

Improving the WSNs Lifetime using Energy Efficient Hierarchical Clustering

Ahad Abbaspor Topraghaleh

Computer Engineering Department, Urmia Branch,
Islamic Azad University, Urmia, Iran
irpaper2016@gmail.com

Abstract

As the number of sensor nodes increases in the Wireless Sensor Network (WSN), their management and utilization become more complicated. Hierarchical clustering can effectively manage large number of the sensors in the WSN by multiple layers of clusters and improves the lifetime of the WSN. This paper presents a hierarchical clustering scheme for large homogenous WSNs. This clustering solution applies K-MEANS clustering algorithm to create two level clustering in WSNs and tries to improve the WSN's lifetime by selecting node which have higher energy than the cluster's average energy. Simulation results indicate that the proposed protocol can improve the WSNs lifetime better than the LEACH and SEP protocols.

Keywords: WSNs, Clustering, K-MEANS, Hierarchical, Lifetime, Energy

1. Introduction

Wireless sensor networks (WSNs) consist of a large number of small, low-power, cheap sensors having limited sensing, data processing, and wireless communication capabilities [1]. Each sensor has the capability to sense data from its sensing environment and to have meaningful information from the entire monitored area, they should operate with each other [2].

In flat WSNs all sensors have equal role and functionality and any data are forwarded to the sink using the multi-hop routes. To provide energy efficient data transmission in the WSN, many flat routing protocols are presented in the literature. However, as the number of sensors in the WSN increases, their management and utilization become more complicated [3].

Clustering is one of the techniques which provide scalability and robustness for the WSN and provides spatial reuse of the bandwidth, simpler routing, and decrease in energy consumption of the WSN by reducing the sensors take part in data forwarding.

In addition, the energy consumed to manage a large number of nodes will be very high, which involves a large number of election algorithms.

In hierarchical networks, sensor nodes are organized into the cluster heads and cluster members. The cluster head or CHs try to manage the lower level of nodes or cluster members and collect data from them. Each CH collects data from its cluster members, aggregates the data, and then transmits the aggregated data to the sink. All of the hierarchical clustering schemes aim at selecting the best CH and clustering the nodes into the clusters to save more battery power of sensor nodes. Since the CHs are responsible for the collecting, aggregating, and transmitting data to the sink, they consume more energy compared to the other cluster members. The hierarchical clustering protocol may execute reclustering and reselecting of CHs periodically in order to distribute the load uniformly among the whole network. Although hierarchical protocols have their native weaknesses such as requiring time synchronization, potential producing non-optimal routing, and utilizing higher overhead for cluster management, they reveal attractive advantages by

dealing with the constraints in WSNs. Compared with flat protocols, hierarchical protocols offer a more feasible solution to handle large-scale networks with their enhancements to better share limited wireless channel bandwidth, balancing node energy consumption and reduce communication expense.

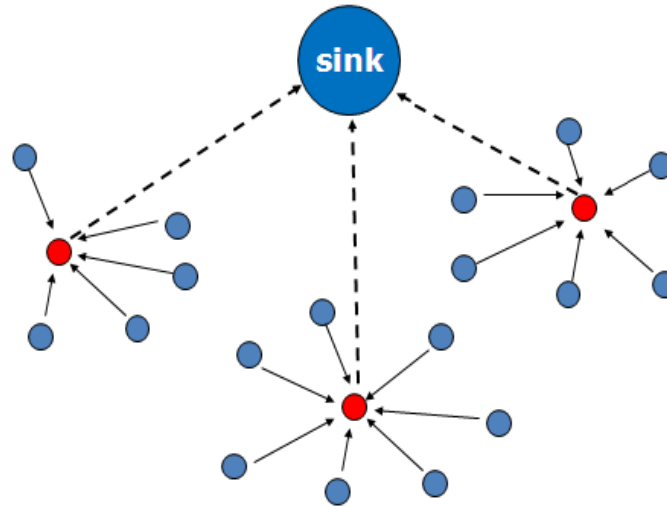


Figure 1. Clustering in WSN

When the network is partitioned into clusters data transmission can be classified into two stages, *i.e.*, intracluster communication and inter-cluster communication. Non-cluster-head nodes first send their data to the cluster heads (CHs) and CHs send the data to BS. Cluster heads co-operate with each other to forward their data to the base station. The goal is to select cluster heads that minimize transmission costs and energy usage. Cluster based routing has been shown to be more energy efficient and increase the network lifetime through data aggregation. Sensor network used to be designed as hierarchical clustering structure to achieve this goal [4].

As the number of sensor in the WSN increases, their management and utilization become more complicated and new clustering and routing schemes are required to deal with them. Hierarchical and multi-level clustering schemes are able to handle large number of the sensor nodes and can manage them in various levels of the clusters. However, most of these schemes are designed with the static sink assumption and often cannot provide effective clustering unless with incurring high clustering overheads to the WSN.

To mitigate these problems, in this paper, a K-MEANS-based hierarchical clustering scheme is proposed to support large-scale homogenous WSNs supported by a mobile sink. By symmetric distribution of the clusterheads in all area of the WSN, this solution is able to reduce the energy consumption of the network and improve the lifetime of the WSN. Simulation results indicate the effectiveness of our proposed solution in reducing the power consumption of the sensor nodes and increasing the WSN's lifetime.

The paper is organized as follows: Section 2 describes the existing hierarchical clustering scheme proposed in the literature. Section 3 briefly introduce the K-MEANS clustering algorithm and section 4 presents our proposed hierarchical clustering solution. Section 5 provides the extensive simulations results of our solution and finally section 6 presents the concluding remarks and future research directions.

2. Related Works

In [5] Lakhlef presents a new hierarchical partitioning scheme named MCCC which takes into account the size of cliques and clusters and is aimed to minimize the number of

hops between the cluster head and its nodes. In fact, this size of clusters and cliques has a direct effect on the energy consumption and this is motivated by the need to have minimum and maximum size for cliques and clusters.

Clustering with taking into account the size of cliques and clusters avoids a significant number of election algorithms because when the cluster/cliue has a moderate size the cluster head may survive longer to manage its cluster. Within the hierarchical scheme the sensor nodes are organized into levels. In each level, there are normal sensor nodes and cluster heads of sensor nodes [8, 12]. We call the super cluster head the cluster head of last level. The super cluster head manages the nodes via the normal cluster heads. In the hierarchical partitioning methods available in the literature (see the next section), the cluster head can belong to another different geographic area in term of number of hops compared to its nodes. Therefore, the change of messages between the cluster head and its nodes is time-consuming. In our algorithm, we impose on the cluster head to belong to the same region by partitioning the network into cliques at the root level (level 0). Then, we partition the network with a hierarchical scheme into cliques, each clique will have a size between k and $2k$. The parameter k is a positive integer that refers to the size of the clusters and cliques. More precisely, the clusters and cliques will have size according to the parameter k , exactly sizes between k and $2k$. The choice of k depends on the memory capacity of the sensor nodes and the energy capacity of the cluster heads. If the sensor nodes can store a great number of nodes and the cluster heads has enough energy to manage a great number of nodes therefore k can be chosen large. Otherwise, it should be chosen small. To generate these cliques of level 0 we must goes through three steps in which we try to have a maximal cliques having size upper than k . The cluster head will be automatically at one hop to nodes, so every node can communicate directly with every node. In another step the cliques of the level 0 are divided into cliques having sizes between k and $2k$. This partitioning is repeated hierarchically every time that the number of cluster heads obtained is greater than $2k$. The procedure is repeated for the cluster heads, which are in the same clique to form new cliques for the next level.

In [6] authors present a two-level hierarchy called TL-LEACH to realize a protocol that better saves the energy consumption. This scheme build a two-level hierarchy and uses random rotation of local cluster base stations (primary cluster-heads and secondary cluster-heads). This permits to better distribute the energy load among the sensors in the network especially when the density of network is higher. TLLEACH uses localized coordination to enable scalability and robustness.

To find an efficient way to generate clusters, the authors in [3] present the a hierarchical agglomerative clustering (HAC) algorithm which in six-step clustering provides a bottom-up clustering by grouping similar nodes together before selecting the cluster head (CH). With automatic CH rotation and re-scheduling, DHAC avoids re-clustering and achieves uniform energy dissipation through the whole network.

In [7] the authors propose a distributed algorithm for organizing sensors into a hierarchy of clusters with to spread energy usage over all sensors and improve the system lifetime of WSN. This scheme applies the multi-hop technique in the intra-cluster communications in order to save energy and integrate it with an adaptive back off strategy to achieve load balancing among sensor nodes and achieves fairly uniform cluster head distribution across the network.

In [8] authors present a distributed weight-based energy-efficient hierarchical clustering protocol called DWEHC. In this scheme, each node first locates its neighbors and calculates its weight based on its residual energy and distance to its neighbors. Then largest weighted node in a neighborhood become a cluster head and its neighbors join the cluster head hierarchy. The clustering process terminates in $O(1)$ iterations, and does not depend on network topology or size.

In [9], the high-efficient multilevel clustering is abstracted as a root tree which has the performances of the minimal relay set and the maximal weight according to graph theory.

A mathematical model is produced for the clustering virtual backbone and based on it, an algorithm called energy-aware multilevel clustering (EAMC) is proposed which reduces the number of relays used for data transmission by minimizing the amount of the cluster heads in the root tree. Furthermore, the algorithm combining with periodical rotation of cluster-heads, enable the energy load be evenly distributed among all the sensors in the network.

3.K-MEANS Clustering

Figure 2 indicates the steps of the K-MEANS clustering algorithm. Generally, K-MEANS is an unsupervised learning algorithms for data clustering proposed by MacQueen in 1967. It is aimed to produce a certain number of the clusters from data and minimizes the intra-cluster distance:

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (1)$$

In which $\|x_i^{(j)} - c_j\|^2$ is the distance measure between a data point $x_i^{(j)}$ and its cluster centroid.

4. TwoLK-MEANS

This section presents our proposed centralized two level clustering scheme for WSNs. Our scheme is based on the LEACH protocol. However, because LEACH protocol utilizes a probabilistic clustering method, it produces asymmetric clusters which are not energy efficient. The figure 3, shows the clustering produced by the LEACH protocol which are asymmetric and consume a lot of energy for intra-cluster communications.

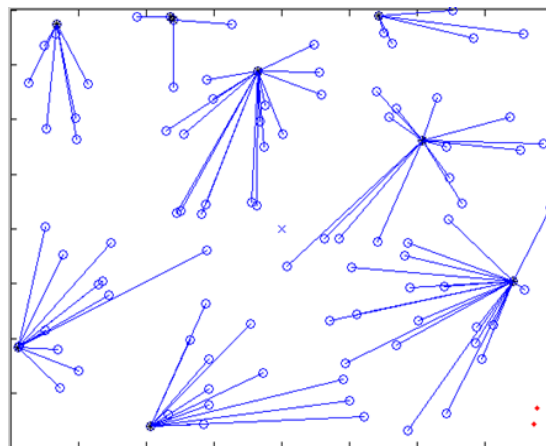


Figure 3. Clustering by LEACH Protocol

To solve this problem in the proposed clustering protocol which is called TwoLK-MEANS is presented which creates symmetric clusters as shown in Figure 4.

Algorithm K-MEANS
Input: information about data
Output: K median

1. Consider K points as initial centroids, in the space used by the data which should be clustered.
2. Assign each data to the nearest centroid.
3. After all data are assigned, recalculate the K centroids.
4. Repeat the steps 2 and 3 until the centroids do not change.

Figure 2. Steps of the K-MEANS Algorithm

This clustering algorithm produces two levels of clusters to reduce the distance of intra-cluster communications. We assume that sensor nodes are distributed randomly and also they have GPS and can inform their location to the base station. When the base station receives the nodes positions, can conduct the proposed clustering algorithm to produce two level clusters. Afterwards, the base station broadcasts the cluster heads list to the network and all sensor nodes can receive this message. After this, the sensor nodes which are selected as cluster head, advertise that they begin to operate as cluster head. Then each sensor node contacts the nearest cluster head to send its data based on the TDMA schedule specified by the cluster head node.

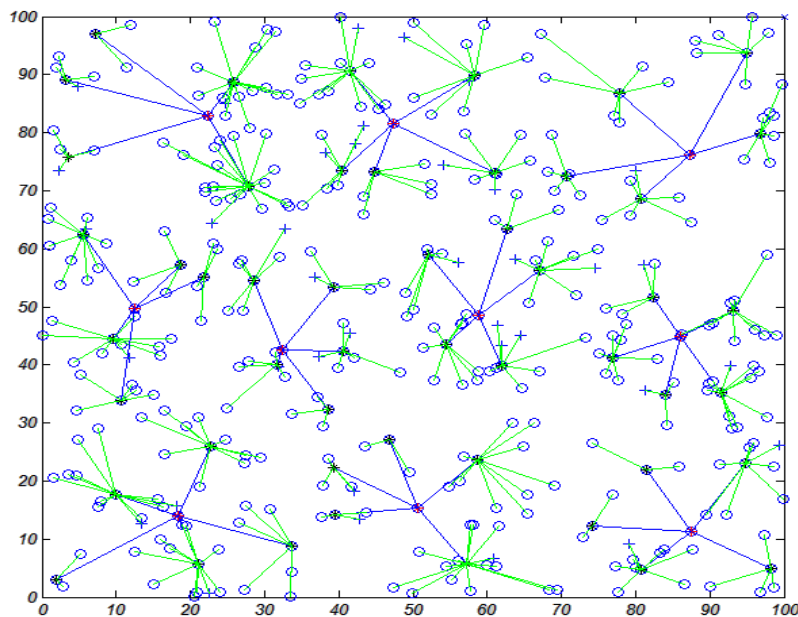


Figure 4. Clustering with the TwoLK-MEANS

1. Sensor send their location information to the base station
2. The base station creates a number of clusters which their numbers are decided based on the WSNs node count.
3. After the center of cluster is founded by using K-means, the closest to the cluster center and most appropriate node which has more energy than CH_Threshold can become the cluster head
4. Then members of each clusters are clustered again using the K-means algorithm.
6. Nodes which are closer to the median and have Energy more than CH_Threshold are selected as cluster head.
7. Sink broadcasts the cluster heads List to the WSN
8. By receiving this list, each cluster head declare itself to its neighboring sensor nodes
9. Like cluster setup phase, sensors join to the closest cluster head
10. Each cluster head broadcasts the TDMA schedule required for its cluster members.
11. Each sensor transmits its data in its time slot to its cluster head

Figure 5. K-MEANS-based Clustering in WSN

Figure 5 exhibits our proposed K-MEANS based two level clustering algorithm for WSNs which contain large number of sensor nodes. Thus we have only modified the Advertisement phase of the LEACH protocol to select better nodes as cluster head.

As exhibited in Figure 4, by using our proposed K-MEANS-based solution, we can create more symmetric two level clusters which reduces the total distance for communication in WSN. Because the energy consumption of sensor nodes is related to the distance which they send their data, the proposed two level clustering approach, reduces the energy consumption of sensor nodes and improves the WSN's lifetime. As shown in figure 6 dead nodes are considered in the clustering process of the proposed algorithm.

In this scheme, the amount of the energy consumption for transmitting a data packet with the size of H bits to the d distance can be computed as follows:

$$E_{tx}(d, H) = \begin{cases} H * (E_{elec} + E_{fs}d^2) & \text{if } d < d_0 \\ H * (E_{elec} + E_{amp}d^4) & \text{if } d \geq d_0 \end{cases} \quad (2)$$

In this equation, E_{elec} is the energy consumption per bit to run the required circuitry. Also, E_{fs} and E_{amp} are the amount of energy per bit dissipated in the RF amplifier according to the distance d_0 which can be obtained as follows:

$$d_0 = \sqrt{\frac{E_{fs}}{E_{amp}}} \quad (3)$$

Also, the amount of energy consumption to receive a packet with H bits can be computed as follows:

$$E_{rx}(H) = H * E_{elec} \tag{4}$$

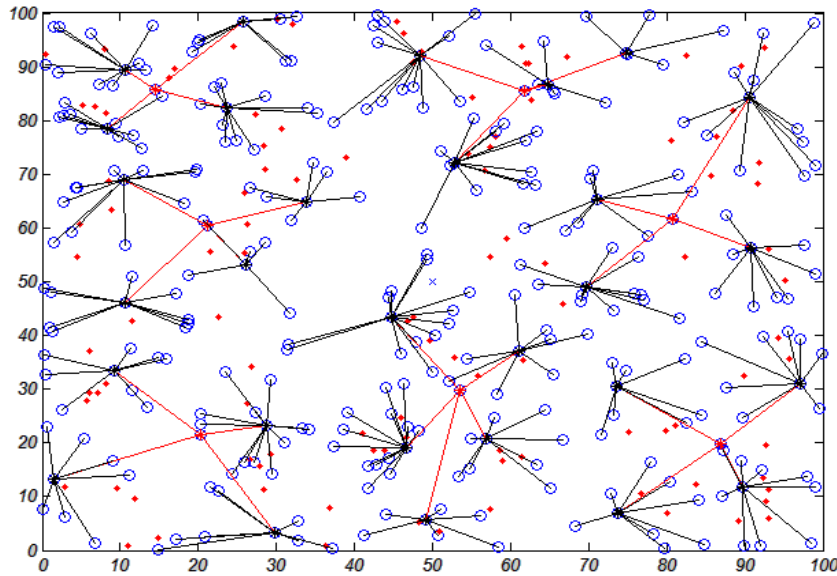


Figure 6. Considering Dead Nodes in Clustering

5. Simulation Results

To indicate the effectiveness of the presented solution, this section tries to provide the extensive simulation results of the TwoLK-MEANS, LEACH and SEP protocols in the MATLAB software. In these three scenarios, mainly the number of dead nodes in each rounds are achieved. Figure 6 shows the dead nodes in the two level clustering process.

We provide our simulations in three different scenarios. At first a scenario with a 200 meter*200meters network area is used and the sink nodes is placed at the location of (200, 200). Moreover, the 200 sensor nodes are distributed in a random fashion in the sensor network and the sensors have 0.1-0.2J battery power. These simulations are performed based on three scenarios which sink node are placed in the different positions. Table 1 provides the important parameters applied in these simulation scenarios to evaluate the presented scheme.

Figure 7 shows the number of dead nodes in each round of SEP and our proposed protocol in the first scenario. By definition, we consider nodes which have zero energy, as dead. The battery of such nodes are used up by various tasks that sensor nodes perform to sense their environment and send the achieved data to the sink. As exhibited by the figure, the proposed clustering approach effectively improves the WSN's lifetime and increases the amount transmitted data by the sensors. Also, as indicated in this figure, our solution improves the First Node Die time at least twice of the SEP protocol. Figure 8 presents number of dead nodes in each round of SEP and our proposed protocol in the second scenario which network nodes have 0.2J energy. As we outlined before, our scheme by transmitting data to cluster heads that are near, consumes less energy and the average energy of network nodes will be higher than the LEACH protocol.

Figure 9 shows the number of dead sensor nodes in each round of SEP, LEACH and our proposed protocol in the second scenario which network nodes have 0.3J battery power.

Table 1. Simulation Parameters	
Parameter	Value
Number of nodes	100 nodes
Network size	200*200
BS location	(100,100)
Data packet size	4000 bit
Initial energy	0.1J and 0.2 J
Aggregation ratio	10%
E_{elect}	50nj/bit
ϵ_{fs}	10pj/bit/m ²
ϵ_{amp}	0.0013pj/bit/m ⁴
E_{DA}	5nj/bit

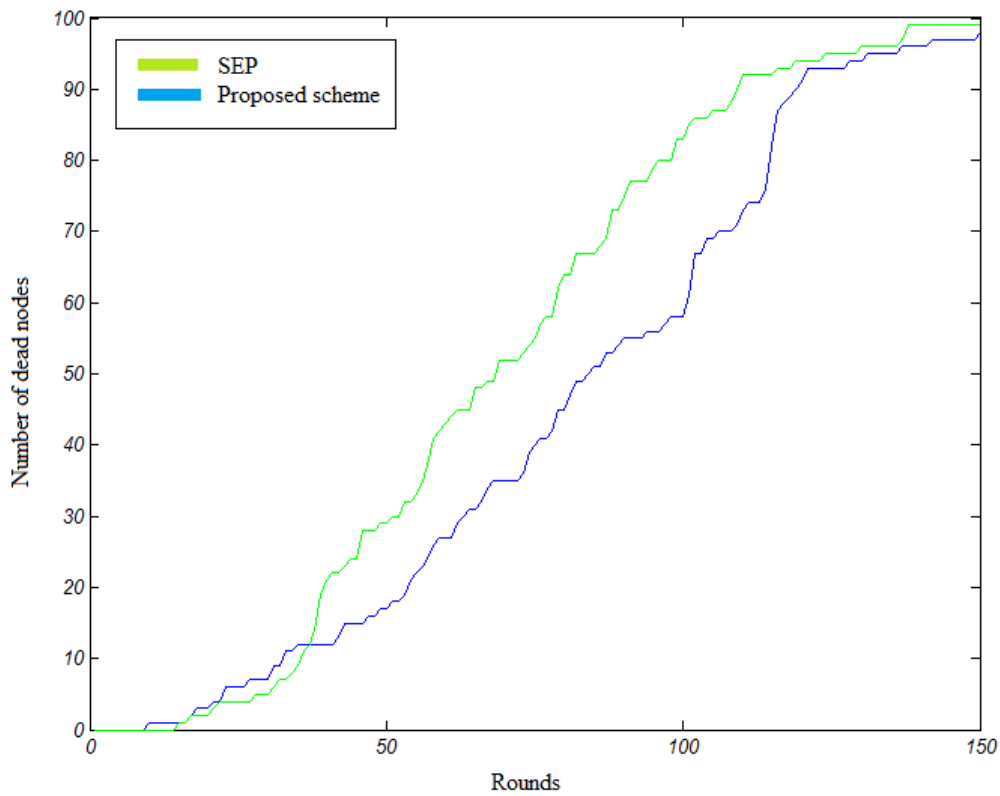


Figure 7. Number of the Dead Nodes in the First Simulation Scenario

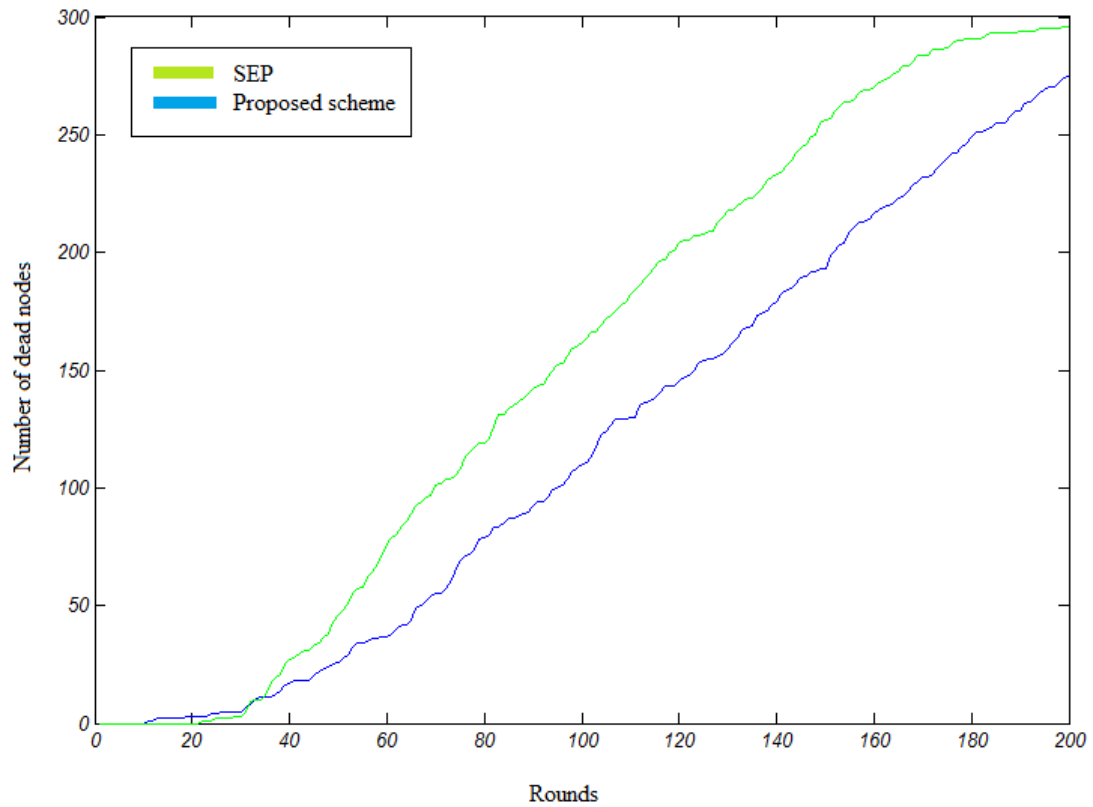


Figure 8. Number of the Dead Nodes in the Second Simulation Scenario

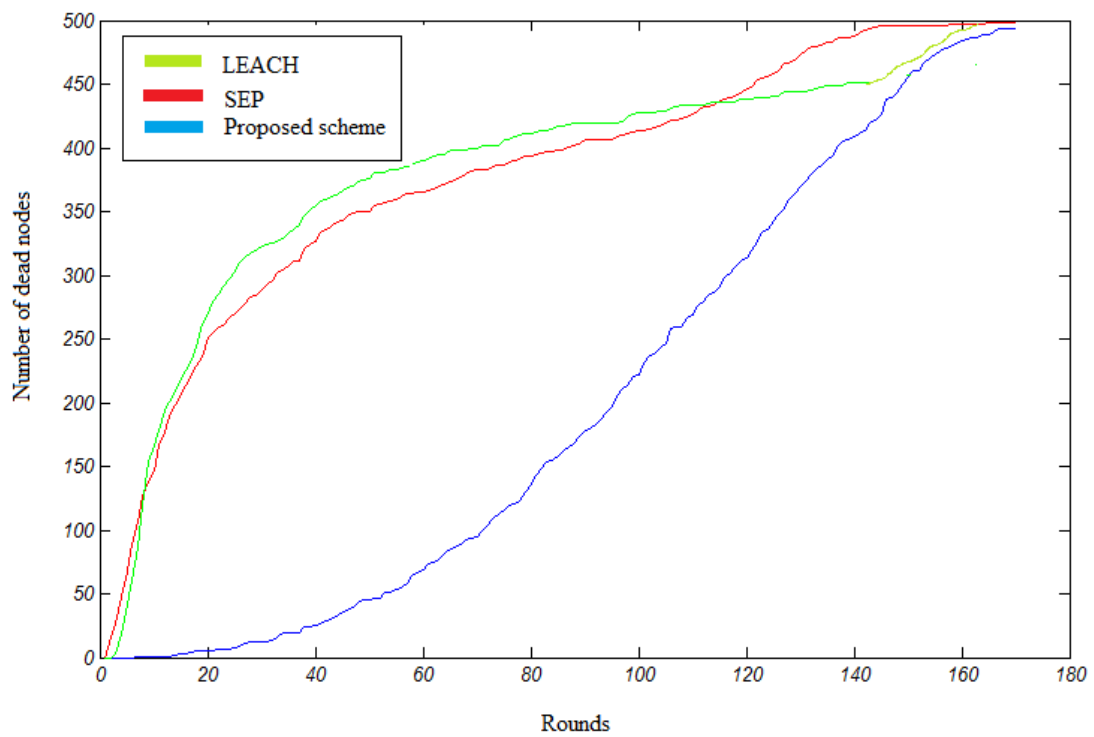


Figure 9. Number of the Dead Nodes in the Third Simulation Scenario

6. Conclusion

Improving the network lifetime is one of the main challenges in the WSNs. Hierarchical clustering is one of the effective solution aimed to organize the large number of WSN's sensors into some clusters and it can increase the lifetime of WSNs, by reducing the distance of intra-cluster and inter-cluster communications.

In this article, a hierarchical clustering scheme is presented for WSN which applies K-MEANS clustering algorithm to create two level nested clusters of sensor nodes. This algorithm is aimed to create symmetric clusters and tries to reduce the total distance required to communicate within each cluster. Extensive simulations conducted in the MATLAB software in various scenarios indicate the effectiveness of the proposed clustering algorithm in mitigating the energy consumption of the WSN and improving its lifetime against the SEP and LEACH algorithms.

References

- [1] H. Zhen, Y. Li and G.-J. Zhang, "Efficient and dynamic clustering scheme for heterogeneous multi-level wireless sensor networks," *Acta Automatica Sinica*, vol. 39, (2013), pp. 454-460.
- [2] S. A. Sert, H. Bagci and A. Yazici, "MOFCA: Multi-objective fuzzy clustering algorithm for wireless sensor networks," *Applied Soft Computing*, vol. 30, (2015), pp. 151-165.
- [3] C.-H. Lung and C. Zhou, "Using hierarchical agglomerative clustering in wireless sensor networks: An energy-efficient and flexible approach", *Ad Hoc Networks*, vol. 8, (2010), pp. 328-344.
- [4] S. Saha and R. Chaki, "Hierarchical Cluster Based Query-Driven Routing Protocol for Wireless Sensor Networks", *Proceedings of the International Conference on Information Systems Design and Intelligent Applications 2012 (INDIA 2012) held in Visakhapatnam, India, (2012) January*, pp. 657-667.
- [5] H. Lakhlef, "A multi-level clustering scheme based on cliques and clusters for wireless sensor networks," *Computers & Electrical Engineering*, vol. 48, (2015), pp. 436-450.
- [6] V. Loscri, G. Morabito and S. Marano, "A two-levels hierarchy for low-energy adaptive clustering hierarchy (TL-LEACH)", *IEEE Vehicular Technology Conference*, (2005), pp. 1809.
- [7] J. Wang, Y.-T. Cao, J.-Y. Xie and S.-F. Chen, "Energy efficient backoff hierarchical clustering algorithms for multi-hop wireless sensor networks", *Journal of Computer Science and Technology*, vol. 26, (2011), pp. 283-291.
- [8] P. Ding, J. Holliday and A. Celik, "Distributed energy-efficient hierarchical clustering for wireless sensor networks", *Distributed computing in sensor systems*, ed: Springer, (2005), pp. 322-339.
- [9] X. Yan, J. Xi, J. F. Chicharo and Y. Yu, "An energy-aware multilevel clustering algorithm for wireless sensor networks", *Intelligent Sensors, Sensor Networks and Information Processing, 2008. ISSNIP 2008. International Conference on*, (2008), pp. 387-392.