

FUZZY-CLUSTERING BASED DATA GATHERING IN WIRELESS SENSOR NETWORK

Arezoo Abasi and Hedieh Sajedi

Department of Mathematics, Statistics and Computer Science, College of Science,
University of Tehran, Tehran, Iran

ABSTRACT

Wireless Sensor Networks (WSN) is spatially distributed, collection of sensor nodes for the purpose of monitoring physical or environmental conditions, such as temperature, sound, pressure, etc. and to cooperatively pass their data through the network to a base station. The critical challenge is to minimize the energy consumption in data gathering and forwarding from sensor nodes to the sink. Cluster based data aggregation is one of the most popular communication protocols in this field. Clustering is an important procedure for extending the network lifetime in wireless sensor networks. Cluster Heads (CH) aggregate data from relevant cluster nodes and send it to the base station. A main challenge in WSNs is to select suitable CHs. Another communication protocol is based on a tree construction. In this protocol, energy consumption is low because there are short paths between the sensors. In this paper, Dynamic Fuzzy Clustering data aggregation is introduced. This approach is based on clustering and minimum spanning tree. The proposed method initially uses fuzzy decision making approach for the selection of CHs. Afterward a minimum spanning tree is constructed based on CHs. CHs are selected efficiently and accurately. The combining clustering and tree structure is reclaiming the advantages of the previous structures. Our method is compared to the well-known data aggregation methods, in terms of energy consumption and the amount of energy residuary in each sensor network lifetime. Our method decreases energy consumption of each node. When the best CHs selected and the minimum spanning tree is formed by the best CHs, the remaining energy of the nodes will be preserved. Node lifetime has an important role in WSN. Using our proposed data aggregation algorithm, survival of the network is improved.

KEYWORDS

Sensor networks; Energy efficiency; Data aggregation; Fuzzy decision making.

1. INTRODUCTION

Wireless sensor network (WSN) is a collection of thousands of low-cost, low power electronically programmable devices, which are deployed in a monitored area in stochastic manner [2]. In this area, there is no opportunity for maintenance and battery replacement for the most of the applications, which use the sensor nodes to surveillance the remote field [1]. Potential applications of sensor networks include Industrial automation, Automated and smart homes, Video surveillance, Traffic monitoring, Medical device monitoring, Monitoring of weather conditions, Air traffic control, Robot control.

The most-distinguishing attributes of nodes used in WSNs are the limited power supply, storage capacity and communication bandwidth required. In WSN, bandwidth utilization and energy saving is a very important criterion for any existing and new applications. Normally, data collected from WSNs are large which makes it essential to eliminate redundant data, minimize the number of transmissions, and improve the energy consumption. The effort to reduce the number of data packet transmission with the in-network processing is called data aggregation [3].

Our sensor's battery is limited. The lifetime on each node depends on the power that has significantly affected the relationship between the nodes. One of the accurate requirements of these nodes is the efficient use of the saved energy. Multiple algorithms have been designed for impressive handling of nodes energy in WSNs using several clustering schemes [4, 5]. Optimal data aggregation can save nodes energy. In this sensor network, data are gathered by the sensor nodes from our study area. There is a data transmission method that merges data from several sensor nodes into one pack, which is data aggregation. Decreasing the disjointed communication at different levels and in turn to reduce the total energy consumption is the main aim of data aggregation.

There are dissipated different amounts of energy to process raw data. There are two popular protocols: Cluster based data aggregation [6] and Tree based data aggregation [7]. Some of WSNs consists of clusters, in which each cluster has a CH. CHs have a significant impress in network lifetime. An ideal CH is the one, which has the highest residual energy, maximum number of neighbour nodes around the CH and the shortest distance from the base station[8]. Whatever the selected CH is more similar to the ideal CH, network lifetime is increased.

We can use Multiple Attribute Decision Making (MADM) approach to select CHs with multi criteria [9]. This method quantitatively selects alternatives based on their multiple criteria. The main problem is the difficult estimation of the exact values of all the criteria. Synchronous consideration of all criteria in CHs selections can be used MADM approach. In case of multi criteria, fuzzy based MADM methodologies are efficient and impressive [10, 11].

In this paper, we proposed a hybrid approach called DFC data aggregation, which gathers and combines data and avoids redundant data transformations, therefore successively reduces power consumption and bandwidth.

Proposing DFC data aggregation, we preserve the advantages and minimize the disadvantages of the clustering and tree based approaches. We use DFC data aggregation to extend the lifetime of WSNs and energy consumption of sensor nodes. The optimized CHs are selected to spread energy efficiently using multi criteria. CHs are selected based on the residual energy, the number of neighbour nodes and distance from the base station. After cluster formation, CHs receive data from member nodes in clusters, aggregate data and send it to the base station. A spanning tree covers all the sides as vertices and consists no cycles. The tree is constructed in the procedure that the node with the smallest identifier is chosen as the root [10, 12]. All the nodes with the shortest path conjunct to the selected root. The protocol requires that each node exchange configuration message in a specific format, which contains its own identifier, its chosen root, and the distance to this selected root. Each node updates its configuration message upon identifying a root with a smaller identifier or the shortest-path neighbour. In addition, the neighbour for which the shortest route configuration message comes from is chosen as the parent of a node whenever it is detected.

In this paper, we employ multi criteria decision-making approach, Fuzzy Analytic Hierarchy Process (FAHP) and hierarchical fuzzy in clusters on WSNs [13, 14]. The Analytic Hierarchy Process (AHP) considers a set of assessment criteria, and a set of alternatives among which the best decision is to be made. The AHP generates a weight for each evaluation criteria according to the comparisons of the criteria. The superior the weight, the more significant the corresponding criterion. The AHP method could improve the network lifetime significantly.

In this research, we also analyse other methods, including Low Energy Adaptive Clustering Hierarchy (LEACH) [14], Tree Dara Aggregation (CTDA) [16], Modified Cluster based and Tree based Data Aggregation (MCTDA) [16], and Cluster based and Tree based Power Efficient Data

Collection and Aggregation (CTPEDCA)[17]. We compared our proposed method with these mentioned methods in terms of energy consumption and the amount of energy remaining in each sensor network lifetime.

Simulation conclusions illustrate that our proposed approach is more efficient than LEACH, CTDA MCTDA and CTPEDCA algorithms considering energy consumption.

The remainder of this paper is organized as follows. Section 2 is about previous work. In Section 3, we describe our system model. The proposed algorithm is explained in Section 4. Section 5 describes the experimental results and discussions and finally Section 6 concludes the paper.

2.PREVIOUS WORK

Energy efficiency is one of the key design requirements in battery-powered wireless sensor networks. One important solution for it is to minimize the number of transmitted message in the network [18].

Several works have been done on data aggregation in wireless sensor network that reduce the power consumption. Clustering in WSNs is an effective procedure to decrease the energy consumption of sensor nodes. In cluster based routing algorithms for wireless networks, LEACH is famous because it is simple and efficient. In LEACH, CH nodes are selected randomly and all the non-CH nodes are formed based on the received signal power from the CHs. In LEACH each node can become a CH, there is no pattern in electing CHs and all nodes have the same chance to be a CH, thus LEACH is not efficient. CHs are selected randomly and the energy is divided between all the nodes equally. CHs aggregate all received data from all nodes in the clusters [15].

LEACH forms clusters based on the received signal strength and use the CHs as portals to the sink. All the data processing like data fusion and aggregation are locally accomplished into the cluster. CH is selected periodically among the nodes of the cluster. LEACH forms distributed clusters, where nodes make independent decisions without any concentrated control. In LEACH, each CH has a straight communicates with the base station no matter the distance is close or not. When the network is massive, the communication between CHs and the base station consumes much energy for the long distance transmission. In LEACH, the size of clusters can be increased if the number of CHs is reduced. This makes induced excessive delays introduced by the number of nodes in the same cluster [19, 20].

The work in [21] presents the hybrid approach for cluster-based aggregation, which adaptively selects the appropriate data aggregation function. This paper shows an improvement in energy consumption with the velocity of the target. Dynamic clustering shows better performance when velocity of the target is high.

CTDA is a hybrid cluster and tree based algorithm and is proposed for data aggregation. It employs a data aggregation mechanism in the CH to lessen the amount of data transmitted. Therefore, CTDA decreases the energy dissipation in communication. CTDA decreases data transfer volume so it enhances energy efficiency and attains the purpose of saving energy of the sensor nodes. CTDA decreases the number of nodes, which directly send data to the base station. In WSN with constrained energy, it is inefficient for sensors to select CHs randomly. CTDA method does not perform any calculation in choosing the CHs and select CHs randomly. It is non-optimal to selected CHs by chance because it imposes an additional burden to the network. CTDA does not consider the amount of remaining energy in the nodes and it increases the wasted energy and decreases the lifetime of the network [16].

In MCTDA method, minimum spanning tree does not do data aggregation and only data of CHs by tree structure is sent to the base station [22].

CTPEDCA uses the full distribution in hierarchical WSNs. CTPEDCA is based on clustering and Minimum Spanning Tree routing strategy for CHs and the time complexity is small. CTPEDCA can balance the energy consumption of all the nodes, particularly the CH nodes in each round and extend the lifetime of the networks. In each round, CTPEDCA allows only one CH communicate directly to the base station. In CTPEDCA, a CH with the maximum remaining energy is selected as the base, CH0. CH0 constructs a minimum spanning tree between all CHs and broadcasts tree information for all the CHs. If the number of CH is K , $K-1$ CHs send data only to CH0 and CH0 transmit data to the base station. The disadvantage of this method is the network is dependent on the CH0. CH0 is placed under pressure and needs a lot of energy. If CH0 is failed, the network also failed. When the base station is too far, this method is not useful [17].

In WSN, improving the energy performance and maximizing the networking lifetime are the main challenges. For this reason a hierarchical clustering scheme, called Location Energy Spectral Cluster Algorithm (LESCA) is proposed in [23]. LESCA specifies the number of clusters in a WSN automatically. It is based on spectral classification and considers the remaining energy and some properties of nodes. LESCA uses the K -way algorithm and proposes new features of the network nodes such as average energy, distance to the base station, and distance to cluster centers in order to determine the clusters and to elect the cluster's heads of a WSN. If the clusters are not constructed in an optimal way and/or the number of the clusters is greater or less than the optimal number of clusters, the total consumed energy of the sensor network per round is increased exponentially.

3. MODEL DESCRIPTION AND ASSUMPTIONS

In this work, we consider a multi-hop WSN consisting of n stationary and location-aware sensor nodes, denoted by $\{s_1, s_2, \dots, s_n\}$, which are distributed randomly throughout an area. The network contains the sink node denoted by s_0 in a preassigned location that collects data from all sensor nodes.

In our proposed study, we consider the following assumptions:

- All the nodes know their location and nodes are distributed randomly in the experimental area.
- The base station has no energy constraint and is located at the top of the area.
- The initial number of CHs is constant and does not change over time.

The superiority of protocols is changed because there are different presumptions about the radio features, such as energy dissipation in transmitter and receiver modes. In our plan, a simple model is used for the radio energy dissipation, which is the transmitter, power amplifier, and receiver dissipates energy to run the radio electronics [24]. The distance between the transmitter and the receiver is used for the free space (d_2 power loss) and the multipath fading (d_4 power loss) channel models.

In general, the free space (fs) model is used when the distance is less than a threshold d_0 and if more than the threshold d_0 , the multipath (mp) model is used [24].

Therefore, when n bit data message is transmitted over a distance d to achieve an acceptable signal, the energy expended by the radio ETX can be expressed as Eq. (1).

$$E_{TX}(n, d) = \begin{cases} n E_{elec} + n \varepsilon_{fs} d^2 & d \leq d_0 \\ n E_{elec} + n \varepsilon_{mp} d^4 & d \geq d_0 \end{cases} \quad (1)$$

where, ε_{fs} is the energy consumed by the amplifier to transmit at a shorter distance. ε_{mp} is the energy consumed by the amplifier to transmit at a longer distance. E_{elec} is the energy dissipated in the electronic circuit to transmit or receive the signal, which relied on agents such as the digital coding, modulation, filtering and spreading of the signal. E_{rx} is the radio energy consumed to receive this message, which is calculated by Eq.(2).

$$ERX(n) = n * E_{elec} \quad (2)$$

4.DYNAMIC FUZZY CLUSTERING (DFC) DATA AGGREGATION

This paper proposes an algorithm for data aggregation called Dynamic Fuzzy Clustering (DFC) data aggregation. DFC data aggregation uses the concepts of cluster and tree based algorithms. The main idea of the cluster based routing is to lessen the amount of data transmission via engage the data aggregation mechanism in the CH. DFC data aggregation decreases the energy dissipation and saves the residual energy of the nodes. DFC data aggregation has three steps:

- CHs selection
- Cluster construction
- Tree formation of CHs

Our proposed method is inspired from two approaches named Pareto Optimal Solutions [25-26] and Fuzzy TOPSIS. At the beginning of the network, we select CHs based on Fuzzy TOPSIS [8]. The clusters are formed based on the distance between nodes and CHs. Then, the tree is organized due to CHs situation. This process continues until the first CH dies or the CH energy gets lower than a defined threshold. In this case, we determine CHs based on Fuzzy TOPSIS again. We determine the CHs based on maximizing the amount of energy efficiency. Although the initial number of CHs is assumed constant, it can decrease clustering the performing of the algorithm. The number of CHs is related to the several parameters, i.e., network topology, residual energy of nodes, and the relative costs of calculation versus communication. The iteration of the mentioned steps creates rounds in DFC algorithm.

In the sequel, we describe the steps of DFC data aggregation.

4.1.CH selection

Multi Criteria Decision Making (MCDM) techniques have been applied to quantitative decision-making problems [27]. MCDM can be divided into two main categories. Multi-attribute decision-making (MADM) approach [24] is one of the main categories of MCDM techniques. On the other hand, multi objective decision-making (MODM) [25] is another main category in MCDM techniques.

In this paper, we use MODM (Pareto optimal technique) and MADM (fuzzy TOPSIS) for selecting CHs.

MODM (Pareto optimal technique)

The Pareto optimal solutions introduced by the economist Vilfredo Pareto [26]. Pareto technique determines the solution space which solutions are non-dominated. Pareto solution space specifies an area which comprising of all conceivable solutions in multi objective decision making problems. The solution space is classified into three groups, namely, completely dominated, neither dominant nor dominating and non-dominated.

MADM (fuzzy TOPSIS)

It is often difficult to determine the exact values of attributes of the sensor nodes [8]. Thus, we use a fuzzy approach to determine the comparative significance of criteria instead of exact values. In this algorithm, five fuzzy linguistic variables are considered. Figure 1 illustrates the fuzzy triangular functions. The triangular membership functions are determined in Table 1.

Table1. Transformation of fuzzy triangular membership function

Rank	Triangular membership function
Very Weak (VW)	(0.00, 0.10, 0.25)
Weak (W)	(0.15, 0.30, 0.45)
Moderate (M)	(0.35, 0.50, 0.65)
Strong (S)	(0.55, 0.70, 0.85)
Very Strong (VS)	(0.75, 0.90, 1.00)

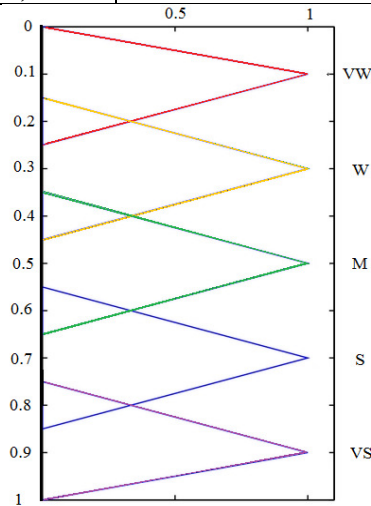


Figure 1. Fuzzy triangular functions

TOPSIS approach has contributed to solving the decision-making problems. In fuzzy TOPSIS approach, decision matrix has “m” alternatives and “n” attributes that could be assumed as a problem of “n” dimensional hyperplane has “m” points whose location is given by the value of

their attributes [10]. i, j , respectively are $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$. The decision matrix is as the following:

$$A = \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \vdots & x_{ij} & \vdots \\ x_{m1} & \dots & x_{mn} \end{bmatrix} \quad (3)$$

The weight of the j th column of matrix A is shown as (4):

$$C = [c_1, c_2, \dots, c_j, \dots, c_n] \quad (4)$$

where x_{ij} and c_j are fuzzy numbers. We have determined 0.5, 0.25, and 0.25 weights to the remaining energy, number of neighbours, and distance from the sink s_0 , respectively. P is a fuzzy decision matrix, which is normalized as the follow:

$$P = [p_{ij}]_{m \times n}$$

F is the weighted, normalized fuzzy decision matrix.

$$F = \begin{bmatrix} c_1 p_{11} & c_2 p_{12} & \dots & c_n p_{1n} \\ c_1 p_{21} & c_2 p_{22} & \dots & c_n p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ c_1 p_{m1} & c_2 p_{m2} & \dots & c_n p_{mn} \end{bmatrix} \quad (5)$$

In order to simplify the above matrix ($f_{mn} = c_n p_{mn}$), we summarize it as follows:

$$F = \begin{bmatrix} f_{11} & f_{12} & \dots & f_{1n} \\ f_{21} & f_{22} & \dots & f_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ f_{m1} & f_{m2} & \dots & f_{mn} \end{bmatrix} \quad (6)$$

The best conceivable solution is the shortest distance from the ideal solution, and the worst conceivable solution is the furthest distance from the ideal solution.

The best and the worst solutions are obtained from the weighted, normalized fuzzy decision matrix given by (6). The Best Solutions are defined as BS_j and WS_j is an acronym for the Worst Solutions:

$$BS_j = \{(\max f_{ij} | i = 1, 2, \dots, m), j = 1, 2, \dots, n\} \quad (7)$$

The worst solutions are defined as:

$$WS_j = \{(\min f_{ij} | i = 1, 2, \dots, m), j = 1, 2, \dots, n\} \quad (8)$$

We select a solution which is the nearest from the best conceivable solution and the furthest from the worst ideal solution. The distances of each alternative from the best solution and the worst solution are the separation measures. Distance of Best Solutions (DBS) and Distance of Worst Solutions (DWS) are given as:

$$DBS_i = \sum_{j=1}^n d(f_{ij}, BS_j) \quad i = 1, 2, \dots, m \quad (9)$$

$$DWS_i = \sum_{j=1}^n d(f_{ij}, WS_j) \quad i = 1, 2, \dots, m \quad (10)$$

Rank indices of TOPSIS are estimated as:

$$Rank_i = \frac{DBS_i}{DWS_i + DBS_i} \quad (11)$$

Superior TOPSIS rank nodes are selected as the CHs. Each selected CH gets a unique identifier (ID).

4.2. Cluster Construction

All the selected CHs disseminated identity message to non-CH nodes in the network. Each node calculates the distance from all the CHs then joins to the cluster, which has the minimum distance from its CH. K specifies the number of CHs. A distance matrix is used for re-clustering nodes based on the distance to the selected CHs. The distance metric used here is the Euclidean metric. The Euclidean distance between CH and a node is relying on their situations. Consider X and Y are two nodes, i and j demonstrates two node locations. Euclidean distance is calculated based on Eq. (12):

$$d(X, Y) = \sqrt{(X_i - Y_i)^2 + (X_j - Y_j)^2} \quad (12)$$

Each element in the distance matrix represents the difference between the CH and the node. After cluster formation, each CH is accountable for gathering the data from all the nodes in the cluster.

When a framework (of data) from all the nodes in the cluster is consummated and aggregation is performed, each CH dispatches the framework to the base station. The proceeds of re-clustering and data transportation is continued for R rounds until all the nodes being dead. If the number of nodes in the cluster gets smaller than the predefined threshold, the cluster is merged with the neighbouring clusters.

4.3. Tree formation of CHs and Data transmission

After cluster formation, the CH sends message to all non-cluster nodes in wireless sensor network which includes the CH ID, location, cluster size (for example the number of nodes in each cluster), and remaining energy. CHs also send their data and location to the base station. Base station prepares a minimum spanning tree based on the position of CH nodes so the minimum spanning tree is between CH nodes and the base station. In this plan, CHs use free area channel model to send data to the base station. In each round, the minimum distance from a vertex to another vertex is chosen based on the location of CH nodes in the tree. Combining data from several sensors used for removing the redundant transmission. Non CH nodes send their data by the framework to the CH while they are in transmission mode, so data transmission is broken into frameworks. Nodes could dispatch their data without any collision in the network. In this research, we assumed that nodes are all the time synchronized by having the base station sent out synchronization pulses to each node. When the CH receives the data from all the non CH nodes, it performs data aggregation to produce a useful data message for sending to the base station. After aggregating data, CHs transmit their resultant data along the tree (by the minimum spanning tree between CH nodes). Finally, the base station receives the final resultant data. Non-CH nodes could leave clusters when its energy is finished. If any non-CH node leaves, the related cluster releases it. If CH node is dead or a new node is joined to the network, the CH selection algorithm should be re-run.

In this paper, we consider two versions of DFC, named DFC-1 and DFC-2. In DFC-1, a node consumes its finite energy budget during the algorithm. We consider a specific threshold in DFC-2 for the CHs. When the amount of energy of a CH passes from the specified threshold, a new CH is selected. In DFC-2, our threshold is achieved when the amount of energy of CH is reduced by half.

5.SIMULATION RESULTS AND ANALYSIS

In this section, we demonstrate detailed simulation experiments to evaluate the performance of the proposed algorithm. We adopt MATLAB as the platform tool, which is popularly used in the simulation experiments of wireless sensor networks [18].

In our proposed algorithm, the number of nodes is set to 100. The sink is situated far away from the area. In Cluster based approach, we consider ten CHs ($K=10$) in the network, which divide the nodes into ten clusters. We do an experiment in which different numbers of CHs are evaluated by the three criteria. The numbers of studied CHs are 5, 8, 10, and 15. We have compared them in remaining energy, energy consumption and the number of dead nodes. After R rounds, the most optimal CHs have more remaining energy, minimize the amount of consumed energy and the number of dead nodes. We set $R=140$. In Figure 2, experimental results show that K CHs are the most optimal conditions in comparison with other CHs. The selected optimal CHs have the lowest wasted energy and dead nodes, these CHs can keep more energy.

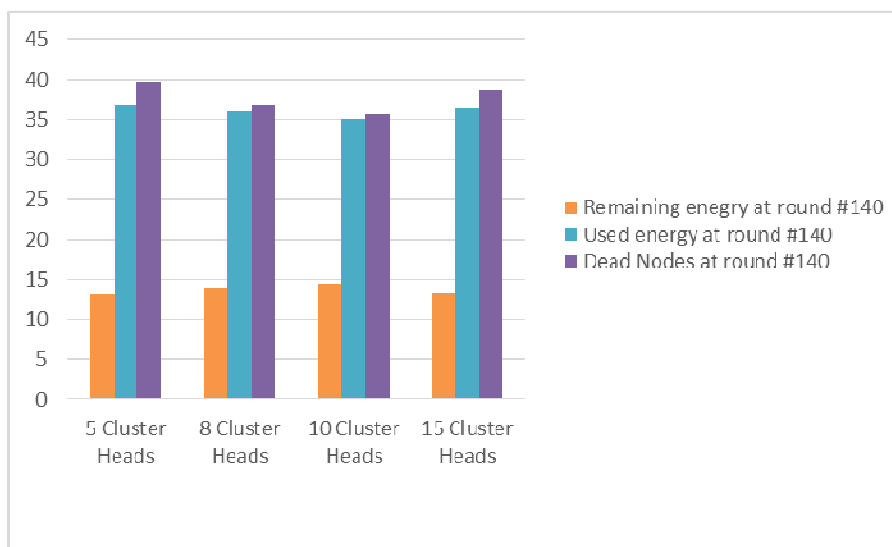


Figure 2. The effect of number of clusters in the DFC based on the number of dead nodes and used energy and remaining energy at round 140.

For selecting the best CHs, we have used Pareto optimal solution. Pareto optimal CHs are considered three criteria containing the remaining energy of the node, the minimum distance from the sink, and the number of adjacent nodes. Our properties of the criteria are normalized in the range [0, 1].

Membership function is applicable for converting the quantities into linguistic variables, afterward variables are converted into a fuzzy triangular membership function. We specify the fuzzy best solutions and fuzzy worst solutions. According to these quantities, we calculate the separation rate and rating indices for the selecting node. The lifetime of the network is extended in the period of the number of cycles until the first node in the network runs out of its complete energy. CHs are chosen for each node until all the nodes expand their whole energy. In a Tree based data aggregation approach, an aggregated tree is constructed based on a minimum spanning

tree which source nodes are thought out as leaves, so data are forwarded by the parent node for each node. The tree-based procedure has a low distance between each node and its parents, thereby wasted energy is diminished. Nevertheless, the depth of the tree is high. This hybrid method uses the advantages of the clustering and the tree structures while minimizing the disadvantages of them. A comparison of our proposed method against LEACH, CTDA, and CTPEDCA is represented that the present protocol is more effective than other mentioned methods in WSNs. We use four different routing protocols: LEACH protocol, CTDA algorithm, MCTDA algorithm, and our propounded method DFC algorithm for evaluating the performance of our proposed protocol with the rests. The simulation conclusions show the proposed approach has better efficiency than LEACH, CTDA, and CTPEDCA. We use a uniform simulation environment to facilitate comparison of energy savings and consume energy between protocols. Hundred sensor nodes are randomly spread in an area, which are shown in Figure 3 with coloured dots. The base station is placed far away from the area, at coordinates (50,200) which is shown in Figure 3 with an orange diamond.

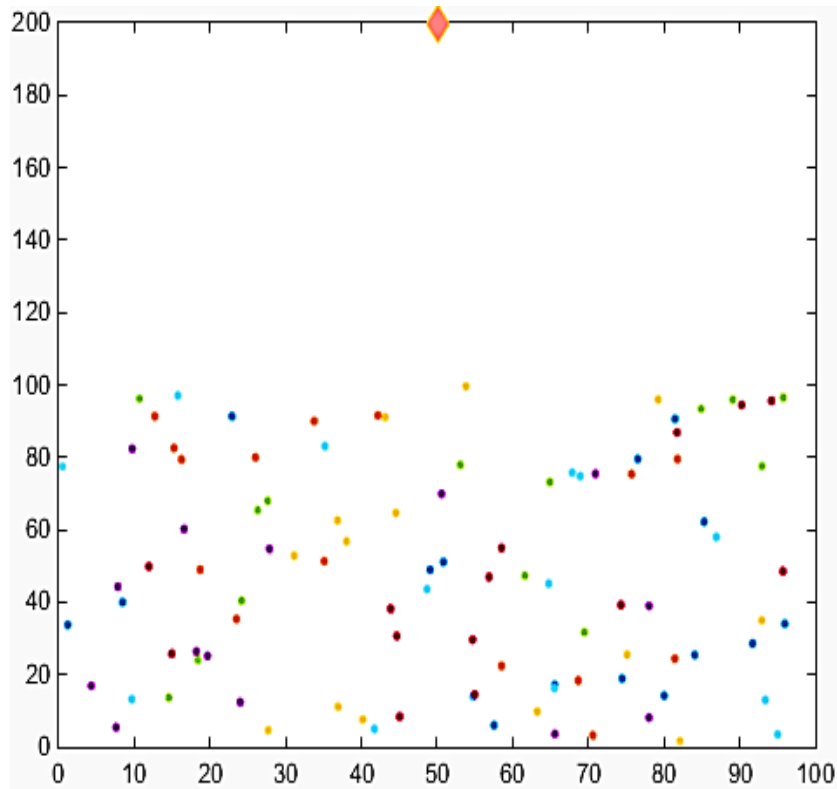


Figure 3. A simulated wireless sensor network with nodes and a base station at the top.

In Table 2, the parameters used to develop the simulation in our experiments are listed.

Table 2. Simulation parameters

Parameter	Value
Network Size	100 × 100 m
Number of the nodes	100
E_{elec}	50nJ / bit

ϵ_{fs}	10 pJ/bit/m ²
ϵ_{mp}	0.0013 pJ/bit/m ⁴
BS location	(50,200)
EDA (data aggregation)	5nJ / bit / signal
Control Packet size	800
Data Packet size	4000
R	140
K	10

A node is considered "dead" when it spent all its energy in the transferring process and unable to send and receive the data. The simulation results of dead nodes are shown in Figure 4.

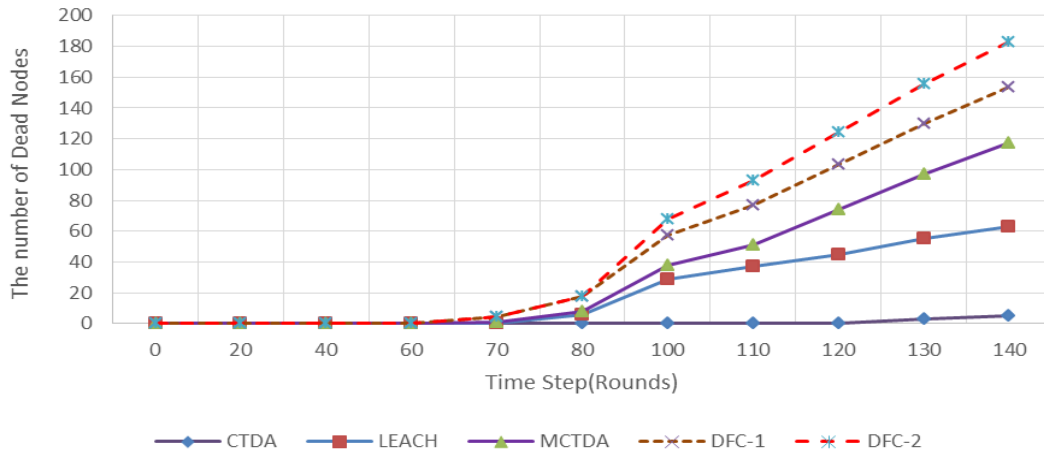


Figure 4. The number of dead nodes during the simulation.

Although the number of dead nodes in CTDA is low, but CTDA has many disadvantages. CTDA selects CHs randomly and it does not have any calculations to select the CHs. CTDA may select a CH with very low energy or choose a CH with the least number of neighbours. The number of dead nodes in DFC-1 and DFC-2 is less than LEACH and MCTDA. This pros is because of the CHs are calculated and elected based on three criteria the remaining energy, distance to the base station and the number of neighbours around.

According to the short distance between nodes in the proposed approach, network lifetime is increased. Furthermore, to decrease node solubility, DFC-1 and DFC-2 algorithms are more energy efficient all over the simulation. In DFC-2, we define a threshold for the amount of energy in CH, when the node's energy is less than the threshold, the new CH is replaced.

The simulation results of residual energy are illustrated in Figure 5. Our findings show that the remaining energy is increased. Choosing the correct CH in the proposed method makes shorter distance between nodes. Nodes are selected as CHs, which have the largest number of neighbours. Thus, less energy are wasted so each node can hold more energy. Energy consumption of the nodes is reduced.

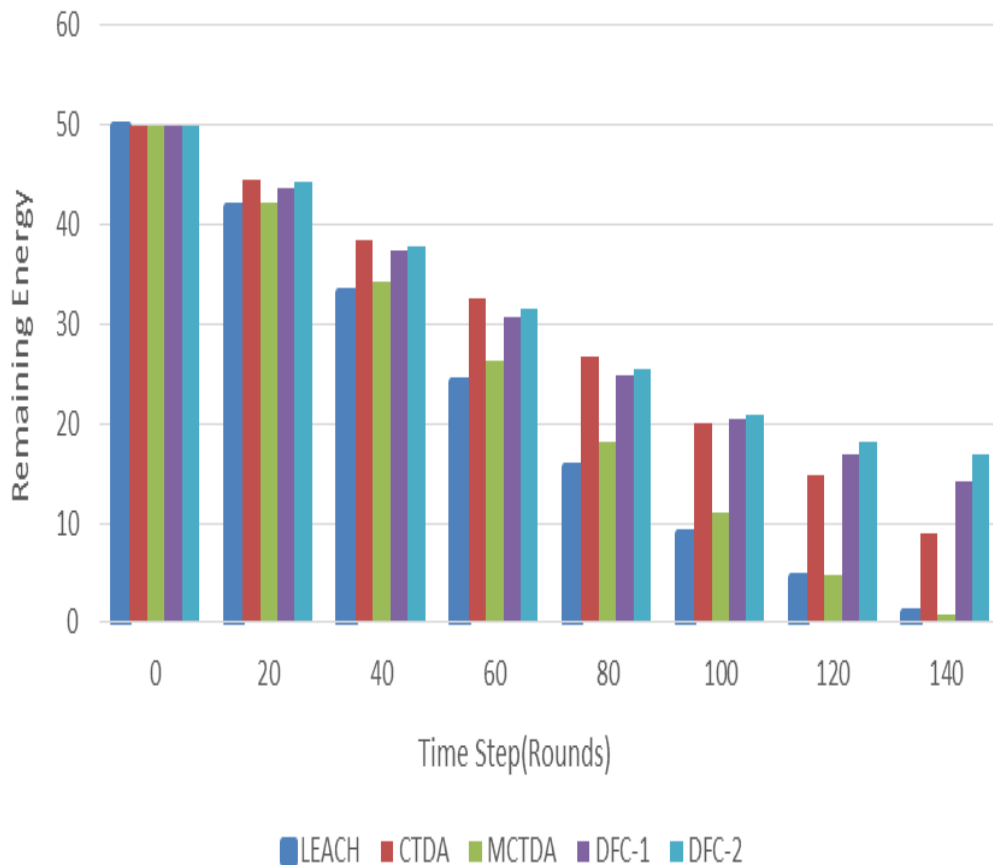


Figure 5. Remaining Energy of the nodes.

Figure 6 demonstrates the total dissipated energy by using LEACH, CTPEDCA, CTDA, and the proposed algorithm during the network simulation.

According to the results in Figure 6, CHs with the highest rank send data to the base station, which decrease the total energy consumption. In LEACH method, usually CHs attempt to send data to the base station due to the great distance through a multipath expunction channel model and consumption of energy is high. In MCTDA, accumulate data does not operate between the nodes in the tree and all data packets of CHs are sent to the base station, so energy consumption is high.

In DFC method, the amount of energy consumption is less because the CHs are selected intently. The minimum spanning tree constructed by elite CHs, for this reason DFC saves more energy than other mentioned algorithms.

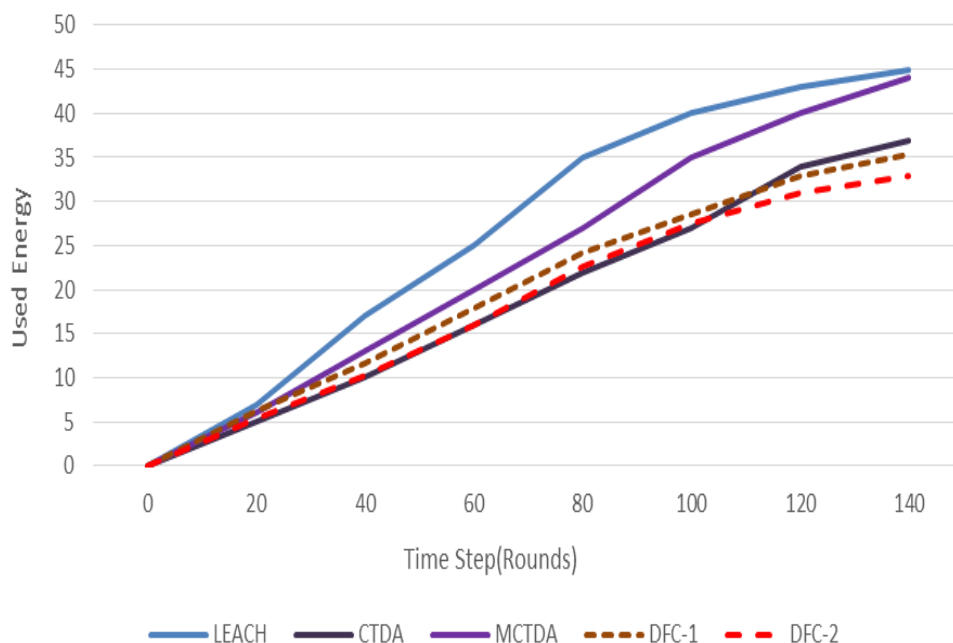


Figure 6. Total Consumed Energy.

6. CONCLUSIONS AND FUTURE WORKS

A fundamental challenge in the design of WSNs is the proper utilization of resources that are scarce. In this paper, we employ Fuzzy TOPSIS method for finding the best CHs in WSNs. Three criteria contain remaining energy, distance of the nodes from the base station and the number of neighbour nodes. These criteria are discussed in order to optimize the number of CHs. The tree-based method constructs a minimum spanning tree distance between CHs and the base station, which lead to decreasing energy dissipation. We proposed an energy effective algorithm in this paper, called DFC. DFC is a cluster and tree based data aggregation. Our proposed algorithm is compared with LEACH, CTDA and MCTDA protocols. The conclusions of this simulation demonstrate that the DFC is a considerable energy saving node which increase the network lifetime compared to the above- mentioned protocols. In the future, we will work on the extension of the DFC for mobility and heterogeneity of both nodes and sink.

REFERENCES

- [1] Akyildiz, I. F. Su, W. , (2002), "Wireless Sensor Networks: A Survey.", J. Computer Networks., Vol. 38, pp. 393-422.
- [2] Chanak, P., Banerjee, I., (2016) , "Fuzzy rule-based faulty node classification and management scheme for large scale wireless sensor networks", Expert Systems with Applications, Vol. 45, pp. 307-321.
- [3] S. Mantri, D., Rashmi Prasad, N., Prasad, R., (2015), "Bandwidth efficient cluster-based data aggregation for Wireless Sensor Network," Computers & Electrical Engineering, Vol. 41, pp. 256-264.
- [4] Azad P, Sharma V. (2013), "Cluster Head Selection in Wireless Sensor Networks under Fuzzy Environment", ISRN Sensor Networks, pp. 1-8.
- [5] Abbasi, A. A, Younis M. (2007) "A survey on clustering algorithms for wireless sensor networks", Computer Communications, Vol. 30, No. 14-15, pp. 2826-2841.

- [6] Asemami M, Esnaashari M., (2015) “Learning automata based energy efficient data aggregation in wireless sensor networks”, *Wireless Networks*. pp. 256–264.
- [7] Selvin S, Kumar S. (2012), “Tree Based Energy Efficient and High Accuracy Data Aggregation for Wireless Sensor Networks”, *Procedia Engineering*, Vol. 38, pp. 3833-3839.
- [8] Baykasoğlu A, Gölcük İ. , (2015), “Development of a novel multiple-attribute decision making model via fuzzy cognitive maps and hierarchical fuzzy TOPSIS”, *Information Sciences*, Vol. 301, pp. 75-98.
- [9] Rathod M, Kanzaria H.(2011) , “A methodological concept for phase change material selection based on multiple criteria decision analysis with and without fuzzy environment.”, *Materials & Design*, Vol. 32, No. 6, pp.3578-3585.
- [10] Yang, Taho, and Chih-Ching Hung, (2007), “Multiple-Attribute Decision Making Methods For Plant Layout Design Problem”, *Robotics and Computer-Integrated Manufacturing* , Vol. 23, No.1, pp. 126-137.
- [11] Rajagopalan, R. Varshney, P. K.(2006) “Data Aggregation Techniques In Sensor Networks: A Survey”, *J. IEEE Communications Surveys & Tutorials*, Vol. 8, No. 4, pp. 48-63.
- [12] Abdullah L, Najib L.(2014), “A new type-2 fuzzy set of linguistic variables for the fuzzy analytic hierarchy process”, *Expert Systems with Applications*, Vol. 41, No.7, pp. 3297-3305.
- [14] Sun C. (2010), “A performance evaluation model by integrating fuzzy AHP and fuzzy TOPSIS methods”, *Expert Systems with Applications*, Vol. 37, No. 12, pp. 7745-7754.
- [15] Akkari W, Bouhdid B, Belghith A.,(2015), “LEATCH: Low Energy Adaptive Tier Clustering Hierarchy”, *Procedia Computer Science*, Vol. 52, pp. 365-372.
- [16] Sajedi H, Saadati Z., (2014), “A Hybrid Structure for Data Aggregation in Wireless Sensor Network”, *Journal of Computational Engineering*, Vol. 2014, pp. 1-7.
- [17] Wang W, Wang B, Liu Z, Guo L, Xiong W.,(2011), “A Cluster-based and Tree-based Power Efficient Data Collection and Aggregation Protocol for Wireless Sensor Networks”, *Information Technology J.*, Vol. 10, No. 3, pp. 557-564.
- [18] Cheng,H., Su, Z., Xiong, N., Xiao, Y., (2016), “Energy-efficient node scheduling algorithms for wireless sensor networks using Markov Random Field model”, *Information Sciences*, Vol. 329, pp. 461–477.
- [19] Richard, W. G., (2009), *Extending LEACH routing algorithm for Wireless Sensor Network*. Data Communications Engineering, Makerere University.
- [20] Batra, N., Jain, A., & Dhiman, S., (2011), “An optimized energy efficient routing algorithm for wireless sensor network”, *International Journal of Innovative Technology and Creative Engineering.*, Vol. 1.
- [21] Woo-Sung, J., Keun-Woo,L., Young-Bae, K., Sang-Joon, P., (2011), “Efficient clustering-based data aggregation in wireless sensor networks”, *Journal of Wireless Networks*, Vol. 17, No. 5. pp. 1387-400.
- [22] Ranjani, S.; Krishnan, S.; Thangaraj, C., (2012), “Energy-Efficient Cluster Based Data Aggregation for Wireless Sensor Networks”, In *Proc. of International Conference on Recent Advances in Computing and Software Systems*, pp. 174-179
- [23] Jorio A, El Fkihi S, Elbhiri B, Aboutajdine D., (2015), “An Energy-Efficient Clustering Routing Algorithm Based on Geographic Position and Residual Energy for Wireless Sensor Network”, *Journal of Computer Networks and Communications*, pp.1-11.
- [24] Heinzelman, W. Chandrakasan, A. Balakrishnan, H., (2000), “Energy-Efficient Communication Protocol for Wireless Microsensor Networks”, In *Proc. of IEEE International Conference on System Sciences*, pp. 1-10.
- [25] Chauhan A, Vaish R., (2013), “Pareto Optimal Microwave Dielectric Materials”, *Advanced Science, Engineering and Medicine*, Vol. 5, No. 2, pp. 149-155.
- [26] Kasprzak E, Lewis K. ,(2001) “Pareto analysis in multiobjective optimization using the collinearity theorem and scaling method”, *Structural and Multidisciplinary Optimization*, Vol. 22, No. 3, pp. 208-218.
- [27] Zanakis S, Solomon A, Wishart N, Dublisch S., (1998) “Multi-attribute decision making: A simulation comparison of select methods”, *European Journal of Operational Research*, Vol. 107, No. 3, pp. 507-529.

AUTHORS

Arezoo Abasi was born in Tehran, Iran in 1994. She studies Computer Science at University of Tehran since 2012. She was awarded Khaje Nasir price in 2010. Her interests include artificial intelligence, soft computing and wireless sensor network.

Hedieh Sajedi received a B.Sc. degree in Computer Engineering from Poly Technique University of Technology in 2003, and M.Sc. and Ph.D degrees in Computer Engineering (Artificial Intelligence) from Sharif University of Technology, in 2006 and 2010, respectively. She is currently an Assistant Professor at the Department of Computer Science, University of Tehran, Iran. Her research interests include Computer Networks, Machine Learning, and Signal Processing.