

## Towards Energy Efficient k-MEANS Based Clustering Scheme for Wireless Sensor Networks

<sup>1</sup>Moazam Bidaki, <sup>2</sup>Reza Ghaemi, <sup>3</sup>Seyed Reza Kamel Tabbakh

<sup>1</sup>Department of Computer Engineering, Neyshabur Branch, Islamic Azad University, Neyshabur, Iran,  
[Bidaki@iau-neyshabur.ac.ir](mailto:Bidaki@iau-neyshabur.ac.ir)

<sup>2</sup>Department of Computer Engineering, Quchan Branch, Islamic Azad University, Quchan, Iran, [r.ghaemi@iauuq.ac.ir](mailto:r.ghaemi@iauuq.ac.ir)

<sup>3</sup>Department of Software Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran, [rezakamel@computer.org](mailto:rezakamel@computer.org)

### Abstract

Clustering is one of the effective methods which are applied in the wireless sensor networks to preserve the precious battery power of sensor nodes. Various clustering approaches have been proposed in the literature and LEACH is one of clustering and routing schemes which can be used to manage the sensor network. In this paper, we present an improved version of LEACH protocol which uses Kmeans clustering algorithm for selecting better nodes for clusterhead role. Our proposed solution tries to create symmetric clusters and by reducing the average intra-cluster communications distance, improves the network lifetime. It also, considers the sink location when the sink node is positioned far away and selects nodes as clusterhead which will consume less energy for handling cluster members. Thus, because, sensed data are transmitted to the closer clusterheads, our proposed clustering solution can prolong the network lifetime.

**Keywords:** Wireless Sensor network, Clusterhead, Kmeans, Lifetime, Energy.

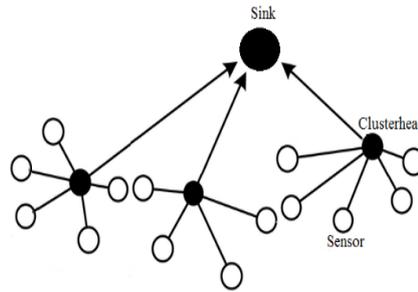
### 1. Introduction

WSN or Wireless Sensor Network is a set of autonomous sensors which are often randomly distributed to monitor temperature, sound and pressure of an environment. WSNs are often deployed in unattended and hostile environments and the sensors cooperatively pass their data through the WSN to a base station or sink node[1, 2]. But, sensors have limited power and processing capabilities, as a result, providing energy efficient solutions to reduce energy consumption is of high importance [3].

One of the techniques that is used for decreasing the energy consumption of sensor networks is the clustering of sensors into disjoint and non-overlapping groups [4-7]. Clustering algorithms based on some metric such as mobility, stability of node or energy level, select some nodes as clusterheads. Each clusterhead should advertise its role and accept its neighbors as cluster members [8, 9]. Afterwards, clusterhead reads its members to forward their sensed data to the sink by consuming less energy [10, 11]. Generally, the clustering solutions presented in the literature have various objectives, for example they may provide load balancing, fault-tolerance, maximize network lifetime or connectivity[12]. Moreover, clustering algorithms provide other advantages like scalability, reducing the routing overheads by data aggregation and mitigating communication bandwidth.

In [13] LEACH or Low-Energy Adaptive Clustering Hierarchy is presented by Heinzelman et al. It considers a dense WSN of homogeneous, sensors which is aimed to

sense and report data to a sink node. LEACH operate in rounds and creates clusters of sensors by randomly elected sensors as clusterheads. Because, the sink node is often far away, the clusterheads must spend significant energy for this transmission. To reduce this problem, LEACH utilizes randomized rotation of cluster-heads to evenly distribute the energy load among the sensors in the network [14, 15]. LEACH uses localized coordination to enable scalability and robustness for dynamic networks, and incorporates data fusion into the routing protocol to reduce the amount of information that must be transmitted to the base station [16, 17]. Therefore, it can decrease the energy consumption and is able to distribute energy dissipation evenly throughout the sensors to increase the network lifetime [18, 19].



**Figure 1. Clustering in Wireless Sensor Networks**

To improve the problems of LEACH protocol, in this paper we present a centralized KMeans-based clustering scheme for WSN. In this scheme, nodes which are approximately positioned in the center of cluster members are selected as clusterheads. This clustering solution reduces the total distance of intra-cluster communications and mitigates the power required for data transmission. Moreover, by considering sink location in clusterhead selection process, it consumes less power and improves the WSN's lifetime.

The rest of this article is structured as follows: Section 2 presents a brief overview of the LEACH protocol and its features, section 3 presents our proposed K-Means based centralized clustering algorithm. Moreover, section 4 provides the evaluation and simulation results of KM-LEACH and the LEACH protocol. Finally section 5 provides the concluding remarks.

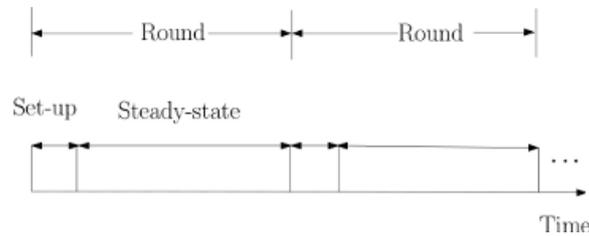
## 2. LEACH Protocol

LEACH is a clustering and routing protocol presented by Heinzelman et al. and various extensions and improvements of this protocol have been presented in the literature. Normally, the LEACH protocol is organized in some rounds and as figure 2 indicates, each round has a set-up phase and a steady-state phase which the set-up phase, itself consists of the following subphases:

- Advertisement subphase.
- Cluster setup subphase.
- Broadcast schedule subphase

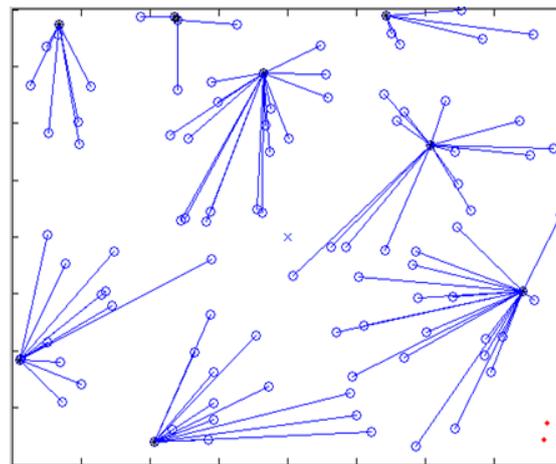
In the Advertisement phase, sensor nodes independently determine to be clusterhead by choosing a random number between 0 and 1. If the produced random number is less than a threshold  $T(n)$ ?, then sensor node becomes a clusterhead only for one round. The applied threshold is computed as follows:

$$T(S) = \left\{ \begin{array}{l} \left( \frac{P}{1 - P \left( R \bmod \frac{1}{P} \right)} \right) \quad \forall n \in G \\ 0 \quad \forall n \in G \end{array} \right\} \quad (1)$$



**Figure 2. Phases of the LEACH Protocol**

After each node determined to be clusterhead in current round, it informs the neighboring nodes by broadcasting an advertisement packet. The neighbor nodes which do not want to be clusterhead, pick the advertisement packet with the strongest Received Signal Strength (RSS) and in the cluster-setup phase, they join to the clusterhead using a CSMA protocol. In the broadcast schedule subphase the clusterhead knows the number of members and constructs a TDMA schedule, picks a CDMA code randomly and broadcasts this information to its members.



**Figure 3. A Sample Clustering by LEACH Protocol**

However, as indicated in the figure 3, the randomly created clusters by LEACH, may not be ideally perfect. For example, multiple clusterhead may be placed near to each other which this reduce the effectiveness of clustering and the cluster members must send to clusterheads positioned far from them. In addition, often clusters produced by LEACH are asymmetric and clusterhead is not located in the center of cluster member nodes. Such clusters increase the total distance between nodes and the clusterhead and this causes more energy to be consumed in the intra-cluster communications. On the other hand, when sensors want to send their data to the clusterheads, should send to clusterheads located far from them. To mitigate this issue, the nodes should be selected as clusterhead

which are positioned near the center of the cluster members. Intra-Cluster Distance or ICD for the  $i$ th cluster can be computed as follows:

$$ICD_i = \sum_{j=1}^{N_i} \sqrt{(x_{Sj} - x_{CHi})^2 + (y_{Sj} - y_{CHi})^2} \quad (2)$$

Where in this equation  $Sj$  is the  $j$ th sensor node and the  $CHi$  is the  $i$ th clusterhead. Moreover, the LEACH protocol does not consider the sink location in clusterhead selection process and finally a node may be selected as clusterhead which consumes more power to communicate with the sink.

In the next section, we present our proposed KMeans-based solution to improve the LEACH protocol and increase its efficiency and network lifetime.

### 3. K-Means Clustering

K-means is an unsupervised learning algorithms for data clustering proposed by MacQueen in 1967. It produce a certain number of clusters from data and tries to minimize the following an objective function:

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (3)$$

Where  $\|x_i^{(j)} - c_j\|^2$  is a chosen distance measure between a data point  $x_i^{(j)}$  and the cluster center  $c_j$ , is an indicator of the distance of the  $n$  data points from their respective cluster centers. Figure 4 indicates the KMeans clustering algorithm.

- . Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
  - . Assign each object to the group that has the closest centroid.
  - . When all objects have been assigned, recalculate the positions of the K centroids.
  - . Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

**Figure 4. KMeans Algorithm**

### 4. KM-LEACH

In this section, we utilize the proposed KMeans-based clustering method to provide an improved version of the LEACH protocol. We assume that sensor nodes are distributed randomly and also they have GPS and can report their positions to the sink. After sink gets the position of all nodes, it can create the required number of clusters by using KMeans clustering algorithm, based on the number of network nodes. Then, it broadcasts the clusterheads list to all sensors of the network. After this, the sensor nodes declared as clusterheads, begin to operate as clusterhead and broadcast their presence on the network.

The rest of our proposed protocol is the same as the LEACH and nodes join to the closest CH and CH broadcasts the TDMA scheduling for nodes.

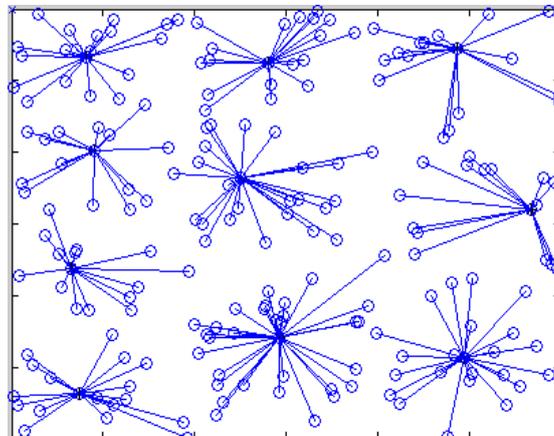
In our proposed solution, we use KMeans algorithm to cluster the WSN nodes. By using clustering algorithm symmetric clusters can be approximately created and this reduces the intra-cluster communication's distance. Figure 5 exhibits our proposed KMeans-based clustering algorithm for WSN.

1. Sensor send their location information to the sink
2. Sink creates a number of clusters based on the number of WSN nodes
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\_EThreshold can become the clusterhead
4. Sink broadcasts the clusterheads List to the WSN
5. By receiving this list, each clusterhead declare itself to its neighboring sensor nodes
6. Like cluster setup phase, sensors join to the closest clusterhead
7. Each clusterhead broadcasts the TDMA schedule required for its cluster members.
8. Each sensor transmits its data in its time slot to its clusterhead

**Figure 5. KMeans-based clustering in WSN**

Thus we have only modified the Advertisement phase of the LEACH protocol to select better nodes as clusterhead.

As exhibited in figure 6, by using our proposed KMeans-based solution, we can create more symmetric clusters and this reduces the total distance for communication in WSN. As a result, our solution reduces the energy consumption of nodes and improves the network lifetime.



**Figure 6. Clustering with our Proposed KMeans-based Solution**

## 5. Simulation Results

This section provides the extensive simulation results of the KM-LEACH, LEACH protocol and direct transmission in the MATLAB software. In these simulation scenarios, we evaluate the following parameters in the LEACH protocol and our proposed solution:

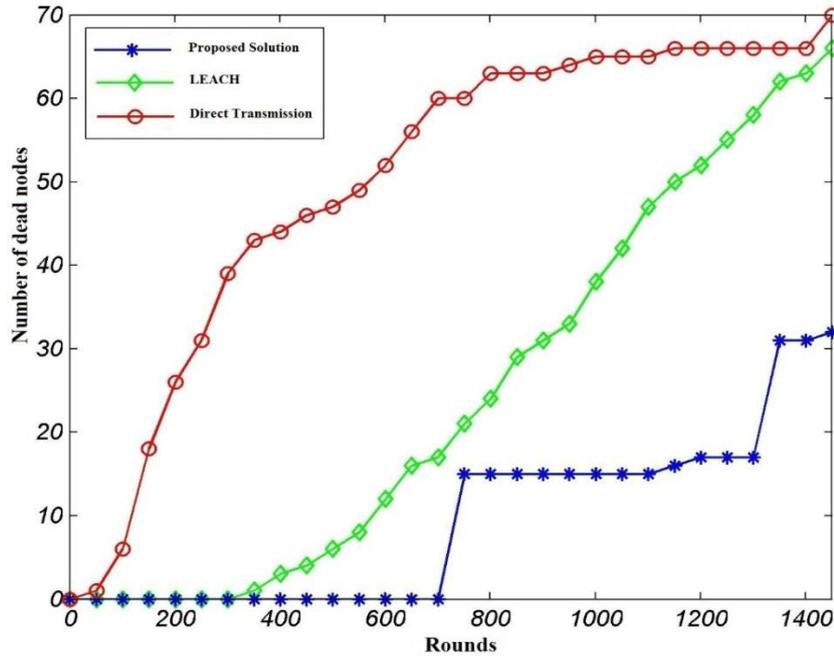
- Number of dead nodes in each round
- Total energy of sensors in each round

We provide our simulations in three different scenarios. In the first scenario, a 200\*200meter simulation area is considered and sink is positioned at the location of (200, 200). Also, 200 sensor nodes are placed randomly in the simulation area which try to send their data to the sink node. These nodes are randomly distributed in the WSN and have 0.4J battery power. These simulation are conducted in three scenarios which different positions are considered for sink node and different number of sensor nodes with different amount of the initial energy is assumed. Table 1 shows the parameters which have been applied in the first simulation scenario.

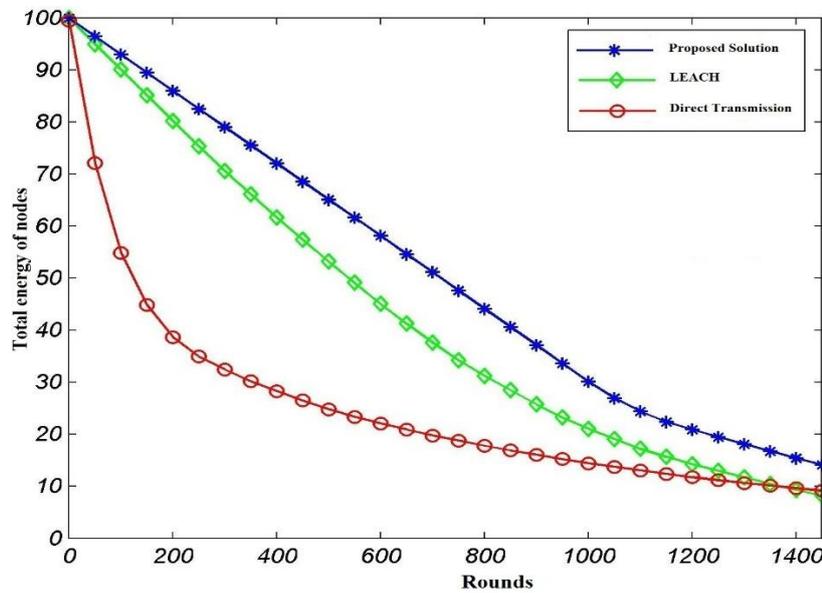
Figure 7 shows the number of dead nodes in each round of LEACH, direct transmission and our proposed protocol in the first scenario. When network nodes consumes their energy and their battery deplete, we consider them as dead nodes. As indicated in this figure, our solution improves the network lifetime and this increase the amount sensed and transmitted by sensors. Also, as indicated in this figure, our solution improves the First Node Die time at least twice of the LEACH protocol. Figure 8 presents the total energy of the sensor nodes and indicates that our solution can decrease the energy consumption of the sensor nodes throughout the simulation which will result in more network lifetime.

As we outlined before, our scheme by transmitting data to clusterheads that are near, consumes less energy and the average energy of network nodes will be higher than the LEACH protocol.

<b>Table 1. Simulation Parameters</b>	
<b>Parameter</b>	<b>Value</b>
Simulation Area	(200×200 )m
Sink location	(200, 200)
Number of nodes	200 nodes
Initial energy	0.4J
P(in LEACH)	0.1
ETX	50×0.000000001J
ERX	50×0.000000001J
Efs	10×0.000000000001J
EDA	5×0.0000000001J



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



**Figure 8. Total Energy of the Sensor Nodes in the First Scenario**

In the second scenario, the initial energy of the sensors are considered 1.2J and 100 sensor nodes are applied in a 300\*300 meter area. Also, sink node is positioned in the point of the (150,300) location. Figure 8 indicates the total energy of the sensor nodes in the WSN. These simulations are conducted in the 1450 rounds.

Figure 9 indicates the total energy of the sensor nodes in each round of LEACH, direct transmission and our proposed protocol in the second scenario. As indicated in this figure, our proposed solution consumes less power for transmitting the data from sensor nodes to the sink and almost in all rounds the average of sensors' energy is maintained better than the LEACH protocol.

Table 2. Simulation Parameters	
Parameter	Value
Simulation Area	(300×300 )m
Sink location	(150×300)
Number of nodes	100 node
P(in LEACH)	0.1
Initial energy	1.2J

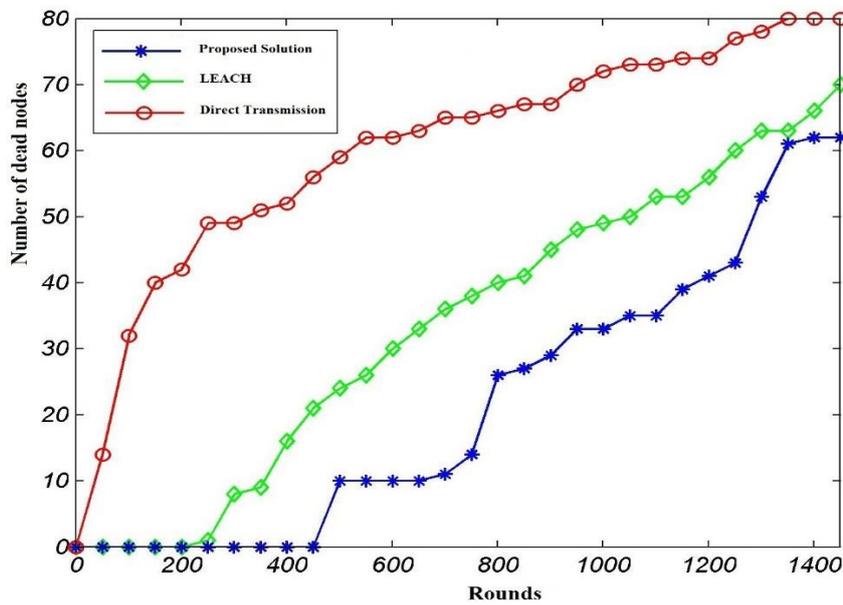


Figure 9. Number of the Dead Nodes in the Second Scenario

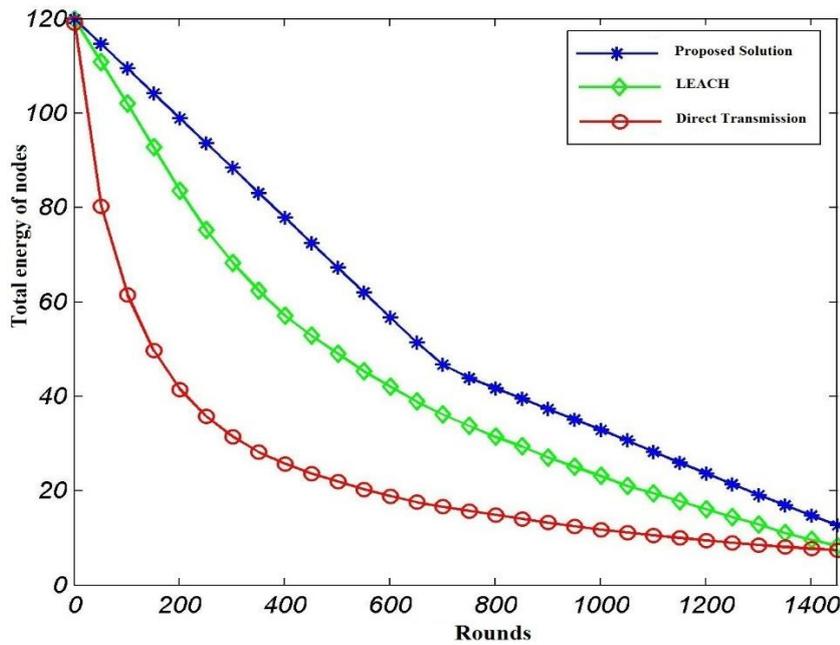


Figure 10. Total Energy of the Sensor Nodes in the Second Scenario

In the third scenario, the initial energy of the sensors are considered 1J and 100 sensor nodes are applied in a 250\*250 meter area. Also, sink node is positioned in the point of the (200, 300) location. The figures 11 and 12 indicate the results of the simulation for the third scenario.

Figure 11 indicates the number of dead nodes in each round in the third scenario. As shown in this figure, our solution improves the network stability time and total lifetime of the WSN. Figure 12 indicates the total energy of the sensor nodes in the WSN. These simulations are conducted in the 1400 rounds.

Table 3. Simulation Parameters	
Parameter	Value
Simulation Area	(250×250 )m
Sink location	(200, 300)
Number of nodes	100 node
P(in LEACH)	0.1
Initial energy	1J

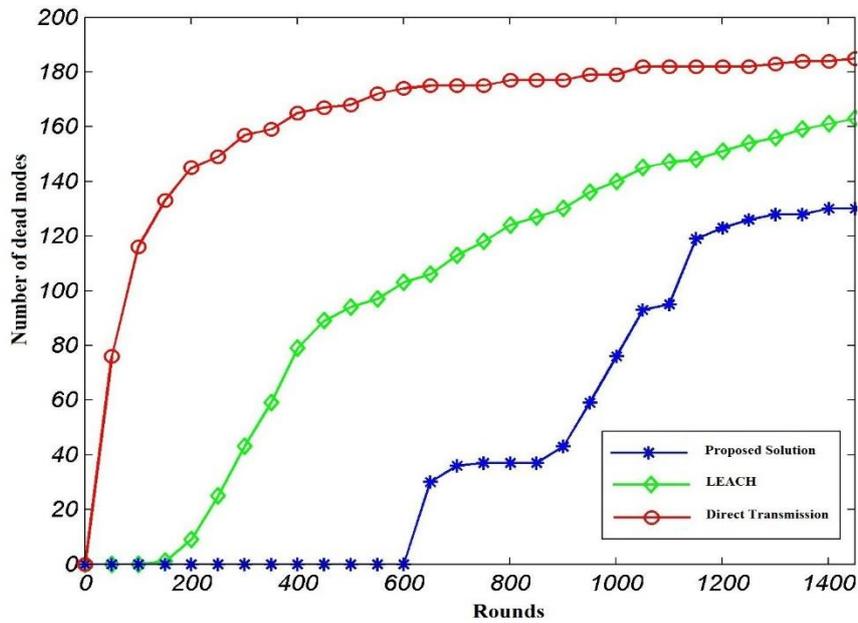


Figure 11. Number of Dead Nodes in the Third Scenario

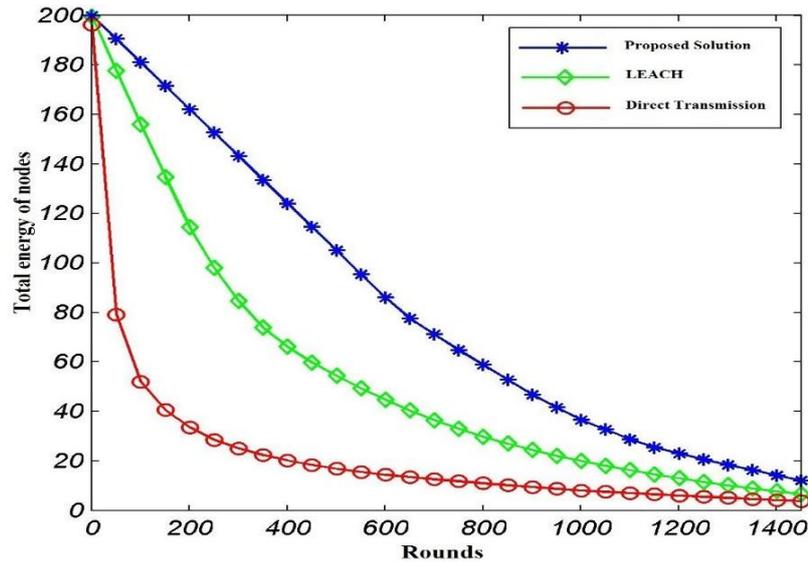


Figure 12. Number of the Dead Nodes in the Third Scenario

Thus as shown in the previous three scenarios, our solution reduces the number of dead nodes and achieves this by reducing the required energy for transmission of sensed data. As outlined before, because of the symmetric clusters which have been produced by our scheme, sensors send the sensed data to closer clusterheads than the LEACH and this mitigates the energy consumption for data transmission.

## 6. Conclusion

Wireless sensor networks consist of many small, inexpensive, resource constrained devices that communicate wirelessly in a multihop network. Because of limited battery of each sensor, improving WSN's lifetime is of high importance and various solutions have been proposed in the literature. Network clustering reduce the energy consumption of the sensor nodes because they transfer their data through one or more clusterhead nodes to the sink. However, selection of the right node for clusterhead task directly affects the lifetime of whole WSN.

In this paper, we presented a KMeans-based clustering algorithm named KM-LEACH which tries to select nodes near the centroid founded by the KMeans as clusterhead and have the required energy to perform the required tasks. The simulation results indicate that our proposed KM-LEACH can effectively reduce the power consumption in the sensor nodes and prolong the WSN's lifetime.

## References

- [1] Vieira, M.A.M., et al. Survey on wireless sensor network devices. in Emerging Technologies and Factory Automation, 2003. Proceedings. ETFA'03. IEEE Conference. 2003. IEEE.
- [2] Yick, J., B. Mukherjee, and D. Ghosal, Wireless sensor network survey. Computer networks, 2008. **52**(12): p. 2292-2330.
- [3] Tyagi, S. and N. Kumar, A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks. Journal of Network and Computer Applications, 2013. **36**(2): p. 623-645.
- [4] Abbasi, A.A. and M. Younis, A survey on clustering algorithms for wireless sensor networks. Computer communications, 2007. **30**(14): p. 2826-2841.
- [5] Katiyar, V., N. Chand, and S. Soni, Clustering algorithms for heterogeneous wireless sensor network: A survey. International Journal of Advanced Networking and Applications, 2011. **2**(4): p. 745-754.
- [6] Kumar, V., S. Jain, and S. Tiwari, Energy efficient clustering algorithms in wireless sensor networks: A survey. IJCSI International Journal of Computer Science Issues, 2011. **8**(5).

- [7] Min, X., et al., Energy efficient clustering algorithm for maximizing lifetime of wireless sensor networks. *AEU-International Journal of Electronics and Communications*, 2010. **64**(4): p. 289-298.
- [8] Masdari, M. and M. Tanabi, Multipath Routing protocols in Wireless Sensor Networks: A Survey and Analysis. *International Journal of future generation communication and networking*, 2013. **6**(6): p. 181-192.
- [9] Masdari, M., S.M. Bazarchi, and M. Bidaki, Analysis of secure LEACH-based clustering protocols in wireless sensor networks. *Journal of Network and Computer Applications*, 2013. **36**(4): p. 1243-1260.
- [10] Boyinbode, O., H. Le, and M. Takizawa, A survey on clustering algorithms for wireless sensor networks. *International Journal of Space-Based and Situated Computing*, 2011. **1**(2-3): p. 130-136.
- [11] Chamam, A. and S. Pierre, A distributed energy-efficient clustering protocol for wireless sensor networks. *Computers & electrical engineering*, 2010. **36**(2): p. 303-312.
- [12] Bidaki, M. and M. Masdari, Reputation-Based Clustering Algorithms in Mobile Ad Hoc Networks. *International Journal of Advanced Science and Technology*, 2013. **54**: p. 1-12.
- [13] Heinzelman, W.R., A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. in *System sciences*, 2000. Proceedings of the 33rd annual Hawaii international conference on. 2000. IEEE.
- [14] Katiyar, V., et al. Improvement in LEACH protocol for large-scale wireless sensor networks. in *Emerging Trends in Electrical and Computer Technology (ICETECT)*, 2011 International Conference on. 2011. IEEE.
- [15] Tong, M. and M. Tang. LEACH-B: an improved LEACH protocol for wireless sensor network. in *Wireless Communications Networking and Mobile Computing (WiCOM)*, 2010 6th International Conference on. 2010. IEEE.
- [16] Abdulsalam, H.M. and L.K. Kamel. W-LEACH: Weighted Low Energy Adaptive Clustering Hierarchy aggregation algorithm for data streams in wireless sensor networks. IEEE.
- [17] Abdulsalam, H.M. and L.K. Kamel. W-LEACH: Weighted Low Energy Adaptive Clustering Hierarchy aggregation algorithm for data streams in wireless sensor networks. in *Data Mining Workshops (ICDMW)*, 2010 IEEE International Conference on. 2010. IEEE.
- [18] Farooq, M.O., A.B. Dogar, and G.A. Shah. MR-LEACH: multi-hop routing with low energy adaptive clustering hierarchy. in *Sensor Technologies and Applications (SENSORCOMM)*, 2010 Fourth International Conference on. 2010. IEEE.
- [19] Tao, L., Z. Qing-Xin, and Z. Luqiao. An improvement for LEACH algorithm in wireless sensor network. in *Industrial Electronics and Applications (ICIEA)*, 2010 the 5th IEEE Conference on. 2010. IEEE.

