International Journal of Academic Research in Computer Engineering
Print ISSN: 2476-7638 and Online ISSN: 2538-2411
Vol. 2, No. 1, Pages. 46-52, May 2018

	International Journal of Academic Research in Computer Engineering Volume 2, Number 1, Pages 46-52, May 2018 www.ijarce.org	IJARCE

A New Efficient Data Gathering Approach using Fuzzy Clustering in WSNs

Aziz Hanifi[1*]

1. Payame Noor University, Department of Computer Engineering, P. O. Box 19395-3697, Tehran, Iran.

Receive Date 2017.06.11; Accepted Date: 2018.01.10, Published Date: 2018.05.15

*Corresponding Author: Aziz Hanifi (hanifi_aziz@yahoo.com)

Abstract
Data gathering in Wireless Sensor Networks (WSNs) is one of the important operations in these networks. These operations require energy consumption. Due to the restricted energy of nodes, the energy productivity should be considered as a key objective in design of sensor networks. Therefore the clustering is a suitable method that used in energy consumption management. For this purpose many methods have been proposed. Between these methods the LEACH algorithm has been attend as a basic method. This algorithm uses distributed clustering method for data gathering and aggregation. The LEACH-C Method that is the improvement of LEACH, perform the clustering in centralized mode. In this method, collecting the energy level information of every node directly in each period increases the energy cost. In this paper, cluster head selection for increasing the network lifetime has been proposed based on fuzzy theory. The proposed ideas have been implemented over the LEACH-C protocol. Evaluation results show that the proposed methods have a better performance in energy consumption and lifetime of the network in comparison with similar methods.

Keywords: Clustering, Sensor Networks, Fuzzy Theory, Energy Estimates

1. Introduction

WSNs are a class of wireless ad-hoc networks. In these networks, sensor nodes collect data from physical environment. Then process this data and send to the Base Station (BS). So this will allow monitoring and control many types of physical parameters. Sensor nodes have limited energy and in most applications, using new energy sources is not possible. Therefore lifetime of sensor nodes is extremely dependent on energy stored in their battery. Clustering is a designing method that used for management of WSNs. In this method, the network is divided into several independent collections called clusters. Each cluster contains a number of sensor nodes and a cluster head node. Member nodes in a cluster send their data to relative cluster head node. Cluster head node aggregates these data and send to the base station. Therefore, clustering in sensor networks has advantages such as data aggregation support [1], data gathering facilitation [16], organizing a suitable structure for scalable routing [17], and efficient propagation of data in the network [18]. On the other hand, distance between nodes and base station has a direct impact on the nodes energy. So the location of base station node plays an important role in nodes energy consumption level.

Data gathering in WSNs is an important operation of these networks. For this purpose, many methods have been proposed. The LEACH [3] protocol is a basic hierarchical method. This protocol is suitable for monitoring applications. Each node periodically senses the area and sends information. In this algorithm, the clustering method has used for data gathering and aggregation. Clusters and cluster head nodes selected randomly, therefore there is no assurance to select the exact improved number and uniform distribution of cluster head nodes throughout the network. Many improvements in LEACH protocol have been presented. LEACH-C method [4] is an example of these improvements. In LEACH-C, establishment of clusters is done using a centralized algorithm by the base station in the beginning of each period. Base Station uses the received information from nodes for finding the predetermined number of cluster heads and network configuration within the clusters. This information contains position and energy of nodes. Another improvement to this algorithm is the use of estimation. One of these algorithms is LEACH-CE [5]. In the proposed technique energy level collected from all nodes in two primary periods but not collected in the other periods. Instead, the average energy of initial periods is used. Considering that the energy estimation in this method is not precise, this clustering scheme is neither efficient nor suitable. There is some proposed clustering methods that ABCP [19] and ABEE [20] and HMM [11] are examples of them. Existing of time correlated data in all of these methods cause to energy waste. The TINA [7] algorithm and its improvement proposed for not sending correlated data. These algorithms have error in reporting, and are not suitable approaches. LEACH-CEC algorithm [20] uses time series forecasting to implement the idea of not sending time correlated data. Also a model for estimating remaining energy of nodes has been proposed. Finally in order to awareness of base station with the number of correlated data for each node, suggested a method for more accurate estimate of energy. WSNs have a non-deterministic nature and changes in many cases, are unpredictable. For this reason, using fuzzy approaches because of its ability in the modeling of indecisive systems can be useful. High-speed In decision making, working in noisy environments, work with imprecise data, working with heterogeneous and sometimes conflicting parameters, no need to heavy computational load and so on are the feature that fuzzy logic provides to us. Using fuzzy theory to select cluster head for increasing network lifetime will be useful.

The remaining of this paper is organized as follow: In Section 2 related works are reviewed. In Section 3, cluster head selection algorithm using fuzzy theory is proposed. Analysis of experiments offered in Section 4, and we finally in Section 5 summarize and discuss future works.

2. Related Works

2.1 LEACH

One of the most famous hierarchical routing protocols based on clustering, is the LEACH [3] protocol. In this method each of cluster members sends their data to the cluster head. The cluster head aggregates this data and sends to the BS. So the communication cost will be reduced. Fig.1 illustrates this concept:

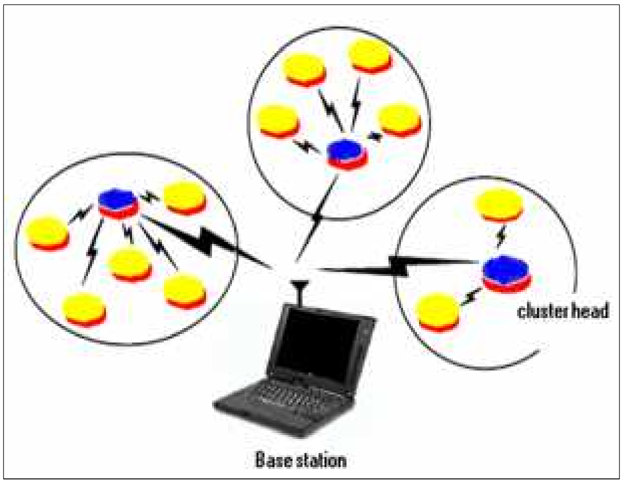

Figure 1. A sensor network with clustering

Cluster forming operations and data transmission in LEACH is done in two phases. These phases shown in Figure (2) and Figure (3):

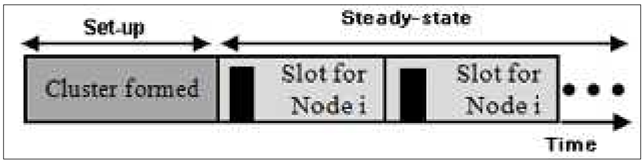

Figure 2. Period of LEACH

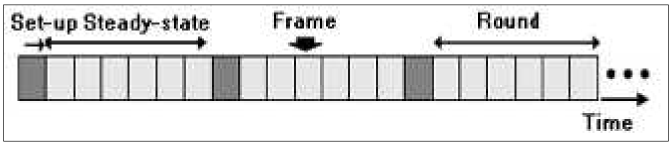

Figure 3. Details of period

Setup phase involves cluster forming and cluster head selection. In this stage, cluster and cluster head randomly selected. After constructing the cluster, cluster head propagates Time Division Multiple Access (TDMA) scheduler to determine data transfer time for member nodes. Then the steady-state phase starts. In the steady-state phase, each member node of a cluster send its data to the cluster head only in its time slot and in the rest of time pieces goes to sleep mode for energy conservation.

In this method, the cluster head consumes more energy for receiving, processing and directly sending this data to the BS node. So for increasing the life time of the network it is necessary to replace role of cluster head between network nodes. Many improvements over the LEACH method have been provided. In these improvements firstly, as far as possible the best clustering and cluster head selection is done, secondly possible as possible overhead of the protocol has been dropped. LEACH-C method is an example of these improvements.

2.2 LEACH-C

In LEACH-C [4], clusters forming in the beginning of every period are done using a centralized algorithm by the base station. The base station uses received information from nodes that includes energy and node status, uses this information during the setup phase for finding predetermined number of cluster heads and network configuration within the clusters. Next classification of nodes in the clusters is done to minimize energy consumption in order to transfer their data to the related cluster head. Results show that LEACH-C overall performance is better than LEACH, because of the optimal forming of clusters by the base station. In addition, the number of cluster heads in each period of LEACH-C is equal to the expected optimal value. While in LEACH the number of cluster heads varies in different periods because of lack of global coordination. As in LEACH-C at the beginning of every period energy of nodes must be sent to BS, therefore nodes early discharged and the network lifetime reduces. Another improvement on this algorithm is the use of energy estimation. The LEACH-CE method is an example of these methods.

2.3 LEACH-CE

In the LEACH-CE method [5], the energy level of all nodes collected only in two primary periods and not be collected in other periods. Instead because of knowing information about energy level of nodes, we can calculate energy consumption average for each node by using information of two primary periods. This means that reducing calculated energy level from the energy level of node, causes predicting current energy level of node. The problem of this algorithm is that firstly energy estimation is not done precisely and secondly if nodes have correlated data, while not sending correlated data means that previous data is valid, so this plan of clustering is not suitable and efficient.

2.4 LEACH-CEC

Three ideas are proposed here: 1. the data time correlation, 2. the distributed clustering model, 3. the exact location of the base station and the hybrid method from them.

In the data time correlation algorithm, Time Series Forecasting method (TSF) used to decide sending or not sending of data. Then in time t in the beginning of each period, base station sends percentage of error e(t) to all nodes. First data sensed by node and sent. Second and third and fourth data sent based on the improved TINA algorithm. Then the node runs time series function to determine the value of prediction of trend line model, to create trend model. In the next times the sensed data compared with predicted value of trend model, if the difference between these two values exceeds a threshold value, data sent to the given node and the node recalculate prediction function of trend model to update the trend line. Otherwise, the sensor node does not send the sensed data with this insurance that sensed data placed in accuracy range of data. So only some data have to be sent that are very different from the trend line model. This help to prevent energy loss.

For doing the best clustering that is necessary to know the energy of nodes. For this, let us divide Leach-c protocol to three phases: set up-state, steady-state, setup-distributed. In the first phase, the nodes send their positions and energy to the base station. In a centralized way, the base station does the first period of clustering. In the steady-state phase, data nodes are sent to the cluster heads and in the end of period, the remaining energy of every node is sent to its corresponding cluster head. In the third phase, clustering is done by head clusters, in this phase, for each node, the data time correlated algorithm is applied as the following. We call the proposed model of energy consumption nodes energy efficient distributed clustering (EDC) and explain as follow.

The location of base station is important in a network in which the density of nodes in a region is more than the others, so it is desired that the base station is located in a region with more density of nodes. We call this method, dynamic base station. In this method, the base station has no information about the position of networks. Thus, according to the method which is presented, we try to locate the base station in a region with more density of nodes. We call the proposed model of energy consumption nodes DBS1 and explain as follow.

BS node should be informed of data time correlation in nodes to estimate precisely energy of them. Therefore cluster head create a table that containing list of all members of the cluster. Cluster head registers every node in to the table that have correlated data and do not sent in certain times. In the end of each period, cluster head sends this table with collected data to the base station. This table contains nodes ID and number of times that these nodes not sent data. Base station uses this information for clustering decisions in centralized methods. Ultimately that cause to energy estimation in centralized methods is more carefully while the best clustering is created and the network lifetime increases. So in total lifetime of the network, first phase has done once but setup and steady-state phases done as in leach-c.

3. Proposed method

In this approach we use the fuzzy theory to determine cluster head. This cluster head determination Carried out using three parameters of Degree, energy and distance. Each node knows its position. The Degree parameter shows Number of neighbors for a node. Each node using Mac layer provides a List of its one hop neighbors and their degree is calculated. The distance for each node is equal to its distance to cluster head. We named this proposed method FUZZY-LEACH-C.

3.1. The process of the proposed method

Cluster head selection:

Suppose that A= {high energy}, B= {high degree} and C= {low distance} and assume that A(E) , B(D) and C(d) Be fuzzy functions related to above sets.

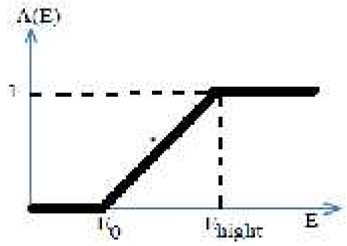

Figure 4. Energy level

$$A=\begin{cases} 1 & E > E_h \\ \frac{E-E_c}{E_h-E_c} & E_0 < E < E_h \\ 0 & E < E_0 \end{cases} \qquad (1)$$

[1] Dynamic base station

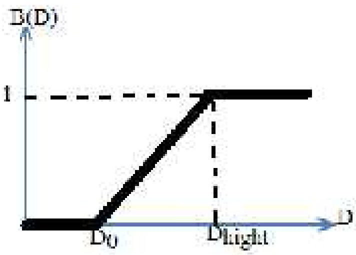

Figure 5. Degree

$$B(D)=\begin{cases} 1 & D > D_{hight} \\ \frac{D-D_0}{D_{hight}-D_0} & D_0 < D < D_{hight} \\ 0 & D < D_0 \end{cases} \qquad (2)$$

At the next stage this function is calculated.

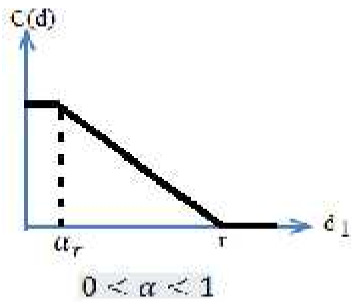

Figure 6. Distance

$$C(d)=\begin{cases} 1 & d < \alpha_r \\ \frac{r-d}{(1-\alpha)r} & \alpha_r < d < r \\ 0 & d > r \end{cases} \qquad (3)$$

d is the distance of node from cluster head in the related cluster. When energy of cluster head node be less than E0 it calculates values of A(En) , B(Dn) and C(d(xn, yn),(xh, yh)) for each node that located in one hop, and (xn, yn) is the coordinates of node, En is the nodes energy, Dn is the degree and dn is the nodes distance to cluster head coordinates that is equal to (xh, yh). Then calculates F(n)= α A(En) + β B(Dn) + ɣ C(d(xn, yn),(xh, yh)) and then obtains max f(ni) and niεC in which C determines the cluster with cluster head coordinates of (xh, yh). If this maximum value be equal to f(nj) then the cluster head informs node nj to introduce itself as cluster head.

In the above formula (α, β and ɣ) are energy, degree and central respectively. In the other hand if we suppose values of ith row of the following matrix as energy, degree and the distance of ith node to its related cluster head respectively (by supposing that N nodes exist in cluster C) so:

$$M=\begin{bmatrix} A(E_1) & B(D_1) & C(d_1) \\ \vdots & \vdots & \vdots \\ A(E_i) & B(D_i) & C(d_i) \\ \vdots & \vdots & \vdots \\ A(E_N) & B(D_N) & C(d_N) \end{bmatrix} * [\alpha \quad \beta \quad \gamma]^T \qquad (4)$$

$$=[\ \alpha A(E_1)+\beta B(D_1)+\gamma C(d_1),\ldots,\alpha A(E_i)+\beta B(D_i)+\gamma C(d_i),\ldots,\ \alpha A(E_N)+\beta B(D_N)+\gamma C(d_N)](6)$$

And di shows the ith node distance to its cluster head.

$$M=\sum_{i=1}^{N}\alpha A(E_i)+\beta B(D_i)+\gamma C(d_i) \qquad (7)$$

M=[Mi] , 1≤ i ≤N Mj=max Mi

Node nj introduce itself as cluster head. We must purpose that in order to calculate distance from cluster head it is not essential to know neighboring nodes location. Because the distance can be calculated using received signal strength and many other methods.

4. Simulation Results

Simulation is done over the Linux Fedora core.10 operating system using network simulator of NS2. LEACH and LEACH-C protocols from uAMP project in MIT University are simulated over NS2. Determined scenarios for a real simulation environment are as follows:

- Sensor network topology using 100 nodes
- Environment dimension is 100m*100m
- Cluster number is 5
- Data transfer speed is 1Mbps
- Base station located in position (100,50)
- Period length is 20 seconds
- Wireless transmission speed is 3*108 m/s
- Primary energy of each node is 2 J
- Receivers and senders follow the model [3] with these parameters:

$$E_{elec}=5.0\times10^{-8}\,J/bit,\ E_{tx}=E_{rx}=5.0\times10^{-8}\,J/bit$$

$$\pounds_{free\text{-}space\,amp}=1.0\times10^{11}\,J/bit/m^2,\ \pounds_{two\text{-}ray\text{-}amp}=1.3\times10^{15}\,J/bit/$$

E_{tx}, E_{rx} Parameters are sending and receiving energy for each Bit. Our experiment accomplished with LEACH, LEACH-C, LEACH-CEC and FUZZY-LEACH-C protocols.

In the NS2 simulator and LEACH and LEACH-C protocols, data generated with uniform distribution. But phenomenon that observed by sensor nodes usually continuously change with time. So data received by sensor node are dependent. Therefor data generated in simulation must have normal distribution. We have run each of protocols 20 times and resulting diagrams are the average of runs.

Figure 7 shows energy consumption in each period. In this figure we compared LEACH, LEACH-C, LEACH-CEC and the proposed FUZZY-LEACH-C protocols. Then calculate energy consumption values for each of them. As simulation results show our proposed FUZZY-LEACH-C algorithm has better operation in contrast to other methods.

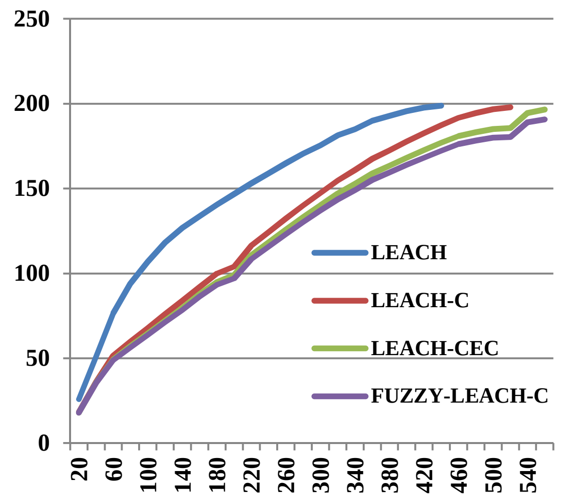

Figure 7. Energy Consumption Value in Total of Network Topology

Figure 8 shows number of alive nodes in different times. In this figure, 4 methods mentioned above surveyed in each period from the viewpoint of alive nodes number. As seen in figure 8 in our proposed FUZZY-LEACH-C algorithm, number of alive nodes in each time slot is more than other methods. In the leach protocol sensor nodes death stared in time 300, but in distributed protocols this time is 100.

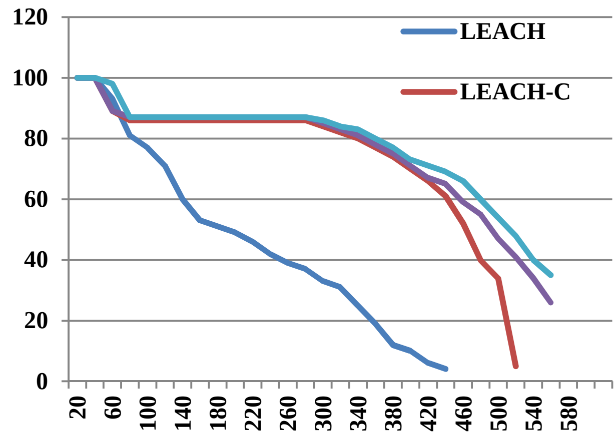

Figure 8. Number of Alive Nodes

5. Conclusion and Future Work

In this paper we solve the problem of cluster head selection in all of discussed methods. So using a fuzzy method, cluster head selection is done efficiently as network lifetime increases. Also in LEACH protocol we have eliminated periodic transfer of nodes data. Using proposed FUZZY-LEACH-C algorithm there is no need that all nodes send their data to base station in every period. Nodes send their position to base station only when network starts. Base station creates network topology and using fuzzy method determines cluster head in each period. In this paper using fuzzy algorithm, we have improved network lifetime in LEACH-CE protocol and restricted energy waste. In future works we will try to do nodes clustering in mobile sensor networks using fuzzy theory.

References

1. S. Bandyopadhyay, E.J. Coyle, An energy efficient hierarchical clustering algorithm for wireless sensor networks, Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No.03CH37428), Vol. 3, pp. 1713-1723, 2003.
2. M. Ali and S.K. Ravula, Real-Time Support and Energy Efficiency In Wireless Sensor Networks, School of Information Science, Computer and Electrical Engineering (IDE), Technical report, IDE0805, 2008.
3. A.A. Abbasi, M. Younis, A survey on clustering algorithms for wireless sensor networks, Computer Communications, Vol. 30, Issues 14–15, pp. 2826-2841, 2007.
4. W.B. Heinzelman, A.P. Chandrakasan, H. Balakrishnan, An Application-Specific Protocol Architecture for Wireless Micro sensor Networks, IEEE, IEEE Transactions on Wireless Communications, Vol. 1, Issue 4, pp. 660-670, 2002.
5. W.B. Heinzelman, A.P. Chandrakasan, H. Balakrishnan, Energy-Efficient Communication Protocol for Wireless Microsensor Networks, IEEE Published in the Proceedings of the Hawaii International Conference on System Sciences, pp. 1-10, 2000.
6. S.D. Muruganathan, D.C. F. Ma, R.I. Bhasin, A.O. Fapojuwo, A centralized energy-efficient routing protocol for wireless sensor networks, IEEE Communications Magazine, Vol. 43, Issue: 3, pp. S8-13, 2008.
7. E.U. Munir, M. Aslam, T. Shah, M.M. Rafique, An advanced heterogeneity-aware centralized energy efficient clustering routing protocol for wireless sensor networks, International Green Computing Conference, pp. 1-10, 2014.
8. S. Banerjee and S. Khuller, A Clustering Scheme for Hierarchical Control in Multi-hop Wireless Networks, Proceedings of IEEE INFOCOM, 2001
9. W.P. Chen, J.C. Hou and L. Sha, Dynamic Clustering for Acoustic Target Tracking in Wireless Sensor Networks, Proceedings of the 11th IEEE International Conference on Network Protocols, pp. 284-294, 2003.
10. C.K. Liang, Y.J. Huang and J.D. Lin, An Energy Efficient Routing Scheme in Wireless Sensor Networks, 22nd International Conference on Advanced Information Networking and Applications, IEEE 2008.
11. J. Amiri, M. Sabaei and B. Soltaninasab, A New Energy Efficient Data Gathering Approach in Wireless Sensor Networks, Communications and Network, Vol. 4 No. 1, pp. 61-72, 2012
12. M. J. Handy, M. Haase, D. Timrnermann, Low Energy Adaptive Clustering Hierarchy with Deterministic Cluster-Head Selection, IEEE, 2002.
13. J.M. kim, H.K. Joo, S.S. Hong, An Efficent clustering Sheme through Estimate in Centralized Hierarchical Routing Protocol, IEEE,2006.
14. A.S. Mohamed, B. Jonathen, TINA: A Scheme for Temporal Coherency-Aware in Network Aggregation, 3rd ACM int. workshop on data engineering for wireless and mobile access, USA, pp. 69-76, 2003.
15. X. Dai, F. Xia, Z. Wang, Y. Sun, An Energy-Efficient In-Network Aggregation Query Algorithm for Wireless Sensor Networks, First International Conference on Innovative Computing, Information and Control - Volume I (ICICIC'06), Vol. 3, pp. 255 - 258, 2006.
16. X. Wang, J.J. Ma, S. Wang and D.W. Bi, Time Series Forecasting Energy-efficient Organization of Wireless Sensor Networks, Sensors, 7(9), pp. 1766-1792, 2007.
17. G.H. Riahy, M. Abedi, Short term wind speed forecasting for wind turbine applications using linear prediction method, Renewable Energy, Vol. 33, Issue 1, pp. 35-41, 2008.
18. M. Ali and S.K. Ravula, Real-Time Support and Energy Efficiency in Wireless Sensor Networks". School of Information Science, Computer and Electrical Engineering Halmstad University. January 2008.
19. P. Hu, Z. Zhou, Q. Liu, F. Li, The HMM-based Modeling for the Energy Level Prediction in Wireless Sensor Networks, 2nd IEEE Conference on Industrial Electronics and Applications, pp. 2253-2258, 2007.
20. C.K. Liang, Y.J. Huang and J.D. Lin, An Energy Efficient Routing Scheme in Wireless Sensor Networks, 978-0-7695-3096-3/08 2008 IEEE DOI 10.1109/WAINA.2008.199
21. E.H. JUNG, S.H. LEE, J.W. CHOI, A Cluster Head Selection Algorithm Adopting Sensing-Awareness and Sensor Density for Wireless Sensor Networks, IEICE TRANS. COMMUN., Vol. E90–B, No. 9 SEPTEMBER 2007
22. T.-C. Hou and T.-J. Tsai. An access-based clustering protocol for multi-hop wireless ad hoc networks, IEEE Journal on Selected Areas in Communications. Vol. 19, Issue 7, pp. 1201-1210, 2015.
23. X. Hong and Q. Liang, An access-based energy efficient clustering protocol for ad hoc wireless

sensor network". 15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, (PIMRC 2004), Vol. 2, pp. 1022- 1026, 2004.