 39:9427–9434
43. Kim D, Wang W, Wu W et al (2013) On bounding node–to–sink latency in wireless sensor networks with multiple sinks. Int J Sens Netw 13:13–29
44. Safa H, Moussa M, Artail H (2014) An energy efficient genetic algorithm based approach for sensor-to-sink binding in multi-sink wireless sensor networks. Wirel Netw 20:177–196
45. Jiang H, Sun R (2014) Energy optimized routing algorithm in multi-sink wireless sensor networks. Appl Math 8:349–354
46. Deng R, He S, Chen J (2018) An online algorithm for data collection by multiple sinks in wireless-sensor networks. IEEE Trans Control Netw Syst 5:93–104
47. Masdari M, Naghiloo F (2017) Fuzzy logic-based sink selection and load balancing in multi-sink wireless sensor networks. Wirel Pers Commun 97:2713–2739
48. Cheng L, Niu J, Cao J et al (2014) QoS aware geographic opportunistic routing in wireless sensor networks. IEEE Trans Parallel Distrib Syst 25:1864–1875
49. Ahmed A, Bakar KA, Channa MI et al (2017) Energy-aware and secure routing with trust for disaster response wireless sensor network. Peer–Peer Netw Appl 10:216–237
50. Golsorkhtabar M, Nia FK, Hosseinzadeh M, Vejdanparast Y (2010) The novel energy adaptive protocol for heterogeneous wireless sensor networks. Proceedings of 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), IEEE: 178–182
51. Chand S, Singh S, Kumar B (2014) Heterogeneous HEED protocol for wireless sensor networks. Wirel Pers Commun 77:2117–2139


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

**Sandeep Verma** is currently a research scholar in ECE department in Dr. BR Ambedkar National Institute of Technology, Jalandhar. His research interest includes energy efficient Wireless Sensor Networks and IoT based architectures. He is the potential reviewer of IEEE Access Journal.

**Dr. Neetu Sood** received the B.Tech. degree (with Honour) in Electronics and Communication from Sant Longowal Institute of Engineering and Technology, India 2000. In year Jan. 2002, she completed her M.E. Degree (With Honour) in Electronics and Communication Engineering from Thapar University, Patiala, India. She completed her PhD in ECE from Dr. B R Ambedkar National Institute of Technology, Jalandhar. Currently, she is working as an Assistant Professor at Dr. B R Ambedkar National Institute of Technology, Jalandhar. Her research interest includes Wireless Sensor Network, simulation of wireless systems based on OFDM and simulation of fading channels.

**Dr. Ajay K Sharma** received his BE in Electronics and Electrical Communication Engineering from Punjab University Chandigarh, India in 1986, MS in Electronics and Control from Birla Institute of Technology (BITS), Pilani in the year 1994 and PhD in Electronics Communication and Computer Engineering in the year 1999. He has worked as a Director at NIT Delhi and also had a charge as a Director at NIT Hamirpur. Currently, he is Vice Chancellor of I.K. Gujral Punjab Technical University, Kapurthala, Punjab. He has supervised 18 Ph.D. and 48 M.Tech thesis.