Journal of Engineering and Applied Sciences 13 (Special Issue 14): 11081-11090, 2018
ISSN: 1816-949X

Clustering in a Wireless Sensor Network using Fuzzy Logic Method

Rula Sami Khudhair
College of Material's Engineering, University of Babylon, Babylon, Iraq

Abstract: In the IT world, the wireless sensor network poses a very high position. According to its various applications, it is useful in many industries. In this network, sensors form the main body and transmit the information to the server by checking the environment but an important challenge is the limited energy of the sensors. A remarkable solution to this challenge is using clustering to send information by the cluster heads. In this study, we present a new clustering method using fuzzy logic that can perform clustering with the least energy consumption. In this method, we perform clustering according to four criteria including energy, number of neighbors, centrality and distance from the main server. This combination has been able to provide a highly optimized cluster that selects the best and the most suitable nodes as a cluster head in each step. The simulation results and the comparison of the proposed method show that the proposed method is better than previous methods and has improved 12% of energy consumption which is a very important factor.

Key words: Neighbors, information, network, clustering, sensors, heads

INTRODUCTION

Recent advances in electronics and wireless communications have created the ability to design and manufacture low-power, small size sensors with affordable prices and various applications. These small sensors which are capable for performing tasks such as receiving various environmental information, processing and transmitting them have led to the emergence of ideas for the creation and expansion of new generation of networks known as Wireless Sensor Networks (WSNs) (Van Greunen and Rabaey, 2003).

Wireless sensor networks today have many applications. In these applications, sensors are often remotely deployed and independently operated. In these unprotected environments, the sensor cannot be charged, so, energy constraints are the most important problem that should be considered. In a large wireless receiver network, sensors are often grouped into clusters. A cluster is required for sensor networks applications where in a large number of ad-hoc network sensors are deployed to measure targets. If any node in the network begins to communicate and transmit data separately over the network, it will experience large data traffic and collisions. This phenomena will lead to rapid energy depletion from the wireless sensor network. Clustering is a method used to overcome these issues. In clustered networks, some nodes are considered as cluster heads in each cluster. Sensor nodes in each cluster send their corresponding cluster data to the cluster head and the cluster head collects the data and sends them to the central station.

Clustering facilitates the efficient use of the limited energy of sensor nodes and extends the lifetime of the network. Given that the sensor nodes in the clusters send messages at short intervals, that is these messages are only published within the cluster, so, the cluster heads are the nodes which send messages at longer distances in relation to the other sensor nodes within the cluster (from the cluster head to the main station) thus, they consume more energy (Akkaya and Younis, 2005).

The alternating re-election of the clusters based on the remaining energy, provides a possible solution for balancing the energy consumption of each cluster. In fact, clustering reduces the efficiency of data transfer by reducing the sensors trying to send data to the main station. Collecting data at clusters via. within cluster communications also helps to eradicate duplicate data. Clustering is proposed because of its network scalability, energy storage and network topology stability. The clustering reduces the overhead communication between sensor nodes, although, some clustering algorithms have disadvantages such as overhead on the cluster head selection path, appointment and cluster formation process.

Importance of the research: Wireless sensor networks are one of the most important tools for gaining information and understanding the environment. Despite the advances made in these types of networks, sensor nodes due to their large number, small size and contingency placement methods, they depend on low-power batteries to provide their own energy.

Normally, due to the use of these types of networks in harsh and inaccessible environments, there is no possibility of recharging or switching sensor nodes. Therefore, one of the most important issues in wireless sensor networks is the issue of extreme energy constraints. Also, because the efficiency of sensor networks is highly dependent on the network lifetime and network coverage, it is vital to consider energy storage algorithms in the design of long-life sensor networks. Clustering is one of the main approaches to designing efficient and scalable wireless sensor network protocols. The use of clusters reduces communication overhead due to data transmission and as a result, reduces energy consumption and interference between nodes. In many applications, cluster organization is a natural way of grouping together close nodes to use related data and delete plugins data. By aggregating and combination of the nodes data in the cluster head, the total volume of data sent to the base station is significantly reduced and less energy and resources are consumed (Soro and Heinzelman, 2009).

Literature review: To reduce energy losses and increase network lifetimes, different clustering protocols are proposed. In this study, we try to briefly introduce several clustering protocols in which researchers seek to save on energy consumption of wireless nodes.

By Baranidharan and Santhi (2016) the use of a spanning tree is suggested to create clusters with proper properties. In this algorithm, energy is not emphasized as the main parameter.

By Natarajan and Selvaraj (2014), a clustering algorithm has been proposed to extend the lifetime of the nodes, focusing on focused clustering methods that are implemented at the main station. This algorithm works for single cluster mode and aims to increase the time that the first node in network dies off. The location of the cluster heads are determined in a way by which energy consumption is minimized. For this purpose, the proposed algorithm selects the circular environments around the main station, based on the weights of the nodes and then selects the cluster heads near a series of optimal points. Also, in order to make decision about the connection of the normal nodes to the cluster heads, it determines the distance from the cluster head and the remaining energy of the cluster head.

By Abbasi and Younis (2007), fuzzy logic is used to determine the nodes of the cluster heads. Given the local data of the nodes such as the energy level, the distance to the main station and the local density with the help of the fuzzy logic system, determines the probability of choosing a node as the cluster head. Also, for optimal routing, the max-min ant colony has been used.

By Ran *et al*. (2010), a multi-level hierarchical structure has been suggested, so that, cluster heads are selected based on their degree and their residual energy.

By Mazumdar and Om (2017), fuzzy logic method is used to find clusters. In this algorithm, three fuzzy variables are used to select cluster heads. Node energy, node focus and node centering are these three parameters. In this method, the main station initially collects the necessary information from all the nodes and then selects a node as the cluster head based on fuzzy rules. In this method, there is only one choice for each round while more CH is needed to balance energy consumption and improve network lifetime.

By Kim *et al*. (2008), CHEF is suggested and CHs were selected based on fuzzy logic. The difference here is that in this method more than one cluster head in each round is selected locally. The fuzzy set contains the energy of the nodes and their local distance. CHEF chooses a random number for each sensor and if this number is less than the predefined threshold of popt, then the chance of that node is definite, therefore, there may be eligible nodes losing their chance in a random manner.

By Rana *et al*. (2015), fuzzy logic has been used to form clusters using the FIS system in cluster formation. Criteria such as the energy level of the nodes, the distance of each cluster head node to the main station and the distance from each node to the cluster head have been considered.

By Kandpal *et al*. (2015), a leach-based protocol is described. In this protocol, the following steps are performed, respectively:

- Clustering phase
- Steady state phase and selection of the next round cluster head

At stage A, based on the clustering phase of the LEACH protocol, firstly, the number of well-readable cluster heads and the optimal number of members of the optimal cluster are calculated:

$$K_{opt} = \sqrt{\frac{N}{2\pi}}\sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp} d^4_{to\,BS} - E_{elec}}}$$

Based on the above formula for the number of 100 sensor nodes in the simulation environment and based on the LEACH standard energy model, five sets of clusters are appropriately determined. Given the following relationship:

$$Member_{opt} = \frac{N}{K} - 1 \quad (2)$$

Where:
N = The number of network nodes
K = The number of optimal cluster

Memberopt is the number of the members of optimal cluster. For high values, the value of N/K-1 is considered equal to 19 sensor nodes as a member of each cluster head. The components of phase A are as follows: determine and calculate the number of optimal cluster heads in the network, send a message from the sink node to the network.

Receiving the sink message on each node and send the geographic information and ID of each node and send this information to the sink. When the sink node receives the messages of the network nodes, based on the maximum return time, it chooses the first cluster head. Time is directly related to the distance because the speed of the transmission of information packets is constant. Then, the sink node sends the chosen cluster head node's ID to the whole network.

Chosen cluster head sends its message to the network. Network nodes receive the first message from the cluster head and sends their local information and ID to it. If a node becomes a cluster member it can no longer change its cluster. The cluster head node selects the number of N/K-1 nodes from nodes with minimum return times. The minimum time indicates that the closer nodes are became as the members of the cluster. Then, the cluster head node sends the identity of the cluster member nodes to the sink.

The cluster head node chooses the maximum return time, from the all received messages in the previous step as the second cluster head. Then, it sends the new cluster head ID to the sink. The steps 5-8 are repeated, so that, the optimal number of cluster heads, K_{opt}, achieve.

When the number of selected cluster heads reaches to K_{opt}, sensor nodes within each cluster will pick up their cluster head. In the steady state, in the next round of selection, this protocol selects the cluster head on the basis of the maximum remained energy.

By Mirzaie and Mazinani (2017), a multiplicative algorithm adapted to the use of Adaptive MCFL fuzzy logic is presented. In addition to clustering nodes in different steps using different clustering algorithms, the proposed algorithm avoids the selection of new cluster heads by relying on previous cluster heads which reduces the number of messages and saves energy. This algorithm works in three phases as follows:

Algorithm 1; Multiplicative algorithm:

The first step
Determining neighbors and specifying their number for each node
Changing "remaining energy" and "number of neighbors" inputs to Fuzzy mode for each node
Comparing the fuzzy output of each node with its neighbor's fuzzy output
Selecting the node with the highest output as the cluster head with all the radio neighbors
Sending data from the node to the cluster head and from the cluster head to the station base
Second stage
Selecting the previous round's clusters heads for each cluster in the current round
Third step
Determining neighbors for each node
Fuzzy Inference Using Parameters such as "Remaining Energy" and "cluster head Distance in the Previous Round" for each node
Comparing the fuzzy output of each node with its neighbors' fuzzy output
Selecting the node with the highest output as the cluster head
Sending data from the node to the cluster head and from the cluster head to the base station

MATERIALS AND METHODS

Our proposed method chose the cluster head according to four items: the remaining energy of the node: a sensor with a maximum energy level because the cluster head node overhead is larger than the other nodes.

Centrality of the node: The nearest sensor to the center of gravity of the cluster in fact, the node whose average distance from the other cluster nodes is minimal. In fact, the centrality of the cluster head reduces energy consumption for communication within a cluster (between member nodes and the cluster head).

Number of neighbor nodes: When a node is at the center of attention, it means that many signals pass through it to reach a cluster head and it is better that this node itself becomes as the cluster head.

Distance from the base station: As the distance from the base station increases, one node should spend more energy on sending the data to the base station, therefore with increasing distance, the chance of selecting nodes as the cluster head decreases. Also, nodes near the base station in addition to aggregating intra cluster data, also, transfer data from other cluster heads to the base station and consume more energy, here, as taking into account the different radio radii, a kind of load balance has been created. These four items are inputs that the fuzzy system uses to select the cluster head.

Fuzzyfier: As it mentioned above, fuzzyfier uses four items. Related membership function for each item is shown in Fig. 1. The remaining energy membership function is shown in Fig. 1. The centrality membership

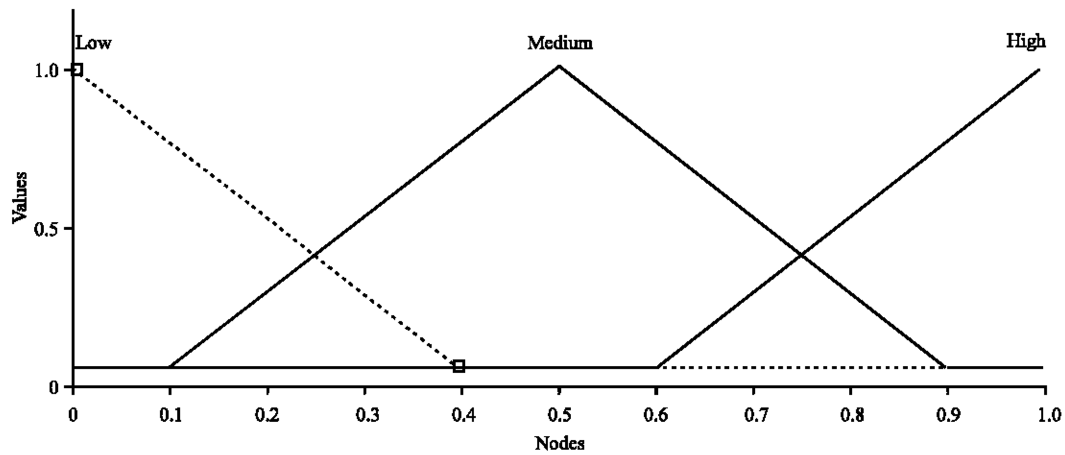

Fig. 1: The remaining energy membership function

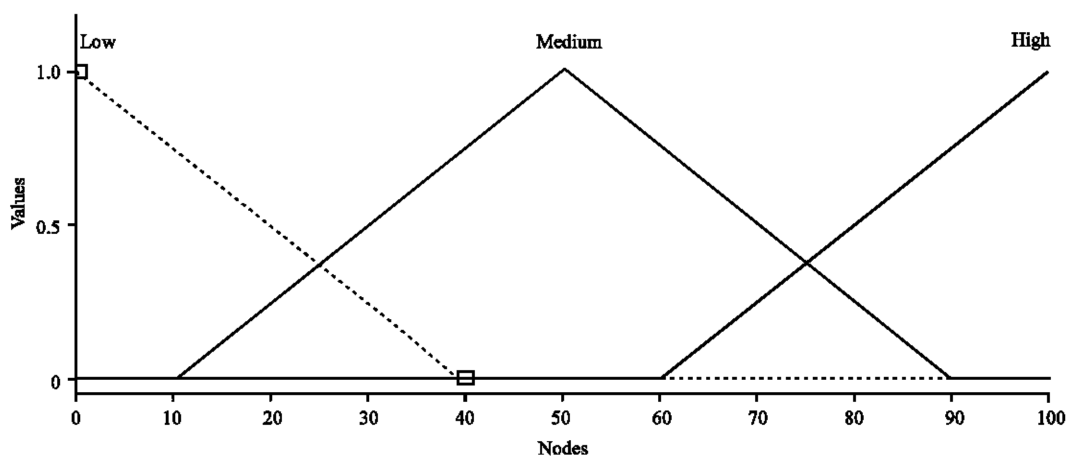

Fig. 2: The centrality membership function of the node

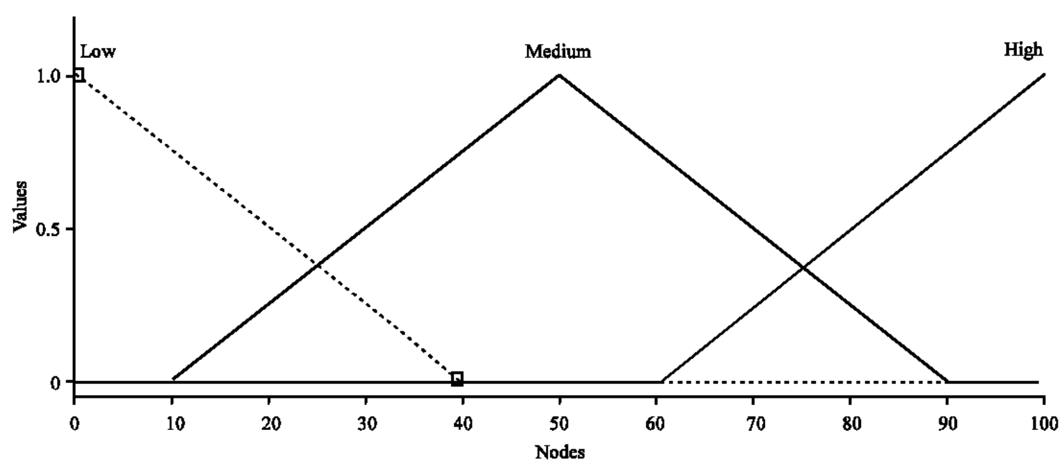

Fig. 3: The membership function for the number of the neighbors

function of the node is shown in Fig. 2. The membership function for the number of the neighbors is the shown in Fig. 3. The membership function for the distance from the base station is the shown in Fig. 4.

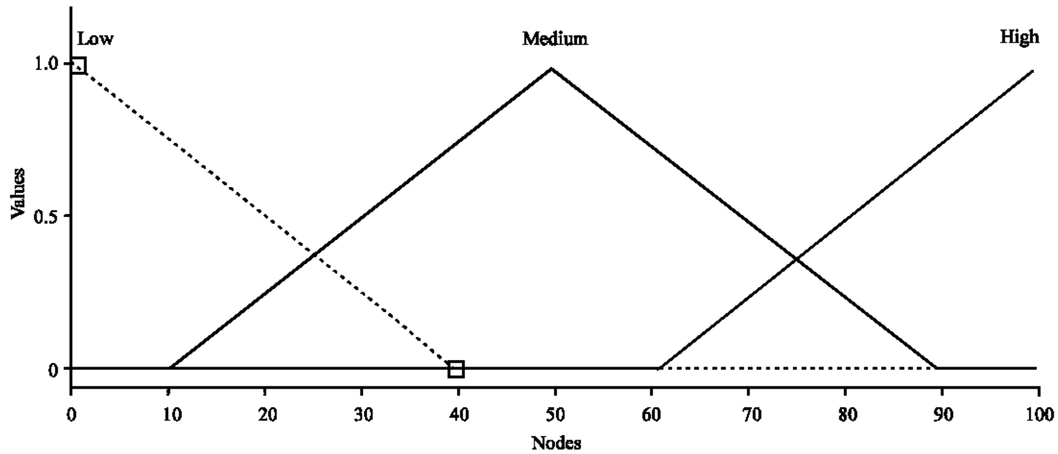

Fig. 4: The membership function for the distance from the base station

Table 1: Some of fuzzy rules

Probability of becoming cluster head	Centrality	Distance from base station	Number of neighbors	Energy
Medium	Medium	High	High	Low
Very low	Low	Low	Low	Low
Medium	High	Medium	Medium	Low
Very high	High	Medium	High	Medium
High	High	Low	Low	Medium
Very high	Medium	Low	High	High
Medium	Low	High	Medium	High

Fuzzy inference step: According to the four inputs used in the fuzzy inference step, the Mamdani engine is used and some of the parameters of this engine are shown in Table 1.

D fuzzy step: At this stage, the output of each rule is represented as a number and nodes with a high numeric chance are selected as the header according to the following scenario.

After calculating the chance of each node, it must be determined which node will be selected as the cluster head. To do this, nodes with chances greater than zero will compete and each node will wait as long as T time. The value of this time is calculated from Eq. 1. After this time, the node sends its introduction Message (M) as the cluster head to all its neighbors:

$$T = \frac{t_m}{P} \quad (3)$$

Where:

P = The value of the odds for each node

t_m = A constant value for the entire network which is based on the network structure.

The larger the p is the shorter it will wait. The time the node waits, one of the following scenarios may occur: the node receives the message from another node: this indicates that there was another node that had a higher chance and a shorter waiting time. The node which is sending the message is selected as the cluster head. The nodes who received this message will no longer send their message.

The node will not receive any messages during this time, so, the node sends its message to all its neighbors and is selected as the cluster head.

The advantage of this random time is that the nodes do not enter into controversy and calculating the amount of power. It is noteworthy that chance calculation operations consume energy. A node with a higher chance will wait shorter and then send its message.

Selecting the members of a cluster: Once the cluster has been identified, the other nodes should select their cluster head. If a node receives a message from several cluster heads, it selects a cluster head that has more power and sends the connection message to the cluster head and the connection is made.

Deactivate nearby nodes: After clustering to extend the network life and cover more environment, sensors that are close to each other or have a high environmental contribution can be active in turn, so that, according to a scheduling algorithm at any one moment, one of them is active, the advantage of this method is that the sensors that are close to each other and transmit the same information do not work simultaneously and will cover the information transition more longer which will increase the quality of coverage. The process of deactivating the adjacent nodes is as follows: each cluster is divided into four parts (four quadrants of trigonometry).

If the size of each quarter is s, the size of the environment covered by each sensor is v, then at least n = s/v is sensors are required per quarter. Now, if the number of sensors is greater than n, the sensors are activated, respectively, each sensor turns off after sending each message and the next sensor is turned on.

Changing the cluster head: Since, most operations are performed by the cluster head, it consumes more energy and should be replaced after a while. In this study, the following solution is used to select the new cluster head. The main server is aware that the cluster head should be changed and inform all the cluster heads.

From the specified time that the server specifies, all cluster heads and nodes that have calculated their odds will send their message in accordance with the previous algorithm. New cluster heads will be introduced. New cluster heads will be introduced near the cluster where the cluster head should be replaced. The members of each cluster are determined by the previous algorithm. The nodes are placed in new clusters.

Optimized routing based on clustering: We use the central router controller to select the best path between distributed cluster heads. The central controller determines which of the distributed cluster heads is selected to move the packets. To do this, the following algorithm is applied.

There are several cluster heads, each with a number of sensors. A graph is formed in the central controller. The vertices of this graph are head clusters. If there is a relationship between the two clusters, there will be an edge between these two vertices. Each edge has a numeric metric that is calculated based on the three characteristics of bandwidth, traffic and distance which indicates how much time each packet needs to be moved between these two vertices. This metric is calculated by cluster heads. How to calculate this metric is explained. The central controller select the edge between two vertices which is optimally located.

If the traffic and delay in a cluster are high, the cluster head will be suspended, informing the central controller that it is not able to receive the new packet and thus preventing the acceptance of the further packets and following delay.

According to the above, the packet is always fairly distributed in the network and the algorithm works toothily to always choose the best and least delayed path.

How to calculate the link value between two cluster heads: To calculate the value of each active link, three attributes are considered:

- Bandwidth
- The traffic
- Nearest distance

The first criterion is the closest route and if the path has more traffic than what is defined by the threshold, then the next path will be replaced. For this mechanism, a completely dynamic formula is presented based on the three above features. This numeric formula ranges from 0-1 for each link and the closer the number is to one, the link will have a higher rating for selection. The formula is as follows:

$$P = D+B+T \quad (4)$$

Where:
D = Distance
T = Traffic
B = Bandwidth

As stated above, the distance poses the highest score and its value is between 0 and 0.5, B and T are a number between 0 and 0.25 and each of these three is calculated according to the following equations.

Calculation of parameter D: This parameter corresponds to the distance between the two clusters which is a number between 0 and 0.5. To calculate D, we have all distance between adjacent cluster heads, D value for the least distance is equal to 0.5 and other values will be calculated accordingly.

Calculation of parameter T: As mentioned, this parameter is related to traffic and the traffic status of all links is available in the controller. This parameter varies between 0 and 0.25. In a cluster, the least traffic in the link equals the value of 0.25 and based on this link, the value of t for other links is also computed.

Calculation of parameter B: This parameter represents the bandwidth and data transfer speed of a link; Its value varies between 0.0 and 0.25. Like the previous two parameters, the highest bandwidth in the cluster has the value of 0.25 and the remaining links can be set based on this link.

RESULTS AND DISCUSSION

Route selection method: After all the links are scored and the sender and receiver are identified, all routes within the network between the sender and receiver are checked based on points and the route that has the highest score is to select for transferring the data. How to calculate the rating points is expressed in the following:

$$P_{Routing} = \frac{\sum_{i=1}^{N} P_i}{N} \quad (5)$$

Where:

P_i = The rating of each link

N = The total number of links

The route with the highest score is selected as the transfer path.

Simulation parameters: MATLAB Software is used to implement proposed algorithms. In order to investigate the efficiency of the algorithm, the parameters including coverage, the number of live nodes and the residual energy content are considered. The proposed algorithm is compared with three other methods. As discussed in this study, the optimal algorithm should have more live nodes and more energy remaining in different periods of implementation and also have high packet delivery rates. Since, the primary dispersion pattern of the sensors has a great influence on the determination of these parameters in each round of implementation of the simulations the same basic pattern has been applied to all algorithms. Parameters are described in Table 2.

Clustering energy consumption: One of the things that consumes a lot of energy is clustering because many messages are transferred in this process. In this section, the energy used for clustering has been evaluated.

The results of this simulation are shown in Fig. 5. The proposed method uses less energy than previous clustering methods which increases network energy lifetime.

Energy saving according to scheduling: One of the strategies for reducing energy consumption and maintaining network lifetime that increases coverage on the network is to use a scheduler for sending data. The sensor nodes that are placed near together and the environmental data being sent are similar to each other can be controlled by a scheduler. This will save energy in the nodes. The simulation scenario for this solution is to compare the amount of remaining energy in the network at a constant time for the new method and the conventional method. The simulation results are shown in Fig. 6.

Table 2: Simulation parameters

Parameters	Values
Network size	200-1000 m
Number of sensors	50-500
Each node's energy	0.1-1
Number of sinks	1
Sensor's range	0.5 m

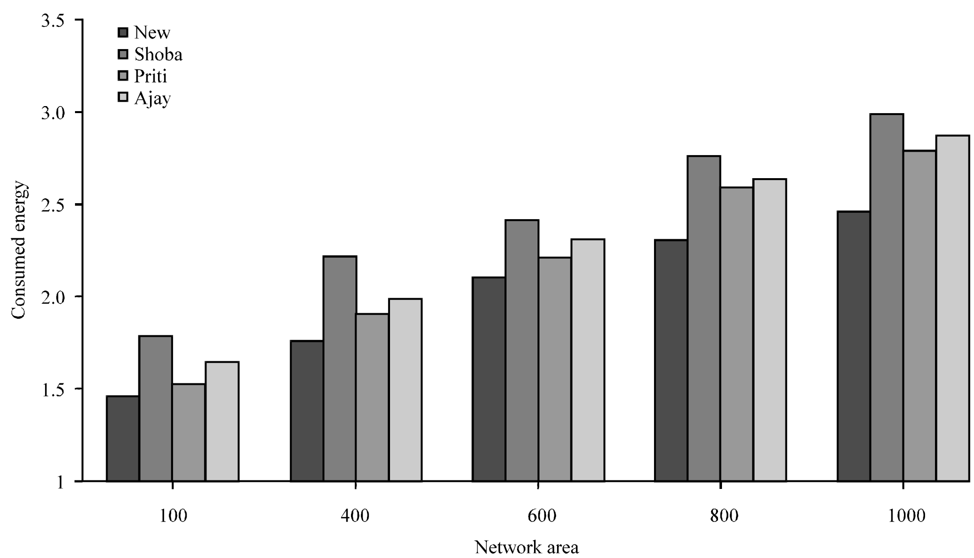

Fig. 5: Consumed energy for clustering

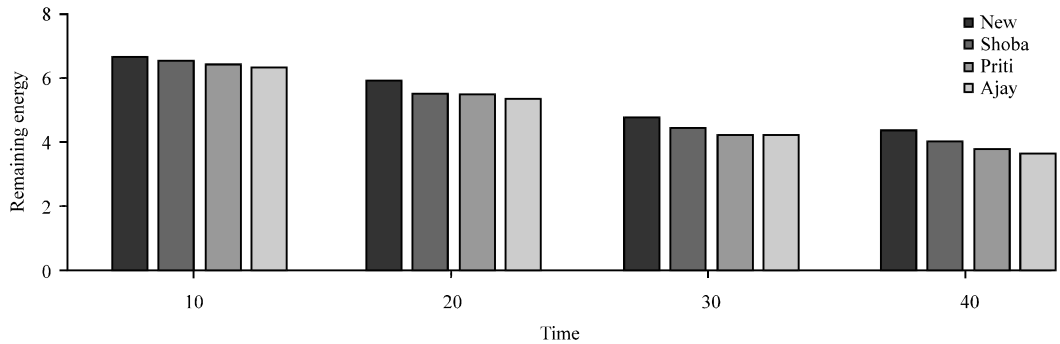

Fig. 6: Amount of remaining energy

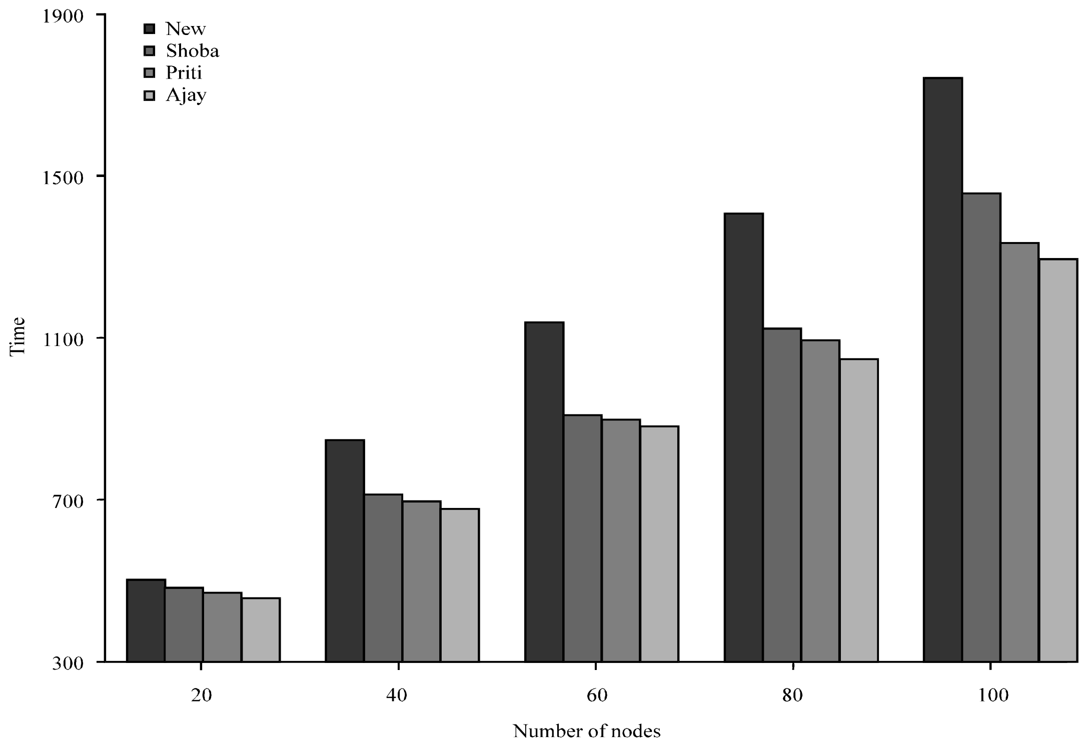

Fig. 7: Network lifetime

Network lifetime: In this scenario, we plan to investigate how long the entire network can last. In this situation, several random nodes are selected to send their packet to the base station, the simulation ends when all nodes lose their entire energy (Fig. 7). As Fig. 7 shows in the proposed method, network lifetime has been increased.

Message delivery time to the central station: In the simulation scenario, the goal is to get the time which last to carry a message from the node to the base station. Each node contains a message that sends it to the cluster head and the cluster head sends the received message to the base station. Since, the main path (spine) in the proposed method is at the optimal point, the clusters can communicate more quickly with it. In Fig. 8, the time of the proposed method is compared with the previous methods. In this scenario, the network size is constant and the number of nodes varies. The speed of the message sending is considered one meter per second.

As can be seen from Fig. 8 in the proposed method, messages have been sent more faster to the central station.

Fair load distribution: The next criterion is to evaluate the proposed method of distributing load and network traffic on links between cluster heads. The

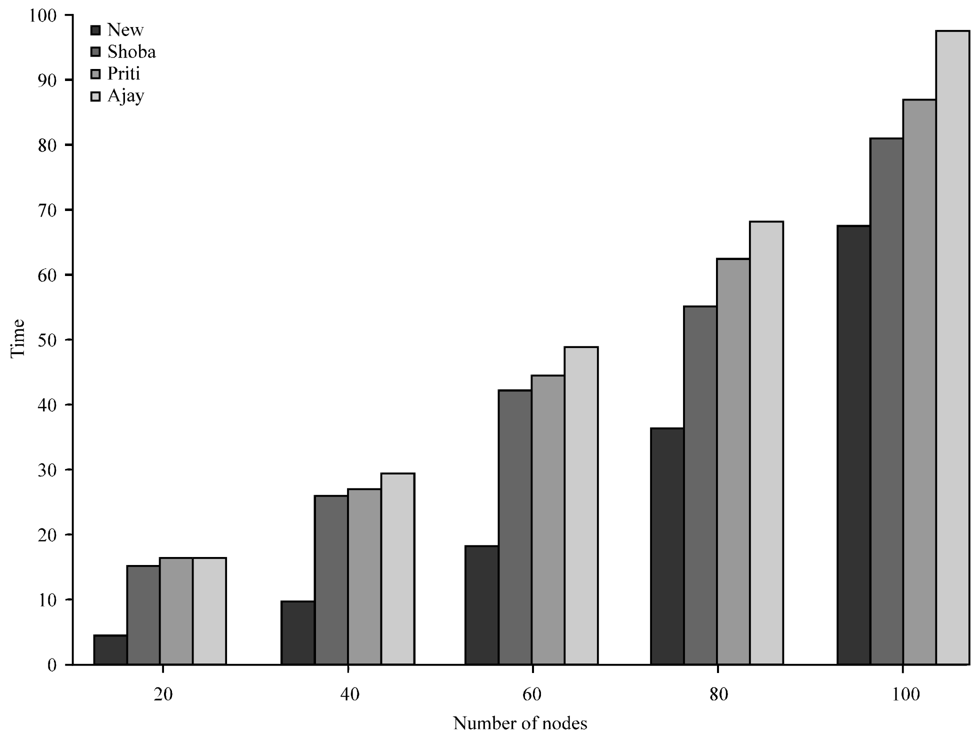

Fig. 8: Comparing the spent time for sending a message to the base station

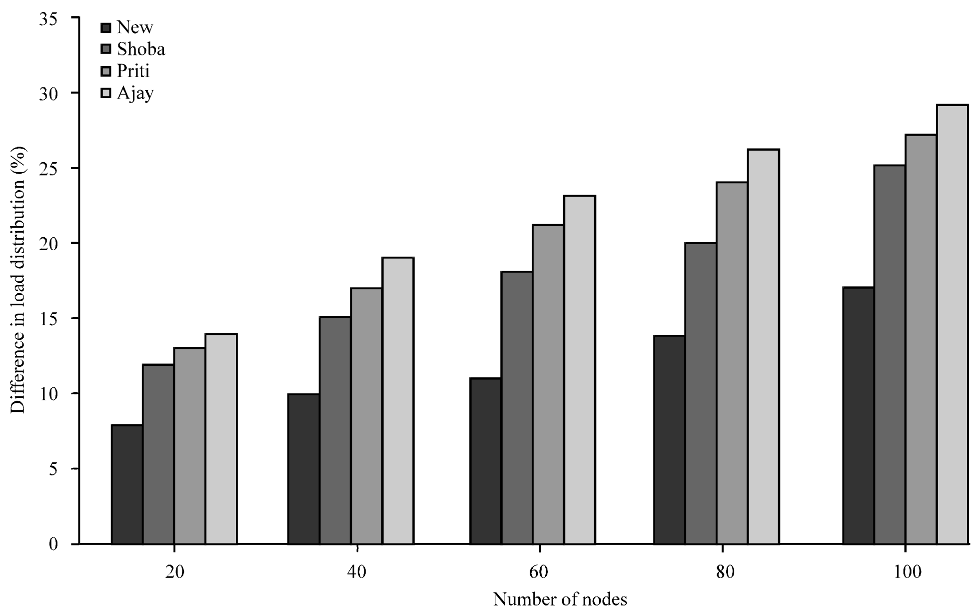

Fig. 9: Load distribution with the variable number of nodes

more traffic is distributed equitably, the better network performance and network lifespan. The simulation scenario in this case is as follows:

- Variable number of clusters
- Fixed number of packets
- The sender and receiver are selected randomly

The average load distribution on the links is shown in Fig. 9. To calculate the load distribution, the difference between the maximum and minimum load is displayed.

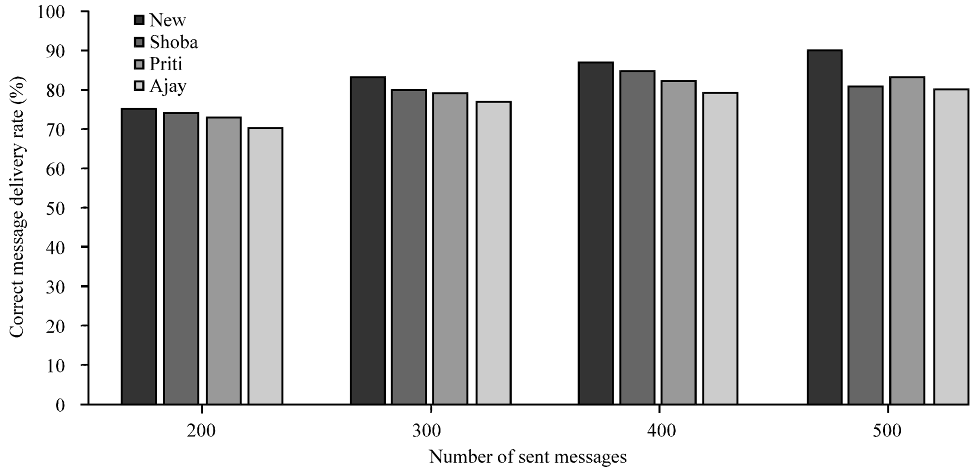

Fig. 10: Correct sending rate of pockets in the network

The simulation results show that proposed method has a better distribution because it operates in each step optimally and this results in fair traffic distribution.

Correct sending rate: One of the most important evaluation criteria is to examine the correct packet sending rate in the network. In order to simulate and evaluate this criterion, the fixed size of network is considered and in each step the number of sent packets is increased. The result is shown in Fig. 10. In Fig. 10 shows, the correct delivery rate has improved in the proposed method.

CONCLUSION

The simulation results show that the proposed method improved energy saving which results in increasing the life of the network. Also, scheduling has made the messages delivery to the base station and reducing the traffic.

REFERENCES

Abbasi, A.A. and M. Younis, 2007. A survey on clustering algorithms for wireless sensor networks. Comput. Commun., 30: 2826-2841.

Akkaya, K. and M. Younis, 2005. A survey on routing protocols for wireless sensor networks. Ad Hoc Networks, 3: 325-349.

Baranidharan, B. and B. Santhi, 2016. DUCF: Distributed load balancing unequal clustering in wireless sensor networks using fuzzy approach. Applied Soft Comput., 40: 495-506.

Kandpal, R., R. Singh and H.L. Mandoria, 2015. Analysis on enhancements in LEACH protocol for WSN. SSRG Intl. J. Comput. Sci. Eng., 2: 1-8.

Kim, J.M., S.H. Park, Y.J. Han and T.M. Chung, 2008. CHEF: Cluster head election mechanism using fuzzy logic in wireless sensor networks. Proceedings of the 10th International Conference on Advanced Communication Technology (ICACT'08) Vol. 1, February 17-20, 2008, IEEE, Gangwon Province, South Korea, ISBN:978-89-5519-135-6, pp: 654-659.

Mazumdar, N. and H. Om, 2017. Distributed fuzzy logic based energy-aware and coverage preserving unequal clustering algorithm for wireless sensor networks. Intl. J. Commun. Syst., Vol. 30, 10.1002/dac.3283

Mirzaie, M. and S.M. Mazinani, 2017. Adaptive MCFL: An adaptive multi-clustering algorithm using fuzzy logic in wireless sensor network. Comput. Commun., 111: 56-67.

Natarajan, H. and S. Selvaraj, 2014. A fuzzy based predictive cluster head selection scheme for wireless sensor networks. Proceedings of the 8th International Conference on Sensing Technology, September 2-4, 2014, Liverpool F.C., Liverpool, UK., pp: 560-566.

Ran, G., H. Zhang and S. Gong, 2010. Improving on LEACH protocol of wireless sensor networks using fuzzy logic. J. Inform. Comput. Sci., 7: 767-775.

Rana, S., A.N. Bahar, N. Islam and J. Islam, 2015. Fuzzy based energy efficient multiple cluster head selection routing protocol for wireless sensor networks. Intl. J. Comput. Network Inf. Secur., 4: 54-61.

Soro, S. and W.B. Heinzelman, 2009. Cluster head election techniques for coverage preservation in wireless sensor networks. Ad Hoc Networks, 7: 955-972.

Van Greunen, J. and J. Rabaey, 2003. Lightweight time synchronization for sensor networks. Proceedings of the 2nd ACM International Workshop on Wireless Sensor Networks and Applications, September 19, 2003, San Diego, CA., USA., pp: 11-19.