

The Application of Meta-Heuristic based Clustering Techniques in Wireless Sensor Networks

Maryam Mirzapour Moshizi, Vahid Khatibi Bardsiri and Elham Heydarabadipour

*Department of Computer Engineering
Bardsir Branch, Islamic Azad University, Bardsir, Iran
Kvahid2@live.utm.my*

Abstract

Wireless sensor networks are normally utilized for monitoring and controlling of specific environments. They are made out of a large number of low-cost sensor nodes which are densely separated in distributed environments. The information collected by sensors should be transmitted to a base station. Recent developments indicate that the interest in applications of wireless sensor networks has increased and expanded on a large scale. One of the most important issues in this type of networks is limited energy. Clustering is a suitable method for increasing network lifetime that protects the limited sensor resources by energy saving. Meta-heuristic algorithms have been widely used for clustering of wireless sensor networks. In most complicated problems, it is required to evaluate a multitude of possible modes, to determine an accurate answer. Meta-heuristic algorithms can provide an answer in acceptable time constraints. They play an effective role in solving such problems by optimizing the energy consumption of the networks that has a significant impact on the network lifetime. In this paper, the application of clustering algorithms in wireless sensor networks has been investigated. The selected clustering methods are based on the meta-heuristic algorithms.

Keywords: *Wireless Sensor Network, clustering, Meta-Heuristic*

1. Introduction

The sensor network is one of the tools for getting environmental information that has attracted the attention of researchers. In these networks, the nodes energy relies on batteries with limited capacity. Due to the use of these types of networks in harsh and inaccessible environments, recharging or replacing the sensor nodes are impossible, so one of the most important issues in wireless sensor networks is energy management. Clustering is one of the effective ways to consume energy efficiently in such networks. More energy efficient leads to prolong the networks lifetime. In most papers, the network lifetime is the time that is spent to die the first or last node of network. The main issue of clustering is selecting the cluster heads (CH) and clusters. Many algorithms have been proposed for clustering of networks [1]. LEACH¹ is one of the famous proposed clustering algorithms in wireless sensor networks. It is a comparison benchmark of several algorithms. Through clustering, the data collected from each cluster nodes before sending to the base station (BS) is locally processed in CH and then in a new package will be sent to BS. Figure 1 shows an example of a wireless sensor network.

¹ Low-Energy Adaptive Clustering Hierarchy

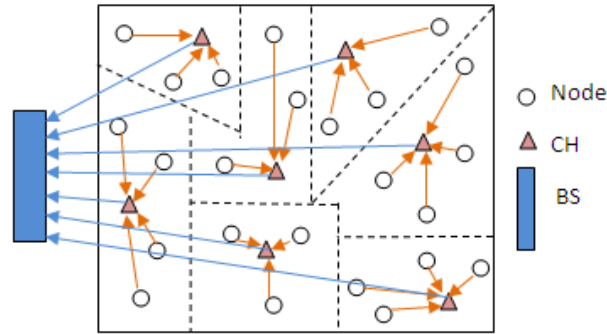


Figure 1. Application of Clustering in Wireless Sensor Networks

2. Comparison Criteria of Clustering Algorithms

- The number of clusters : In many clustering methods, it is determined by the number of CHs so that we can say the number of clusters is clear and definite, otherwise the number will vary
- Node Stability: When the number of clusters varies, the clustering is non-adaptive and stable, otherwise it is stable clustering .Stable clustering means that sensors don't move in clusters, and number of clusters is fixed over the network lifetime [2].
- Dynamics of the nodes: When the nodes are dynamic, the status and number of cluster's nodes in the clusters would change, otherwise, with fixed nodes and fixed clusters only position of CHs may be changed [2].
- Clustering Method: In some algorithms, at first, clusters are formed, and then CHs are selected among nodes, but for the others, at first CHs are selected and then clusters.
- Awareness of the environment: Most of clustering methods are based on information obtained from their environment and neighbors.
- The dynamics of CHs: When the selection of CHs is dynamically performed, the nodes of clusters are constantly changed.
- CHs selection: CHs can be randomly predetermined or selected from the set of nodes or by calculation and evaluation of nodes during clustering.
- Energy Consumption
- Other purposes of the clustering

3. Clustering Algorithms in Wireless Sensor Networks

Sensor networks have a long history and many applications in the real world [3]. Methods for collecting environmental data have already changed. In the past, information was collected by the close nodes connected with each other and BS by wires while nowadays small distributed nodes collect and process data via wireless technology data [4]. This technology includes software, electronic and wireless communication engineering [5]. In general, wireless sensor networks consist of hundreds or thousands of sensors. Clustering is an effective technique to manage the sensors. In this section, we look into a number of clustering algorithms with the mentioned criteria.

1- **GABEEC**²: In this clustering method, genetic algorithm is used. Two-stages of clustering are considered as follows:

- Set up: In this phase, clusters are created and don't change during the cycles. The number of clusters is fixed at each cycle. This step is carried out only once.

- Steady state: In this phase, all nodes can communicate with their own CH. Each node has a TDMA³ table for communicating with CHs. CHs get their information from all nodes and send them as a packet to BS. When CHs send their information to BS, the cycle will be completed. At the end of each cycle, BS reviews all the CHs and their nodes energy. If CH's energy was less than the average energy of its nodes, the node with highest energy will be selected as a new CH and the previous CH comes as a node into the cluster. In this way, the clusters don't change, just CHs may be changed in each cycle. In this algorithm, CHs are displayed with 1 and the other nodes with 0. A network is shown with chromosome which is a collection of bits.

First, genetic algorithm starts with a population that includes pre-defined number of chromosomes. Each chromosome in genetic algorithm is evaluated based on some features and by using a special method. The best chromosomes are selected and two operators of Crossover and Mutation imposed on them so that better population is achieved.

This algorithm was compared with LEACH, HCR, HCR- GA and the result showed increase in the network lifetime and energy saving [6].

2- **ACO-C**⁴: The purpose of this algorithm is to save energy of network by ant colony optimization algorithm which is suitable for CHs. Hypothesis is as follows:

- All the nodes are stable and feel the environment in the same way and always have data to send to the BS. Generally each node can act as a CH.

- The site of BS is fixed and can be close to the network or far from it.

In this paper, two models are considered: 1-empty space 2- multi-way channel model for energy consumption. If the distance between the nodes is less than the threshold (d_0), the free space model and otherwise the multi-way channel model is employed. ACO finds the shortest path between the source of food and ant's nest. Ants communicate with each other and use pheromones to determine the path they must follow.

Here, this algorithm is used for clustering of N sensors in K clusters. For this purpose, R agents call ACO algorithm. Each agent by using pheromones creates solutions called S with the length of K and each element represents a selected node as CH. The cost function is calculated for these solutions and the best solutions with the lowest cost are selected with the name of L. Then a local search is performed on L. The cost function value will be updated for each node, which is the effect of pheromones. This cycle continues and eventually gives the best L which is the best CHs.

This protocol runs over several cycles in BS and each cycle consists of two phases. In the Set-Up phase, the clusters are formed, and all the nodes send energy information and position to BS. In Stead-State phase, BS chooses nodes that have more energy than the average energy of the network as the possible CHs and then runs the ACO algorithm to find the best solution in terms of cost function.

² Genetic Algorithm Based Energy Efficient Clusters

³ Time Division Multiple Access

⁴ Clustering Based on Ant Colony Optimization

After selecting the best solutions, BS introduces CHs by a message to nodes. This algorithm distributes CHs in all over the network and essentially CHs are located in the cluster center.

The algorithm is compared with LEACH, LEACH-C and PSO-C, the result is increasing the network lifetime and energy saving and fast sending [7].

3- **ACO-MRP**⁵ : In this algorithm, a multi-protocol routing based on dynamic clustering and Ants Colony Optimization algorithm is proposed. This approach can maximize the lifetime of network and reduce energy consumption of nodes.

Some measures, such as energy consumption, communication between nodes, residual energy and path length are important criteria in the design of routing in the MRP.

In the first phase, a node is selected as CH and then clusters are formed. Secondly, Ants Colony Optimization algorithm (ACO) is used for routing between CHs and nodes. Finally, the CHs dynamically select paths for data transmission with a probability that depends on the choice of the criteria. This choice is performed with a data transmission evaluation function. MRP can get the maximum lifetime of the network with two ways:

- Reduce data transfer by clusters which causes reduction of energy consumption.
- Using multiple paths for load distribution in the network

In routing, each node has information from its neighbors that is kept in a table related to each cluster. The simulation results indicate that MRP can increase the lifetime of network and reduce the average of energy consumption [8].

4- **ANCH**⁶: The goal of this algorithm is determining the appropriate number and size of clusters to increase the network lifetime. If the number of clusters is low, the large amounts of energy will be used to send data from nodes to CHs. On the other hands, if the number of clusters is too large, the large amounts of energy will be used to exchange data between CH and BS. ANCH is a balance between these two things.

The algorithm is based on the well-known LEACH algorithm. LEACH uses cycles to determine CHs to balance the energy consumption of the network. Each cycle consists of two phases: Set-Up and Steady-State. Clusters are organized in the Set-Up phase and transfer data from nodes to the CHs in Steady-State phase. In this algorithm, for each node is considered a parameter x and in each cycle the value of x is a random number between 0 and 1. A sensor is selected as CH so that the value of x is less than the threshold. In this algorithm, each node can be selected as CH once. ANCH has two steps: ANCH1 and ANCH2.

ANCH1: In this algorithm, it has been tried that the location of CHs be close to the center of the clusters so that the costs of data transfer be equal.

Here the parameter d is considered as closeness. If in a specified cycle distance between two CHs is less than d , two CHs are very close to each other and one of them should be ignored. After selection of the first CH, LEACH process is followed. The CH that is produced at this stage compares its distance with the previous CH. If the distance is less than d , the new CH is ignored and remains as a candidate CH for the next cycle. In this section, for each node a parameter P is considered as a probability to become a CH for a node. This step produces clusters with the same size, but their number is too high.

ANCH2: In this phase, threshold is variable and will increase cycle by cycle. This process reduces the amount of CH and optimizes the distance between CHs and BS. In the last cycle, optimized CHs are selected.

⁵ A Multipath Routing Protocol Based on Ant Colony Optimization

⁶ A New Clustering Algorithm

The algorithm is compared with LEACH algorithm, the result is improving the network lifetime and saving energy [9].

5- **PSOA**⁷ : In this paper, some linear and nonlinear formulas for network routing and clustering are presented and then the PSO algorithm is used for routing and clustering. In clustering based on PSO, the energy consumption of nodes is partially normalized. In this clustering, all areas are intelligently coding and are used for a multi-purpose function.

This algorithm is compared with LDC, GLBCA and CB algorithms and the result is to increase the network lifetime and to reduce CHs energy consumption [10].

6- **AHC**⁸ : This algorithm consists of two methods of hierarchical clustering and K-mean clustering. Its aim is to achieve the optimal number of clusters and to determine CHs for a distribution network in one place. In this method, high-energy nodes are selected as CH. Hierarchical clustering algorithm acts as follows:

It involves grouping two adjacent clusters recursively until the optimal clustering remains. With this method, different clusters are formed at different distances. Clustering method is as follow:

- Each node is considered as a cluster
- Compute the adjacency matrix
- Repeat these steps until only one cluster remains:
 1. Combining two adjacent clusters together on the basis of specified criteria
 2. Updating the adjacency matrix

This method has disadvantages such as its applicability for low data.

K-mean clustering algorithm is as follows:

This algorithm creates K clusters including at least one node. This method optimizes the cluster centers. Clustering method is as follows:

- selecting k nodes as the clusters center
- Repeat these steps until no changes happen to the clusters center or the maximum number of iterations is reached:
 - Setting all nodes which are closer to the center in D Set.
 - Updating the center of each cluster
 - Calculating the main position of the nodes in each cluster.

In terms of computing for large amounts of data, this method is faster than hierarchical algorithms.

Hybrid algorithm uses two above algorithms as follows:

1. Let each object be a cluster
2. Calculating adjacency matrix
3. Computing proximity matrix
4. Computing average distance of nodes to base station
5. Repeat:
 - 5.1. Merging two closest clusters based on distance
 - 5.2. Updating proximity matrix
 - 5.3. Updating number of clustered nodes
6. Computing the frequency of different cluster formed during Merging.
7. Computing the upper limit for number of clusters using formula 1:

⁷ Particle swarm optimization approach

⁸ An Hybrid Clustering

$$\sqrt{\frac{n}{2}} \quad (1)$$

8. Optimal number of cluster is that one which has maximum frequency and which lies within the upper limit
9. Performing k-means clustering on the computed optimal number of clusters
10. Computing the center of each cluster using formula 2:

$$\frac{\sum_{k=1}^n x(k)}{n} \quad (2)$$

11. Declaring the node closest to the center of each cluster as CH

As it is seen, this algorithm optimizes the number of clusters and determines the most appropriate CHs for each cluster [11].

7- **SIC-C⁹** : The proposed algorithm is based on the structure of insect colonies. In this paper, three centralized and decentralized clustering algorithms such as LEACH, PEGASIS and PEDAP are examined and each of three types of algorithm increases the network lifetime.

The proposed method has been compared with these three algorithms, the simulation results show that the proposed algorithm improves the network lifetime and impacts on the network coverage and reduces delay in data collection and saves energy [12].

8- **ANC¹⁰** : Here the clustering issue is dividing N nodes into K clusters. Each set has one leading node M_i ($i=1,2,\dots,K$). Leading nodes are the CHs. In this paper, the network is described as a digraph $G(V,E)$ in two-dimension plane. V is the set of nodes and E is the communication between nodes. E_{ij} means that V_i and V_j are connected. The cluster set is denoted as $C=\{C_1,C_2,\dots,C_{nc}\}$ and Each node V_i belongs to one of the clusters.

- V_i can get its position and its neighbors, based on the information of itself or one of its neighbors. V_j is introduced as a candidate for CH and then investigates whether it would be a cluster member or CH.
- The heuristic information on each node changes if that node is chosen as a cluster head candidate.
- The above steps are repeated several times until some CHs selected.
- R_{single} determines the minimum radius range. If the nodes can communicate with each other directly within R_{single} , they are the neighbor of each other.

The algorithm steps are as follows:

1. The nodes are presented as a graph and the neighbors are identified by R_{single} .
2. Calculating P_{ij} , for each node V_i . (P_{ij} is the probability of selection of V_j as CH by V_i).
3. Each node V_i introduces its CH candidate and all of them are kept as part of the collection of Precluster and then pheromone information of each node change.
4. While number of iterations is less than the maximum, repeat steps 2 and 3. Otherwise, go to step 5.

⁹ Clustering Based on Social Insect Colonies

¹⁰ A Novel Clustering

5. Each node V_i chooses one of its neighboring nodes (V_j) that has the maximum amount of pheromones as CH. If in this step several nodes with equal value exist, the node with the last j will be selected.
6. After this step, the message of setting CHs will be sent to all of its neighbors.

The complexity of this algorithm is $O(n)$. This algorithm has high accuracy and in each cluster there will be only one CH. In this method, by using pheromones parameter, the best node is selected as CH. This algorithm compared with LEACH and GAP is improved. Communication costs are reduced in this way and average residual energy of CH has increased.

Table 1. Comparison and Classification of Algorithms

Purposes	Energy Consumption	CHs selection	Awareness	Clustering Method	Dynamics of CH	Dynamics of the nodes	Cluster stability	Number of clusters	Clustering Model
Increase the network lifetime	Reduction	Calculated	Not have	The formation of clusters	Dynamic	Fixed	Stable	Fixed	GABEEC
Increase the – network lifetime Send data faster	Reduction	Calculated	have	The formation of clusters	Dynamic	Fixed	Stable	Fixed	ACO-C
Increase the network lifetime	Reduction	Selected	have	Select CH	Dynamic	Variable	Unstable	Variable	ACO- MRP
Increase the network lifetime	Reduction	Random	Distance to adjacent nodes	Select CH	Dynamic	Variable	Unstable	Variable	ANCH
Increase the network lifetime	Reduction	Random	have	Select CH	Dynamic	Variable	Unstable	Variable	PSOA
Increase the – network lifetime Optimize the number of clusters	Reduction	Calculated	Distance to adjacent nodes	The formation of clusters	Dynamic	Variable	Unstable	Variable	AHC
Increase the – network lifetime Network Reduce –Coverage delays in data collection	Reduction	Calculated	have	Select CH	Dynamic	Variable	Unstable	Variable	SIC-C
Increase the – network lifetime Reduce the communications cost	Reduction	Select by nodes	Not have	Select CH	Dynamic	Variable	Unstable	Variable	ANC

4. Analysis and Discussion

According to Table 1, in most algorithms, the number of clusters and the number of CHs varies. In clustering algorithms, CHs consume high energy and thus the arrangement of the clusters and CHs change in every period. Changing the arrangement of clusters makes clustering unstable and dynamic. Due to the high energy consumption of CHs, in the fifth column, it can be seen that the choice of CHs for all clustering algorithms is dynamic and fixed nodes are not selected as CHs. According to the sixth column, it is observed that the clustering of different protocols have two main forms , clusters may be formed and then CHs are selected from the cluster nodes or CHs may initially be selected and based on algorithms or the environment information, the clusters are formed. As noted, some nodes have information about neighbors. This property is seen in the seventh column of Table1. In the third part of the paper, the clustering algorithms were briefly described and it was observed that the CHs selection methods are not the same. Some CHs are predefined while some of them are randomly selected and finally some CHs are identified by calculation. In the eighth column of Table1, the selection method in each

clustering algorithm is represented. Finally, it can be seen that all algorithms reduce energy consumption and save energy and increase the network lifetime.

5. Conclusion

Wireless sensor networks include a large number of cheap sensors and are made of limited battery and computational limitations. These sensors collect data and send it to a central unit via wireless communication. The sensors can also communicate with each other. Energy efficiency in wireless sensor networks is very important, saving energy leads to increasing the network lifetime. Clustering protocol for increasing the lifetime of the network not only reduces costs but also saves sensor energy. In prior studies, numerous methods have been proposed for clustering of nodes in wireless sensor networks and meta-heuristic algorithms have also been utilized to increase the efficiency and lifetime of the network along with the clustering methods. In this paper, eight clustering algorithms based on meta-heuristic methods were reviewed such that nine criteria were evaluated for each algorithm. In clustering algorithms, CHs consume high energy and thus the arrangement of clusters in every period seems better. According to Table1, it was confirmed that changing the arrangement of clusters makes the clustering process unstable and dynamic. Due to the high energy consumption of CHs, CHs selecting in all cluster nodes are dynamically carried out and fixed nodes are never selected as CHs. In total, it can be seen that all reviewed algorithms reduce energy consumption and save energy while increase the network lifetime.

References

- [1] O. Boyinbode, H. Le, A. Mbogbo, M. Taldzawa, and R. Poliah, "A survey on clustering algorithms for wireless sensor networks", 13th International Conference in Network Based Information Systems (NBIS), (2010), Takayama .
- [2] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks", computer communications, vol. 30, (2007).
- [3] R. Mittal and M. P. S. Bhatia, "Wireless sensor networks for monitoring the environmental activities", Computational Intelligence and computing research, (2010).
- [4] J. Portilla, A. Otero, V. Rosello, J. Valverde, Y. E. Krasteva, E. de la Torre and T. Riesgo, "13.13. Wireless Sensor Networks From Real World to System Integration – Alternative Hardware Approaches", Comprehensive Materials Processing, vol. 13, (2014).
- [5] S.- H. Yang, "Wireless Sensor Networks", Principles, Design and Applications, (2014).
- [6] S. Bayrakli and S. Z. Erdogan, "Genetic Algorithm Based Energy Efficient Clusters (GABEEC) in Wireless Sensor Networks", Procedia Computer Science, vol. 10, (2012).
- [7] M. Ziyadi, K. Yasami and B. Abolhassani, "Adaptive Clustering for Energy Efficient Wireless Sensor Networks based on AntColony Optimization", Seventh Annual Communication Networks and Services Research Conference, (2009), Moncton, NB.
- [8] J. Yang, M. Xu, W. Zhao and B. Xu, "A Multipath Routing Protocol Based on Clustering and Ant Colony Optimization for Wireless Sensor Networks", Sensors, vol. 10, (2010), Basel, Switzerland.
- [9] M. M. Zanjireh, A. Shahrabi, and H. Larijani, "ANCH: A New Clustering Algorithm for Wireless Sensor Networks", 27th International Conference on Advanced Information Networking and Applications Workshops, (2013), Barcelona.
- [10] P. Kuila and P. K. Jana, "Energy efficient clustering and routing algorithms for wireless sensor networks: Particle swarm optimization approach", Engineering Applications of Artificial Intelligence, vol. 33, (2014).
- [11] G. Kumar, H. Mehra, A. R. Seth, P. Radhakrishnan, N. Hemavathi and S. Sudha, "An Hybrid Clustering Algorithm for Optimal Clusters in Wireless Sensor Networks", Conference on Electrical, Electronics and Computer Science, (2014), Bhopal.

- [12] C.-T. Cheng and C. K. Tse, "A Clustering Algorithm for Wireless Sensor Networks Based on Social Insect Colonies", IEEE SENSORS JOURNAL, vol. 11, no. 3, (2011).
- [13] J. Dai and S. Liu, "A Novel Clustering Algorithm in Wireless Sensor Networks", International Conference on Industrial and Information Systems, (2009), Haikou.
- [14] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks", IEEE Transactions on Wireless Communications, vol. 1, no. 4, (2002).

