IJCSN International Journal of Computer Science and Network, Volume 5, Issue 2, April 2016
ISSN (Online) : 2277-5420 www.IJCSN.org
Impact Factor: 1.02

Fuzzy C-Means Clustering Algorithm for Optimization of Routing Protocol in Wireless Sensor Networks

[1] Melaku Tamene, [2] Kuda Nageswara Rao

[1, 2] Dept. of CS & SE, Andhra University
Visakhapatnam, AP, INDIA

Abstract - Optimization of energy is the fundamental requirement at all levels of system design in wireless sensor networks. Routing protocol based on clustering is not only scalable with the dimension of networks but also offers efficient management of energy. Configuration of suitable clusters is the primary focus in design of cluster based routing protocols for wireless sensor networks. Of the most decisive factors to setup suitable clusters, energy and distribution of nodes into clusters are the dominant ones. In this paper, an efficient cluster formation is presented based on the fuzzy c-means clustering algorithm. The eligibility of nodes to act as cluster head is defined with respect to the average residual energy so as to avoid premature collapse of networks. The optimization problem consists of finding the most favorable set of cluster leaders from the eligible set so that the communication distance of nodes from cluster leaders is minimized. The protocol is implemented in OMNeT++ simulation environment and experimental studies reveal that the proposed protocol defeats LEACH, LEACH-C and CHEF protocols with respect to the various network performance metrics.

Keywords – *Wireless Sensor Networks, Fuzzy Clustering, Network Lifetime, Protocol, Optimization.*

1. Introduction

The great advance and growth of semiconductor technology makes the realization for integration of hundreds of thousands of transistors into a single chip. Micro-electro-mechanical systems (MEMS) technology is evolved from the fabrication process of VLSI manufacturing techniques that combines silicon based microelectronics with micromachining technology [1]. MEMS laid the foundation for system-on-chip (SOC) technology in which the micro-electro-mechanical systems and other devices are interconnected into a single microchip. Wireless sensor networks (WSNs) are emerged from SOC technology. The decrease in the size and cost of sensors resulting from such technological advancement has fuelled an interest to deploy large set of sensor nodes in the target environment. These tiny sensor nodes have integrated sensing units, microcontroller, transceiver and the power source [2-3]. The variety of mechanical, thermal, biological, chemical, optical, and magnetic sensors may be attached to sensor node in order to sample the ambient conditions and transform them into electric signals in which processing of such signals reveals some properties about events appeared in the vicinity of sensor. The application of wireless sensor networks ranges from habitant monitoring, surveillance, health monitoring to process control in industry [4-10]. Depending on the time criticality of applications, the nature of data reporting in WSNs can be proactive or reactive. In reactive networks, any sudden change in the sampled event triggers data communication while events are reported periodically or as per request of user in proactive networks. Nodes are communicated over the wireless link and the nature of data flow towards the base station is basically derived from the coordination of distributed nodes.

The presence of severe energy constraint in sensor node downs the salient features of WSNs. Wireless transmission and reception of packets in the transceiver is the dominant source of energy depletion compared to dissipation in sensing and computations. Of the many routing strategies, routing protocols based on cluster formation is a proficient technique to extend the network lifetime in WSNs [11-12]. In clustering, sensor nodes elect their own leader named by cluster head that will be in charge of handling tasks such as resource allocation, data compression and reporting to the base station. The state of being selected as cluster head is usually accompanied by bulky energy loss compared to regular nodes, so cluster head election shall be done under rotation scheme to maintain energy balance among nodes. Nominating the appropriate cluster heads is one of the chief research challenges in cluster-based protocols.

One of the earliest hierarchical routing algorithms is low energy adaptive clustering hierarchy (LEACH) protocol [13]. In LEACH, the cluster head is elected upon random selection of number between 0 and 1 and then comparing the generated number with predefined threshold model. The elected cluster heads broadcast the advert message within the transmission region using CSMA MAC protocol and regular nodes hearing this announcement pass a decision on the best cluster head to join upon the received signal strength. Regular nodes also use CSMA MAC protocol to inform about their membership to the respective cluster head. In steady state phase, each regular node sends data to its cluster head based on pre-allocated TDMA schedule. The nomination of cluster heads in LEACH is quite random that the number of elected cluster heads in each round varies in stochastic manner and the distribution of cluster heads within WSN area is not usually uniform. Due to random selection of cluster heads, the one with the lowest energy may be nominated as cluster head which clearly reveals energy unaware nature of LEACH protocol. The authors in [14] present modified version of LEACH termed as centralized LEACH (LEACH-C) protocol. In LEACH-C, the base station selects the suitable cluster heads based on metaheuristic simulated annealing algorithm. Similarly, the authors in [15] modify LEACH by appending residual energy parameter in the probabilistic threshold model so as to prohibit lower energy nodes from election process. In spite of promoting higher energy nodes for being elected as cluster head, the random nature of threshold equation still makes uncertain election process.

The paper in [16] describes a hybrid, energy efficient, distributed (HEED) clustering protocol. Initially, the tentative cluster heads will be probabilistically elected based on their residual energy and these nodes announce their wish of being cluster head embedding cost metric in their advertisement message. The temporary cluster heads not only set their cluster head probability but also double this value in between HEED iterations. The qualification for being selected as final cluster head relies on the extent of terminating HEED execution. Then, regular nodes hearing the final cluster head advertisement will decide to which cluster they join based on the given cost metric.

Multi-hop routing techniques avoid massive energy loss incurred in long distance transmission, but sensor nodes nearer to the base station will face early depletion of energy (hot spot problem) as they relay packets from nearby hops in addition to their own packets. The authors in [17] propose energy efficient unequal clustering (EEUC) mechanism to manage energy cavity problem in hot spot area. The EEUC protocol is based on the idea that nodes nearer to the base station shall have smaller cluster size to compensate burdens imposed on forwarding packets coming from nearby hops and the size of cluster shall be varied accordingly depending on the distance from the base station.

The authors in [18] present grid based clustering algorithm based on the idea of cross-level transmission to minimize unbalanced load distribution in the network. Initially, the WSN area shall be divided into square grids and clusters are built in each virtual block. The cluster head in each virtual block is in charge of forwarding packets to neighbor and cross-level grids based on mathematically determined values of forwarding probability.

The paper in [19] describes the fuzzy c-means induced clustering approach to improve the network lifetime of wireless sensor networks. The base station executes the fuzzy c-means algorithm to nominate feasible cluster heads for the first round and any other election of cluster heads in subsequent rounds will be decided locally without the participation of base station. Initially, energy and distance of node from the centroid of a cluster as well as base station are used for nomination in which the clusters built at this stage will be fixed onwards for the entire network operation. Then, in subsequent rounds, the acting cluster heads are updated locally in each cluster based on the residual energy of node and its distance to base station. The algorithm reduces the energy dissipation involved in cluster reconstruction at the base station.

The authors in [20] present the concept of fuzzy theory for distributed selection of cluster heads. The protocol is termed CHEF (cluster head election mechanism using fuzzy logic) in which the sum of distance of a node with respect to its neighbors and residual energy are used to determine the selection probability of being cluster head. In each cluster setup phase, node takes any number in the range of 0 and 1 and then advertizes its state if the number is less than the predefined threshold value. The competition for being selected as final cluster head depends on the comparison of fuzzy computed values of the selection probability of nodes.

In this paper, the Fuzzy C-Means Clustering (FCM) algorithm has been applied to optimize the distribution of nodes into clusters for cluster based routing protocols. The FCM algorithm partitions the network into predefined number clusters in which the degree of belongingness of a node to a particular cluster falls somewhere between 0 and 1. The absolute belongingness of a node to either of the clusters in hard (crisp) clustering techniques does not have practical importance in many engineering problems. The fuzzy clustering uses the concept of partial belongingness and nowadays it has been applied in different fields of study such as image processing, pattern recognition and cluster analysis. The protocol is implemented using OMNeT++ simulation tool and it offers satisfactory results compared to LEACH, LEACH-C and CHEF protocols.

The organization of the rest of this paper is as follows. Section 2 describes system model and assumptions. Section 3 discusses the proposed protocol. The experimental results are presented in section 4. Finally, conclusions are made in section 5.

2. System Model and Assumptions

The radio energy dissipation model in [21] has been used to estimate the energy consumption of nodes in radio communications. That is, the energy dissipated in radio hardware to transmit l bit packet to a distance d is given as follows.

$$ETx(l,d)=\begin{cases} l\,Eelc+l\,efs\,d^2 & ,d<dco \\ l\,Eelc+l\,emp\,d^4 & ,d\geq dco \end{cases} \quad (1)$$

Where the term $dco=\sqrt{efs/emp}$ denotes the threshold (crossover) distance and $Eelc$ represents the energy consumption per bit in radio electronics due to digital coding, modulation, filtering and spreading of signal. The parameters efs and emp are energy dissipation factors in the power amplifier for free space and multipath fading channel propagation models respectively. The radio also expends energy while receiving packets. The energy dissipated in radio electronics to receive l bit packet can be written as follows.

$$ERx(l)=l\,Eelc \quad (2)$$

Moreover, the energy consumption in microcontroller for data aggregation (Eda) is also considered in the model. Sensor nodes are assumed homogenous in terms of supply energy, processing power and memory size. The base station is equipped with unlimited power supply. Nodes are considered immobile after deployment and they have variable transmission power levels. Assume that nodes can locate themselves whenever required.

2.1 Optimum Number of Clusters Mathematical Background

Network topology control based on clustering is an efficient solution to maximize the network lifetime in energy constrained wireless sensor networks. The underlying principle in clustering is grouping nodes into fixed number of clusters based on their similarity measured in terms of the spatial position of nodes. With small number of clusters, regular nodes are expected to transmit with high power level to reach the cluster heads. On the other hand, the data aggregation rate will be reduced for large number of clusters. Hence, limiting the number of clusters to an optimal value is the desirable phenomenon to extend the network lifetime in wireless sensor networks. Assume $ro \times ro$ region over which N nodes are uniformly deployed. Considering that the base station is usually located far from nodes in which the separation distance most likely exceeds the threshold value, the energy consumption in cluster heads follows d^4 power loss model. Presumably the distance between regular nodes and cluster head is less than the threshold distance, so d^2 power loss model is applied to estimate the energy dissipation for such nodes. Let the symbol d_{toch} represents the average distance between regular nodes and cluster head. Assume the average distance between cluster head and base station is denoted as d_{tobs}. The energy consumption of cluster heads (Ech) and non-cluster heads ($Enon-ch$) in each data gathering round can be mathematically written as follows.

$$Enon-ch=l\,Eelc+l\,efs\,d_{toch}^2 \quad (3)$$

$$Ech=\left(\frac{N}{k}-1\right)l\,Eelc+\frac{N}{k}l\,Eda+l\,Eelc+l\,emp\,d_{tobs}^4 \quad (4)$$

Where k is the number of clusters and Eda is the data aggregation cost per signal. Considering that nodes are uniformly distributed across the region, the expected squared distance of nodes to cluster head can be given as follows.

$$E\left[d^2{}_{toch}\right]=\int\int\left(x^2+y^2\right)\rho(x,y)dxdy=\frac{ro^2}{2\pi k} \quad (5)$$

The symbol $\rho(x,y)$ in the above equation denotes node distribution in WSN field. The total energy dissipated in the network ($Etot$) per each round is owing to dissipation in cluster heads and regular nodes.

$$Etot=l\begin{pmatrix} N\,Eelc+N\,Eda+k\,emp\,d_{tobs}^4 \\ +N\,Eelc+N\,efs\frac{1}{2\pi}\frac{ro^2}{k} \end{pmatrix} \quad (6)$$

Now, the optimum number of clusters can be simply obtained by taking the derivative of $Etot$ with respect to

k and equating to zero. Accordingly, the optimum number of clusters can be computed as follows.

$$k_{opt} = \frac{\sqrt{N}}{\sqrt{2\pi}} \sqrt{\frac{efs}{emp}} \frac{ro}{d_{tobs}^{2}} \tag{7}$$

In spite of assumptions taken to compute the optimum number of clusters, the simulation studies to find optimum value of k in many research articles nearly match with the analytical result. The design approach in this paper is based on this predefined number of clusters.

3. Proposed Protocol Design

The distribution of nodes into clusters has a paramount influence on the functionality of nodes in wireless sensor networks. The average distance of nodes from the respective cluster heads is reduced for good distribution of nodes in the clusters. In addition to minimizing the communication cost within the cluster, unbalanced load distribution can also be controlled efficiently if the distribution of nodes into clusters is optimized. Hence, the optimization problem consists of finding the set of nodes that can be configured as cluster head such that there is well distribution of nodes in the clusters. Consider the network consists of N nodes.

The problem is formulated as grouping N nodes into c clusters such that the expected distance of nodes from the center of clusters is minimized. The base station is in charge of computing the cluster centers on the reception of the geographical location of nodes in WSN field. Due to the fuzzy nature of many problems in science and engineering, the fuzzy clustering method has been applied to find the cluster centers. Assume the network is partitioned into c clusters (c is defined with respect to the optimum number of clusters). Consider the degree of belongingness of node i to cluster j is represented as uij. The problem of optimizing the distribution of nodes into clusters involves minimization of the following objective function.

$$fobj = \sum_{j=1}^{c} \left(\sum_{i=1}^{N} uij^{m} \, d_{ij}^{2} \right) \tag{8}$$

Where the symbol dij defines the Euclidean distance of node i from the centroid of cluster j and m is the fuzzy control parameter. The fuzzy c-means clustering algorithm is used to find the centroid of clusters. Let the variable $\mathbf{U}f$ consists of vectors of degree of belongingness of each node to the given cluster. The FCM algorithm begins with initializing the values of uij. The pseudocode of the FCM clustering algorithm is depicted in Algorithm-1. The values of membership degrees are updated during the course of iteration according to the following equation.

$$uij = \frac{1}{\sum_{\ell=1}^{c} \left(\frac{dij}{di\ell} \right)^{2/m-1}} \tag{9}$$

The centroid of each cluster is computed by taking the mean of all data points (the position of nodes in WSN field) weighted by their degree of belongingness to the cluster as shown in the following equation.

$$cj = \frac{\sum_{i=1}^{N} uij^{m} \, pos\,(node\ i)}{\sum_{i=1}^{N} uij^{m}} \tag{10}$$

The termination criteria for the FCM algorithm is when the difference between contents of $\mathbf{U}f$ in successive iterations is less than a set threshold value or a predefined maximum number iterations is reached.

Once the FCM algorithm converges, the base station configures the cluster heads by making the use of information available in vector of cluster centers (centroid). Initially, the base station computes the average residual energy of nodes and those nodes whose energy falls beyond the average are marked as eligible for being elected as cluster leader. Then, nodes which belong to the eligible set and very close to the centroid of clusters are configured to become cluster leader.

Finally, the base station announces configuration details (information about TDMA transmission schedule and the corresponding cluster leader for each regular node) back to nodes. Upon receiving the message, the cluster members send the data to their cluster heads. Since making the use of FCM algorithm in cluster formation minimizes the distance of regular nodes to the cluster leaders, the data transmission cost of regular nodes is optimized in the proposed protocol. Considering that the cluster heads are not only cope with data aggregation but also deal with data transmission to the base station , the selection of cluster leaders from the eligible set (defined with respect to residual energy) avoids premature death of nodes . The operation of protocol in each data gathering round proceeds in similar fashion until the network is fully collapsed.

Algorithm-1: Fuzzy C-Means (FCM)
Input: position of nodes
Output: cluster centers
begin
    initialize $\mathbf{U}f$
    repeat
        for cluster $j = 1$ to c **do**
            $\mathbf{c}j \leftarrow$ compute cluster centroid
        end for
        update $\mathbf{U}f$
    until the algorithm converges
    return $\{\mathbf{C}\}$
end

4. Experimental Results

The protocol is implemented using objective modular network testbed (OMNeT++) simulation tool [22]. A total of N (N=100) sensor nodes are placed randomly across $100\,m \times 100\,m$ WSN area. The parameters used in the simulation are initial energy= 2J, base station location= $(175, 50)\,m$, control packet size =25 bytes, data packet size=500 bytes, $Eelc$ = 50nJ/bit, energy loss for data aggregation (Eda) =5nJ/bit/signal, efs =10pJ/bit/m^2, emp =0.0013pJ/bit/m^4 and TDMA frames per round= 6. The number of working nodes per each round, the variance of residual energy among them and the total energy consumption of nodes are examined in the simulation.

Network lifetime is a vital metric to measure the performance of protocols in sensor networks. The definition of network lifetime varies with the requirements of application but the standard definition is the amount of time that the network is considered fully operative from application point of view. For some applications, the network is considered non-operative (non-functional) if the delay to deliver the data exceeds the threshold. For some other applications, the network is believed to be non-functional when the coverage loss falls below the desired level. There are also applications for which the network is considered non-functional when the network is partitioned due to energy depletion of some nodes. Having considered that energy conservation is the primary issue in WSNs, the last definition is most acceptable and evaluation of the network lifetime in this experiment is based on this definition. From this point of view, the network lifetime can be generally taken as the time elapsed till the first node runs out of energy or the time till some fraction of nodes exhaust their energy. Most of sensors sample similar events due to their placement in the proximity of each other. Considering that some of the neighbors of failed nodes can capture correlated data, the failure of only a few nodes does not suddenly collapse the network. Hence, the time at which the network loses 50% of nodes has been taken as the definition for network lifetime. The time till half or 50% of nodes die (HND) is an appropriate metric in large scale networks compared to other metrics such as the time till first node dies (FND) and the time till last node dies (LND). The number of functional sensor nodes per each round is depicted in Fig. 1. The simulation result reveals that the proposed protocol improves the network lifetime by 50%, 13.4%, 10.18% compared to LEACH, LEACH-C and CHEF protocols respectively. The FCM based protocol also defeats LEACH, LEACH-C and CHEF protocols in terms of FND metric.

The energy dissipation among nodes should be balanced in WSNs to avoid premature death of nodes. The state of load balance in the network has been evaluated by using the variance of residual energy of nodes in each round. The simulation result depicted in Fig. 2 reveals that FCM based protocol dominates LEACH, LEACH-C and CHEF protocols with respect to load balance in the network. The total energy consumption of nodes per each round is also examined in the simulation. In Fig. 3, 20 rounds are taken randomly to compare the total energy consumption of nodes for each protocol. The simulation result reveals that the proposed protocol has lower energy consumption of nodes than its counterparts.

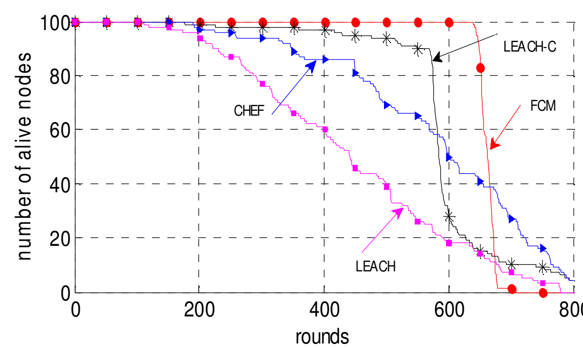

Fig. 1 Alive nodes versus time

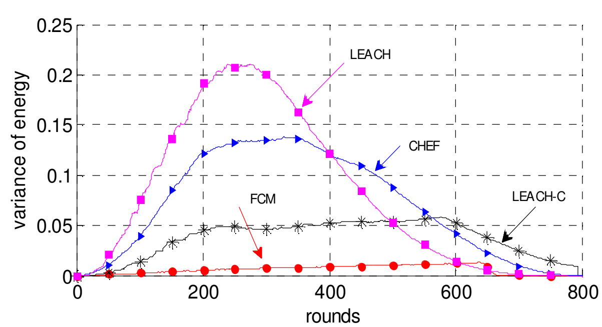

Fig. 2 Variance of residual energy among nodes

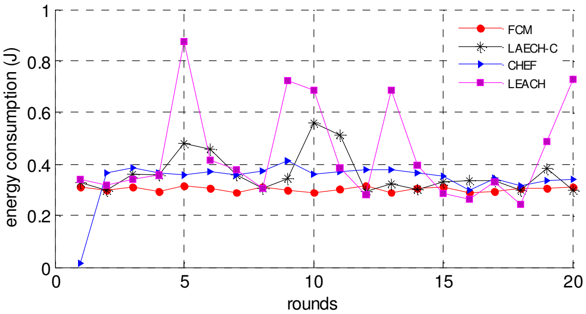

Fig. 3 Total energy consumption of nodes versus time

5. Conclusion

The presence of constraint of energy in sensor nodes challenges the design and development of protocols for wireless sensor networks. Considering that energy efficiency is one of the first and most research problems in wireless sensor networks, energy efficient routing protocol is presented using the fuzzy c-means clustering algorithm. The FCM based routing protocol has been tested with respect to the network lifetime, load balance and total energy consumption of nodes. The simulation results prove that the proposed protocol performs better compared to LEACH, LEACH-C and CHEF protocols.

References

[1] Z. Li, B. Liu and W. Wang, MEMS processing and Fabrication Techniques and Technology-Silicon–based Micromachining, Microsystems and Nanotechnology, Springer Berlin Heidelberg, pp. 287-352, 2012.

[2] G. Asada, M. Dong, T.S. Lin, F. Newberg, G. Pottie, W.J. Kaiser, and H.O. Marcy, “Wireless integrated network sensors: Low power systems on a chip,” in Proceedings of the 24th European Solid-State Circuits Conference (ESSCIRC), 1998, pp. 9–16.

[3] K. Sohrabi, V. Ailawadhi, J. Gao and G.J. Pottie, “Protocols for self-organization of a wireless sensor network,” IEEE Personal Communications Magazine, Vol. 7, No. 5, 2000, pp. 16–27.

[4] I.F. Akyildiz, Su. Weilian, Y. Sankarasubramaniam and E. Cayirci, “Wireless sensor networks: a survey,” Computer Networks, Vol.38, No. 4, 2002, pp. 393–422.

[5] K. Sohraby, D. Minoli, and T. Znati, Wireless Sensor Networks: Technology, Protocols, and Applications, John Wiley and Sons, Inc., Publications, 2007.

[6] J. Yick, B. Mukherjee, and D. Ghosal, “Wireless sensor network survey,” Computer Networks, Vol. 52, No. 12, 2008, pp. 2292–2330.

[7] V. Potda, A. Sharif, and E. Chang , “Wireless sensor networks: a survey,” in Proceedings of International Conference on Advanced Information Networking and Applications Workshops, WAINA ’09, 2009, pp. 636–641.

[8] M.P. Durisic, Z. Tafa, G. Dimic, and V. Milutinovic, “A survey of military applications of wireless sensor networks,” in Proceedings of Mediterranean Conference on Embedded Computing (MECO 2012), 2012, pp. 196–199.

[9] M. Welsh, Sensor networks for medical care, Technical Report TR-08-0, Harvard University Technical Report, 2005.

[10] K. Lorincz, B. Chen, G.W. Challen, A.R. Chowdhury, S. Patel, P. Bonato, and M. Welsh ,“Mercury: A wearable sensor network platform for high-fidelity motion analysis ,” in Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems (SenSys’09), 2009, pp. 183–196.

[11] W. Jin, S. Lei, J. Cho, Y.K. Lee, S. Lee, and Y. Zhong, “A Load balancing and Energy aware Clustering Algorithm in Wireless Ad-hoc Networks,” in Proceedings first International Workshop on RFID and Ubiquitous Sensor Networks, 2005, pp.1108-1117.

[12] O. Younis, M. Krunz, and S. Ramasubramanian, “Node Clustering in Wireless Sensor Networks: Recent Developments and Deployment Challenges,” IEEE Network (special issue on wireless sensor networking), Vol.20, No. 3, 2006, pp.20-25.

[13] W. Heinzelman, A. Chandrakasan and H. Balakrishnan,“Energy efficient communication protocol for wireless microsensor networks,” in Proceedings of IEEE Hawaii International Conference on System Sciences, 2000 pp. 3005-3014.

[14] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, “Application specific protocol architecture for wireless microsensor networks ,” IEEE Transactions on Wireless Communications, Vol. 1, No. 4, 2002, pp. 660–670.

[15] M.J. Handy, M. Haase, and D. Timmermann , “Low energy adaptive clustering hierarchy with deterministic cluster head election,” in Proceedings of international Workshop on Mobile and Wireless Communications Network, 2002, pp.368-372.

[16] O. Younis, and S. Fahmy, “HEED: a hybrid, energy efficient, distributed clustering approach for ad hoc sensor networks,” IEEE Transactions on Mobile Computing, Vol.3, No.4, 2004, pp.366–379.

[17] Le Chengfa, Ye Mao, Chen Guihai, and Wu Jie, “An energy-efficient unequal clustering mechanism for wireless sensor networks,” in Proceedings of IEEE international Conference on Mobile Ad hoc and Sensor Systems, 2005.

[18] L. Wenwei, Z. Yihua and P.Jian, “A Grid based routing algorithm with cross-level transmission to prolong lifetime of wireless sensor Networks,” Chinese Journal of Electronics, Vol.19, No.3, 2010.

[19] Jiejie Chen ,“Improving Life Time of Wireless Sensor Networks by Using Fuzzy C means Induced Clustering,” in Proceedings of World Automation Congress (WAC), 2012, pp. 1-4.

[20] J. Myoung Kim, S. Park, Y. Han and T. Chung, “CHEF: Cluster head election mechanism using fuzzy logic in wireless sensor networks,” in Proceedings of 10th International Conference on Advanced Communication Technology, 2008, pp. 654–659.

[21] D. Smithgall, “Toward the 60 gm wireless phone,” in Proceedings of the 1998 Radio and Wireless Conference (RAWCON), 1998.

[22] OMNeT++ Network Simulator. [Online]. Available: http://www.omnetpp.org/.

Author Profile:

Melaku Tamene received B.Sc. degree in Electrical Engineering from Bahirdar University, Ethiopia (2008) and M.Tech. degree in Electronics and Computer Engineering from Addis Ababa University, Ethiopia (2010). From July 2012, he works as Ph.D. scholar in the Department of Computer Science and Systems Engineering, Andhra University, India. His current research area is intelligent network protocols design and analysis in wireless sensor networks.

Kuda Nageswara Rao is a professor in Andhra University, India. He received B.E. degree in Electronics and Communication Engineering from GITAM University, India, M.Tech. degree in Computer Science and Engineering from Andhra University, India and Ph.D. degree from J.N.T University. His research interest includes computer networks, TCP/IP, internet technologies, telematics, data communications, wireless networks, wireless sensor networks and cloud computing.