

A Hybrid Approach of Fuzzy C-Means Clustering and Neural Network to Make Energy-Efficient Heterogeneous Wireless Sensor Network

Amit Kr. Kaushik

Manav Rachna University, Faridabad, Haryana, India

Article Info

Article history:

Received Aug 31, 2015

Revised Nov 20, 2015

Accepted Dec 12, 2015

Keyword:

FCM

Neural network

Sensor node

Wireless sensor network

ABSTRACT

The Wireless sensor network has been highly focused research area in recent times due to its wide applications and adaptability to different environments. The energy-constrained sensor nodes are always under consideration to increase their lifetime. In this paper we have used the advantages of two approaches i.e. fuzzy c-means clustering and neural network to make an energy efficient network by prolonging the lifetime of network. The cluster formation is done using FCM to form equally sized clusters in network and the decision of choosing cluster head is done using neural network having input factors as distance from basestation, heterogeneity and energy of the node etc. Our Approach has successfully increased the lifetime and data capacity of the network and outperformed different approaches applied to the network present in literature.

*Copyright © 2016 Institute of Advanced Engineering and Science.
All rights reserved.*

Corresponding Author:

Amit Kr. Kaushik

Manav Rachna University

Faridabad, Haryana, India

Email: amitkaushik@mru.edu.in

1. INTRODUCTION

Wireless Sensor Network (WSN) is a self-configuring network which is highly adaptable to different environment scenarios. They are highly useful network in human-inaccessible environment for monitoring purposes. It is composed of a set of sensor nodes (which are also called motes). The lifetime of WSN is always a topic of research because the sensor nodes are constrained in term of battery life and it is not feasible to recharge the battery at regular basis because of their deployment in remote and hostile places. So we need those routing protocols or approaches which can maximize the data capacity and also prolong the lifetime of sensor node. So the challenge is to develop low-power communication approaches with low-cost on-node processing and self-organized connectivity/protocols. Several protocols were developed to make the communication energy-effective to increase lifetime of the networks. The wireless sensor networks can be categorized into homogenous and heterogeneous networks according to the type of nodes network is using. The homogenous network is the network where all nodes have equal energy and all having equal probability to become cluster head. The heterogeneous networks have different kind of nodes having different energy and probabilities of becoming clusterheads. In this paper we will work on heterogeneous networks. Different protocols have been for homogenous and heterogeneous networks. These protocols used different approaches to improve the communication and transmission of the packets in the network. But they lack different things such as to have an optimal distribution of nodes in network, equally sized clusters, role of residual energy in election of cluster head etc. The non-optimal distribution often makes the dissipation of the energy more rapidly in the cluster than the optimal distribution in the network. The unequal-sized clusters lead to the lesser data-capacity of the network as compared to the equal-sized clusters. Different protocols have been made to make the routing efficient such as LEACH [1, 2], [3], SEP [4], DEEC [5] etc. These all

protocols were based on concept of clustering in the network to conserve the energy of the nodes. This was firstly applied in the LEACH protocol. LEACH protocol was developed as clustering based routing protocol in which the cluster-head is elected using probability-based thresholding mechanism. LEACH [1] elects the cluster head by using a thresholding expression as

$$T(i)=p/(1-p(r\text{mod}1/p)) \quad (1)$$

Where i is node, p is the probability of choosing the cluster head and r is round number. A random no 's' is generated between 0 and 1 and is compared with $T(i)$. If $T(i)$ is greater than s then node i become cluster head otherwise not. But as it was based on probability based cluster-head election which can lead to non-optimal and unequal-sized cluster and also did not take account the residual energy factor in determination of cluster heads.

The [3] was to make improve the LEACH protocol by removing its discrepancies by taking the factor of residual energy in determination of cluster head.

The SEP [4] introduced the heterogeneity in the network by making two types of nodes i.e. normal and advanced having different energies and probabilities to be getting elected as cluster head. This heterogeneity is extremely helpful to make the lifetime of network better [8].

The DEEC [5] used the ratio of residual and average energy in calculating the threshold for the election of cluster head which improved the SEP further. But this also framed non-optimal cluster in the network.

The EDEEC [6] increased the heterogeneity by one more node i.e. normal, advanced and super node in the network. The normal has lowest energy and probability to become cluster head in the cluster. This algorithm combined aspects of heterogeneity and ratio of residual and average energy per round to improve the stability period and lifetime of the network.

TADEEC [7] routing algorithm used the best of heterogeneity and TEEN [8] to increase the chances of high energy nodes to become cluster heads more than the low energy nodes. Also it removed the redundancy of the data at cluster heads and sensor nodes by implementing the two thresholds i.e. hard and soft threshold so that same data don't have to be propagated to the cluster heads and basestation [9]. This used four heterogenous nodes in the network i.e. normal, advanced, super and supadvanced having their increasing energies and probabilities of choosing as cluster heads in the network. This outperformed the LEACH, SEP, DEEC and EDEEC in terms of stability period, lifetime and data capacity of the network.

The [10] used the well known clustering technique based on fuzziness i.e. fuzzy c-means clustering to frame the clusters in the network. The optimal and equalized clusters are framed by the FCM algorithm. The Fuzzy c-means clustering basically associates to each node a corresponding membership value to each cluster number which have to be formed. The clusters are formed according to the highest degree of belongness (also known as degree of relationship) to a particular cluster number. After the clusters are formed the cluster heads are chosen based on the maximum residual energy node among the cluster members and also its proximity to corresponding cluster center. The cluster heads further communicate the data collected through the cluster member sensor nodes to the basestation.

[11] Realized the fuzzy c-means clustering algorithm on wireless sensor network. This paper realized the 50 sensor nodes in the hardware having TinyOS operating system. This hardware implementation successfully proved the effectiveness of FCM in wireless sensor network and this implementation outperformed the LEACH.

[12] Proposed a fuzzy-based simulation system for sensor networks and calculation of the lifetime of a sensor by considering the remaining battery power, sleep time rate and transmission time rate.

[13] Used the fuzzy logic mechanism in heterogeneous network where it applied the energy, heterogeneity and proximity to basestation factors in fuzzy inference system in determination of cluster heads.

2. RESEARCH METHOD

In this paper, we have applied two methods for betterment of the wireless sensor networks. The fuzzy c-means clustering which has been used to cluster the sensor nodes and neural networks which has been used to take the decision of election of cluster head among the cluster members in the cluster. The FCM algorithm is illustrated in figure 1. The degree of belongness of node i to cluster j is given by u_{ij} . For each node i and cluster j , u_{ij} is calculated and maximum value is selected to which the node belongs. In a cluster the cluster head is the dominant node as the further communication to the basestation would be done by the cluster head. Several factors have to be considered while electing the cluster heads. In a heterogeneous

network the sensor nodes have different energies and probabilities of getting elected as cluster heads. The heterogeneity factor plays a important role in election of the cluster head.

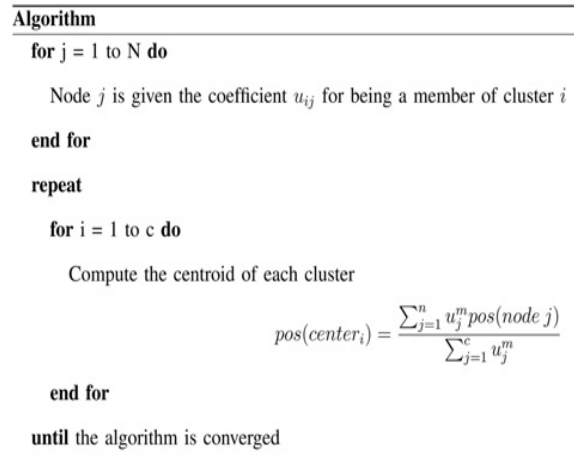


Figure 1. The FCM algorithm

In this paper the heterogeneity factor is calculated from the probability of getting elected as cluster head and energy of the node, So higher energy node would have larger probability of elected as cluster head and ultimately would have higher heterogeneity factor. The residual energy also has a significant role in election of cluster head. The residual energy is the energy left in the node after communication has been done. A node having high residual energy would have more chances to be chosen as cluster head than low energy node as communication needs dissipation of energy. The Energy required to send the packets is proportional to the distance between the sensor nodes. So the node which is at smallest distance from the basestation among all cluster nodes should be elected as cluster head in the cluster. So in this paper these three factors i.e. Heterogeneity of the node, residual energy of the node and distance to the basestation are considered while electing the cluster head. The neural network works as artificial human brain which basically take the decision based on the given inputs and applied weights. The weight applied to an input signifies the importance of that input in taking the decision or calculating the output of the network. In this approach, the neural network has taken these factors as input for each node in the corresponding cluster and applied the weights on these factors and sum of product of input factors and weights take decision of electing as cluster head or not. The inputs has different roles in calculation of the output as one input has positive role while other has negative role. The weights are decided based on the how positively or negatively factor is affecting the decision or output of the network. The weight is negative when its value negatively affects the decision of choosing as cluster head such in case of distance of node from basestation as we have to suppress the role of product of input and weight in summation and positive in case of heterogeneity and energy as their high values lead to better cluster heads.

3. RESULTS AND DISCUSSION

In this project, we have simulated our approach in matlab. Area is taken as 100*100 square meters where the basestation is located at (50,50). The sensor nodes are randomly distributed in the field around the basestation. The localization technique has not been used in our network. The nodes are distributed manually around the basestation and locations of the sensor nodes have been known in advance during the network processing. The nodes form the cluster and the elected cluster head among cluster members aggregate the data from all the members and send it further to the basestation. We have used the heterogeneous network where we have four types of nodes i.e normal, advanced, super and supadvanced. These nodes have energies in this way [7]:

normal= E_0 , advanced = $E_0(1+a)$, super= $E_0(1+b)$

supadvanced= $E_0(1+c)$ where $a < b < c$

The distribution of nodes is done using m and m_0 [7]. As

$normal = n * (1 - m)$;

$advance = n * m$; $super = (m_0 * m * n) / 2$; $supadvanced = (m_0 * m * n) / 2$;

We have used two mobility models i.e. free space and multipath model as used in [14].

$$E_{tx} = E_{elec} * l + E_{fs} * l * d^2 \quad \text{if } d > d_0 \quad (2)$$

$$E_{tx} = E_{elec} * l + E_{mp} * l * d^4 \quad \text{if } d < d_0 \quad (3)$$

Eq. (1) is applied in free space network model and Eq (2) is applied in multipath network model where d_0 is calculated as:

$$d_0 = \sqrt{\frac{E_{fs}}{E_{mp}}} \quad (4)$$

And l is the size of the message, E_{tx} is the energy required to transmit the message and d can be the distance between the nodes or node and basestation. The E_{fs} is the amplification energy for the free space model while E_{mp} is the amplification energy in multipath network model. The heterogeneity of each node is given as

normal: $P * E(i) / ter$;
 advanced: $P * (1+a) * E(i) / ter$;
 super: $P * (1+b) * E(i) / ter$;
 supadvanced: $P * (1+c) * E(i) / ter$;
 where
 $ter = (1+a * m - m * m_0 * (a - ((b+c)/2)))$;

In these $c > b > a$; where $c=1$; also

The simulation parameters have been in table 1. The table lists out all the values for the initial energy, the nodes percentage in network, the size of packets etc.

The results of our approach has been shown in figure 3, 4 and 5. The sensor nodes form clusters and cluster members send their sensed data to their corresponding cluster heads. The cluster heads then further give this data to the basestation. The data sending has been monitored in two ways i.e. the data sent by the nodes per round to the cluster heads and data sent per round by the cluster heads to the basestation. These throughputs have been shown in figure 6 and 7. In this paper we have compared our proposed approach to the two different approaches presented in literature survey i.e. FCM only and LEACH algorithm. We have taken different parameters on which we have compared our algorithm with these present approaches. These are alive nodes i.e. nodes which are alive in the networks and it has been traced in each round, dead nodes which have died in the networks in each round, data capacity of the clusters and throughput of the network i.e. the packets sent per round to the basestation and cluster heads. The analysis taking different parameters based on figures given has been done in the following points.

1. In figure 3, we have plotted the data that has been sent by the cluster members to their corresponding cluster heads with the increasing rounds. We have observed that data has highly aggregated at cluster head in our approach. These data aggregation is dependent upon the liveliness of the cluster member nodes and Our approach is trying to not elect the low energy node as cluster head as major communication is done by the this node. The throughput for cluster level data has also been observed and plotted in Figure 6. This throughput is the how many packets being generated by a single node in a network. This parameter show the average data send by a cluster member node as part of a networks. In both observations we have outperformed the other two techniques employed in past.
2. The second parameter being taken for checking the efficiency of our approach is to count how many nodes are still alive in the network as efficient protocol always maintain high alive in network. This parameter is calculated on level of energy node are having as node having no battery or sending or receiving capabilities are considered dead as they are able to do further communication. We can see that percentage of alive nodes in our algorithm outperformed the FCMonly and LEACH algorithms in figure 4. In our algorithm even more than 50% nodes are alive in 10k rounds.
3. The figure 5 show the same case of dead nodes so only 20-30 % nodes have been died in 10k rounds where in LEACH all nodes have been dead and FCMonly approach has 60-70% nodes have been dead.
4. The data sent to the basestation by the cluster heads has also been monitored and it is observed that the throughput of network using our approach is higher than other approaches. The cluster heads are aggregating the data collected from 10 cluster members of their corresponding cluster.
5. The optimal and similar sized clusters capability of our approach can be checked from figure 8. The nodes in each clusters have been plotted as number of nodes in each cluster in similar in nature and also non-optimality issue has been resolved at higher extent. In our network we have taken 10 clusters and FCM is used to assign the nodes to each cluster. Very few clusters are showing non-similarity in count

of their cluster member nodes. In our network in each cluster has been assigned average 8-10 cluster nodes

Table 1. Simulation Parameters for the network

Parameters	Value
Area	100*100 square meters
Basestation	(50,50)(in m)
Initial Energy	0.5J
Transmission Energy	50nJ/bit
Receiver Energy	50nJ/bit
No Of Nodes	100
Free space Amp Energy	10pJ/bit/m ²
Multipath Amplification Energy	0.0013pJ/bit/m ⁴
Message Size(B)	4000 bits
Round	10000
Aggregation Energy	5nJ/bit/packet
P	0.10
m	0.5
m0	0.2

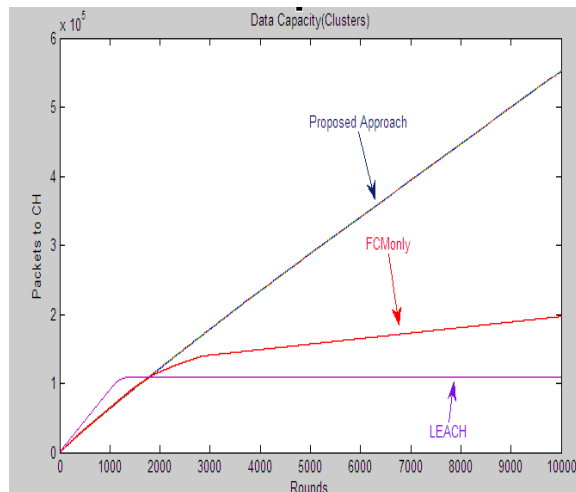


Figure 3. The data capacity of the network using different approaches at cluster level

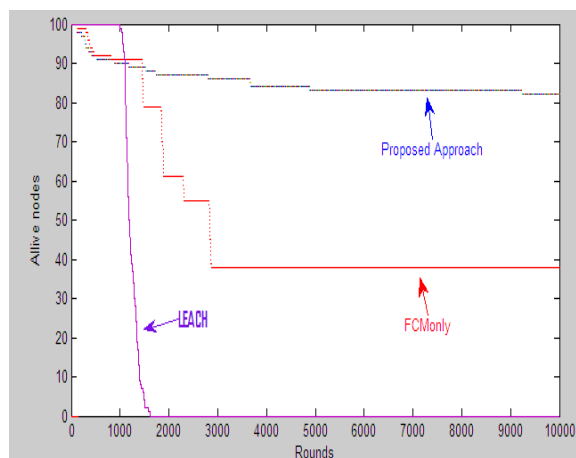


Figure 4. Alive nodes vs rounds in different approaches

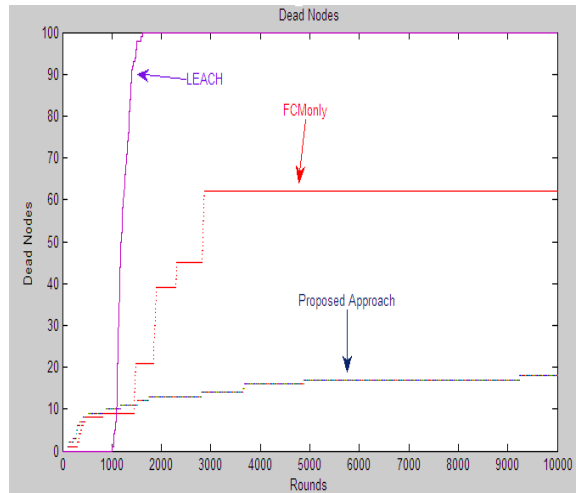


Figure 5. Dead nodes vs rounds in different approaches

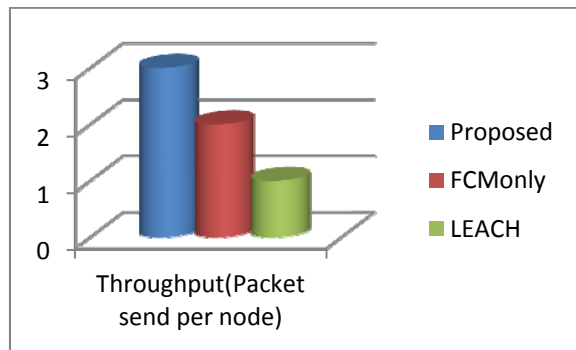


Figure 6. The throughput of the network observed at basestation

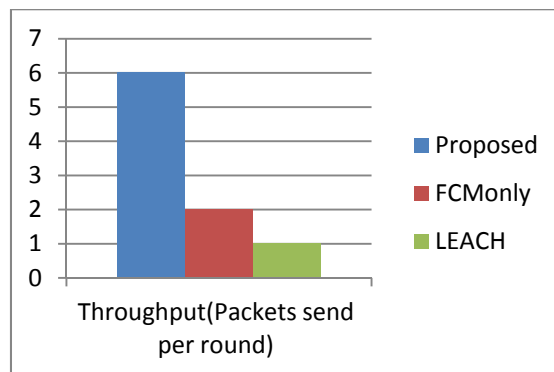


Figure 7. The throughput observed per node in the network

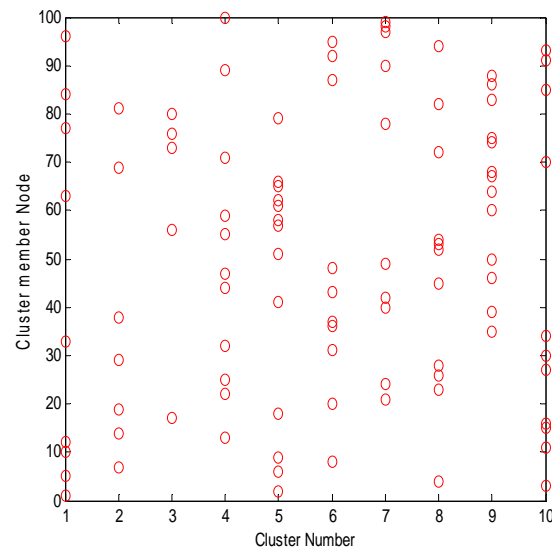


Figure 8. The Cluster member nodes in each cluster

4. CONCLUSION

The proposed approach used the fuzzy C-means clustering. In each of the cluster our algorithm tried to have similar sized clusters of sensor nodes and also neural networks, using the different factors have elected the cluster heads. Our approach worked on optimality of cluster and persistency of nodes in network and successfully outperformed the present approaches. The throughput at both cluster and network level is highly appreciating and showing effectiveness of using this hybrid approach. Different factors play a dominant role in selecting cluster head as only one factor can not judge a node as cluster head and also the probabilistic perspective not assure us about selecting a good cluster head so we can say our approach has taken care all these factors and worked on these and successfully approached that efficient routing technique. Our future work will be to combine the other clustering and decision approaches in sensor networks and see how they affect our approach.

REFERENCES

- [1] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless sensor networks", in the Proceeding of the Hawaii International Conference System Sciences, Hawaii, January 2000.
- [2] W.R. Heinzelman, A.P. Chandrakasan, H. Balakrishnan, "An applicationspecific protocol architecture for wireless microsensor networks", IEEE Transactions on Wireless Communications, 1 (4) (2002) 660-670.
- [3] Rupesh Mehta, Abhishek Pandey & Pratik Kapadia, "Reforming Clusters Using C-LEACH in Wireless Sensor Networks", In International Conference on Computer Communication and Informatics (ICCCI -2012), Jan. 10 – 12, 2012, Coimbatore, INDIA.
- [4] G. Smaragdakis, I. Matta, A. Bestavros, "SEP: A Stable Election Protocol for clustered heterogeneous wireless sensor networks", in: Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA 2004), 2004.
- [5] L. Qing, Q. Zhu, M. Wang, "Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks", ELSEVIER, Computer Communications, 29, 2006, pp 2230- 2237.
- [6] Parul Saini, Ajay.K.Sharma, "E-DEEC- Enhanced Distributed Energy Efficient Clustering Scheme for heterogeneous WSN", in: 2010 1st International Conference on Parallel, Distributed and Grid Computing (PDGC - 2010).
- [7] Anamika Chauhan and Amit Kaushik, "TADEEC: Threshold Sensitive Advanced Distributed Energy Efficient Clustering Routing Protocol for Wireless Sensor Networks", International Journal of Computer Applications (0975 – 8887), Volume 96 No. 23 June 2014.
- [8] Amit K. Kaushik, "Performance Evaluation of Proactive and Reactive Routing Protocols in Wireless Sensor Networks", In International Journal of Computer Applications (0975 – 8887), Volume 110 – No. 16, January 2015.
- [9] A. Manjeshwar and D.P. Agarwal, "TEEN: a routing protocol for enhanced efficiency in wireless sensor networks", In 1st International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing, April 2001.
- [10] D.C. Hoang, R. Kumar, Member and S.K. Panda, "Fuzzy C-Means Clustering Protocol for Wireless Sensor Networks", In IEEE International Symposium on Industrial Electronics (ISIE), 2010.

- [11] Duc Chinh Hoang, Rasjesh Kumar, Sanjib Kumar Panda, “Realisation of a cluster-based protocol using fuzzy C-means algorithm for wireless sensor networks”, In IET Wireless. Sens. Syst., 2013, Vol. 3, Iss. 3, pp. 163–171.
- [12] Qi Wang, Kulla, E., Mino, G., Barolli, L. “Prediction of Sensor Lifetime in Wireless Sensor Networks Using Fuzzy Logic”, In IEEE 28th International Conference on Advanced Information Networking and Applications (AINA), 2014.
- [13] Pal R. , Sharma, A.K., “FSEP-E: Enhanced stable election protocol based on fuzzy Logic for cluster head selection in WSNs”, 2013 Sixth International Conference of Contemporary Computing (IC3).
- [14] T. Rappaport, Wireless Communication: Principles and Practice, Englewood, NJ: Prentice Hall, 1996.
- [15] MATLAB 7.4.0(R2007a) www.mathworks.com

BIOGRAPHY OF AUTHOR



Amit Kumar Kaushik is working as Assistant professor in department of computer science and technology in Manav Rachna University, Faridabad, Haryana, India. He completed his M.Tech in Information systems from Delhi Technological University, New Delhi and B.Tech from UIET, Kurukshetra University. His research interests include the wireless sensor networks, Distributed Systems and LTE networks.