JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

FCM Based Efficient Clustering Algorithm for Wireless Sensor Network

Kumari Nishu (*M.Tech CSIT Student)*
Dept. of Computer Sciences & Information Technology
Vaugh Institute of Engg. and Technology, SHUATS, Allahabad, India
kumarinishu.er@gmail.com

Vivek Prakash Singh (*M.Tech CSIT Student)*
Dept. of Computer Sciences & Information Technology
Vaugh Institute of Engg. and Technology, SHUATS, Allahabad, India
vivek777prakash@gmail.com

Dr. Hari Mohan Singh *(Assistant Professor)*
Dept. of Computer Sciences & Information Technology
Vaugh Institute of Engg. and Technology, SHUATS, Allahabad, India
Harimohansingh@shauts.edu.in

Abstract **- Wireless sensor network (WSN) brings a new paradigm of real-time embedded systems with limited computation, communication, memory, and energy resources that are being used for huge range of applications where the traditional infrastructure-based network is mostly infeasible. The sensor nodes are densely deployed in a hostile environment to monitor, detect, and analyze the physical phenomenon and consume considerable amount of energy while transmitting the information. It is impractical and sometimes impossible to replace the battery and to maintain longer network life time. So, there is a limitation on the lifetime of the battery power and energy conservation is a challenging issue. Hierarchical or cluster based routing methods are well-known techniques with special advantages related to scalability and efficient communication. Clustering technique is also utilized as an energy-efficient routing in Wireless Sensor Networks where the nodes with the highest residual energy can be used to gather data and send the information. We proposed and analyzed a Fuzzy C-Means (FCM) based Efficient clustering algorithm for WSNs. A comparison among FCM has been done. Simulation result shows that proposed protocol can reduce energy consumption and improve the network lifetime.**

Keywords—*Wireless Sensors Network, Fuzzy C-Means Clustering*

I. INTRODUCTION

Wireless Sensor Networks (WSNs) consist of hundreds to thousands of tiny sensor nodes equipped with sensing, data processing, and communication units [1]. These sensor nodes are used to collect information about ambient environment, eg., temperature, humidity, light, vibration, acoustic, etc. The measurement is pre-processed and useful data is transferred to other sensor nodes or to a Base Station (BS). Due to these capabilities, WSNs can be applied in various potential applications such as target tracking, habitat monitoring, healthcare monitoring, surveillance, etc. However, to make WSNs feasible to be employed, a number of requirements in the design and operation of the network need to be satisfied. Since sensor nodes are powered by limited energy source, energy conservation is commonly considered the most key challenge in order to guarantee the connectivity of the network and extend the lifetime of the sensor nodes, especially when the deployment field is inaccessible and battery cannot be replaced. Even if unlimited energy source like solar, wind, etc., is utilized, efficient operation of sensor networks is necessary because of the fluctuation and intermittent nature of these sources. It is also recognized that usually communication task consumes the most energy during the network operation. Many routing protocols have been proposed to obtain efficient-energy communication for the WSNs for recent years. In [2], routing techniques are classified into three categories based on network structure which are flat, hierarchical and location-based routing protocols:- In flat routing protocol, all sensor nodes are typically assigned equal roles and has the same functionality; sensor nodes collaborate to perform the sensing task as well as communication task.-In hierarchical protocol or cluster based routing method, there are usually two types of sensor node: Cluster Head (CH) and non-CH nodes. Non-CH nodes mainly carry out sensing task and only send the information to the CH when necessary, while CHs collect data from other nodes and send to the end users.- In location-based protocol, routing data needs the information of sensor nodes locations in the deployed field.

Among these, hierarchical or cluster-based protocols are well-known techniques with special advantages related to scalability and efficient communication. The concept of hierarchical routing is also utilized to perform energy-efficient routing in WSNs. By grouping the nodes into clusters as shown in Fig.1 with the assistance of data aggregation and fusion techniques, efficient usage of energy resource is obtained because the overall amount of data transmitted to the BS is significantly decreased, intra-cluster communication enables to reduce the transmission distance of non-CH nodes and then reduce energy consumption. Furthermore, duty cycling of the non-CH nodes can be carried out by the CH within the cluster, therefore, member nodes are allowed to enter sleep mode for a longer time. However, gathering and processing information as well as transferring data to the BS cause higher energy dissipation at the CH, and thus lead to shorter lifetime of the CH and the network connection cannot be maintained anymore. In order to balance the energy usage of sensor nodes and CH, the solution of rotating CH roles among sensor nodes within the cluster is popularly used. The selection of CH is based on residual energy of the sensor nodes, any nodes with residual energy higher than a threshold can be chosen to become CH.

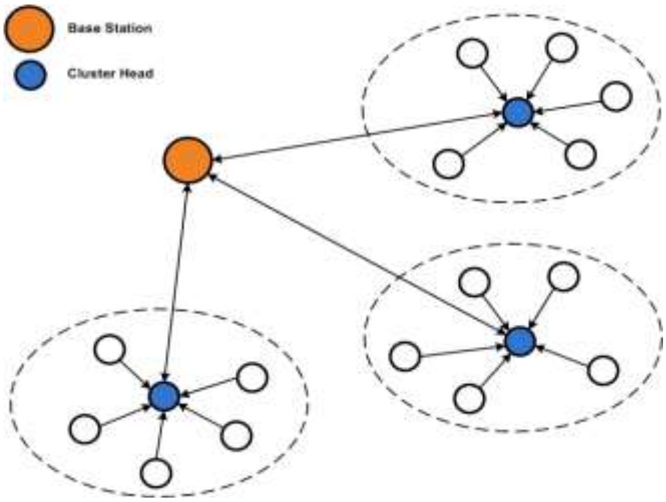


Fig. 1. Architecture of the Hierarchical Network

Low Energy Adaptive Clustering Hierarchy (LEACH) [3] is a typical cluster-based protocol using a distributed clustering formation algorithm. The cluster heads are selected with a predetermined probability, other nodes choose the nearest cluster to join, based on the strength of the advertisement message they received from the cluster heads. After forming the clusters, cluster heads compress data arriving from the sensor nodes and send an aggregated packet to the BS in order to reduce the amount of information sent to the BS. To avoid intracluster and intercluster collision, Time Division Multiple Access (TDMA) or Code-Division Multiple Access (CDMA) can be utilized, thus a better efficient communication is achieved.

II. PRELIMINARIES

We first present the assumption and model of the network under consideration.

A. Assumptions: We consider scenario of application in which sensor nodes are deployed randomly in order to continuously monitor the environment. The information collected by sensor nodes is sent to a base station located far from the deployment field. Each sensor nodes can operate either in sensing mode to monitor the environment parameters and transmit to the base station or cluster head mode to gather data, compress it and forward to the BS.

III. FCM CLUSTERING PROTOCOLS

FCM clustering protocols is centralized clustering algorithms, the base station computes and allocates sensor nodes into clusters according to the information of their location and the cluster head is assigned to the node having the largest residual energy. We consider a network of N sensor nodes which is partitioned into c clusters: $C_1;C_2; :::;C_c$. The purpose of the cluster formation in this protocol is to minimize the following objective function:

$$J_m = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{m} \|x_i - v_i\|^2 \quad \ldots\ldots\ldots\ldots\ldots(1)$$

$$\sum_{i=1}^{c} u_{ij} = 1 \ \forall j = 1,\ldots\ldots\ldots, n \quad \ldots\ldots\ldots (2)$$

In equation (1), c is the number of the clusters, n is the data number, m is the fuzzy rate (m is a real number more than 1 and in most cases m is elected as 2), x_i is the kth data, v_i is the ith cluster, u_{ij} is the membership grade of the jth data in ith cluster. In all clustering algorithms of Fuzzy clustering, the number of the clusters is defined and the clusters take their primary value. Then, using equations (2) and (3), the values are updated and this procedure continues till the difference between the data is fixed.

$$V_i = \sum_{j-1}^{n}(u_{ij})^m x_j \,/\, \sum_{j-1}^{n}(u_{ij})^m \quad \ldots\ldots\ldots\ldots\ldots (3)$$

$$U_{ij} = 1/ \sum_{i-1}^{c}\left(\|x_j-v_i\| / \|x_j-v_i\|^{2/(m-1)}\right) \quad \ldots\ldots (4)$$

FCM clustering is usually used for finding the structure in the data, which are not labelled. In this situation, it is trying to put the data into different clusters to reach a target function, which holds the minimum value. So, selecting the distance Fuzzy functions, it is possible to cluster the data optimally.

The distance between the sensor node and the center point is Euclidean distance. By achieving minimization of the spatial distance, the energy balance among sensor nodes is optimized.

Our FCM based Efficient clustering algorithm include 3 phases: clustering calculation, Efficient selection of cluster head and data transmission. The operation of the protocol is partitioned into rounds. In each round, the cluster heads collect data from all cluster members and transfer to the BS.

A. Clustering calculation

We consider the application scenario that N sensor nodes are deployed randomly into a field with an area of $M*Mm^2$. After being spread out, these sensor nodes send a message to the base station with the information of their geographical location; based on this information the base station will calculate the cluster centers and allocated sensor nodes into cluster using FCM algorithm. FCM algorithm is first proposed by Bezdek [11] to be used in cluster analysis, pattern recognition, image processing. In the case of our application, FCM algorithm is applied to cluster the sensor nodes. Each node is assigned a degree of belonging to cluster head rather than completely being a member of just one cluster. Therefore, the nodes close to the boundary of a cluster may become members of the cluster with a degree approximating the degree of belonging to the neighbor clusters.

The convergence is achieved when the difference between the coefficients in two iterations is less than a threshold or a large number of iterations is reached.

After forming the clusters, the BS chooses the nearest nodes to cluster centers to become CH. Once the cluster creation is complete, base station send the information of the cluster head and to which cluster a node belongs to all of the nodes.

To identify the number of clusters, the following formula [5] can be used:

$$\mathrm{Copt} = \sqrt{N} \therefore \sqrt{2 * \pi\sqrt{\mathrm{EFs} \therefore \mathrm{Emp}}}\,\mathrm{M} \therefore \mathrm{d2toBS}$$

B. Efficient Selection of Cluster head

After the cluster is created, the non-cluster head nodes send data toward the base station through the cluster heads. The process of selecting clusters is repeated every round of exchanging data among sensor nodes. Only at the first stage, the cluster head of each cluster is chosen by the base station; after that the current cluster head makes decision of selecting which node will become the cluster head at the next round. During the transmission from the sensor nodes to CH, residual energy of each nodes are attached to the data packet, this information assists the CH choose the node with the highest residual energy and nearest to the cluster center to be cluster head at the next round. Based on the number of the alive nodes within the cluster, the new CH creates a TDMA schedule to allocate the time when cluster members can transmit.

It is a hierarchal protocol where sensor nodes independently elect itself as a center point based on probability of membership function to become a center point. Such selection of centroid (center point) is normally placed in center of the cluster as shown in the **Figure 2.**

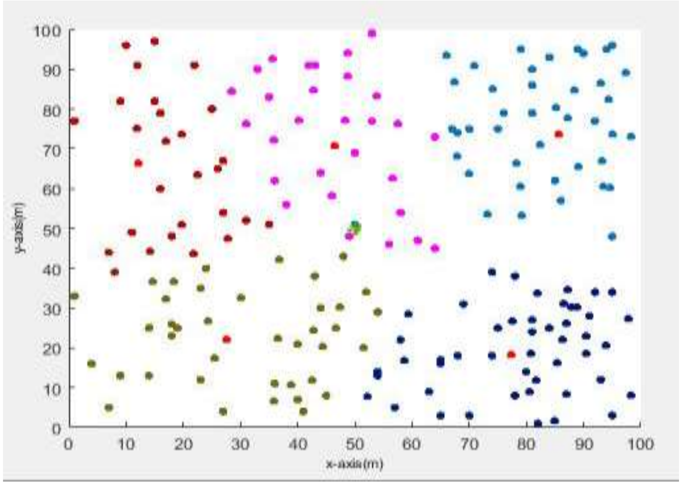


Fig 2: Network clustering using FCM

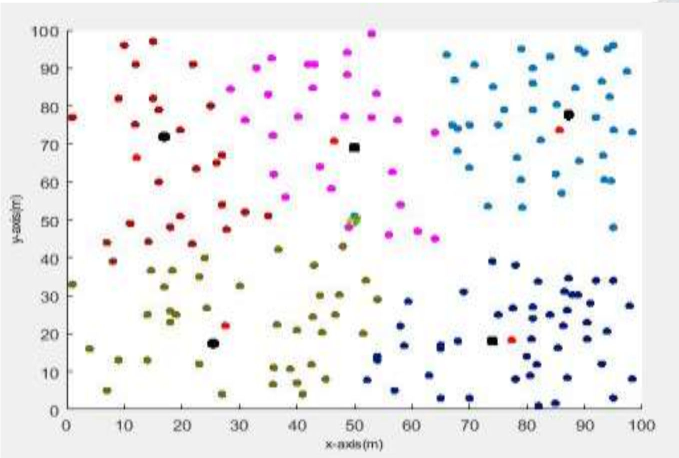


Fig 3: Efficient selection of Cluster Heads

The step by step working procedure of the proposed algorithm is depicted in **Figure 3.3** and algorithm is shown in **Figure 3.** Initially optimal value of k (number of clusters) is determined based on number of live sensor nodes and area of the observation field. After that Fuzzy C-means algorithm is used to find the centroid of all the clusters(k). After that sensor nodes are selected around each centroid of the clusters with higher energy and minimum distance. Then Data Aggregation is performed with the selected cluster head. The important point is that after this when residual energy of selected CHs is less than threshold energy, next node around the centroid will be selected. This process is repeated 4-6 times and after that re-clustering performed to get new centroid at different locations.

1. Algorithm for Re Clustering and Re-selection of Cluster Head

// Initialization

1. Determine k-1 number of Clusters

2. Apply FCM Algorithm to perform network clustering and identify centroid of each cluster (k)

3. For each cluster

Select *n* no of nodes with high energy and minimum distance around each centroid and insert all these into a Priority Queue, PQ

4. Check the PQ, if PQ is Empty, Then

4.1 POP next cluster head (s) from the Priority Queue

4.2 perform Data Aggregation and data sinking at the base station

5. If (residual energy(CH) < Threshold energy), then go to step 1 until sensor nodes are live in the network.

IV. SIMULATION RESULT AND DISCUSSION

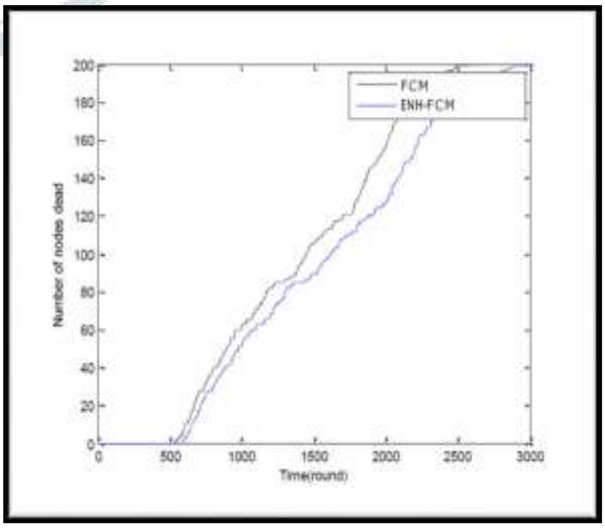
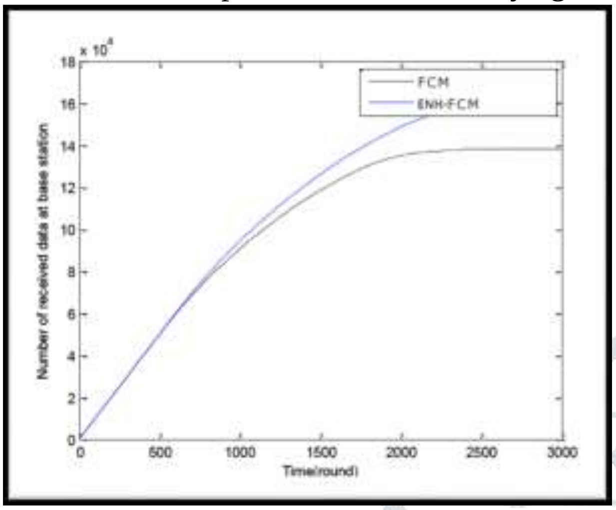


Fig 4. No. of dead nodes FCM v/s Enhanced FCM

Figure 4 is the comparison of FCM algorithm and Proposed FCM based efficient clustering algorithm. The figure represents a graph for probability at 0.5. The number of nodes chosen for the graph are 200 nodes for 3000 rounds. We can see in the FCM protocol nodes start dying at the



round no. 500 however, in Proposed algorithm nodes are dying at the round no.650. This shows that the proposed protocol is better than existing protocol as it improves the network lifetime and is energy efficient.

Fig 5: Number of received data at base station FCM v/s Enhanced FCM

Figure 5 is the data package comparison of FCM algorithm and Proposed FCM based efficient clustering algorithm. From the figure, in the beginning, there is no much disparity about the amount of data package between the two protocols. However, with the time running, we found that the amount of data package sent by the Proposed algorithm is gradually larger than that of FCM. This graph is represented for probability at 0.5 for 200 nodes for 3000 rounds, which as a result, we see no. of received data packet at the base station is same at the starting rounds but after 600 rounds there are some difference in FCM and in Proposed FCM based efficient clustering algorithm.

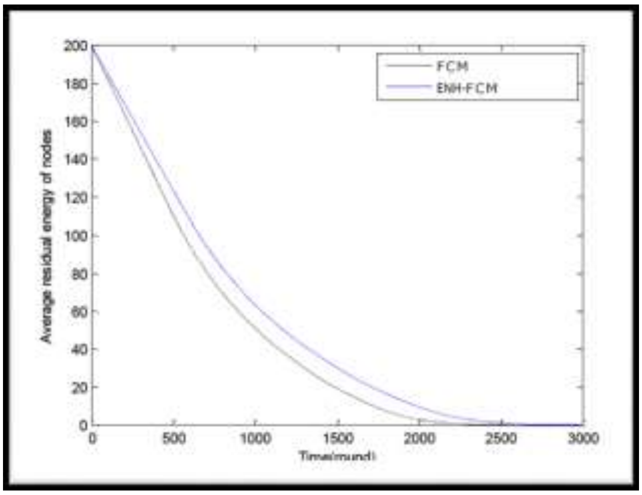


Fig 6. Average residual energy of nodes FCM v/s Enhanced FCM

The comparison of residual energy between the FCM and Proposed FCM based efficient clustering algorithm represented in this figure 6 for probability at 0.5 for 200 nodes for 3000 rounds. This shows that the average residual energy of nodes in case of FCM algorithm is lesser than average residual energy of nodes in proposed algorithm, which means proposed protocol is more efficient than existing protocol.

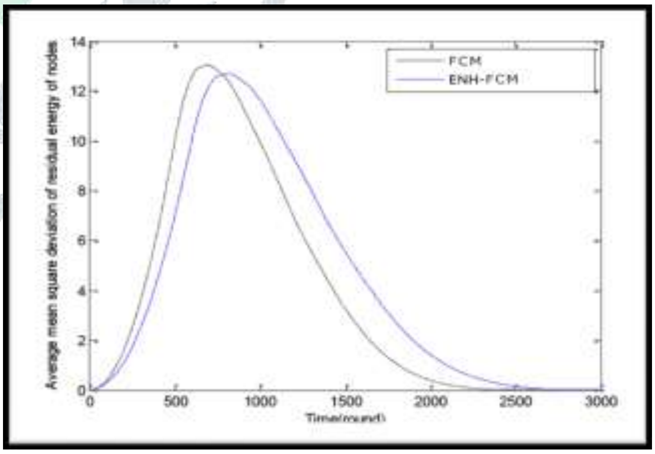


Fig 7. Avg. mean square deviation of residual energy of nodes FCM v/s Enhanced FCM

The graph shows the comparison between FCM and Proposed FCM based efficient clustering algorithm and the probability is set up at 0.5 for 200 nodes for 3000 rounds. The average mean square deviation of residual energy of nodes is highly consumed at the 700 rounds and finished at the 2400 rounds whereas in Proposed algorithm average mean square deviation of residual energy of nodes are highly consumed at the 900 rounds and finished at the 2600 rounds. This thus shows that the proposed protocol is alive for long time than FCM thus improving the network lifetime.

V. CONCLUSION

In this paper, we present a centralized cluster-based protocol, FCM based Efficient clustering algorithm for WSNs. The problems that we were facing with the existing FCM clustering protocol consider positioning of CH, probability based selection of CH in heterogeneous network and uneven distributed of CH. Therefore, after the detailed study of FCM protocol, we proposed FCM based efficient clustering algorithm. This proposed algorithm shows better result as the number of packets received in proposed work are 1,70,000 whereas in FCM algorithm, it is 1,31,000 as there is a difference of 39000 packets and the proposed work increases the network lifetime by approximately 10% as there is an increment of 3000 rounds in FCM based efficient clustering algorithm.

This is more beneficial and scalable than the existing FCM algorithm. Thus, the work discusses the election of cluster head to analyses the performance of the network. Based on the information received from the nodes and taking into account all the factors, the modifications have been proposed in FCM based efficient clustering algorithm to improve its performance to prolong the network protocol namely FCM based Efficient Clustering algorithm for WSNs provides a better communication than conventional FCM algorithm lifetime and balancing the energy consumption in network. With the facilitation of data aggregation, cluster head rotation and TDMA schedule techniques in clusters, energy consumption is balanced among all the sensor nodes and the amount of data transmitted to the BS is reduced remarkably. Our simulation results show that by applying FCM based Efficient clustering algorithm WSNs the power consumption is reduced and the life time of the network is extended significantly when compared with FCM.

REFERENCES

[1] I. F. Akyildiz and W. Su and Y. Sankarasubramaniam and E. Cayirci, Wireless sensor networks: a survey, Computer Networks, vol 38, pp 393-422, 2002.

[2] Jamal N. Al-Karaki and Ahmed E. Kamal, Routing Techniques in Wireless Sensor Networks: A Survey, IEEE Wireless Communications, pp.6-28, 2004.

[3] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, Energy efficient communication protocols for wireless sensor networks, In Proceedings of the 33rd Annual Hawaii International Conference on System Sciences (HICSS), Hawaii, USA, January 2000.

[4]Bezdek, J.C, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.

[5]Ahmed, Mohamed N.; Yamany, Sameh M.; Mohamed, Nevin; Farag, Aly A.; Moriarty, Thomas (2002). "A Modified Fuzzy C-Means

Algorithm for Bias Field Estimation and Segmentation of MRI Data" (PDF). IEEE Transactions on Medical Imaging. 21 (3):193199. CiteSeerX 10.1.1.331.9742. doi:10.1109/42.996338. PMID 11989844.

[6]Alireza, Kashani; Kashani, Amir; Milani, Nargess; Akhlaghi, Peyman; Khezri, Kaveh (2008). Robust Color Classification Using Fuzzy Reasoning and Genetic Algorithms in RoboCup Soccer Leagues. Robocup. Lecture Notes in Computer Science. 5001.pp. 548–555. doi:10.1007/978-3-540-68847-1_59. ISBN 978-3-540-68846-4.

[7]Akyildiz. I., Su.W. ,Sankarasubramaniam. Y and Cayirci. E., (2002), “A survey on Sensor t. 40 (2016) 495–506.

Networks,” IEEE Communications Magazine, Vol. 40, No.8, pp. 102-114.

[8]Ahmed Ismail Shihab (2000). "Fuzzy clustering algorithms and their application to medical image analysis". CiteSeerX 10.1.1.116.3037.

[9]Banerjee, Tanvi (2014). "Day or Night Activity Recognition From Video Using Fuzzy Clustering Techniques". IEEE Transactions on Fuzzy Systems. 22: 483–493 – via IEEE.

[10]B. Baranidharan , B. Santhi ,DUCF: distributed load balancing unequal cluster- ing in wireless sensor networks using fuzzy approach, Appl. Soft Compu