

Clustering Algorithm to Reduce Power Consumption in Wireless Sensor Network

Ayat Kazemi¹, Ehsan akhtarkavan²

¹ Department of Computer Engineering, science and research branch, Islamic Azad University, Boroujerd, Iran
Kazemiayat_ka@yahoo.com

² Electronic and Computer Engineering department, Islamic Azad University, Garmsar branch
Akhtarkavan@iau-garmsar.ac.ir

Abstract

The wireless sensor networks have attracted much research attention in recent years and are used in many applications including military environment, health monitoring, environment monitoring and other fields. In these applications sensors with limited energy are employed. These sensors measure conditions such as humidity, temperature and light in the environment surrounding them and then transform these data into electrical signals. There is a question “Which ones of attributes processed can be seen around these sensors?” power consumption are important factors for a sensor network. In this paper a new clustering algorithm is proposed that can be used for minimizing the power consumption and prolong the network life time.

Keywords: *Wireless Sensor Network, Genetic Algorithm, Clustering, K-Means algorithm, lifetime.*

1. Introduction

Nowadays, after design and installation, greater attention is paid to progressive power management, which deals with reduction in energy consumption within a sensor network and has attracted a lot of interest for research studies all over the world. In addition, in recent years, it has been shown that it is possible to utilize powerful intelligent tools such as genetic algorithm in achieving better energy efficiency in remote sensory networks. Many methods have been used in reducing energy consumption in sensory bundles for the general situations in remote sensory networks. In some cases this is done on the bases of which background protocol layer is used in the design stage. For instance, algorithm presented in [1,2], have done comprehensive review studies on energy reduction protocols in MAC layers.. Many methods have been based on reduction in the level communication within layers known as routing protocols. In [3, 4], have also presented a comprehensive study of various routing protocols and their classification. Many of the modern and new algorithms for

routing problem in the wireless Sensor networks have been proposed and various review studies have been conducted in this field. These routing mechanisms, have considered the natural and substantial features of the Sensor networks in addition to the special applications and structure of these networks. Application and maintenance for these networks are not easy and simple issues, because the energy limitations and instant changes in the situations of the nodes (such as their failures), lead to continuous and repetitive and unpredictable changes in the topological structure of the network. In order to minimize the energy consumption, the proposed routing algorithms are using from some of the known and common routing approaches such as data aggregation, and inter-network processing, clustering, allocating different roles to the nodes and data-based methods. Almost all of the routing algorithms can be categorized in to three general types based on the structure of the network: Plain, Hierarchical and Location-based. In the plain routing protocols for the plain networks, usually all of the nodes have the same roles, and in order to function as a sensing operation, they cooperation with each other. Because the number of these nodes is high, then a global ID cannot be defined and allocated for each of the nodes. This feature leads to data-based routing, in which the base station, sends its request to specific locations, and waits for the answers sent from the nodes which are located in those selected locations. Because the data are requested through investigation, then in order to determine the data specifications, the feature-based naming is necessary. The primary works and studies on this data-based routing, the presented algorithm in [5,6], proved that through negotiations and removing the excessive and extra data, energy saving can be achieved. These two protocols were the motivations for designing many other protocols which have the same idea. In location based routing, the sensor nodes are managed according to their locations. The distances between the adjacent nodes are estimated according to the input signals powers. The relative coordination for the adjacent nodes is obtained by data

communicating between the adjacent nodes. The most significant and important protocols of this type are presented in [7,8]. In the hierarchical routing (Cluster-based routing), the nodes with more energy can be used for processing and reporting the data, while the nodes with less energy are used in order to operate as sensors close to the target point. In reality, the hierarchical routing can be very effective in scalability, life and energy performance of the whole system by clustering and allocating specified roles to the head-clusters, in which the single-route structure is avoided. Hierarchical routing is based on the aggregation and merging the data in order to reduce the number of the sent messages to the base stations, and this method in a very effective approach in energy saving. Usually, the hierarchical routing is a double layer routing in which one layer is used for selecting the head clusters and the other layer is used in order to operate routing function. Although, most of the technologies in this type of routing are not related to the routing, but most of them are about who and when one should send, process or aggregate the data, channel allocation and ,etc. Potentially, the clustering routing methods are the most effective approaches in energy saving in the sensor networks and in recent years they have been used widely. The most important and common protocol from this type of routing methods, is LEACH and LEACH-M [9, 10]. Because of the importance of LEACH protocol and its direct relation to the present study, in this article, we have briefly discussed about that in the following.

✓ LEACH Clustering Protocol

Low Energy Adaptive Clustering Hierarchy (LEACH) protocol is the most recent and known protocol based on the clustering method in the wireless sensor networks, in which the clusters are formed in distributed and diffused way. The most important objective of LEACH protocol is having local base stations (Head Clusters) in order to lower the energy consumption resulted from sending data to more distant base station. LEACH protocol, selects few sensor nodes randomly as the head-clusters, and organize the local nodes as the local clusters. Assigning the nodes to the associated head-cluster is done according to their distance. Therefore, the only extra load for them is intra-cluster communications. Head cluster nodes, need more energy in comparison to the other normal nodes. Therefore selecting the fixed head cluster nodes, leads to fast energy consumption and their early death. The energy balance of the head clusters, with a periodic form, assigns the head cluster role to the different nodes. Also, using from aggregation or merging data in head clusters, decreases the amount of the sent data to the base station, and results in energy saving. The function and performance of LEACH protocol is divided to many rounds. Each round starts with the installation phase (Formation of head clusters), in

which the clusters are arranged. After the installation phase, the transformation phase starts, in which the normal nodes send their data to the head clusters and the head clusters after aggregation or merging the data, send the aggregated data to the base station, in order to decrease the amount of the sent data to the base station. Timing for sending the sensor data in the LEACH protocol is done with the Code Division Multiple Access (CDMA) or Time Division Multiple Access (TDMA). Selecting the head clusters is done through a probability function. Each node, will select a random number between 0 and 1, and if the selected value is less than $T(n)$, then the node will be selected as the current Head Cluster. $T(n)$ is calculated by Eq.(1)

$$T(n) = \begin{cases} \frac{P}{1 - P \left(r \bmod \frac{1}{P} \right)} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

In which , P is the probability of being selected as the Head Cluster, r is the current round number, and G is the set of the nodes which in the current round have not been selected as the Head Cluster. Based on the simulation model, it is proved that only 5 % of the nodes need to be selected as the head cluster. The main advantage of LEACH protocol is the cycling mechanism for the head cluster role and the data aggregation, which enables this protocol to increase the life of the network, but is has some disadvantages too:

First, it assumes that all of the network's nodes are capable of sending data to the base station and have enough computational power in order to support from the MAC protocol. Therefore, in the large scale networks, this protocol is not implementable. Also it assumes that the nodes always have data to be sent, and the close node have the related data. This protocol assumes that all of the nodes in all rounds start with the same energy level, providing that the head cluster consumes the amount of energy which is relatively the same as the other node's energy. The most important disadvantage of LEACH protocol is this fact that it is not clear that how the pre-defined number for the head clusters (e.g. P) will be distributed evenly in the network. In reality, there is no guarantee for the location or the number of the head clusters in each round. Therefore, it is possible that the selected head clusters be aggregated and concentrated in some regions of the network. The solution for this problem can be using from a Centralized Clustering Algorithm.

2. Proposed algorithm

Routing algorithm based on clustering is one of the most important ways to reduce energy consumption in wireless sensor networks. In this part, a novel clustering algorithm is presented. Proposed algorithm is decentralized, meaning

that the formation of clusters and assign the roles to the nodes, are performed by the base station. Base station is considered processed and has no restrictions on energy resources. LEACH algorithm is divided into several rounds of the algorithm. With each installation phase (clusters) begins to occur in clusters organized. Follow each step of the installation process begins during which data is transferred to the cluster heads and normal nodes are added. Each cluster heads, the compound or composition data received from member nodes and the base station sends the data to the template pack. The base station is responsible for clustering nodes in the network and assigning appropriate roles to them. After determining current round, the base station includes a message ID to each node sends each cluster. If a node ID cluster heads with its ID matches the node is a cluster heads, otherwise it is a normal node. The base station per cluster, Timeshare table for multiple accesses (TDMA) and the table to create the effect of cluster heads. Thus, the energy cost for the formation of clusters is required only for the proposed base station and no control packet is not sent by the network nodes. The other algorithm's assume that the base station location and energy level of the nodes in the network has sufficient knowledge (for example, it is assumed that each of the nodes must contain GPS). Another important algorithm assumes that the network nodes are distributed randomly in the network space. All nodes have processing power, communication and energy reserves equal (at the start of the algorithm). The proposed algorithm has following steps:

2.1 Clustering by k-Means algorithm

K-means algorithm divides the data set into k subsets (clusters) so that all members of each subset have minimum distance from their corresponding subset. K-means randomly takes k objects (input instance) as cluster centers. The algorithm then assigns other objects (input instances) to suitable clusters based on the minimum distance from specified cluster centers. After that the mean of each cluster is calculated again and is regarded as new cluster center. This process keeps lasting insofar as there appears no further change in cluster centers. The following equation (Eq. (2)) is minimized in k-means.

$$E_{K-means} = \frac{1}{C} \sum_{k=1}^C \sum_{x \in Q_k} \|x - C_k\|^2 \quad (2)$$

In the above equation, C, Q_k and C_k k are the numbers of clusters, K^{th} cluster and center of cluster, receptively. Also, the following relation is used to determine optimum

value for C that was called Davies-Bouldin criterion that is calculated by Eq. (3)

$$I_{DB} = \frac{1}{C} \sum_{k=1}^C \max_{l \neq k} \left\{ \frac{S_c(Q_k) + S_c(Q_l)}{d_{cl}(Q_k, Q_l)} \right\} \quad (3)$$

Where:

$$S_c(Q_k) = \frac{\sum_i \|x_i - c_k\|^2}{|Q_k|} \quad (4)$$

$$d_{cl}(Q_k, Q_l) = \|c_k - c_l\|^2 \quad (5)$$

In which, C is the number of clusters, S_c is the intra-cluster sparsity and d_{cl} is the distance between center of two clusters. As a result the number of clusters which minimizes "Davies-Bouldin" index is considered as optimum number of clusters. At the end of this step, base station knows the optimum number of clusters and member of each it.

2.2 cluster head selection phase

We are used genetic algorithm with new fitness function to select cluster head in each cluster. The proposed fitness function must be calculated for all nodes belong to each cluster. The node with a minimum fitness function shall be selected as cluster head of current cycle. The main criteria for fitness function that applied in this proposed method are as follows:

- ✓ Number of round that node is selected as cluster head
- ✓ The sensor with maximum energy level. Since computational load of the cluster head node is more than other nodes, a node with adequate energy should be selected as the cluster head. Otherwise the connections between that cluster nodes and the base station will be cut off.
- ✓ The nearest sensor to the base station
- ✓ The nearest sensor to the center of gravity of a cluster

The fitness function is calculated by Eq. (6)

$$\cos t(i) = a \left(\frac{E_0 - E_i}{E_0} \right) + \beta \left(\frac{D_{toBS}_i - D_{toBS}_{min}}{D_{toBS}_{max} - D_{toBS}_{min}} \right) + \lambda \left(\frac{D_{toGC}_i - D_{toGC}_{min}}{D_{toGC}_{max} - D_{toGC}_{min}} \right) + \varpi \left(\frac{CHfreq_i}{CHfreq_{max}} \right) \quad (6)$$

α , β , λ , and ϖ are the coefficients determined experimentally and based on the degree of importance of the related measure in determination of a cluster at the lowest communication at the product of their summation should be one.

2.3 Data Transmission phase

After the establishment of clusters and selection of cluster heads of each respective cluster, the time comes for dispatching the sensed data to the related cluster heads by ordinary nodes. Cluster heads send the data package to the base station after the application of community functions or combination of data. Then the energy consumption of all nodes is determined. Heinzelman has introduced a model for radio hardware energy consumption that is represented in Fig.1 [8].

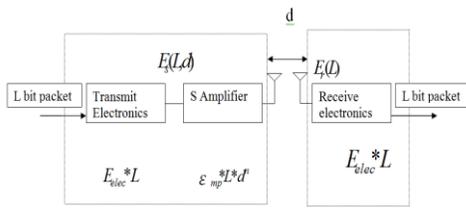


Fig 1. Radio energy consumption model [8]

In this model, each node consumes E_s energy to send L bits of data to d distance from itself. The consumed energy is calculated by Eq.(7) and Eq.(8).

$$E_s = \begin{cases} L \times E_{elec} + L \times \epsilon_{fs} \times d^2 & d < D_0 \\ L \times E_{elec} + L \times \epsilon_{mp} \times d^d & d \geq D_0 \end{cases} \quad (7)$$

Where:

$$D_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}} \quad (8)$$

Where E_{elec} is the energy required for the activation of electronic circuits. The value of E_{elec} depends on the some parameters, such as modulation, digital channel encoding, modulation techniques and spreading of signals.

D_0 is a threshold value when the distance between the sender and receiver is greater than threshold D_0 , the sender can use a multi path fading model, otherwise, the free space model will be used for the channel. Some energy is also consumed by the receiver for receiving L bits:

$$E_r = L \times E_{elec} \quad (9)$$

It is supposed that, in each round, a cluster head receives only one packet from each of the nodes of the related cluster.

3. Simulation and results

The proposed algorithm and LEACH algorithm are simulated by means of Matlab software. The parameters used in this simulation are presented in Tab.(1)

Table. 1 Parameters of simulation

Parameter	
Number of nodes	100
Area	100*100
Location of B.S	(100,100)
Initial Energy	1J
E_{elec}	50nJ/bit
ϵ_{fs}	10Pj/bit/ m^3
ϵ_{mp}	0.0013Pj/bit/ m^4
EDA	5Nj/bit/Signal
Packet Size	4000 bits

Proposed algorithm has been examined with three criteria. Results show that selection of cluster head node with respect to the maximum energy levels generally show better efficiency compared to the other two criteria (nearest node to the base and nearest node to the center node) which has great difference with the results of these two criteria. Hence the best efficiency in proposed algorithm is compared in a graph with the previous two algorithms namely LEACH-M and LEACH [9, 10]. The comparison has been made using 3 parameters in clustering algorithms of sensor network.

- First Dead Time: the number of rounds that the first node of network falls due loss of energy.
- Half Dead Time: the number of rounds in which half of network node (50 nodes) fall due to loss of energy.
- Last Dead Time: the number of rounds that the last node of network falls due loss of energy.

The results of three algorithm and the results of proposed algorithm with different criteria of cluster head selection is shown in following tables, also the results are obtained using numerical averaging of statistical society (20 runs for each algorithm and with the random data).

Table 2. Results of comparison of three algorithms

Scene N=100	Algorithm	First node dead	Half node dead	Last node dead
	LEACH	112	599	814
	LEACH-M	126	763	938
	Proposed Method (Fitness final)	458	2105	2397

Table 3. Results of proposed algorithm with a different criteria to selection of cluster heads

Scene N=100	Algorithm	First node dead	Half node dead	Last node dead
	Proposed Method (number of cluster head)	132	784	1350
	Proposed Method (distance between cluster head)	78	612	1037
	Proposed Method (internal distance in cluster)	89	698	945
	Proposed Method (Fitness final)	194	1354	1630

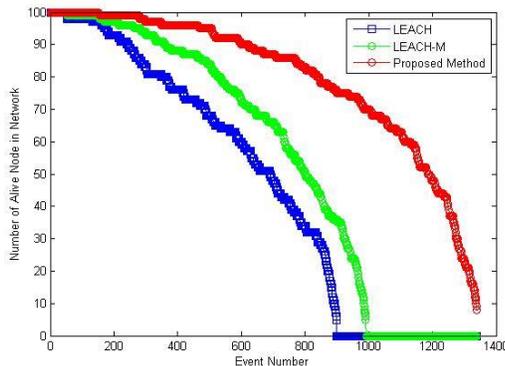


Fig. 2 Comparing the number of alive nodes between Leach, Leach-M and proposed algorithm

Fig.2 shows the number of alive nodes in each round in the network. In Fig.2, the highest curve is related to the new proposed method. This output indicates that the lifetime of the network has increased. The next curves are related to the output of LEACH-M and Leach algorithm, respectively. Fig.3 shows the final energy consumption of the network. The algorithm finishes when all the nodes existing in the network get out of energy. As it can be seen, the proposed algorithm has a higher efficiency than LEACH and LEACH-M algorithm. This matter results in the increase of the lifetime of the network and coverage preservation in the network.

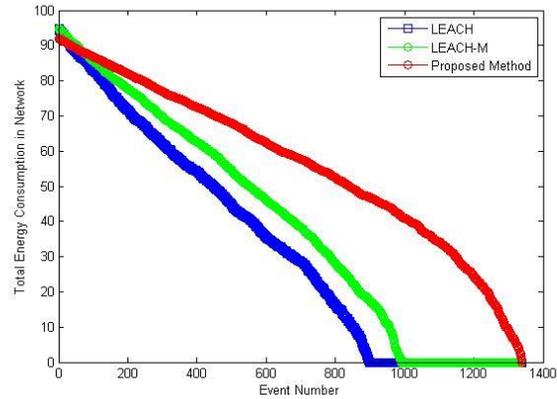


Fig. 3 Comparing the network lifetime between Leach, Leach-M and proposed algorithm

4. Conclusion

Optimal energy consumption and coverage preservation in wireless sensor network is of such a great importance that can lead to the increase of network life time. In this paper, new method is proposed, which using K-Means algorithm to forming the clusters and genetic algorithm to select the cluster head in each cluster. As it can be seen the proposed algorithm compared with LEACH and LEACH-M has a good result in the number of live nodes in each round which leads to the increase of network life time

References

- [1] A. Langendoen K ".Medium Access Control in Wireless Sensor Networks" , In: Book Chapter in "Medium Access Control in Wireless Networks, Volume II: Practice and Standards", Nova Science Publishers,2008.
- [2] A. Demirkol, I. Ersoy, C. Alagoz, "MAC Protocols for Wireless Sensor Networks: a Survey", In: IEEE Communications Magazine, (2006).
- [3] A. Al-karaki J.N, Kamal A.E. "Routing Techniques in Wireless Sensor Networks: A Survey", in: IEEE Wireless Communication, pp.6-28, (2004).
- [4] A. Demirkol I, Ersoy C, Alagoz F. "MAC Protocols for Wireless Sensor Networks: a Survey", In: IEEE Communications Magazine, (2006).
- [5] Heinzelman W, Kulik J, Balakrishnan H. 'Adaptive Protocols for Information Dissemination in Wireless Sensor Networks'. In: Proc. 5th ACM/IEEE Mobicom, Seattle, WA, pp. 174–85, (1999).
- [6] Intanagonwiwat C, Govindan R, Estrin D. "Directed Diffusion: a Scalable and Robust Communication Paradigm for Sensor Networks", In: Proc. ACM Mobi-Com, Boston, MA, pp. 56–67, (2000).
- [7] Xu Y, Heidemann J and Estrin D. 'Geography informed Energy Conservation for Ad-hoc Routing', In: Proc. 7th Annual ACM/IEEE Int'l. Conf. Mobile Comp.and Net, pp. 70–84, (2001).

- [8] Yu Y, Estrin D, Govindan R. 'Geographical and Energy-Aware Routing: A Recursive Data Dissemination Protocol for Wireless Sensor Networks', In: UCLA Comp. Sci. Dept. tech. rep., UCLA-CSD TR-010023, (2001).
- [9] W.Heizelman; A.chandrakasan; H.balakrishnan. An Application-specific protocol Architecture for wirelees Microsensor Networks. IEEE Transaction on wireless communication , vol1, no4, pp:660-670, 2002.
- [10] Q.Liao, H.Zhu,"An Energy Balanced Clustering Algorithm Based on Leach Protocol",Prcoeeding of the International Conference on Systems Engineering and Modeling,2013.