TWO WAY CLUSTERING APPROACH FOR ENERGY CONSUMPTION OF CLUSTER HEAD AND DECREASE PROBABILITY OF FAILURE NODES IN WSNS

Kush Bhushanwar*

Praveen Bhanodia*

Makrand Samvatsar*

Abstract

This paper addresses the problem of energy conservation in clustering algorithm for wireless sensor network. Since the major source of energy consumption in the sensor node is the wireless interface, considerable energy can be saved if the transceivers are completely shut down for a period of time. These sleeping times must be carefully scheduled, or network functionality could be compromised. Here it is classic clustering algorithms in wireless sensor networks and finds two main reasons causing unnecessary energy consumption, which are fixed operation periods and too much information exchanged in cluster-heads selection. Here it is proposed clustering methods with less communication overhead for clustering based on federal management in k-means algorithm effective clustering and distributed algorithm. Once the cluster heads and associated clusters are created, cluster-head-to-gateway routing is used to transfer the data to the base station to reduce the energy consumption of cluster head and decrease probability of failure nodes. to conclude, the Simulation results have shown that our approach achieves lower average delay, more energy saving, more network lifetime and higher delivery ratio than the other protocols.

Keywords— Partitioned clustering, residual energy, clustering algorithm, solar-aware LEACH, wireless sensor network.

* Department of Computer Science & Engineering, Patel College of Science & Technology, Indore (PCST)

I. INTRODUCTION

In this paper addresses the problem of energy conservation in clustering algorithm for wireless sensor network. Since the major source of energy consumption in the sensor node is the wireless interface, considerable energy can be saved if the transceivers are completely shut down for a period of time. These sleeping times must be carefully scheduled, or network functionality could be compromised. A clustering strategy is a distributed protocol that maintains a (minimal) connected backbone of active nodes, and turns into sleeping state the transceivers of non-backbone nodes. Periodically, the set of active nodes is changed to achieve more uniform energy consumption in the network. Clustering is the grouping of similar objects and a clustering of a set is a partition of its elements that is chosen to minimize some measure of dissimilarity [3]. Clustering algorithms are often useful in applications in various fields such as visualization, pattern recognition, learning theory, computer graphics, neural networks, artificial intelligence, and statistics. Practical applications [12] of clustering include pattern classification under unsupervised learning, proximity search, time series analysis, text mining and navigation. Clustering in sensor nodes has been widely pursued by their search community in order to solve the scalability, energy and lifetime issues of sensor networks. Clustering algorithms limit the communication in a local domain and transmit only necessary information to the rest of the network through the forwarding nodes (gateway nodes). A group of nodes form a cluster and the local interactions between cluster members are controlled through a cluster head (CH) (a chosen leader). Cluster [4] members generally communicate with the cluster-head and the collected data are aggregated and fused by the cluster head to conserve energy. The cluster-heads can also form another layer of clusters among themselves before reaching the sink. Issue is placed on partitioned clustering algorithms (as opposed to regular hierarchical clustering), which yield a single partitioning of the data described by a fixed number of parameters [4] [13]. With these parameters being less than the available data, partitioned clustering can afford promising distributed implementation of deterministic approach. A popular federal as well distributed deterministic partitioned clustering approach is offered by LEACH and Solar-aware LEACH, which features simple, highly reliable, and fast-convergent iterations& re-clustering during failure states [5]. Since the sensor nodes are powered by limited power batteries, in order to prolong the life time of the network, low energy consumption is important for sensor nodes. In general, radio communication consumes the most amount of energy, which is proportional to the data size and proportional to the square or the fourth power of the distance. In order to reduce the energy consumption, a clustering and data aggregation approach has been extensively studied [7]. In this approach, sensor nodes are divided into clusters, and for each cluster, one representative node, which called cluster head (CH), aggregates all the data within the cluster and sends the data to BS. Since only CH nodes need long distance transmission, the other nodes save the energy consumption. In order to manage effectively clusters and CHs, distributed clustering methods have been proposed such as LEACH, HEED, ACE and ANTCLUST [5, 6, 7, 8]. LEACH, which is the most popular method, guarantees that every node evenly become CHs but does not take into account battery level and the interrelationship among nodes [5]. HEED, ACE and ANTCLUST achieve better performance than LEACH by taking into account battery level, communication cost, node density, etc. However, they need additional inter-node communications for determining clusters and CHs.

It has been studied many clustering algorithms and found that three basic techniques Hierarchical (clustering) architecture, which are respectively presented in LEACH [1], HEED [6], GAF [7] and P. Santi's algorithm [4]:

1. Selecting cluster-heads periodically: Cluster-heads are selected periodically to evenly distribute the energy load among all the nodes.
2. Virtual grids method: Each node use location information to associate itself with a "virtual grid", in which only one node is active and responsible for processing signals.
3. Consideration of nodes' residual energy: Since cluster heads consume the most energy, residual energy is used to determine whether node can be cluster-head.

II. Literature Survey

Jin-Shyan Lee in et al [1] suggests FL-LEACH protocol that employs fuzzy logic in order to determine the number of CHs that should be used in the WSN. FL-LEACH is a fuzzy inference system that depends on two variables: number of nodes in the network and nodes density. Assuming uniform distribution of the nodes the sensor field, the novelty of the proposed approach is in its ability to determine the number of CHs without getting other information about the network.

Sampath Priyankara in et al [2] proposed a clustering method for wireless sensor networks with heterogeneous node types which select cluster heads considering not only transmission power and residual energy of each node but also those of its adjacent nodes.

Noritaka Shigei in et al [3] compared the performance of the proposed method with that of a well known clustering and routing protocol. Their computational experiments demonstrate that the lifetime and distribution of energy consumptions of the proposed method are better than those of the compared method.

W. Heinzelman, et. al [5] proposed two types of clustering methods with less communication overhead for clustering. The first type, which is based on centralized management, employs vector quantization (VQ) [1] for effective clustering. In the centralized method, BS determines clusters and CHs according to battery level and node location. The second type, which is performed in a distributed autonomous fashion, takes into account battery level and node density. In the distributed method, clustering is performed by the interaction among proximity nodes.

Wireless Sensor Network Replica

This section describes the wireless sensor network replica considered in this paper [18, 19, 20]. The WSN model consists of N sensor nodes and one base station (BS) node as shown in Figure 1. All sensor nodes are identical and are assumed to have the following functions and features:

- Sensing environmental factors such as temperature, pressure, and light
- Data processing by low-power micro-controller
- Radio communication
- Powered by a limited life battery

The Base Station node is assumed to have an unlimited power source, processing power, and storage capacity. The data sensed by sensor nodes are sent to the Base Station node over the radio, and a user can access the data via the Base Station node. In this WSN application, the clock synchronization of sensor nodes is an important issue. Because the time at which a data was sensed is important, which requires low clock skew among all the sensor nodes? We assume that the low clock skew requirement is guaranteed by using a clock synchronization method [5].

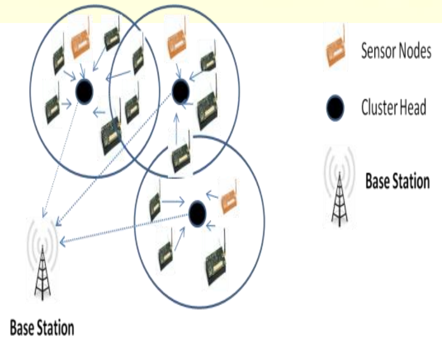

Figure 1: Wireless Sensor Network

The radio communication consumes more energy than the data processing on a sensor node. We assume the following energy consumption model for radio communication. The transmission of a k-bit message with transmission range d meters consumes E_T (k, d) of energy.

$$E_T(k, d) = \begin{cases} k(E_{elec} + \varepsilon_{fs}\, d^2) \text{ for } d <= d0 \\ k(E_{elec} + \varepsilon_{mp}\, d^4) \text{ for } d > d0 \end{cases}$$

$$E_R(k) = k . E_{elec}$$

where E_{elec} is the electronics energy, and ε_{fs} and ε_{mp} are the amplifier energy factors for free space and multipath fading channel models, respectively. The reception of a k-bit message consumes $E_R(k)$ of energy.

III. Proposed Method

Extra effective clustering approaches than LEACH has been proposed such as HEED, HIDCA [6, 7, 8]. Though, they require additional inter-node communications for clustering. In this section, we proposed two types of methods with less inter-node communication for clustering. The primary approach is a federal approach, and the second is a distributed approach.

A. The federal approach

In this method, the Base Station node supervises the clustering by exploit a k-means algorithm which specifies the energy as well as other resource constraints in wireless sensor networks.

The k-Means Based clustering approach: In this energy efficient protocol we make the following assumptions. A large scale network with a large number of nodes exists where the sensors are grouped into clusters. Each cluster has a cluster-head which communicates with the base station and the nodes in its cluster. In a densely deployed large scale sensor network, there is a higher degree of spatial correlation between the data sensed by the sensors in a cluster.

Data aggregation is thus used to eliminate redundancy and minimize the number of transmissions in a cluster in order to save energy. The clusters of sensors are formed in such a way that in a cluster no two sensors are more than some constant distance (d) apart which it is specified according to the type of application.

This assumption is made in order to ensure a higher degree of correlation between the data sensed by the sensors in a cluster. Our protocol uses the k-Means algorithm with certain modifications for in-network data processing and aggregation. The k-Means algorithm is a well known partition based algorithm for clustering of data sets.

We next give a brief description of k-Means algorithm. k-Means is a partition based clustering algorithm for large scale data analysis. In partition based cluster analysis of large data sets, optimal solution is obtained by computationally intensive extensive enumeration of all possible partitions of the data set.

For practical purposes, hence, two we describe our protocol as well as the distributed algorithms that are executed at the sensor nodes and the cluster-head. We define the following terms: k Denotes the number of partitions or groups of a set containing n data values at each sensor obtained by executing k-Means algorithm on the sensed data set at that sensor. k: Denotes the number of partitions or groups of a set data values at each cluster-head obtained by executing the k-Means algorithm at the cluster-head. In this protocol every sensor operates in two phases: Sensing Phase and k Means Phase.

Sensing Phase: The sensing phase is the time interval during which the sensor collects data. As soon as a sensor has sensed substantial amounts of data values it goes into the next phase.

K Means Phase: In this phase k-Means algorithm is executed over the data values collected during the sensing phase. As a result we get a reduced set of k (k n) data items which give a good representation of the n data items sensed by the sensor. Now we present our algorithm in flow chart format. Our algorithm is a federal algorithm which executes independently at the sensor nodes and the cluster-heads.

The algorithm for sensor nodes and cluster-heads and so we present them as follows. For simplicity, we assume that K is equal to k here. The LEACH allows only single-hop clusters to be constructed. On the other hand, in [2] authors proposed the similar clustering algorithms where sensors communicate with their cluster-heads in multi-hop mode. However, in these homogeneous sensor networks, the requirement that every node is capable of aggregating data leads to the extra hardware cost for all the nodes. Instead of using homogeneous sensor nodes and the cluster recongruation scheme, the authors of [17] focus on the heterogeneous sensor networks in which there are two types of nodes: super nodes and ordinary sensor nodes.

The super nodes act as the cluster- heads. The ordinary sensor nodes communicate with their closest cluster-heads via multi-hop mode. The major objective of federal approach is to use the sensor networks, like authors used in [17]. Federal approaches an interconnected set of clusters covering the entire node population. Namely, the system topology is divided into small partitions (clusters) with independent control.

Using a clustering approach, sensors can be managed locally by a cluster head, a node elected to manage the cluster and responsible for relaying data to other cluster heads or the sink. In addition, clustering provides inherent optimization capabilities at cluster-heads, such as data pre-processing. Federal approach is a simpler, but sub-optimal scheme where the nodes employ the mixed communication modes: single-hop mode and multi-hop mode periodically.

This mixed communication modes can better balance the energy load efficiently over WSNs and have already used in [16]. In addition federal approach will tend to preserve its structure when a few nodes are moving and the topology is slowly changing. Otherwise, high processing and communication overheads will be paid to reconstruct clusters.

Within a cluster, it is easy to schedule packet transmissions and to allocate the bandwidth to data traffic. From an energy standpoint, the advantages of our proposed protocol federal approach are as follows: First, by routing all data through the local cluster heads, the nodes avoid high power long distance wireless transmission to the base stations. Only the cluster heads (which are the powerful nodes) have to do it.

A cluster head can reduce the transmission energy expenditure by aggregating the collected data from its cluster before relaying them to the base stations. This reduces the overall network-wide transmission energy expenditure. Since the monitoring applications are often interested only in geographically aggregated data rather than per-node data, aggregation at cluster heads is highly desirable for extending the lifetime of sensor networks [13].

B. Distributed approach

We study Distributed clustering protocols. The performance of two popular schemes, HEED and HIDCA protocols, Node clustering has been widely studied for WSNs and many clustering algorithms have been proposed in the literature, such as LEACH, HEED, and HIDCA. The Highest Identifier Clustering Algorithm (HIDCA), modified from [7], is a primitive clustering protocol. Initially, during the node discovery stage, each sensor node exchange information to determine its neighboring nodes. Then, each node compares its ID with those from its neighbors. If its own ID has the smallest number, the node will become the cluster head and all other nodes will request to join the cluster and hence become cluster members. After the cluster is formed, the cluster head, that is, the node with lowest ID, sends control packets to maintain the operation of the cluster. No cluster head rotation is considered in this protocol. The cluster head keeps serving for the cluster until its battery power is depleted, during which another round of clustering process will take place and the node with the second lowest ID

The Low-Energy Adaptive Cluster Hierarchy (LEACH) combines the MAC (Medium Access Control) and routing functionalities. In LEACH, clusters are generated based on the optimal number of cluster heads, which is calculated using the prior knowledge of uniform node distribution. The cluster head determines a TDMA schedule for each sensor nodes within its cluster. Global synchronization is usually required, which consumes significant amount of network resources. Moreover, the cluster diameter in LEACH is assumed to be unlimited, which may result in the generated cluster members being located far away from the cluster head and each other. In HEED, clusters are generated without any assumption about node distribution. The cluster diameter is limited and fixed, and a cluster head rotation scheme is employed for load balancing. Although HEED can achieve a good load balance in a small area, the traffic loads in different areas are still unbalanced, thus leading to unbalanced energy consumption in the whole network. It should be pointed out that both LEACH and HEED are clusterhead-centric algorithms, which first select cluster heads based on a selection policy, such as the node with the largest residual energy.

IV. LEACH DISTRIBUTED VERSUS SOLAR-AWARE LEACH DISTRIBUTED

A. Evaluation parameter

In this section, the effectiveness of the proposed methods is demonstrated by numerical simulation. The proposed methods are compared with the conventional methods LEACH. In the simulation, N sensor nodes are randomly distributed in the square region of size 100x100, 200x200 and 300x300 m. The base station is 100 meters 150 and 200 away from the center of a side 200x200. Base station is 250 meters, 300 meters and 350 meters away from the center of a side as 300x300. Base station is 350 meters, 400 meters and 450 meters away from the center of a side as 400x400. The parameters used in the simulation are summarized in table 1. The simulation is performed for N = 100, 300 and 1000.

Table I

PARAMETER USED IN SIMULATION

For Energy Model	
d_0	75 m

E_{elec}	50 nJ/ bit
E_{fusion}	5 nJ/ bit
ε_{fs}	100 pJ/ bit/ m^2
ε_{mp}	1.3 fJ/ bit/ m^4
Initial battery level	0.5 Joule
Energy for data aggregation	5 nJ/ bit/ signal
For Packet Model	
Data packet size	800 bit
Broadcast packet size	200 bit
Packet header size	200 bit
For Distributed Method	
R_{inf}	20 meters
R_{end}	55 meters

In this subsection is shown a comparison with the results of the simulations of LEACH and its solar-aware extension. The evaluated results as explained above are related to the number of rounds done until half of the nodes are dead or when the first node is dead, where the latter case is also called network lifetime.

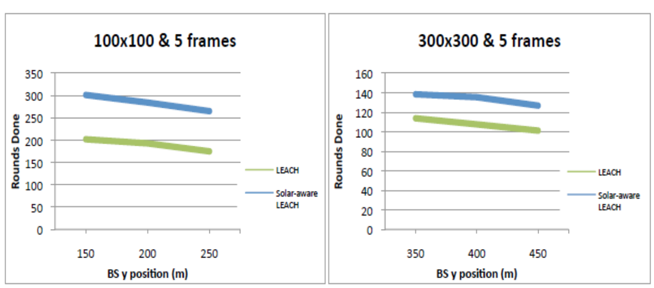

Figure 2: LEACH and Solar-aware LEACH results with 5 frames

B. Half-dead network

It can be seen the differences in the outcomes of both protocols. In the case of a short steady phase, i.e. composed by 5 frames, the solar-aware extension shows a higher number of rounds achieved than the original LEACH distributed version. However both show a similar behavior when the base station (BS) is placed at different distances, getting worse the farther the BS is to the closest node as can be observed in Figure 2.

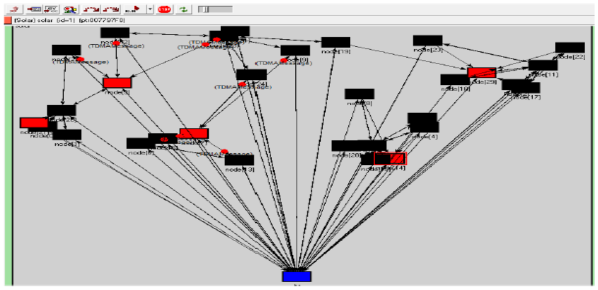

Figure 2: federal management in k-means algorithm effective clustering and distributed algorithm

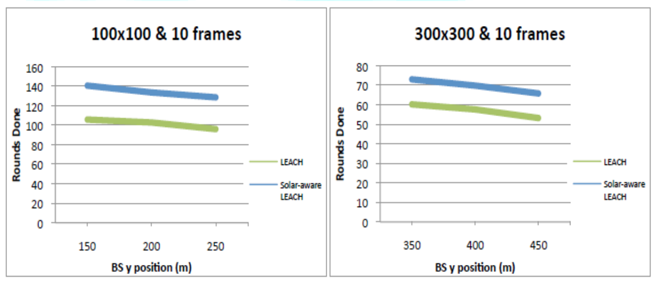

Figure 3: LEACH and Solar-aware LEACH results with 10 frame

When the steady phase has doubled the number of frames, i.e. the duration of the steady phase is doubled, the behavior of both protocols remain the same but decreasing the number of rounds achieved up to almost the half of them, as is shown in figure 4. This situation can be explained as an example of allow-cost set-up phase in energy terms, but a high-cost steady phase due to a non-optimal election of the cluster heads and the direct communication between cluster heads and base station.

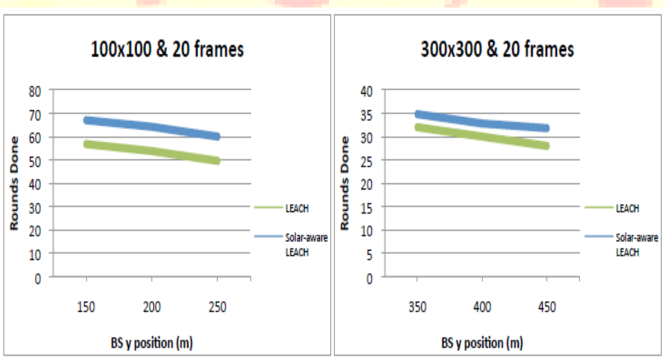

Figure 4: LEACH and Solar-aware LEACH results with 20 frames

In figure 5, it can be observed the results of both protocols with the longest steady phase simulated. The outcomes are really similar to the previous ones as expected, but decreasing the overall amount of rounds achieved. Looking at the figures 5, 6 and can be noticed that the longer the steady phase the smaller the difference in the outcomes between both protocols. It can be explained, as the solar-aware extension is more effective when the steady phase is short and the cluster head election is repeated in a short time. This situation is caused by the fact the election of solar-driven nodes as cluster heads happens on most cases and the duration of the solar state is usually shorter than the steady phase duration. Therefore the longer the steady phase the higher probability of a solar-driven node to turn into a battery-driven one, what could result in higher energy consumption in nodes that have been solar-driven more often than not within the cluster head election rounds.

C. First node dead

In the following figures can be seen the rounds achieved by both protocols when the first node dies. In the case of a short steady phase and a small area network the solar-aware extension gets better results, which achieves even more than 2 times the lifetime of the LEACH-distributed as is shown in figure 6. When the node density decreases or the area network increases, the results of both protocols get closer being still better in the case of Solar-aware LEACH.

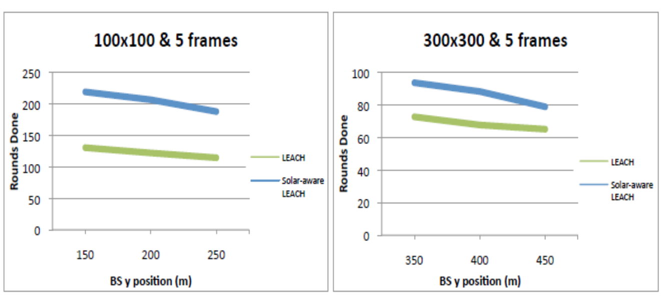

Figure 5: LEACH and Solar-aware LEACH results with 5 frames

If the duration of the steady phase increases, the results of both protocols are really similar as can be observed in figure 6. Even though the Solar-aware LEACH still achieves a longer lifetime, the difference between them is not very noticeable in large area networks, chiefly. Both protocols get worse results the farther is the BS to the closest node.

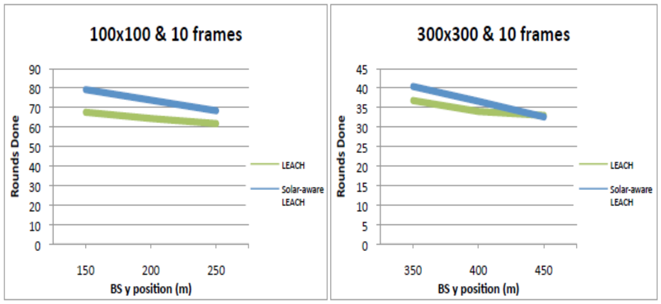

Figure 6: LEACH and Solar-aware LEACH results with 10 frames

V. Conclusion

In this research, it has been described an energy-efficient clustering algorithm in wireless sensor network. Here it is studied classic clustering algorithms in wireless sensor networks and finds two main reasons causing unnecessary energy consumption, which are fixed operation periods and too much information exchanged in cluster-heads selection. Here it is proposed clustering methods with less communication overhead for clustering based on federal management in k-means algorithm effective clustering and distributed algorithm. We propose a new protocol called Federal and Centralized Distributed Energy Efficient Clustering scheme for heterogeneous wireless sensor networks which the Base Station ensure that the high energy nodes becoming a gateways and cluster heads to improve network lifetime and average energy savings. Once the cluster heads and associated clusters are created, cluster-head-to-gateway routing is used to transfer the data to the base station to reduce the energy consumption of cluster head and decrease probability of failure nodes. Finally, the Simulation results have shown that our approach achieves lower average delay, more energy saving, more network lifetime and higher delivery ratio than the other protocols. For future work, a model with heterogeneous wireless sensor nodes with its topology to have good energy efficient and increasing lifetime network may be investigated. It enhances the proposed method to vary the range of the multi-hop zone dynamically by considering residual energy of each node.

REFERENCES

[1] Jin-Shyan Lee, Wei-Liang Cheng, "Fuzzy-Logic-Based Clustering Approach for Wireless Sensor Networks Using Energy Predication" IEEE SENSORS JOURNAL, VOL. 12, NO. 9, SEPTEMBER 2012.

[2] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "A Clustering Method for Wireless Sensor Networks with Heterogeneous Node Types," 18th International Conference on Computer Communications and Networks (IEEE ICCCN 2009), Aug. 2009.

[3] Kyuhong Lee and Heesang Lee," EECR: Energy Efficient Clustering and Routing Approach for Wireless Sensor Networks" 978-1-4577-1088-9/11 IEEE-2011.

[4] Noritaka Shigei, Hiromi Miyajima, Hiroki Morishita, Michiharu Maeda," Centralized and Distributed Clustering Methods for Energy Efficient Wireless Sensor Networks" Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol IMECS 2009, March 18 - 20, 2009, Hong Kong.

[5] W. Heinzelman, et. al., An application-specific protocol architecture for wireless microsensor networks," IEEE Transactions on Wireless Communications, 1(4):660-669, 2002.

[6] Younis, O., Fahmy, S., "HEED: A Hybrid, Energy- Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks," IEEE Trans on Mobile Computing, V3, N4, pp. 366-379, 2004.

[7] Chan, H., Perrig, A., "ACE: An Emergent Algorithm for Highly Uniform Cluster Formation," Proc. 1st Euro Workshop Sensor Networks, pp.154-71, 2004.

[8] Sundararaman, B., Buy, U., Kshemkalyani, A.D., "Clock Synchronization for Wireless Sensor Networks: a Survey," Ad Hoc Networks, V3, pp.281-323, 2005.

[9] R. Krishnan and, D. Starobinski, "Efficient Clustering Algorithms for Self-organizing Wireless Sensor Networks", Ad Hoc Networks, vol. 4, pp. 36-59, 2006.

[10] M. Qin and R. Zimmermann, "An Energy-efficient Voting-based Clustering Algorithm for Sensor Networks",Proceedings of the Sixth International Conference on Software Engineering, Artificial Intelligence, Neworking and ParalleliDistributed Computing and First ACIS International Workshop on Self-Assembling Wireless Networks (SNPD/SAVVN'05), pp. 444451, 2005.

[11] S. Banerjee and S. Khuller, "A Clustering Scheme for Hierarchical Control in Multi-hop Wireless Neworks", Proceedings of IEEE INFOCOM, 2001.

[12] O. Younis and S. Fahmy, "Distributed Clustering in Ad-hoc Sensor Networks: A Hybrid, Energy-Efficient Approach",.

[13] D. V. Paruchuri, "Adaptive Clustering Protocol for Sensor Networks", Aerospace, 2005 IEEE Conf.,pp. 1-8,2005.

[14] Jin-Mookim ,Hayeon-Kyu JOO, Seong-Sik Hong, Woo-Hyun ahn,Hwang-Bin Ryuo," An Efficient Clustering Scheme through Estimate in Centralized Hierarchical Routing Protocol," 2006 International Conference on Hybrid Information Technology (ICHIT'06).

[15] W.R.Heinzelman A. Otandrakasan, and H.Balakrishnan, "Energy Efficient Communication Protocol for Wireless SensorNetworks," Proceedings of the 33rd Hawaii International Conference on System ScienceVolume 2, Jan 2000.

[16] U. Sajjanhar, P. Mitra,"Distributive Energy Efficient Adaptive Clustering Protocol for Wireless Sensor Networks", Proceedings of the 2007 International Conference on Mobile Data Management, pp. 326 - 330,2007.

[17] Elbhiri Brahim,Saadane Rachid,Alba-Pages Zamora, Driss Aboutajdine,"Stochastic Distributed Energy-Efficient Clustering (SDEEC) for heterogeneous wireless sensor networks", ICGST-CNIR Journal, Volume 9, Issue 2, December 2009.

[18] Inbo Sim, KoungJin Choi, KoungJin Kwon and Jaiyong Lee, "Energy Efficient Cluster header Selection Algorithm in WSN", International Conference on Complex, Intelligent and Software Intensive Systems , IEEE, 2009.

[19] Linde, Y., Buzo, A., Gray, R.M., "An Algorithm for Vector Quantizer Design." IEEE Trans on Communications, V28, pp.84-95, 1980.

[20] Heinzelman, W.B., Chandrakasan, A.P., Balakrishnan, H., "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," IEEE Trans on Wireless Communications, V1, N4, pp. 660-670, 2002.