

# Optimization of clustering process in heterogeneous wireless sensor network

Sanjeev Kumar Gupta<sup>1</sup>, Poonam Sinha<sup>2</sup>

Electronics & Communication, Scope College of Engg., Bhopal, India<sup>1</sup>

CS & IT, Barkatullah University, Bhopal, India<sup>2</sup>

**Abstract:** Wireless sensor Networks are resource constrained systems that needs efficient utilization of all resources. With the limited capabilities of sensor nodes in terms of energy resources, processing and communication range, the cluster based protocols (clustering protocols) should be compatible with these constraints. Clustering provides an effective method for enhancing the lifespan of wireless sensor network. Optimization of clustering process improves overall performance in terms of cluster head selection, cluster balancing, aggregation and communication.

**Keywords:** Energy optimization, clustering, communication, aggregation and lifespan..

## I. INTRODUCTION

Wireless Sensor Networks have emerged as an important new area in wireless technology. Recent developments in wireless communication and electronics has enabled the development of low cost, low power, small sized battery operated sensor nodes. They consist of radio transceiver, microcontroller, power supply and actual sensor. Individual sensors have the capacity to detect events occurring in their areas of deployment [1]. The sensing circuitry measures parameters from the environment surrounding the sensor and transform them into an electrical signal, processing of such signals reveals some specific properties about the objects or events happening in the vicinity of the sensor. The sensor stores sensed data and sends to a command center (base station) via radio transmitter either directly or through a data concentration center (gateway). Normally sensor nodes are distributed throughout the sensing field which has to be monitored [2]. They organize themselves into a network through wireless communication and collaborate with each other to complete a common task.

Basic features of sensor networks are limited power, short range broadcast communication, self organizing capabilities, dynamic network topology, multi hop routing and large scale deployment. The advantage of wireless sensor network lies in their flexibility and scalability. As the capability of self organization and wireless communication make them to be deployed in an adhoc manner in remote and hazardous locations without the need of any existing infrastructure. Sensor node can communicate with a far away node in the network through multi hop communication. This allows the addition of sensor nodes in the network to expand the monitoring area and hence proves its scalability and flexibility property. There are many different ways to classify the wireless sensor network. The most common classification is homogeneous and heterogeneous sensor network. This classification is based on characteristic and functionality of the sensor nodes in the network [3]. In homogenous sensor networks all sensor nodes have the same characteristic, hardware and processing capabilities. In heterogeneous sensor networks some sensor nodes are

having higher processing capabilities and complex hardware. This leads to the research on heterogeneous network that prolongs the lifespan and stability of the network.

## II. CLUSTERING IN WIRELESS SENSOR NETWORK

The sensor nodes operate in three modes as sensing, computing and communication and all consume energy. It is widely accepted that the energy consumed in one bit of data transfer can be used to compute a large number of arithmetic operations in the sensor processor. In addition, in a densely deployed sensor network, the physical environment would produce very similar data in nearly all sensor nodes and transmitting such data is more or less redundant. Therefore, all the factors encourage assembling of sensor nodes into groups (clusters) such that data from sensor nodes of a group can be combined together in an intelligent way and transmit only compact data. This process of grouping of sensor nodes in sensor network is known as clustering.

### A. Clustering Process

Clustering schemes consist of four stages: cluster head selection, cluster formation, data aggregation and data communication. Fig. 1 shows the clustering process of one round. In this figure the setup state starts by the cluster head selection stage and proceeds by constructing clusters. The setup state is followed by the data transmission state, which is subdivided into data aggregation and data transmission phases [4].

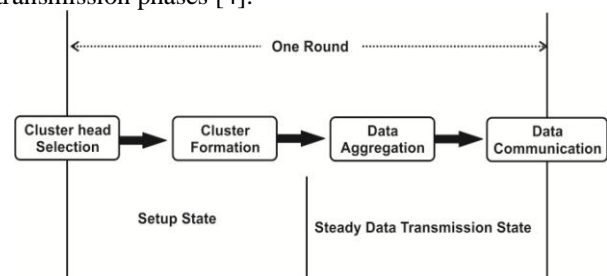


Figure 1. The clustering process of one round.

As sensor nodes are deployed in large numbers, designing and operating such a large size network would require scalable architecture and management strategies. The use of sensor nodes in such environment is energy constrained and their battery cannot be recharged. This leads to designing of energy aware algorithms for extending the lifespan of the sensor networks.

Clustering schemes/protocols have been used in sensor networks in order to achieve the following benefits.

- To minimize the total transmission power.
- To balance the energy consumption among all the nodes.
- To reduce the bandwidth demand and efficient use of limited channel bandwidth.
- To eliminate the redundant and highly correlated data in aggregation process.
- To prolong the battery lifespan of the individual sensors and overall network lifespan by implementing optimized management strategies.
- To reduce the size of the routing table stored at the individual nodes by localizing the route set up within the cluster.
- To increase the scalability of the network.

#### B. Clustering Characteristics

In clustering approaches, there are some characteristics that are related to the internal structure of the cluster [5].

- **Cluster Count:** Cluster count is the number of clusters formed in a round. More number of cluster leads to small size cluster distribution, which is better in terms of energy consumption. In some clustering approaches, the selection of cluster heads are pre assigned which construct fixed clusters or cluster heads can be selected randomly which results in variable number of clusters.
- **Cluster Density:** Cluster density is defined as number of cluster member in the cluster or cluster area. There is a big challenge to minimize the energy consumption of cluster heads in dense clusters. Some of the clustering approaches always use fixed clustering and has sparse density of cluster, but in dynamic clustering approaches cluster density is variable.
- **Cluster Size:** Cluster size is the maximum path length among the member nodes from cluster head. Small sized cluster is better in term of energy consumption because it minimizes the transmission distance and load of cluster head. In some clustering approaches, cluster size is fixed when clusters are fixed throughout the lifespan; otherwise it is variable for each cluster.
- **Message Count:** Message count is the number of message transmission required for cluster head selection. More number of message transmission leads to large amount of energy consumption for cluster head selection procedure. There are many algorithms which are non-probabilistic, and require message transmission for cluster head selection.
- **Stability:** If the members of a cluster are not fixed the clustering schemes are said to be adaptive. Otherwise it is considered as fixed because the cluster count is not varied throughout the clustering process. The fixed cluster count increases the stability of a sensor network.

- **Intra-Cluster Topology:** It indicates the communication within the cluster as direct or multihop. It may be single hop or multihop from sensor node to sensor node or sensor node to cluster head. However, this communication also depends on the sensor's range.

- **Intercluster Head Connectivity:** It indicates the capabilities of sensor nodes /cluster heads communication to base station. If the cluster heads are not having long distance communication capabilities, clustering scheme has to ensure some intermediate provision of routing to base station.

Wireless sensor networks organized into clusters can be broadly classified as homogeneous and heterogeneous networks, depending on the type and functionality of the sensor nodes in the network [6]. All sensor nodes in homogeneous sensor networks have the same characteristics, hardware and processing capabilities. Sensor nodes usually rotate the cluster head roles among themselves, which assures more uniform energy spending among the nodes in the network. In heterogeneous sensor networks, there are generally two or more types of sensors. First the sensor nodes have higher processing capabilities and complex hardware, used generally to create some sort of backbone inside the wireless sensor network [7]. They are designated as the cluster head nodes and, therefore, have to serve as data collectors and processing centers for data gathered by other sensor nodes. Second the participating sensors, with lower capabilities than the previous ones, used to actually sense the desired attributes in the field.

The most interesting research issue regarding clustering protocols is how to form the clusters, how many clusters should be formed in the network, how many members could be in the cluster and how to decide the size of cluster. It is stated that energy consumption of sensor network depends on cluster count, cluster density and cluster size. Cluster density and cluster size are correlated to the cluster count, if cluster count is lower, then cluster draws more energy to manage the more number of cluster members. The sensor network becomes unstable due to dense cluster. Cluster size and cluster density can be reduced by increasing the cluster count. Optimization of cluster count, cluster density and cluster size are the biggest challenge in wireless sensor network is partitioned in clusters in such a way that the stability and network lifespan are enhanced.

### III. OPTIMIZATION OF CLUSTERING PROCESS IN WIRELESS SENSOR NETWORK

In this paper the optimization of clustering process in wireless sensor network includes cluster head selection, cluster formation, data aggregation and data communication that bring together all these to enable development of a wireless sensor network. Each component has some issue to be solved as shown in figure 2. In cluster head selection, optimal number of cluster formation is an important issue. In cluster formation, cluster density and balancing among cluster is another issue. Data aggregation and data communication are strongly correlated to each other. Maximizing data

aggregation and minimizing data communication is also an important issue which is solved by cluster size and inter cluster communication.

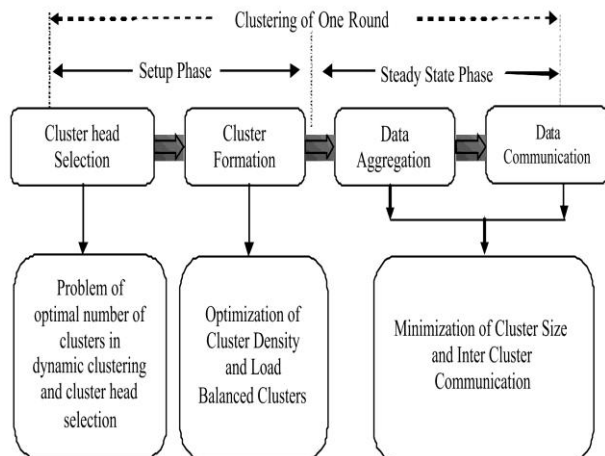


Figure 2. Optimization of clustering process in heterogeneous WSN

#### A. Cluster Head Selection Phase

The first step in clustering schemes is the selection of cluster heads. Cluster head acts as a gateway between the sensor nodes and the base station. The function of cluster head is to perform communication for all the sensor nodes in the cluster and aggregating the data before sending to base station.

Therefore, cluster head selection plays a significant role in the subsequent procedures of clustering schemes to enhance the lifespan and energy-efficiency of the network. Due to this importance, many researchers have focused on optimizing the cluster head selection process. The different strategies implemented in cluster head selection processes are classified as distributed control (self-organized schemes) and centralized control (assisted schemes). In distributed control each sensor node can run their own algorithm and take the decision of becoming cluster head. In centralized control, centralized authority groups the nodes to construct cluster and also select cluster heads for each cluster. The self organized schemes are further classified as probability based and non-probability based. Assisted schemes are also classified as base station assisted and cluster heads assisted schemes. Fig. 3. shows the process of cluster head selection

- **Probabilities Based Clustering Algorithms:** In the category of probability based clustering algorithms, a prior probability assigned to each sensor node is used to determine the initial cluster heads [8]. The probabilities initially assigned to each node often serve as the primary criterion in order to decide individually on their selection as cluster heads. However, other secondary criteria may also be considered during cluster head selection process as the residual energy, initial energy, average network energy etc. Beyond the high energy efficiency, the clustering algorithms of this category usually achieve faster execution or convergence times and reduce volume of exchanged messages.

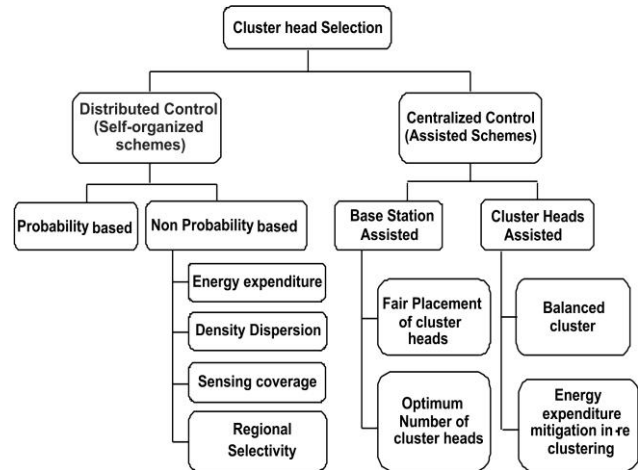


Figure 3. Cluster head selection process.

- **Non Probabilities Based Clustering Algorithms:** In the category of non probability based clustering algorithms, more specific criteria for cluster head selection is considered. They are mainly based on the factors that are considered to improve the stochastic cluster head selection schemes such as energy expenditure [9, 10], density dispersion [11], sensing coverage [12] and regional selectivity. This type of protocols generally require more exchanges of messages and thus leading sometimes to worse time complexity than probabilistic or random clustering algorithms. In energy expenditure schemes the node selected as cluster head consumes more energy than other nodes within the network and demise of a cluster head leads up to losing all the data of an area monitored by its cluster member nodes [13]. In density dispersion based clustering schemes selection of the nodes having the highest remaining energy as cluster head is quite desirable [14]. The dispersion of cluster heads should confirm to the nodes density in distribution; in other words, selection of the nodes from denser areas leads to conserving more energy. In sensing coverage schemes the main objective of implementing sensor nodes is to sense a physical phenomenon and avoiding coverage holes within the network should be a top priority [15]. In regional selectivity based schemes cluster head selection is not the initial stage of the clustering algorithm. First each sensor node finds its neighbouring nodes in a predefined radius or number of hops, or performs a rudimentary regional cluster formation stage based on the position of the nodes within the network, and then the most qualified node in each section is selected as cluster head based on a distributed algorithm [16].

- **Base Station Assisted Clustering Algorithms:** The inexhaustible resources of energy and high processing capabilities of base station are considered as a powerful and reliable source for sensor nodes to which they can shift the burden of cluster head selection and cluster formation phases. This also improves the capabilities of end-user to control the placement and number of cluster heads through the base station in accordance with the characteristics of the network and type of applications. However, these entail the periodic update of the base station with necessary information by sensor nodes. This

scheme has been divided as fair placement of cluster heads [17] and optimum number of cluster heads [18].

• **Cluster Heads Assisted Clustering Algorithms:** Cluster heads can collect the up-to-date states of their cluster members through continual communications in data transmission phases. Using this information, cluster heads can assist in the selection of the next round cluster heads to balance the clusters and to eliminate extra energy expenditures in re-clustering stages. This scheme has been divided as balanced clusters [19] and energy expenditure mitigation in re-clustering.

#### B. Cluster Formation Phase

Cluster formation phase starts from broadcasting the advertisement messages by cluster heads to announce their selection to other nodes, and ends by sending back a join-message to the optimum cluster head by each node.

Cluster formation schemes are grouped into optimal clustering and event-driven clustering. In optimal clustering schemes, the focus is either to manipulate the size of the clusters according to the type of the application and data transmission, or to minimize and balance the energy expenditure in the network by considering the factors such as data correlation, relay traffic and residual energy. On the other hand, the event-driven clustering schemes are proposed to prolong the network lifespan by eliminating dispensable clustering throughout the network and trigger the cluster formation stage only when and where it is needed. The cluster formation phase is given in Fig. 4.

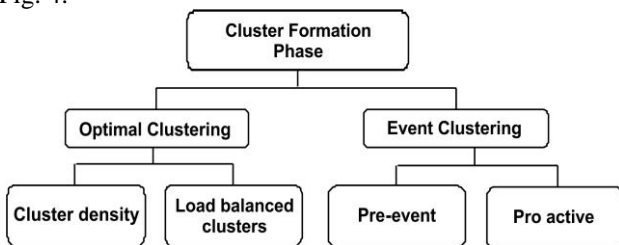


Figure 4. Cluster formation phase.

• **Optimal Clustering:** In optimal clustering cluster formation is based on minimizing the energy expenditure of cluster members [20]. Sensor nodes join the nearest cluster heads by calculation of their distance from cluster heads through the signal strength of the received advertisement messages. This method of clustering does not consider the size of the constructed clusters or even distribution of energy expenditure within the clusters.

• **Event-Driven Clustering:** In this type of cluster formation, researchers have focused on generating energy-efficient clusters and most of these solutions offer pre-event clustering and pro-active clustering [21]. However, cluster formation in the entire field prior to occurrence of an event imposes a significant overhead in terms of energy and processing on the network, while it does not guarantee the better performance of the network in some applications [22].

#### C. Data Aggregation Phase

Data aggregation is the process of aggregating the data from multiple sensors to eliminate redundant transmission and provide fused information to the base station. The main goal of data aggregation algorithms is to gather and

aggregate data in an energy efficient manner. Since sensor nodes are energy constrained, it is inefficient for all the sensors to transmit the data directly to the base station. Data generated from neighboring sensors is often redundant and highly correlated. In addition, the amount of data generated in large sensor networks is usually large for the base station to process. The following are the data aggregation schemes mostly used [23] and fig. 5. shows the data aggregation process.

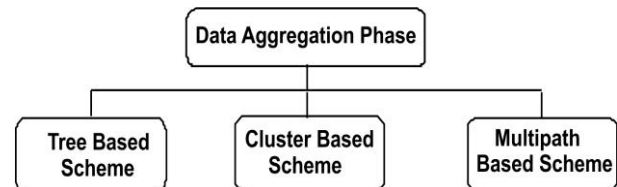


Figure 5. Data aggregation phase.

• **Tree Based Data Aggregation :** This tree based data aggregation scheme is used to achieve distributed data aggregation through some data aggregator nodes in the network. This is assured that the data paths of sensor nodes include these data aggregator nodes [24]. Tree-based data aggregation protocols are used to construct an energy efficient data aggregation tree [25].

• **Cluster Based Data Aggregation :** In this scheme the clusters are formed and in every cluster, a cluster head acts as data aggregation point. LEACH [26] is a clustered approach where cluster heads act as data aggregation points. Another cluster-based data aggregation protocol is HEED [27]. The cluster heads selection in this protocol benefits from the availability of multiple power levels at sensor nodes. A combined metric is composed of the node's residual energy and the node's proximity to its neighbors.

• **Multipath-Based Data Aggregation :** In multipath-based data aggregation scheme, sensor nodes divide their aggregated data into several parts and send the aggregated data to a single point through multiple paths [28]. The main idea behind these schemes is to send duplicate small data to the base station over multiple paths to improve the robustness of the network. Multipath-based data aggregation usually employ ring topology in which sensor nodes are divided into several levels based on the distance from the base station in terms of number of hops [29].

#### D. Data Communication Phase

In this phase, cluster heads as the coordinators of the cluster transmit the aggregated data to the base station for further processing by the end user according to the type of the application. The transmission of a packet from sensor nodes to the cluster head is called intra-cluster transmission and from cluster heads to the base station is called inter-cluster transmission. Intra cluster communication is further divided into single hop and multi hop transmission. Inter cluster communication is further divided into direct and multi hop transmission. In single hop transmission, all the sensor nodes in the respective cluster send sensed data to its respective cluster head directly [30]. Distance is not taken into account in single hop communication [31]. In multi hop transmission all the

sensor nodes of the cluster send sensed data to the nearest neighbor, which is on the way to cluster head [Yin 2008]. In the direct communication cluster head sends aggregated data to base station directly [32]. Distance is not taken into account in direct communication. In the multi hopped communication all the cluster heads send aggregated data to the nearest neighbor cluster head, which comes in the way to base station. Fig. 6 shows the data communication process.

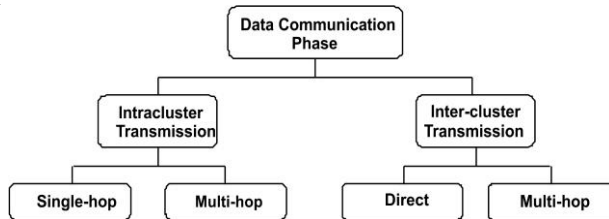


Figure 6. Data communication phase.

#### IV CONCLUSION

In this paper, the optimization of clustering process in wireless sensor network has been explained. The problem of optimum no of cluster heads selection can be optimize by non probability based clustering schemes in wireless sensor network. Cluster formation can be optimized by controlling the redundant nodes and balancing the clusters through variable threshold. To optimize the cluster size and inter cluster communication a threshold can be formulated for cluster head selection which ensures that more number of cluster heads are selected in each round. The proper coordination among grouped cluster heads increase the stability period and network life span.

#### REFERENCES

[1] Akyildiz Ian F., Weilian Su, Yogesh Sankarasubramaniam and Erdal Cayirci, "Wireless sensor networks: a survey", Computer networks 38 (Elsevier), No. 4 (2002): PP. 393-422.

[2] Lewis F. L., "Wireless Sensor Networks", Smart Environments: Technologies, Protocols, and Applications John Wiley, New York, USA, (2004).

[3] Juhana Yrjola, "Summary of Energy Efficiency Communication Protocol for Wireless Micro Sensor Networks", (2005).

[4] Mhatre Vivek and Catherine Rosenberg, "Homogeneous Vs Heterogeneous Clustered Sensor Networks: A Comparative Study", Communications, 2004 IEEE International Conference on 20-24 June (2004), Vol. 6, PP. 3646 – 3651.

[5] Arboleda Liliana MC and Nidal Nasser. "Comparison of clustering algorithms and protocols for wireless sensor networks", Electrical and Computer Engineering, 2006. CCECE'06. Canadian Conference on. IEEE, (2006).

[6] Katiyar Vivek, Narottam Chand, "Clustering Algorithms for Heterogeneous Wireless Sensor Network: A survey", International Journal of Applied Engineering Research Dindigul, vol. 1, No. 2, (2010), ISSN. 0976-4259.

[7] Naeimi Soroush, Hamidreza Ghafghazi, "A Survey on the Taxonomy of Cluster-Based Routing Protocols for Homogeneous Wireless Sensor Network", Sensors 12.6 (2012): 7350-7409.

[8] Sheikhpour Razieh, Sam Jabbehdari and Ahmad Khadem-Zadeh, "Comparison of energy efficient clustering protocols in heterogeneous wireless sensor networks", International Journal of Advanced Science and Technology 36 (2011): PP. 27-40.

[9] Song Dezhen, "Probabilistic modeling of leach protocol and computing sensor energy consumption rate in sensor networks", Texas A & M University (2005).

[10] Li X., Li N. Chen, "An Improved Leach for Clustering Protocols in Wireless Sensor Networks", In Proceedings of the International Conference on Measuring Technology and Mechatronics Automation (ICMTMA 2010), Changsha, China, 13–14 March (2010), pp. 496-499.

[11] Hasbullah, H.; Nazir, B., "Region-Based Energy-Aware Cluster (Rec) for Efficient Packet Forwarding in WSN", In Proceedings of the International Symposium in Information Technology (ITSim 2010), Malaysia, 15–17 June (2010); pp. 1–6.

[12] Tao Y. Zheng Y. L., "The Combination of the Optimal Number of Cluster Heads and energy Adaptive Cluster- Head Selection Algorithm in Wireless Sensor Networks", In Proceedings of International conference on Wireless Communication, Networking and Mobile Computing (WiCOM 2006) Wuhan, China 22-24 Sept. (2006) PP: 1-4.

[13] Mammu A. S. K., Ashwani Sharma and Unai Hernandez-Jayo, "A Novel Cluster-based Energy Efficient Routing in Wireless Sensor Networks", IEEE Computer Society, AINA, (2013).

[14] Tsai, Y.R., "Coverage-preserving routing protocols for randomly distributed wireless sensor networks", IEEE Trans. Wireless. Comm.( 2007), Vol. 6, 1240–1245.

[15] Sun, J.; Liao, F.; Peng, Y., "A Clustering Routing Protocol Based on HCMC in WSNs", In Proceedings of the International Colloquium on Computing, Communication, Control, and Management (CCCM'08. ISECS), Guangzhou, China, 3–4 August 2008; pp. 577–581.

[16] Munaga H. Murthy J., "A Novel Trajectory Clustering Techniques for Selecting Cluster Heads in Wireless Sensor Networks", International Journal Recent Trends Eng. (2009) Vol. 1, PP: 357-361.

[17] Wang, Q.; Xu, K.; Hassanein, H.; Takahara, G., "Design and Stochastic Modeling of Distributed, Dynamic, Randomized Clustering Protocols in Wireless Sensor Networks", In Proceedings of the 24th IEEE International Performance, Computing, and Communications Conference (IPCCC 2005), Phoenix, AZ, USA, 7–9 April (2005), pp. 639–640.

[18] Wang H., Yu, X. Kong, D. Yan, "Route Protocol of Wireless Sensor Networks Based on Dynamic Setting Cluster", In Proceedings of the International Conference on Information Acquisition (ICIA'07), Jeju, Korea, 8–11 July (2007); pp. 112–117.

[19] Heinzelman Wendi Rabiner, A. chandrakasan and H. Balakrishnan, "An Application Specific Protocol Architecture for Wireless Micro Sensor Networks", IEEE Transactions on wireless communication, 2002.

[20] Melodia T, Pompili D., Gungor V. C., "A Distributed Coordination Framework for Wireless Sensor and Actor Networks", In Proceedings of the 6<sup>th</sup> ACM international symposium on Mobile ad hoc networking and computing (MobiHoc 2005), Urbana-Champaign, IL, USA, 25-28, May (2005) PP: 99-110.

[21] Chen X., Niu Z., "A Randomly Delayed Clustering Method for Wireless Sensor Networks", In Proceedings of the IEEE International Conference on Communication (ICC 2006) Istanbul, Turkey, 11-15 June (2006) PP. 3353-3358.

[22] Sharifnejad Mona, and Mansoureh Ghaisabadi Farahani, "Enhancing the Performance of Data Aggregation in Sensor Networks", Information and Communication Technologies, (2006). ICTTA '2nd, vol. 1, PP. 1905-1909.

[23] Madden S., M. J. Franklin, J. Hellerstein, and W. Hong. TAG; "A Tiny AGgregation Service for ad hoc Sensor Networks", In Proceedings of the Fifth Symposium on Operating Systems Design and Implementation (OSDI 02), Boston, MA, USA, December (2002).

[24] Wuhib F., M. Dam, R. Stadler, and A. Clemm., "Decentralized computation of threshold crossing alerts", In 16th IFIP/IEEE International Workshop on Distributed Systems: Operations and Management, DSOM 2005, Barcelona, Spain, October 2005, number 3775 in LNCS, Oct. (2005) pp. 220–232.

[25] Heinzelman Wendi Rabiner and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Micro Sensor Networks", IEEE, Proceeding of the 3<sup>rd</sup> Hawali International Conference on System Science, (2000).

Younis Ossama and Sonia Fahmy, "HEED: a Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks", Mobile Computing, IEEE Transactions on 3.4 (2004), PP. 366-379. Kalpakis, K., Dasgupta, K. and Namjoshi, P., "Efficient algorithms for maximum lifetime data gathering and aggregation in wireless sensor networks, Computer Networks", The International Journal of Computer and Telecommunications Networking, (ELSEVIER), Vol. 42, No. 6, (2003), pp.697-716.

[26] Jason L. Hill, and David E. Culler, "Mica: a wireless platform for deeply embedded networks", IEEE Micro, Vol. 22, No. 6, pp.12-24, (2002).

[27] Nazir B., Hasbullah H., "Energy Balanced Clustering in Wireless sensor Network", In proceedings of the international Symposium in Information Technology (ITSim 2010), Kuala Lumpur, Malaysia, 15-17 June (2010), PP. 569-574.

- [28] Abdulsalam H.M., Kamel L.K., “W-Leach: Weighted Low Energy Adaptive Clustering Hierarchy Aggregation Algorithm for Data Streams in Wireless Sensor Networks”, In Proceedings of the 2010 IEEE International Conference on Data Mining Workshops, Sydney, Australia 13-14 Dec. (2010) PP: 1-8.
- [29] Yin G., Yang G., “An Energy-Efficient Routing Algorithm for wireless Sensor Networks”. In Proceedings of the International Conference on Internet Computing in Science and Engineering (ICICSE 2008), China, 28-29 Jan. (2008) PP. 181-186.
- [30] Vidhya J. and Dananjayan, “Lifetime maximization of multihop WSN using cluster-based cooperative MIMO scheme”, International Journal Computer. Theory Eng. 2.1 (2010): PP. 20-25.