

Energy proficient clustering technique for lifetime enhancement of cognitive radio-based heterogeneous wireless sensor network

Venkatasubramanian Srividhya  and Thangavelu Shankar

Abstract

Utilizing the available spectrum in a more optimized manner and selecting a proper routing technique for transferring the data, without any data collision, from the sensor node to the base station play a major role in any network for increasing their network lifetime. Cognitive radio techniques play a major role to achieve the same, and when combined with wireless sensor networks the above-said requirements can be greatly accomplished. In this article, a novel energy-efficient distance-based clustering and routing algorithm using multi-hop communication approach is proposed. Based on distance, the given heterogeneous cognitive radio-based wireless sensor networks are divided into regions and are allocated with a unique spectrum. Dynamic clustering through distance calculation and routing of data through multi-hop communication is done. The simulation results illustrate that the proposed algorithm has improved energy efficiency and is more stable. The first node death and 80% node death illustrate the improved scalability. Also, the increased throughput aids in maintaining the residual energy of the network, which further solves the problem of load balancing among nodes. All the above results combined with half node death analysis show that the proposed algorithm also has an improved network lifetime.

Keywords

Cognitive radio wireless sensor networks, heterogeneous network, lifetime enhancement, multi-hop clustering, stability

Date received: 18 September 2017; accepted: 4 March 2018

Handling Editor: Xuanli Wu

Introduction

Wireless sensor network (WSN) has paved a gateway to connect back to the wired world. It is a data-centric network that routes the required data to their destination, irrespective of the location from where it is passed. The sensor nodes are randomly deployed in these networks such that the position of these nodes is non-predetermined as given by Akyildiz et al.^{1,2} These lightweight sensors have restricted the amount of energy and are responsible for continuously monitoring the environment that they are deployed in. They are capable of sensing physical parameters such as temperature and humidity in the surrounding and send these data to

the base station (BS) located either outside or inside the network. As the number of nodes sensing a given region increases, the fault tolerance of the network increases. It also leads to precision in the observations made by the sensors in the network. Thus, the

Department of Communication Engineering, Vellore Institute of Technology (VIT), Vellore, India

Corresponding author:

Venkatasubramanian Srividhya, Department of Communication Engineering, Vellore Institute of Technology (VIT), Vellore 632014, Tamil Nadu, India.

Email: v.srividhya@vit.ac.in



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License

(<http://www.creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without

further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<http://www.uk.sagepub.com/aboutus/openaccess.htm>).

distribution of nodes and their density plays a significant role in the scalability of the network.

As each node contains a restricted amount of resources, data fusion and aggregation becomes an indispensable aspect. This forms a base for many algorithms that are structured. In data gathering, the efficiency of a sensor node is measured by decreasing the overheads involved with the data packet transmission and distributing the same load equally among all nodes. Therefore, an efficient data gathering approach is required to reduce the burden imposed on the master node. Data dissemination involves distributing the data in the network in such a way that it reaches the BS by consuming a minimum amount of energy available for the nodes. This forms a basis for formulating different clustering and routing protocols. Also, due to the high-density node deployment, the problem of data collision is very prominent in these routing algorithms. Hence, WSN, when integrated with cognitive radio (CR) technology, helps to overcome these limitations, by providing an opportunistic manner of accessing a shared spectrum by multiple users.

A unique capability of the CR network is to vary its transmitting and receiving parameters, by estimating the adjacent radio conditions in collaboration with other nodes.³ Cognitive radio-based sensor node (CRSN) is best suited for the wide range of application-specific requirements offering several potential benefits. It has been proposed as one of the most promising technologies to address spectrum access and utilization challenges in cognitive radio wireless sensor network (CRWSN).³ Hence, cognition technique when properly incorporated in WSN could bring about a lot of improvement in their performance.

Performance evaluation, by comparison, reveals that the proposed work is improved in terms of network lifetime, stability, and energy efficiency. The rest of the article is organized as follows. Section “Related works” gives an idea about the various algorithms that exist with their merits and the demerits. The next section “Description of the proposed algorithm” describes the newly proposed work, and their simulation results for different scenarios are given in the section “Performance evaluation and result analysis.” The last section “Conclusion” explains the inferences that could be drawn out of the proposed work and the future enhancement which can be made.

Related works

CR technology is an imaginative radio design logic that builds range usage by taking advantage of unused and under-used ranges in progressively changing conditions.⁴ It is stated by Joshi et al.⁵ that the cognitive method is the way toward knowing through

observation, arranging, thinking, acting, and persistently refreshing and updating with a background marked by learning. This distinctive capability of sensing the environmental parameters helps to enhance the lifetime of the network. In a CR communication environment, there are two types of users—the primary users, which are given access to the licensed spectrum, and the secondary users, which are given access to the unlicensed spectrum. The spectrum sensing is carried over by the CR nodes and data transmission takes place efficiently over the spectrum in a more opportunistic manner.

Akan et al.⁶ state that the primary unit which differs between a classical sensor node and CRSN node is the CR transceiver unit as shown in Figure 1. Generally, in any sensor network, all the events are sensed by a sensor node and are communicated to nearby nodes that act as a sink for the sensed information. By this way, all nodes communicate data to the BS through multi-hop routing. This intermediate node which helps to transmit data from a cluster head (CH) to BS consumes more energy due to transmission and reception and dies out fast. Hence, the distance between the source and the BS plays a vital role. Also, while transmitting, the data could undergo collision and thus packet retransmission has to be done often. Hence, an enhanced clustering and routing protocol along with a defined allocation of the spectrum band avoids all the above-mentioned issues. This is accomplished when a CR-based WSN network is formed. As explained by Akyildiz et al.^{1,2} and Cook and Das,⁷ the CR-based WSN nodes change their working parameters for channel access so as to expand the lifetime of the network. Earlier, a conceptual design for CR-based WSN was

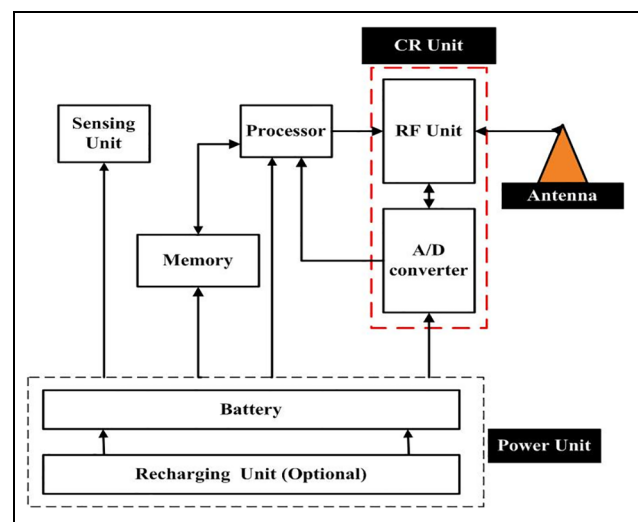


Figure 1. Architecture of cognitive radio-based sensor node.

given where, for the same transmission power, the maximum communication range in a CR channel was almost doubled.⁸ Hence, the throughput of the network was also increased.

Grouping of the nodes in the network into a set of small clusters with one node appointed as a CH is called clustering.⁹ Although clustering is an old concept, choosing an intelligent strategy that could optimize the size of the cluster and can select the most eligible CH is still being researched profoundly. Clustering is helpful in increasing the scalability of the network by reducing the direct communication to the BS from the nodes that are far away. It also helps to accomplish stability and network scalability, as well as to sustain shared tasks, such as channel detecting and channel access, which are the basics of CR operations.¹⁰ The process of electing a CH depends on various factors such as residual energy, number of rounds to which a given node has not become the CH, and finally the optimum number of CHs that are elected by the network for each round. A conceptual design for CR-based WSN was given by Cavalcanti et al., where the author stated that for the same transmit power, the maximum communication range in CR channel is almost twice.⁸ Hence, the throughput of the network could also be increased.

There have been many algorithms designed for node clustering and routing in WSN. These clustering protocols cannot be used in CR-WSNs because the procedure to elect the CHs varies in CR-WSN algorithms as compared to the conventional WSNs. LEACH (Low-Energy Adaptive Clustering Hierarchy) algorithm⁹ is one of the basic homogeneous cluster-based energy optimization algorithms used for improving the lifetime of the nodes in WSN. Even though they work well for small-sized networks, they became less stable when the size of the network was increased causing load imbalance. Also, it did not take into account the residual energy of the node, which was taken into consideration in the DEEC (Distributed Energy-Efficient Clustering) algorithm.¹¹

Later, SEP (Stable Election Protocol) algorithm¹² was proposed, with three-level node heterogeneity (as normal nodes, advanced nodes, and super nodes). This approach helped to improve the lifetime of the network with the same CH election technique used in LEACH. Only one difference should be cited was the modified probability function for becoming the CH and the threshold function. One of the limitations of the LEACH algorithm is its homogeneity, where the nodes that are selected as CHs have to perform more functions than the normal nodes. This caused the energy of those nodes to drain out faster. The performance evaluation of SEP showed an enhanced network lifetime and increased stability as compared to LEACH, but on increasing the network size, the nodes far away from BS started dying out faster. Later, another three-level

heterogeneous network was introduced in EDEEC (Enhanced Distributed Energy-Efficient Clustering) algorithm.¹³ Here, the transmission of data packets incurs energy dissipation depending upon the distance between the sender and the recipient. This process continues until all the nodes exhaust their energy and die. These algorithms increased the network lifetime but did not provide good network stability.

An advanced version of a cluster-based heterogeneous network called MCR (Multi-hop Clustering and Routing) algorithm¹⁴ was found. Here, the elected CHs do not transmit the data directly to the BS. Instead, each CH transmits its data to a nearby CH until it reaches the BS. This strategy helped in increasing the scalability of the network as only a small number of CHs were interacting with the BS and therefore energy dissipation was reduced. Also, the value of heterogeneity was considered to be two which led to an enhanced performance.

In MCR, the election of CH was based on the weighted election probability (WEP) of the nodes where one can make a high-energy node as a CH more frequently. This reduces the burden on the low-power nodes and therefore helped in increasing the network lifetime. Also, by taking high initial energy for the network, throughput and network lifetime increased as it took more time for 50% of nodes to die. As compared to the chain-based algorithms such as PEGASIS (Power-Efficient Gathering in Sensor Information Systems),¹⁵ the reliability increased as the information was passed to the CH through a limited number of hops. Although the MCR algorithm performs well on the grounds of network lifetime and reliability, it fails on the grounds of network stability. Despite these demerits, MCR was able to outperform SEP and LEACH protocols.

Later, another protocol called EEMHR (Energy-Efficient Multilevel Heterogeneous Routing) algorithm¹⁶ was proposed. The main variation in this algorithm was that the heterogeneity of the network was not restricted to two, as was in the case of MCR algorithm. Moreover, the value of the threshold was dependent on the total number of nodes alive and the total number of nodes present initially. The main problem with this algorithm is the increased number of CH at the later stage of the network, which greatly increases the packet loss due to the collision. Also, the stability of the network was poor compared to other similar algorithm, EEMDC (Energy-Efficient Multilevel and Distance Aware Clustering).

The EEMDC algorithm¹⁷ is the first algorithm to consider the factor of distance in assigning the CH probabilities. In this, the complete network was divided into three layers with heterogeneous nodes deployed. These different layers and nodes in these layers had different hop counts. It also incorporated average energy

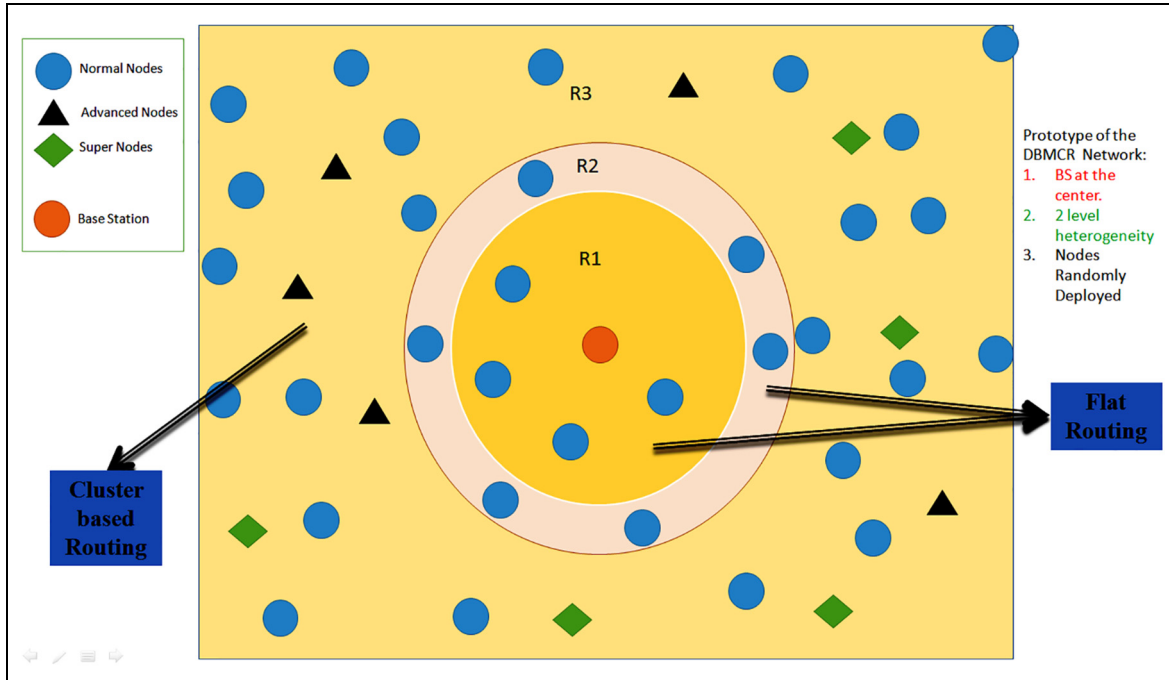


Figure 2. Division of the network regions in DBMCR.

calculation while electing a node as a CH. This strategy helped in optimizing the size of the cluster and the nodes that are nearer to the BS as a preferred option for being the CH. Even though the given algorithm was highly stable and the network lifetime was prolonged, it increased the channel contention caused due to many complex processes involved in selecting the eligible CH and the optimum path to transfer the data. Hence, packet loss due to the collision was very high compared to EEMHR.

When compared with the MCR algorithm, the number of overheads was high in EEMHR and EEMDC. Hence, in order to improve the performance of the MCR algorithm in terms of scalability, stability, throughput, and network lifetime, Tyagi et al.¹⁸ formulated an algorithm called DBMCR (Distance-Based Multi-hop Clustering and Routing). The only difference between MCR and DBMCR is that in order to avoid collision of data packets, the concept of spectrum allocation is introduced. This is one of the major characteristics of CR network which helps to overcome the problems faced by bursty communication of a normal sensor network. By exploiting the spectrum in an opportunistic fashion, CR enables users to sense which portions of the spectrum are available, select the best available channel, and coordinate spectrum access with other users.¹⁹

In this, the WEP was formulated on the basis of the distance of the nodes from the BS. The whole network is divided into three regions based on the distance

threshold as shown in Figure 2. Different spectra are allocated to the first two regions. Flat routing protocol is used in the first and second regions and the third region uses the distance-based clustering technique. The region which is near the BS used direct communication for transmission. First-order radio model was used and the rest of all the design considerations were followed similar to the MCR algorithm. By allocating a spectrum to each region, it was found that the time required for the first node to die was higher as compared to the MCR algorithm and SEP algorithm. A similar trend was observed in the case of the number of rounds before half the number of node dies. In case of the number of alive nodes, it was observed that after certain rounds, DBMCR equals MCR. Later, MCR starts declining faster than DBMCR.

Due to a large number of applications associated with the WSNs, it is necessary to formulate an algorithm that could efficiently deal with the problem of routing and clustering along with reducing data loss due to channel contention. Although existing algorithms have taken care of either network lifetime or effective spectrum allocation among the nodes, the approaches having both the requirements are rare to find. Therefore, a new algorithm that can outperform the existing above algorithms on the basis of stability, scalability, network lifetime and solve the problem of packet loss occurring due to channel contention in the network is proposed. In order to make the proposed algorithm energy efficient and the process of clustering

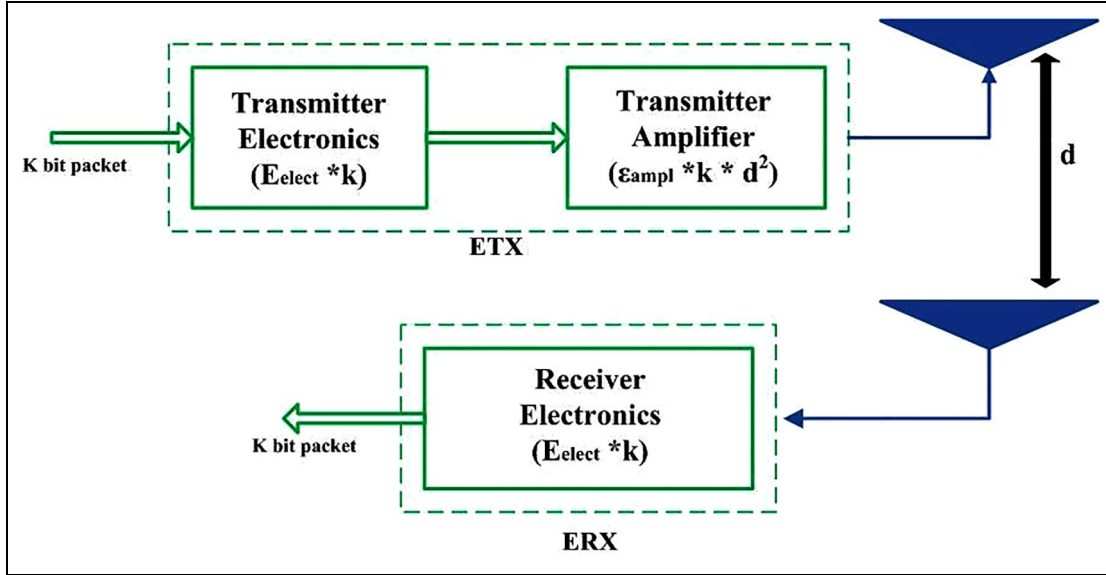


Figure 3. Radio energy model.

more effective, the concepts discussed in the above-said algorithms were amalgamated with CR spectrum allocation technique.

Description of the proposed algorithm

In the proposed algorithm, the advantages of distance-based clustering and multi-hop routing have been amalgamated with effective spectrum allocation strategy, in a heterogeneous CR sensor network environment. Deploying heterogeneous nodes, partitioning the region for better spectrum allocation, electing CH based on the average distance coverage range, and assigning cluster member (CM) based on average threshold distance are the major features that have been introduced in this algorithm. In order to mathematically analyze the proposed model, the nomenclatures used are given in Table 1.

Consider the network area of size $M * M$ with randomly deployed sensor nodes. The area of the network is divided into parts and regions. In deciding the cluster range, the CHs are elected and the calculation of the distance threshold finds the CMs for each CH in the region where clustering takes place. The CHs are responsible for sending data to the BS. The major reason for nodes to lose energy at a high rate is due to collision and packet re-transmission during data transfer. CR spectrum allocation technique helps to overcome this drawback and makes the data transfer technique more energy efficient. Hence, in the proposed algorithm, the region where the clustering and routing concept is deployed is divided into four equal parts. Each part is assigned a unique spectrum from the licensed band for it to transmit data to the BS without any collision.

Table 1. Abbreviations used.

Abbreviation	Meaning
N	Total number of nodes in the network
n	Total number of nodes in region III
$A(R_3)$	Area of region R_3 where clustering occurs
$M * M$	Dimension of the network
E_{elect}	Energy required to elect CH
E_i	Residual energy of a node i
E_{avg}	Average energy of the network
E_{txd}	Energy for transmission of data
E_{rxd}	Energy for reception of data
ϵ_{fs}	Free space amplifier factor
ϵ_{mp}	Multipath amplifier factor
γ	Path loss factor
d_{CH-BS}	Distance between CH and BS
d_i	Distance of node i from BS
d_{avg}	Average distance of nodes from BS
k	Bit rate
k_0	Multipath environment factor
P_{opt}	Optimum probability of becoming CH
P_{ch}	Probability of electing CH
a_i	Heterogeneity factor
T	Threshold of a node to become CH

CH: cluster head; BS: base station.

Energy consumption model

The implementation of proposed algorithm was done using first-order radio model similar to the one proposed in LEACH,⁹ SEP,¹² MCR,¹⁴ and EEMHR¹⁶ algorithms and is shown in Figure 3

$$E_{txd}(k, d) = (E_{elect} * k) + (\epsilon_{ampl} * k * d^\gamma) \quad (1)$$

$$E_{txd}(k, d) = \begin{cases} (E_{elect} * k) + (\epsilon_{ampl} * k * d^2) & (d < d_0) \\ (E_{elect} * k) + (\epsilon_{ampl} * k * d^4) & (d \geq d_0) \end{cases} \quad (2)$$

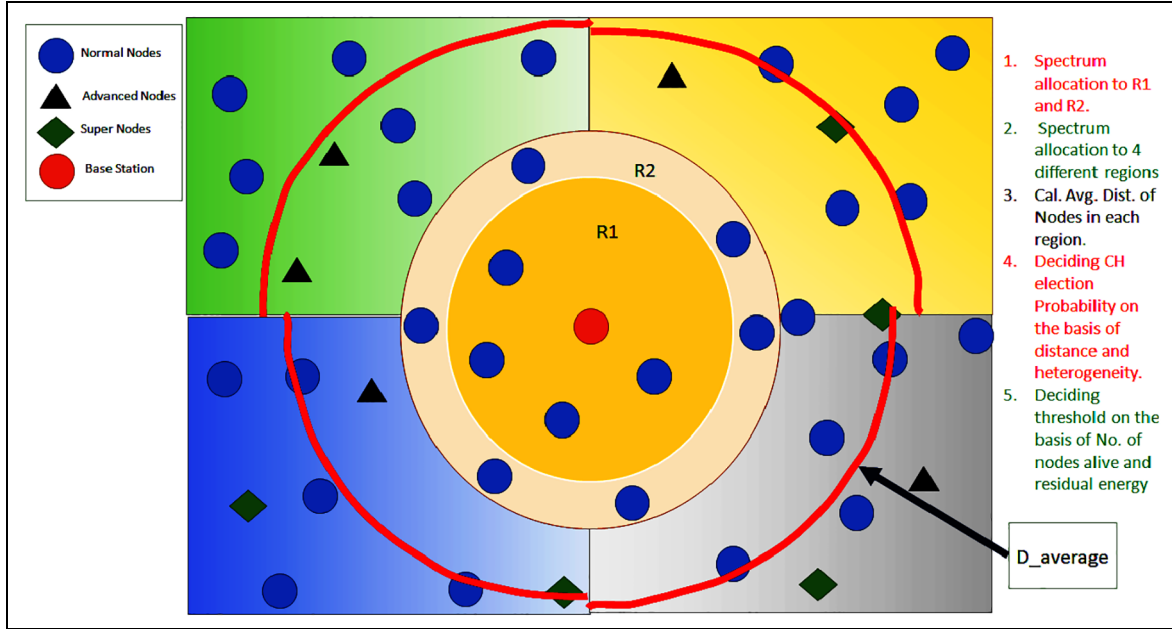


Figure 4. Division of the network region with different spectra in the proposed algorithm.

$$E_{rxn}(k) = E_{elect} * k \quad (3)$$

Each node will use the energies calculated by equations (1)–(3) for transmitting and receiving a “ k ” bit packet.

Division into regions

A network of area $M * M$ meters with randomly deployed nodes is considered. The complete area of the network is divided into four equal parts P_1, P_2, P_3, P_4 , as shown in Figure 4. In order to find the regions R_1 and R_2 , the following calculations were made.

The radius of region 1 (R_1), r_1 , and radius of region 2 (R_2), r_2 , are calculated by considering the following equation

$$\frac{Area(R_1 + R_2)}{Area(R_3)} = \frac{2}{8} \quad (4)$$

Whether the location of the BS is at the center or at the corner, for the above condition to hold true

$$\frac{\pi(r_1^2 + r_2^2)}{M^2 - (r_1^2 + r_2^2)} = 0.25 \quad (5)$$

By considering that the areas of region R_1 and region R_2 are equal, radii r_1 and r_2 are calculated by solving equations (4) and (5). Since the single-hop communication is restricted to an optimum radius, energy dissipation is lesser compared to a multipath environment. Hence, r_2 for this case is 87.7 m. The region beyond R_2 is considered as R_3 region.

Nodes of regions R_1 and R_2 do not involve in the clustering process. Region R_1 transmits the data directly to the BS, while region R_2 communicates partially or totally through region R_1 depending on node densities. That is, if a node which resides in region R_2 wishes to communicate to BS, it can transmit the data directly or acquire the spectrum of region R_1 and transmit through them. One of the most interesting facts in the proposed algorithm is the implementation of single-hop communication in region R_3 . Clustering takes place in this region and each CH can directly communicate to the BS using the spectrum allocated to them.

Spectrum allocation in regions

After dividing $M * M$ meters into four equal parts, the distance of the regions R_1 and R_2 are calculated. Node density of each region is measured by the BS which helps in allocating spectrum to these regions. The first two regions acquire an individual spectrum and region R_3 which is divided into four equal parts acquire a unique spectrum for each part. This division of the network and unique spectrum assignment plays a major role for collision free data transmission. Since one-hop communication of data transfer has been introduced for the CHs of region R_3 , dividing it further into four parts and allocating individual spectrum to each help the CH of each part to transmit their data independently. This avoids packet collision and hence packet retransmission is highly reduced. This plays a major role in increasing the network lifetime further as compared to DBMCR algorithm.

Heterogeneity

In order to define the heterogeneous model of the network, consider “ N ” number of nodes deployed in the network.

The percentage of advanced nodes is taken as “ m ” out of which “ m_0 ” percent of the nodes are super nodes.

The total number of normal nodes is “ $(N(1 - m))$ ”. The total number of advanced nodes is “ $(N - N(1 - m))$ ”.

Out of the total number of advanced nodes, “ $(m_0 * (N - N(1 - m)))$ ” nodes are super nodes and “ $(m * N - m * m_0 * N)$ ” nodes are advanced nodes.

Energy associated with each super nodes is “ $(E_0 * (1 + b))$ ”. Energy associated with each advanced nodes is “ $(E_0 * (1 + a))$ ”, where “ E_0 ” is the initial energy, “ a ” is the energy increment factor for advanced nodes, and “ b ” is the energy increment factor for super nodes.

Therefore, the heterogeneity of the network comes out to be “ $(L = (1 + m(a - m_0 * (a - b))))$ ”.

And, the total increment in the initial energy of the entire network is “ $(L * E_0)$ ”.

Calculation of optimal probability to elect a CH

γ is the path loss factor and is dependent on the distance. Its value is 2 for a free-space model and 4 for a multipath model. There are four scenarios considered for the analysis of this algorithm depending on the position of the BS and the basic energy of the nodes. Consider the BS located at a corner, as shown in Figure 5. Two different initial energies of 0.25 and 0.5 J were taken for simulation. Similarly, in Figure 6,

consider the BS to be at the center, and two different initial energies of 0.25 and 0.5 J were taken for simulation. The mathematical analyses of these scenarios are explained as given below.

First, the optimal value for a multipath environment which is given by equation (6) is considered

$$k_0 = \sqrt{\frac{n\epsilon_{fs}A}{2\pi(\epsilon_{mp}d_{CH-BS}^\gamma - E_{elec})}} \quad (6)$$

k_0 is the multipath environment factor, A is the area where clustering takes place, and M is the total area.

Initially, consider the case when the BS is located at a corner, as shown in Figure 5.

Scenario 1.

- Location of the BS: corner;
- Path loss factor γ : 2;
- Expected value for d_{CH-BS}^2 is

$$E[d_{CH-BS}^2] = \frac{2}{3}M^2$$

The value for multipath environment was found to be given in equation (7)

$$k_0 = \sqrt{\frac{n\epsilon_{fs}A}{2\pi(\epsilon_{fs}(\frac{2}{3}M^2)) - E_{elec}}} \quad (7)$$

Scenario 2.

- Location of the BS: corner;
- Path loss factor γ : 4;

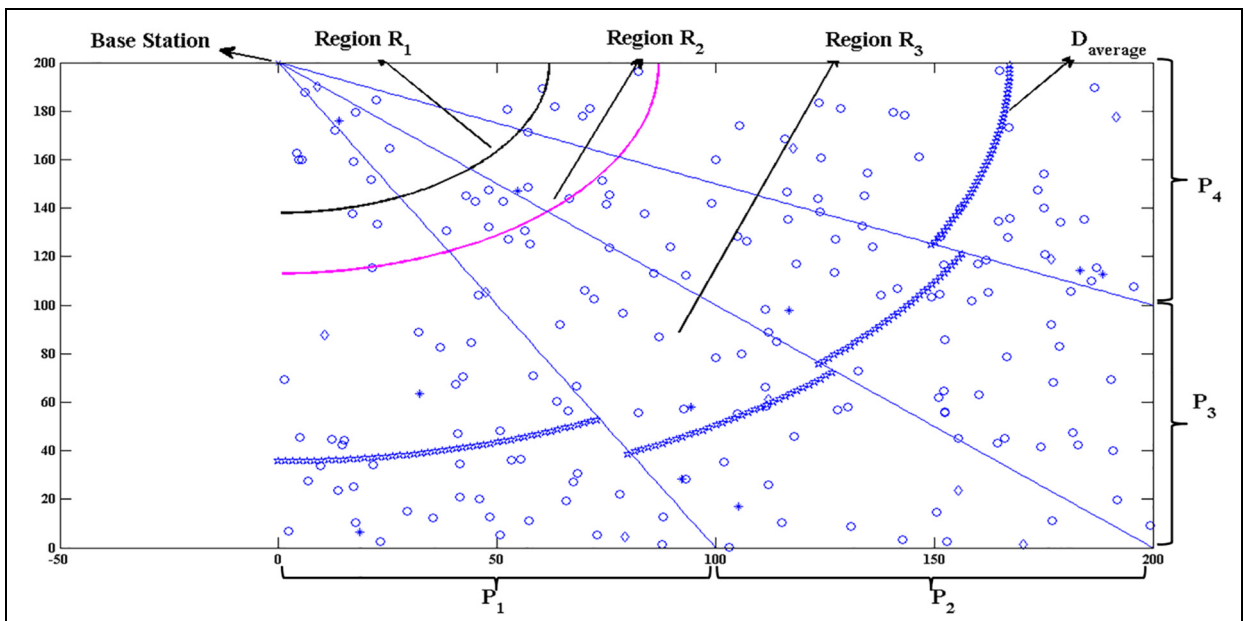


Figure 5. Scenario when the BS is at the corner.

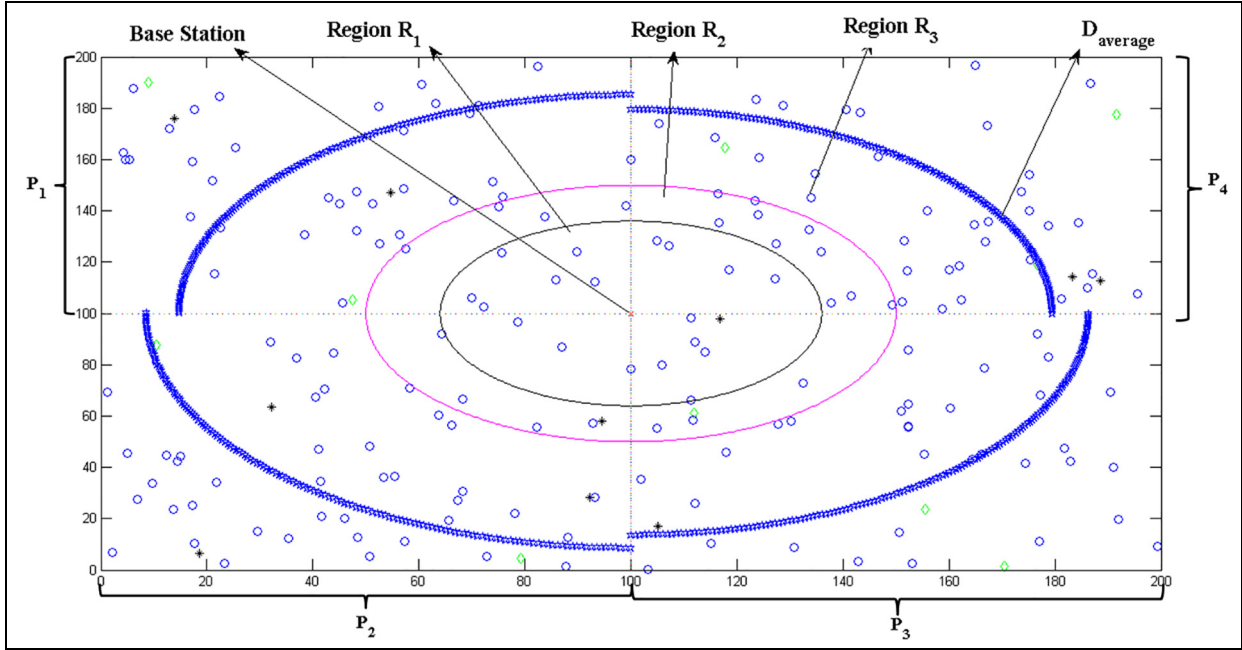


Figure 6. Scenario when the BS is at the center.

Expected value for d_{CH-BS}^4 is

$$E[d_{CH-BS}^4] = \frac{28}{45}M^4$$

The value for multipath environment was found to be given in equation (8)

$$k_0 = \sqrt{\frac{45n\epsilon_{fs}A}{56\pi\epsilon_{mp}M^4}} \quad (8)$$

Next, consider the case when the BS is located at the center, as shown in Figure 6.

Scenario 3.

Location of the BS: center;

Path loss factor γ : 2;

Expected value for d_{CH-BS}^2 is

$$E[d_{CH-BS}^2] = \frac{M^2}{6}$$

The value for multipath environment was found to be given in equation (9)

$$k_0 = \sqrt{\frac{3nA}{\pi M^2}} \quad (9)$$

Scenario 4.

Location of the BS: center;

Path loss factor γ : 4;

Expected value for d_{CH-BS}^4 is

$$E[d_{CH-BS}^4] = \frac{7}{180}M^4$$

The value for multipath environment was found to be given in equation (10)

$$k_0 = \sqrt{\frac{90n\epsilon_{fs}A}{7\pi\epsilon_{mp}M^4}} \quad (10)$$

On the basis of the above formula, the value of optimal probability could be calculated as given by equation (11)

$$P_{opt} = \frac{k_0}{N} \quad (11)$$

Calculation of WEP

After the calculation of optimal probability, it is necessary to find the election probability associated with each type of node, that is, election probability associated with normal nodes, advanced nodes, and super nodes. Since the weight of each type of node is multiplied with the election probability, the complete product is known as WEP. Hence, in order to make the algorithm more efficient, a factor of the distance of nodes from the BS called $D_{average}$ is multiplied. This helps in the reduction in the probability to become a CH for all the nodes far away from the BS. The equations obtained are given below.

Initially, calculate the average distance of all the nodes which are participating in cluster-based communication using equation (12)

$$d_{avg} = \frac{1}{N} \sum_{i=1}^n d_i \quad (12)$$

where d_i is the distance of i th node from the BS and N is the total number of nodes participating in the clustering process.

After average distance calculation, nodes are assigned their WEPs according to the following rule:

If $d_{i,j} < d_{avg,j}$, then

$$P_n = \frac{P_{opt}}{L} \quad (13)$$

$$P_{adv} = \frac{(1 + \alpha)P_{opt}}{L} \quad (14)$$

$$P_{super} = \frac{(1 + b)P_{opt}}{L} \quad (15)$$

Else, if $d_{i,j} > d_{avg,j}$, then

$$P_n = \frac{P_{opt}}{L} * \left(\frac{P_{OS}}{d_{avg,j}} - \frac{d_{i,j}}{d_{avg,j}} \right) * \left(\frac{d_{avg,j}}{d_{i,j}} \right) \quad (16)$$

$$P_{adv} = \frac{(1 + \alpha)P_{opt}}{L} * \left(\frac{P_{OS}}{d_{avg,j}} - \frac{d_{i,j}}{d_{avg,j}} \right) * \left(\frac{d_{avg,j}}{d_{i,j}} \right) \quad (17)$$

$$P_{super} = \frac{(1 + b)P_{opt}}{L} * \left(\frac{P_{OS}}{d_{avg,j}} - \frac{d_{i,j}}{d_{avg,j}} \right) * \left(\frac{d_{avg,j}}{d_{i,j}} \right) \quad (18)$$

where $d_{i,j}$ is the distance of node i with channel j from BS, $d_{avg,j}$ is the average distance of nodes of channel j from BS, P_{opt} is the optimal probability, P_{OS} is the maximum of X and Y co-ordinates of BS, α is the energy incremental factor of advanced nodes, L is the heterogeneity factor (increment in total initial energy of the network due to heterogeneity), b is the energy incremental factor of super nodes, P_n is the probability of a normal node to become a CH, P_{adv} is the probability of a advancement node to become a CH, and P_{super} is the probability of a super node to become a CH. From equations (16) to (18), it is found that the factor $((P_{opt}/d_{avg,j}) - (d_{i,j}/d_{avg,j}))$ is responsible for restricting the CHs within the average distance of that region and the factor $(d_{avg,j}/d_{i,j})$ is responsible for enhancing the effect of distance on the weighted election probabilities.

Threshold calculation

The final step toward a CH election is the calculation of threshold for each type of node. As the threshold value is decided, each node generates a random number which is then compared with the value of threshold of that type of node. If the value of random number is less than the value of the threshold, that node is selected as a CH.

Correspondingly, the value of threshold is calculated below as in equation (19).

If $s_i \in G$, then

$$T_{s_i} = \frac{P_{ch}(i) * E_i * N_j}{\left((1 - P_{ch}(i)) * \left(r_{\text{mod}} \frac{1}{p_{ch}(i)} \right) \right) * E_{avg} * N_{alive,j}} \quad (19)$$

Otherwise

$$T_{s_i} = 0 \quad (20)$$

On the basis of equations (19) and (20), the new threshold is calculated. The clusters are formed and the transmission is initiated.

Choosing a minimum path for the CHs

After the election of the CHs, data are transmitted from each of the CM from their respective CH. These CHs transmit the received data to the BS. If direct transmission fails, then the CH seeks for another CH present at a minimum distance from it to transmit the data. Initially, the location of each CH is broadcasted by the BS to all the sensor nodes. With this information, the processor inside each node helps to calculate this minimum distant CH.

Cluster range

One of the important modifications that is being incorporated in this algorithm is the restriction of a CH's range on the basis of the average distance of the region. Using equation (21), the range of the CH can be decided

$$Range_j = D_{\text{max}} \quad (21)$$

Now, the region with a high value of average distance will have smaller cluster ranges as the CH's location could easily cover far away nodes present in the network. This avoids the connectivity loss in the network among nodes.

Routing

Routing is done by allocating a time slot to each node under the time-division multiple access (TDMA) scheduling and each node waits for its turn to transmit its data, as shown in Figure 7. The CHs which are far away transmit data through a multi-hop transmission, as shown in Figure 8, and the nodes which cannot fall under any cluster transmit the data through a chain process, as shown in Figure 9.

Simulation parameters

MATLAB software^{20,21} is used in order to do the extensive simulation of the proposed algorithm, and Table 2 gives the parametric setting used for the network analysis.

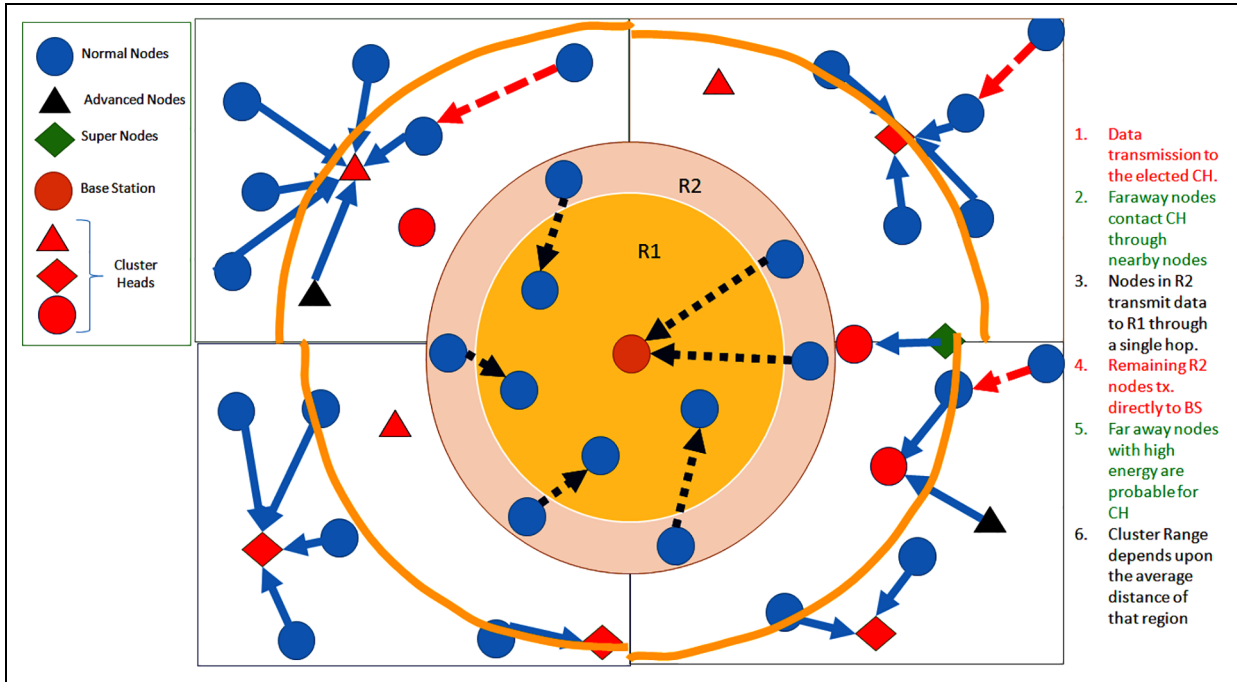


Figure 7. Clustering and routing of data.

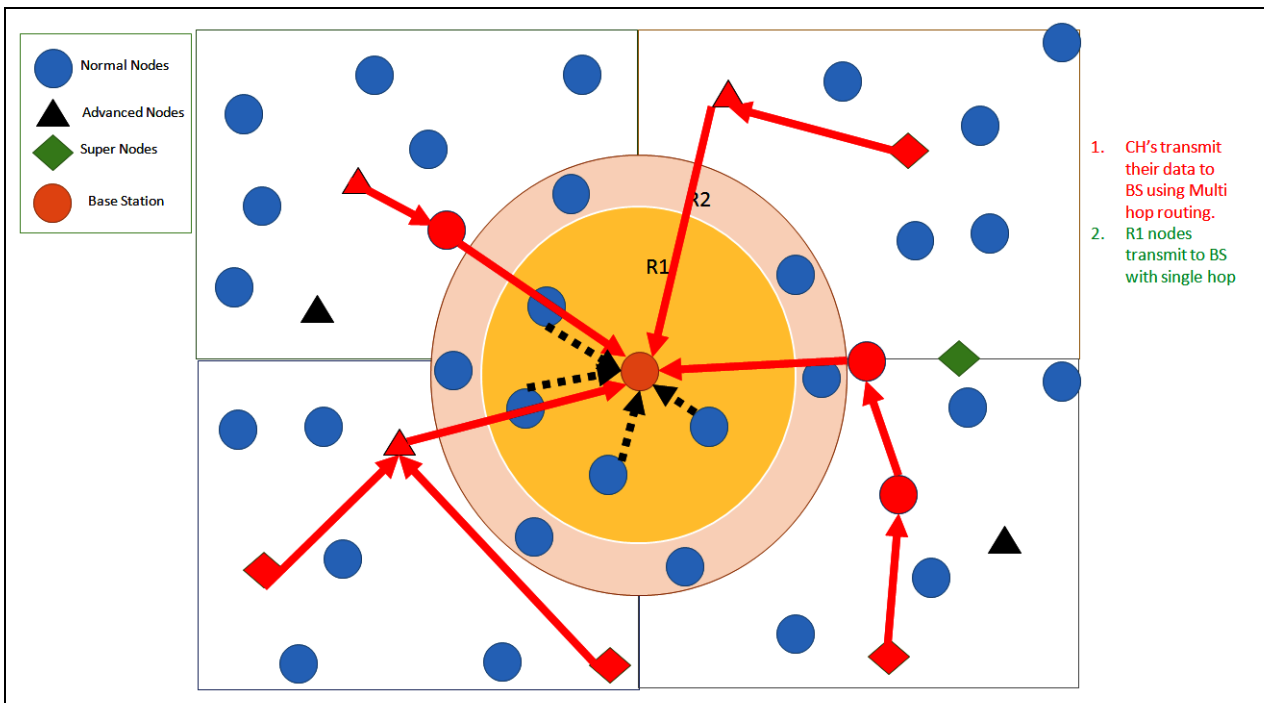


Figure 8. Multihop routing.

Flow chart of the algorithm

The steps involved in creating and simulating the proposed algorithm are given as a flow chart in Figure 10.

Performance evaluation and result analysis

Energy efficiency

Residual energy and throughput are the two main parameters which aid in analyzing the energy efficiency of any network.

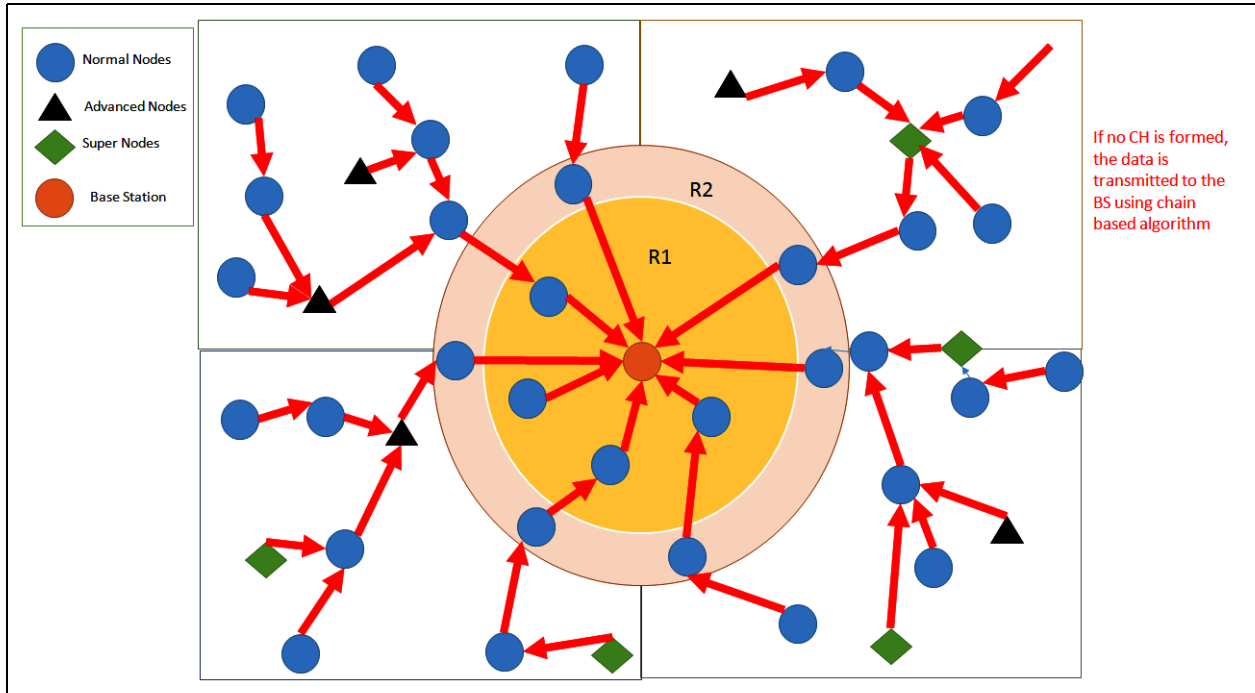


Figure 9. Scenario with no CH formed.

Residual energy. The residual energy analysis gives a comparative study on the amount of energy consumed by the nodes for transmitting and receiving data in a network. This helps in distributing the load among all nodes equally and thereby helps in increasing the network lifetime. The analysis from Figures 11 to 14 proves that the proposed algorithm has better residual energy as compared to SEP, MCR, and DBMCR algorithms.

Throughput. The amount of successful data transmission from a node to the BS is measured from the throughput analysis. It is calculated in terms of the amount of data bits transmitted per time slot. This helps to detect the unsuccessful transmission and request the node for retransmission, which mainly contributes for the energy decrease among the nodes. Figures 15–18 give a comparative analysis and prove that the proposed algorithm has improved throughput compared to SEP, MCR, and DBMCR algorithms.

Improvement of energy efficiency. From Table 3, the percentage increase in residual energy for the proposed algorithm can be calculated. Similarly, the simulation results of residual energy along with the throughput analysis also prove that the proposed algorithm has increased energy efficiency as compared to SEP, MCR, and DBMCR algorithms. Due to this, the amount of successful data transmission without any collision has

Table 2. Simulation parameters.

Parameter	Value
No. of nodes (N)	100 nos
Size of the network	200*200 m ²
Location of the base station	Corner (0, 200) Center (100, 100)
Data aggregation energy	5 nJ/bit/round
ϵ_{amp} (amplification factor)	0.0013 pJ/bit/m ⁴
ϵ_{fs} (free space constant “t”)	10 pJ/bit/m ²
Initial energy	0.5 J
Incrementing factor “a”	
For advanced nodes	1
Incrementing factor “b”	
For super nodes	2
γ path loss factor	4
Packet size	50 bytes
No. of normal nodes	90% of N
No. of advanced nodes	5% of N
No. of super nodes	5% of N

increased. Also, the introduction of heterogeneity and assigning spectrum to different regions helps to distribute the load evenly in the network. This consequently improved the energy efficiency of the proposed algorithm.

Stability

First node death. It is the measure of the death of the first node of a network. Figures 19–22 give a relative

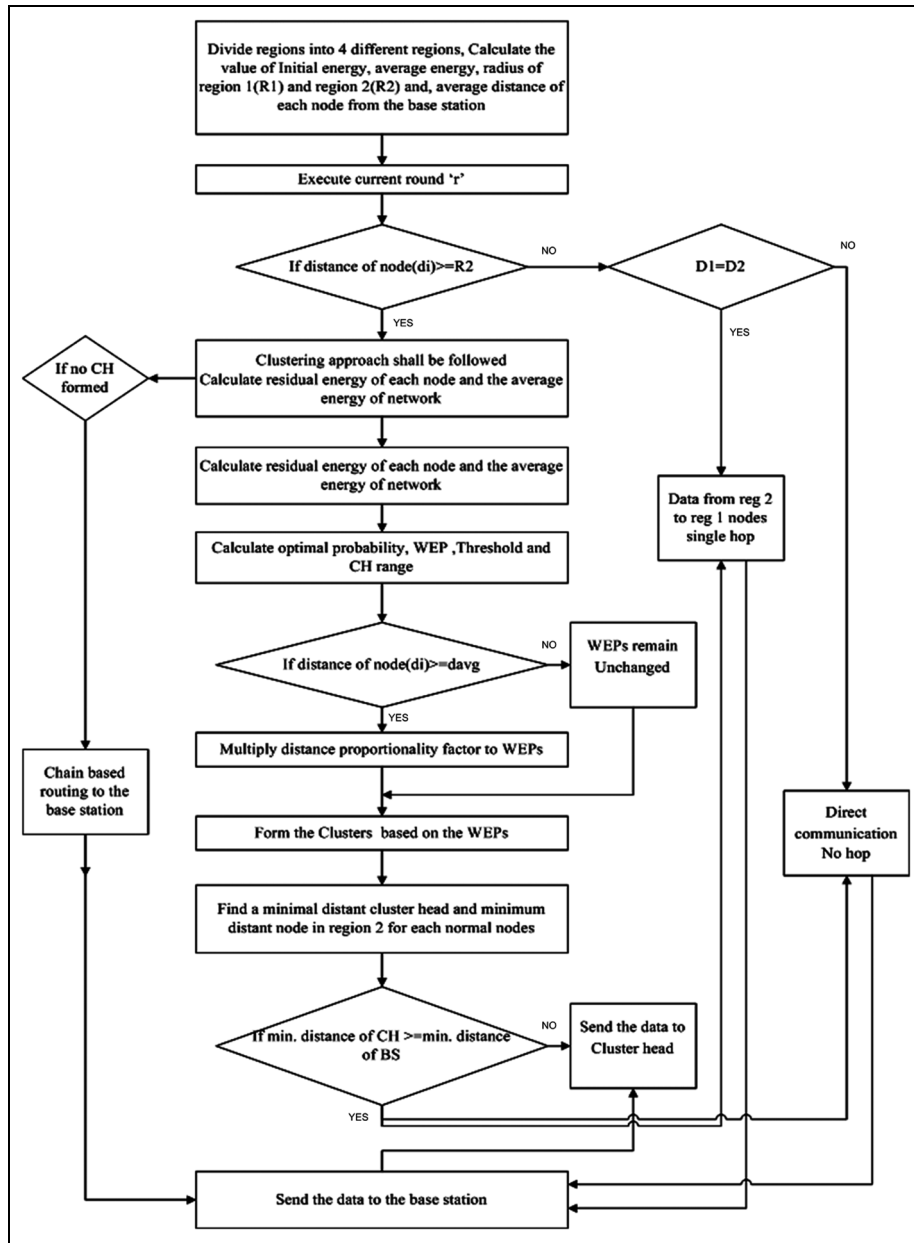


Figure 10. Flow chart for the proposed algorithm.

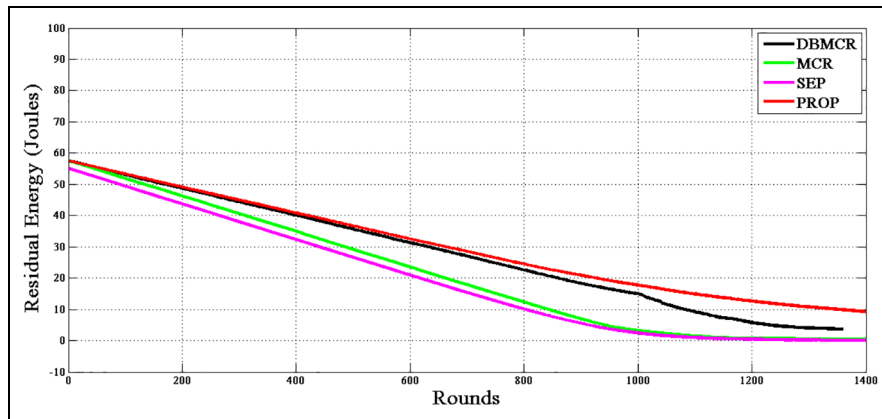


Figure 11. Residual energy of the network for an initial energy 0.5J and BS at the center.

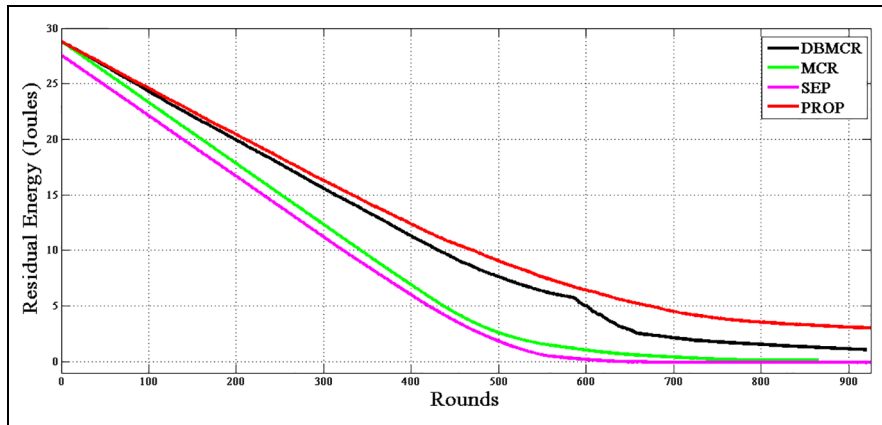


Figure 12. Residual energy of the network for an initial energy 0.25 J and BS at the center.

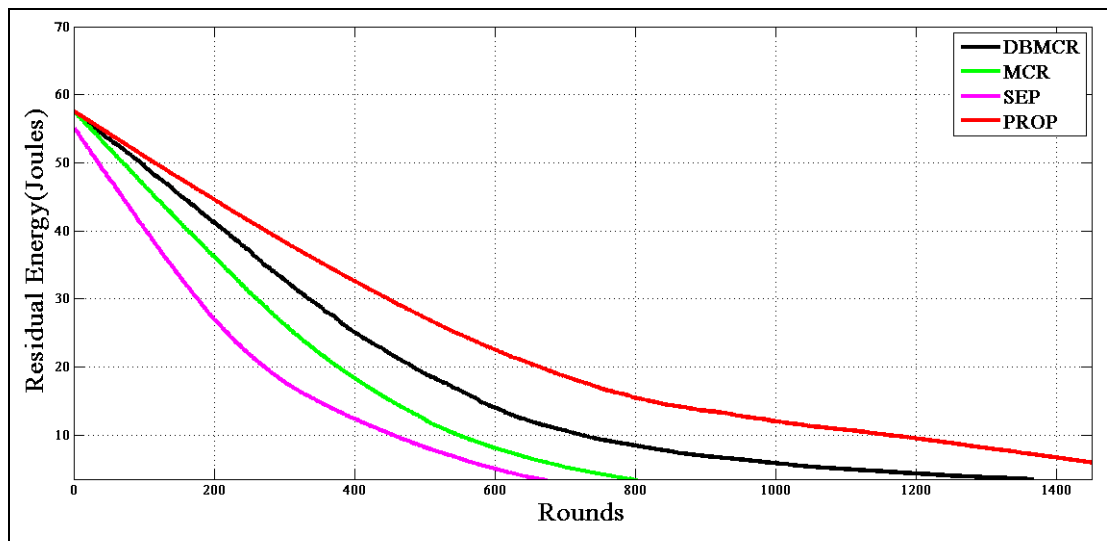


Figure 13. Residual energy of the network for an initial energy 0.5 J and BS at the corner.

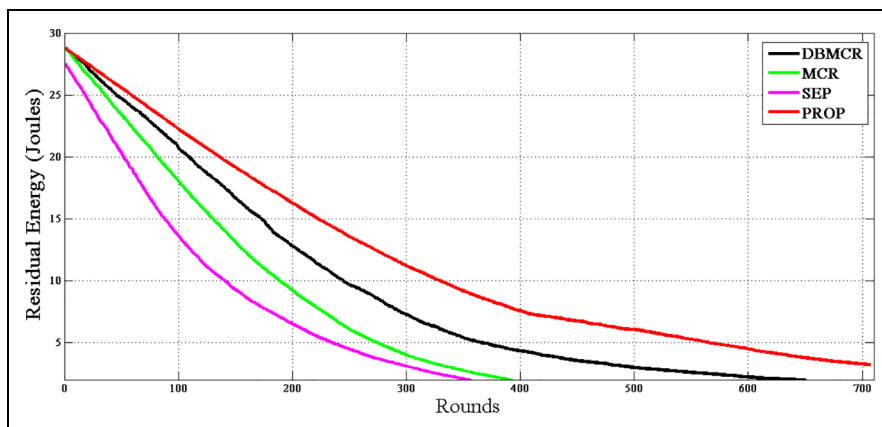


Figure 14. Residual energy of the network for an initial energy 0.25 J and BS at the corner.

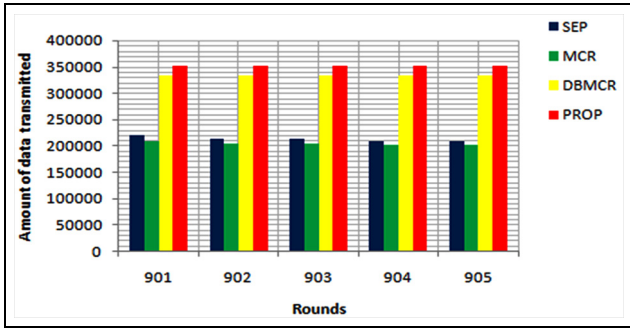


Figure 15. Throughput for an initial energy of 0.5 J and BS at the center.

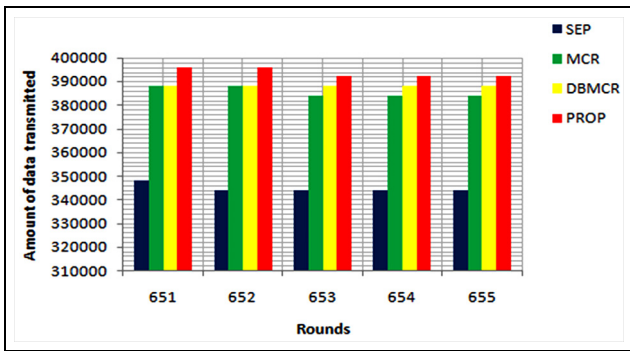


Figure 16. Throughput for an initial energy of 0.25 J and BS at the center.

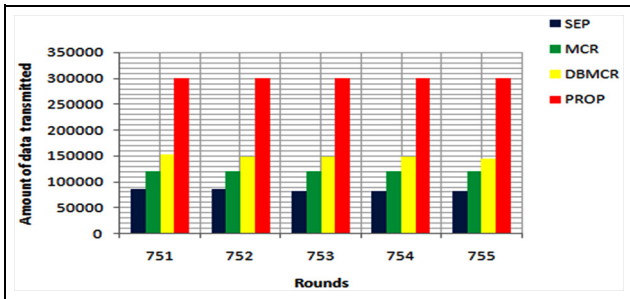


Figure 17. Throughput for an initial energy of 0.5 J and BS at the corner.

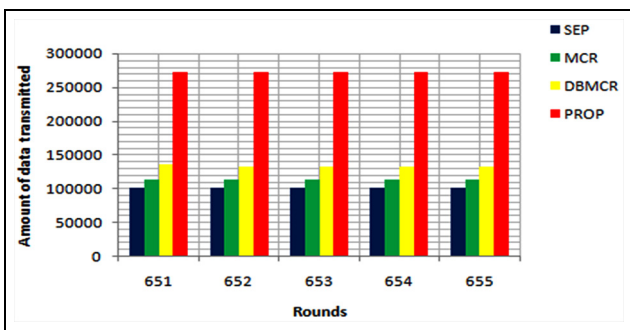


Figure 18. Throughput for an initial energy of 0.25 J and BS at the corner.

Table 3. Analysis of residual energy for proposed algorithm.

Protocol	Location of BS (m ²)	Initial energy (J)	Improvement of proposed algorithm (%)
SEP	(200*200)	0.25	136.67
MCR			121.88
DBMCR			47.91
SEP	(200*200)	0.50	163.16
MCR			114.29
DBMCR			33.92
SEP	(100*100)	0.25	100.00
MCR			93.87
DBMCR			46.15
SEP	(100*100)	0.50	83.91
MCR			43.75
DBMCR			42.86

SEP: Stable Election Protocol; MCR: Multi-hop Clustering and Routing; DBMCR: Distance-Based Multi-hop Clustering and Routing.

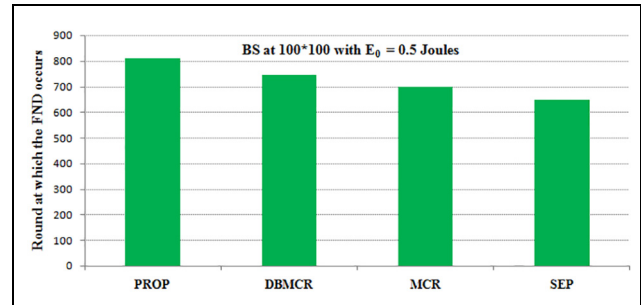


Figure 19. First node death (FND) with initial energy 0.5 J and BS at center.

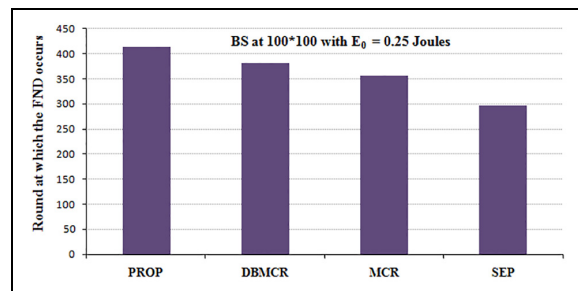


Figure 20. First node death (FND) with initial energy 0.25 J and BS at center.

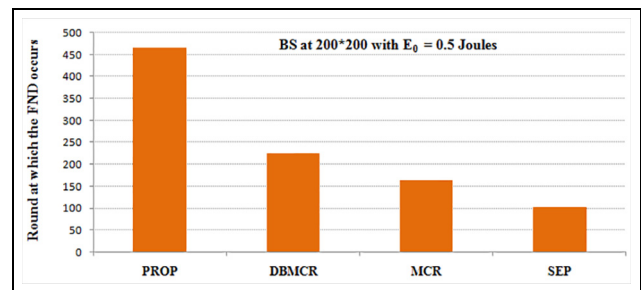


Figure 21. First node death (FND) with initial energy 0.5 J and BS at corner.

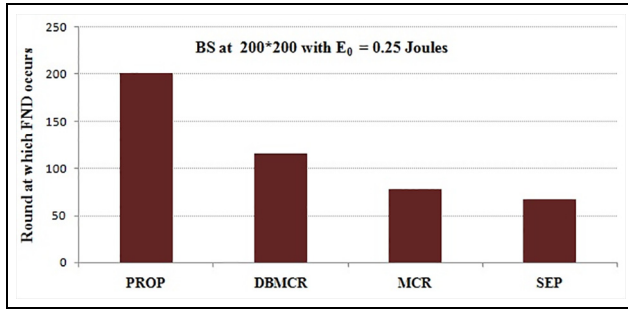


Figure 22. First node death (FND) with initial energy 0.25 J and BS at corner.

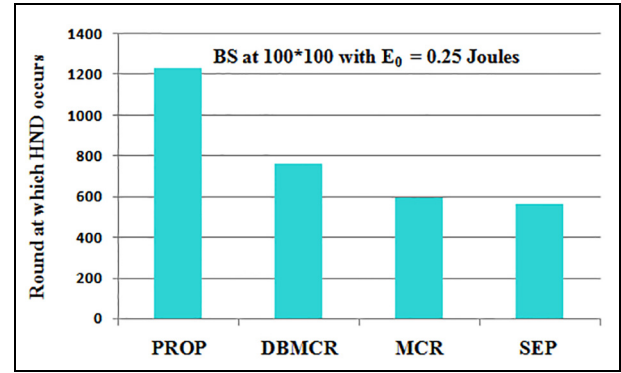


Figure 24. Rounds for 80% node death with an initial energy of 0.25 J and BS at the center.

Table 4. Analysis of stability for proposed algorithm.

Protocol	Location of BS (m ²)	Initial energy (J)	Improvement of proposed algorithm (%)
SEP			25.31
MCR	(200*200)	0.25	22.87
DBMCR			25.4
SEP			56.9
MCR	(200*200)	0.50	32.2
DBMCR			5.5
SEP			111.6
MCR	(100*100)	0.25	96.6
DBMCR			44.5
SEP			202.9
MCR	(100*100)	0.50	180.2
DBMCR			101.16

SEP: Stable Election Protocol; MCR: Multi-hop Clustering and Routing; DBMCR: Distance-Based Multi-hop Clustering and Routing.

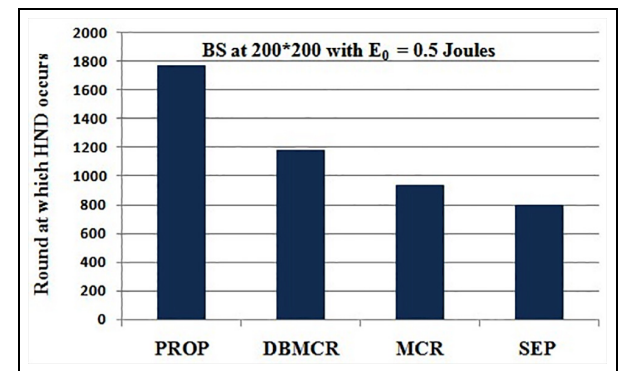


Figure 25. Rounds for 80% node death with an initial energy of 0.5 J and BS at the corner.

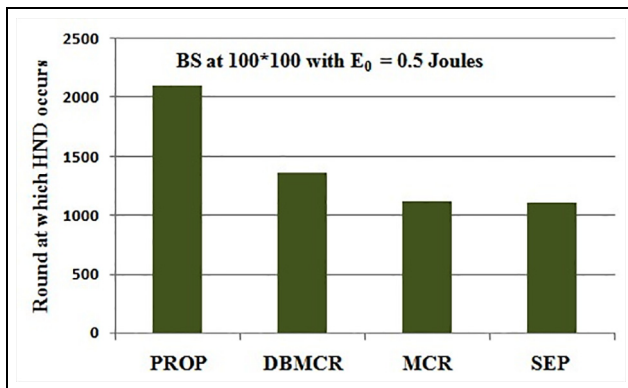


Figure 23. Rounds for 80% node death with an initial energy of 0.5 J and BS at the center.

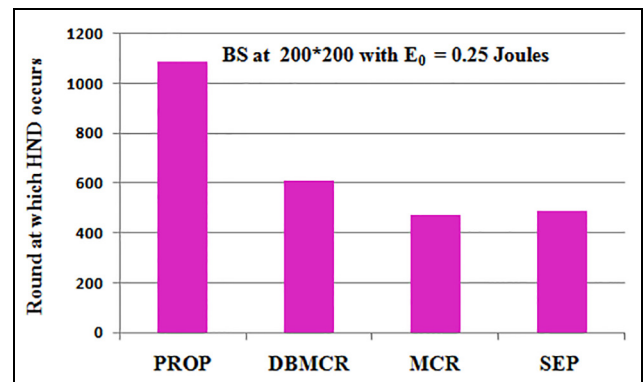


Figure 26. Rounds for 80% node death with an initial energy of 0.25 J and BS at the corner.

analysis of the rounds for which the first node death (FND) happens for the proposed algorithm in comparison with SEP, MCR, and DBMCR algorithms.

Improvement of network stability. In order to assess the overall network performance, the percentage increase

in the stability is given in Table 4. From this analysis, it is found that the proposed algorithm has delayed the FND and has kept the nodes alive for a longer duration compared to others. This delay in the FND signifies that the proposed algorithm has better stability. Dividing the regions and clustering the nodes based on distance, along with the transmission of data through

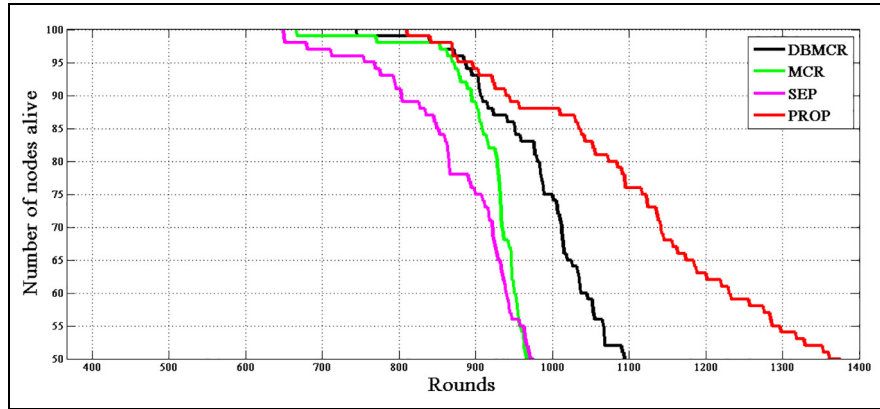


Figure 27. Number of nodes alive for initial energy of 0.5 J and BS at the center.

Table 5. Analysis of FND for proposed algorithm.

Protocol	Location of BS (m ²)	Initial energy (J)	Improvement of proposed algorithm (%)
SEP	(200*200)	0.25	201
MCR			155.7
DBMCR	(200*200)	0.50	74.1
SEP			366
MCR	(200*200)	0.50	184
DBMCR			108
SEP	(100*100)	0.25	38
MCR			28.9
DBMCR	(100*100)	0.50	7.8
SEP			24.76
MCR	(100*100)	0.50	15.6
DBMCR			9.0

SEP: Stable Election Protocol; MCR: Multi-hop Clustering and Routing; DBMCR: Distance-Based Multi-hop Clustering and Routing; FND: first node death.

the allocated spectrum, has greatly avoided the collision of data transmitted and has retained the energy of the nodes. Also, the nodes that are lying outside the average distance circle in the given region will have least probability of becoming the CH. Therefore, by reducing their chances of becoming a CH, the number of rounds before which the first node's death occurs is increased. Hence, these factors have improved the stability of the network as compared to SEP, MCR, and DBMCR algorithms.

Scalability

Scalability of a network is measured by calculating the number of rounds till 80% of nodes die from FND. From Figures 23 to 26, the rounds at which 80% of nodes die are analyzed. By comparing both the simulation results, the FND and 80% of nodes death, it was found that the proposed algorithm was able to prolong

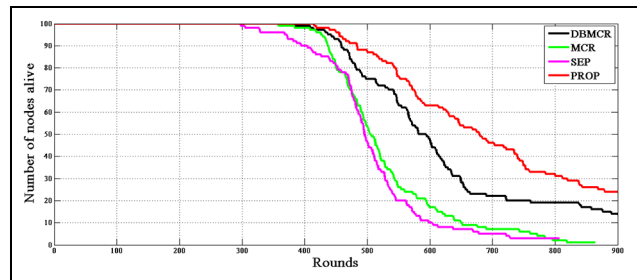


Figure 28. Number of nodes alive for initial energy of 0.25 J and BS at the center.

the death of 80% of nodes for a longer duration than FND. Hence, the scalability of the proposed algorithm is found to be improved as compared to SEP, MCR, and DBMCR algorithms.

Improvement of network scalability. As stated earlier, FND and 80% node death measure the scalability of the network. Table 5 gives the percentage improvement of the proposed algorithm in comparison with SEP, MCR, and DBMCR. Also, with the simulation result of 80% node death in comparison with FND, it is quite evident that the proposed method is able to keep the network alive for more rounds. This proves that the network is highly scalable and has the ability to keep the nodes alive for a longer duration.

Network lifetime

To verify the proposed algorithm on the basis of network lifetime, the simulation results for the number of nodes alive are given in Figures 27–30.

Improvement of network lifetime. The study of Table 6 shows that the round for which the half node death

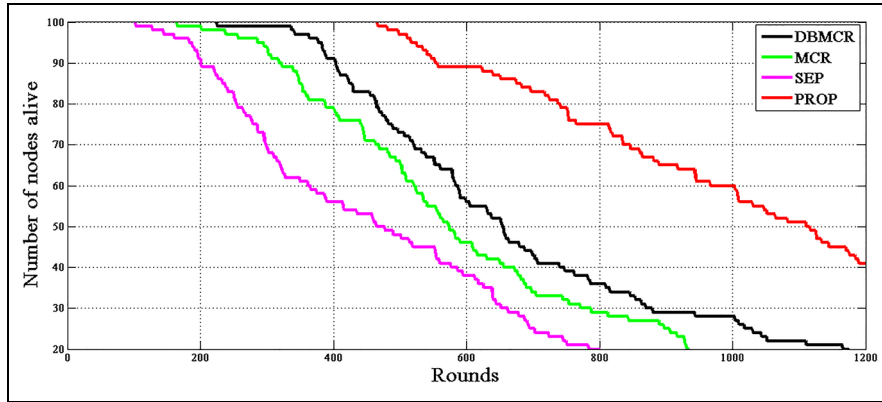


Figure 29. Number of nodes alive for initial energy of 0.5 J and BS at the corner.

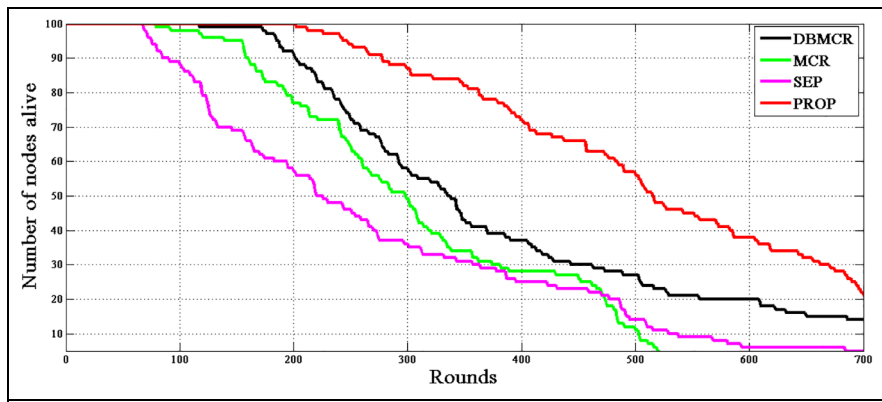


Figure 30. Number of nodes alive for initial energy of 0.25 J and BS at the corner.

Table 6. Analysis of HND for proposed algorithm.

Protocol	Location of BS (m ²)	Initial energy (J)	Improvement of proposed algorithm (%)
SEP	(200*200)	0.25	90.7
MCR			66.13
DBMCR			51.4
SEP	(200*200)	0.50	136.4
MCR			85.17
DBMCR			56.4
SEP	(100*100)	0.25	33.13
MCR			28.11
DBMCR			4.4
SEP	(100*100)	0.50	36.30
MCR			32.50
DBMCR			23.70

SEP: Stable Election Protocol; MCR: Multi-hop Clustering and Routing; DBMCR: Distance-Based Multi-hop Clustering and Routing; HND: half node death.

Table 7. Analysis of network lifetime for proposed algorithm.

Protocol	Location of BS (m ²)	Initial energy (J)	Improvement of proposed algorithm (%)
SEP	(200*200)	0.25	201
MCR			155.7
DBMCR			74.1
SEP	(200*200)	0.50	366
MCR			184
DBMCR			108
SEP	(100*100)	0.25	38
MCR			28.9
DBMCR			7.8
SEP	(100*100)	0.50	24.76
MCR			15.6
DBMCR			9.0

SEP: Stable Election Protocol; MCR: Multi-hop Clustering and Routing; DBMCR: Distance-Based Multi-hop Clustering and Routing.

(HND) occurs is high compared to other algorithms. Since the FND also plays a significant role, its analysis is also considered. It is also evident that after the FND, the energy of all the other members takes a gradual decline in the proposed method. The distribution of

load equally among all nodes and allocation of different spectra have retained the energy and have delayed their FND and HND.

From Table 7, it is observed that the proposed algorithm was able to give a better percentage improvement of the network lifetime as compared to SEP, MCR,

and DBMCR algorithms by a huge extent. The energy-efficient approach of finding an average distance for CH formations and calculating a threshold distance for electing cluster members has greatly helped in reducing the energy used by the nodes for transmitting data to longer distance. Along with this, the division of regions into different parts and assigning the available licensed spectrum has limited the communication of nodes present in those regions to their respective range without any collision.

Conclusion

The proposed algorithm works more efficiently and improves the network stability by a significant proportion with respect to SEP, MCR, and DBMCR. By apportioning the considered area into different regions and parts and by varying the intra-cluster distance range of each CH, the energy dissipation due to data aggregation faced by each CH was reduced significantly. Evidently, using the proposed algorithm, the lifetime of the network has been extended by an average of 74.13% with SEP, 52.9% with MCR, and 33.9% with DBMCR algorithms. The proposed work could be further extended by considering the clustering process with additional parameters such as channel availability and CH selection time along with simulation of the algorithm on an uneven three-dimensional (3D) terrain, which could fetch more realistic results.


Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Venkatasubramanian Srividhya  <https://orcid.org/0000-0001-9911-8671>

References

1. Akyildiz IF, Su W, Sankarasubramanian Y, et al. Wireless sensor networks: a survey. *Comput Netw* 2002; 38(4): 393–422.
2. Akyildiz IF, Su W, Sankarasubramanian Y, et al. A survey on sensor networks. *IEEE Commun Mag* 2002; 40(8): 102–114.
3. Khattab A, Perkins D and Bayoumi M. Cognitive radio networking preliminaries. In: Khattab A, Perkins D and Bayoumi M (eds) *Cognitive radio networks*. New York: Springer, 2013, pp.11–20.
4. Khalid L and Anpalagan A. Emerging cognitive radio technology: principles, challenges and opportunities. *Comput Electr Eng* 2010; 36(2): 358–366.
5. Joshi GP, Nam SY and Kim SW. Cognitive radio wireless sensor networks: applications, challenges and research trends. *Sensors* 2013; 13(9): 11196–11228.
6. Akan OB, Karli OB and Ergul O. Cognitive radio sensor networks. *IEEE Network* 2009; 23(4): 34–40.
7. Cook D and Das SK. *Smart environments: technology, protocols and applications*, vol. 43. Hoboken, NJ: John Wiley & Sons, 2004.
8. Vijay G, Bdira EBA and Ibnkahla M. Cognition in wireless sensor networks: a perspective. *IEEE Sens J* 2011; 11(3): 582–592.
9. Heinzelman WR, Chandrakasan A and Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks. In: *Proceedings of the 33rd annual Hawaii international conference on system sciences*, Maui, HI, 7 January 2000, p.10. New York: IEEE.
10. Yau K-LA, Ramli N, Hashim W, et al. Clustering algorithms for cognitive radio networks: a survey. *J Netw Comput Appl* 2014; 45: 79–95.
11. Qing L, Zhu Q and Wang M. Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks. *Comput Commun* 2006; 29(12): 2230–2237.
12. Smaragdakis G, Matta I and Bestavros A. *SEP: a stable election protocol for clustered heterogeneous wireless sensor networks*. Boston, MA: University Computer Science Department, 2004.
13. Saini P and Sharma AK. E-DEEC-enhanced distributed energy efficient clustering scheme for heterogeneous WSN. In: *Proceedings of the 2010 1st international conference on parallel distributed and grid computing (PDGC)*, Solan, India, 28–30 October 2010, pp.205–210. New York: IEEE.
14. Kumar D, Aseri TC and Patel RB. Multi-hop communication routing (MCR) protocol for heterogeneous wireless sensor networks. *Int J Inf Technol Commun Converg* 2011; 1(2): 130–145.
15. Lindsey S and Raghavendra CS. PEGASIS: power-efficient gathering in sensor information systems. In: *Aerospace conference proceedings*, Big Sky, MT, 9–16 March 2002, vol. 3, p.3. New York: IEEE.
16. Tanwar S, Kumar N and Niu JW. EEMHR: energy-efficient multilevel heterogeneous routing protocol for wireless sensor networks. *Int J Commun Syst* 2014; 27(9): 1289–1318.
17. Mehmood A, Khan S, Shams B, et al. Energy-efficient multi-level and distance-aware clustering mechanism for WSNs. *Int J Commun Syst* 2015; 28(5): 972–989.
18. Tyagi S, Tanwar S, Kumar N, et al. Cognitive radio-based clustering for opportunistic shared spectrum access to enhance lifetime of wireless sensor network. *Pervasive Mob Comput* 2015; 22: 90–112.
19. Wang B and Liu KR. Advances in cognitive radio networks: a survey. *IEEE J Sel Top Signa* 2011; 5(1): 5–23.
20. Moler CB. *Numerical computing with MATLAB*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2004.
21. Moore H. *MATLAB for engineers*. Upper Saddle River, NJ: Pearson Prentice Hall, 2007.