

Balanced Unequal Clustering Algorithm under Noisy Environment

Ravendra Singh ⁺

Dept. of CS & IT, MJP Rohilkhand University, India

Abstract. The transmission link failures are common in wireless sensor network because the sensor nodes are deployed in hostile environment, so it is necessary to maintain dynamic topology to monitor the link quality, which leads to reduce the network lifetime due to increase in communication overheads. Hence energy efficient BUCAN (Balanced Unequal Clustering Algorithm under Noisy environment) has been proposed here to extend the network lifetime of WSNs. The stable and energy efficient approach is proposed here to form unequal clusters using BinInv method to solve the hotspot region problem, thereafter selects the optimal cluster head (CH) to prolong the network lifetime, which utilizes fuzzy logic for handling uncertainties in CH radius estimation, taking into account the factors such as the remaining energy, surrounding noise and the localization of nodes. At last, an efficient hierarchical routing scheme has been proposed to deliver the dynamically compressed data to base station which estimates the link quality by considering signal strength of the multi hop route. The simulated result demonstrates that our proposed algorithm has significant effectiveness in terms of balancing energy consumption as well as maximizing the network lifetime under the same routing criteria better than other traditional algorithms.

Keywords: balanced energy, unequal clustering, noisy environment, wireless sensor network.

1. Introduction

The limitation of power in wireless sensor networks (WSNs) is because of the two operating factors. First, sensors consume lots of energy while operating with full intensity which cannot be supported by a small battery for longer duration. Second, battery cannot be replenished in a sensor network due to inaccessibility of physical locations. Therefore energy management is a critical task in wireless sensor networks due to severely power limited sensors. In order to achieve high energy efficiency and guaranteed lifetime [9], sensor nodes are organized into clusters for aggregating data, which also improves the scalability, resource utilization, power control and promotes local decisions to minimize the need for central organization. Every cluster has a CH which arranges intra-cluster transmission schedules, collects data from its cluster members, combine several correlated data signals into a smaller set of information and forwards data to base station (BS), so they may die earlier because most of the energy of WSN is consumed during communication. Therefore, there is a need to distribute energy consumption adequately among sensor nodes by rotating the CH in cluster in each epoch for balancing remaining energy to prolong the network lifetime for real-time applications.

To address this problem, fuzzy logic can be an alternative model that has the potential for dealing with conflicting situations and imprecision in data using heuristic human reasoning without the need of complex mathematical models [1]. There have been substantial amount of research on clustering protocols for WSNs and presented various algorithms. The CHEF [3] and EAUCF [8] are such clustering protocols which help in reduction of redundancy and improvement over the lifetime of the network. In this line the proposed BUCAN algorithm shows the improvement over competition range for selection of better CH as well as best possible multi-hop route for delivering the compressed data to BS, which results in prolong of WSN lifetime. The rest of the paper is organized as follows: design issue and related work and is given in Section 2,

⁺ Corresponding author. Tel.: + 9412148941.
E-mail address: rsiet2002@gmail.com.

the proposed algorithm and research methodology behind is explained in Section 3, Simulation, results and analysis is discussed in Section 4 and finally Conclusion given at last.

2. Design Issues and Related Work

In order to solve the problems in cluster formation, CH selection and forwarding the gathered data to BS in WSNs, researchers have proposed a variety of methods. LEACH [2] is a most popular clustering algorithm that utilizes randomized rotation to balance energy consumption of cluster-heads over the network. LEACH-C [5] is a centralized clustering algorithm in which BS has the right to select the clusters based upon the annealing algorithm to find optimal number of clusters. These approaches require a large number of messages to be exchanged to improve the global network performance. Other Artificial Intelligence-based protocols require high computational capabilities and a complete knowledge of the WSN prior to deploying the sensor nodes. On the other hand, fuzzy logic-based approaches performed very well in energy efficient routing and clustering heuristics in WSNs. A fuzzy logic approach proposed in paper [6] is used to select a cluster head by considering energy and distance to efficiently route data in WSNs. Another fuzzy logic-based energy-efficient algorithm CHEF [3] selects the cluster heads in a distributed way over energy and distance as input descriptors. Liang [4] presented fuzzy based event detection approach for improving the accuracy of cluster-head election to efficiently route data in WSNs. The F-MCHEL [7] elects a CH with maximum residual energy (RE) as a master CH to forward the aggregated data to BS which is not realistic approach for large network. Article [10] is the survey of existing clustering protocols.

The fuzzy based unequal clustering algorithm EAUCF [8] adjusts the cluster-head radius considering the residual energy and the distance to the base station parameters of the sensor nodes and author claims that EAUCF algorithm is best suitable for real-time applications those require guaranteed lifetime but it also degrades the performance if selected CH placed in noisy environment, which is untouched by the researchers in best of my knowledge. Our previous work [12] is similar to EAUCF, the main difference between is their competition radius estimation methods, which is further modified in this paper to increase the lifetime of WSN. Here surrounding noise has been taken into account to calculate the signal strength for improving the competition range for selection of better CH as well as best possible multi-hop route for delivering the compressed data to BS.

3. Proposed Algorithm

Problems of energy consumption and connectivity are critical for randomly distributed wireless sensor networks. Connectivity is measured in terms of successful packet transmission to the CH and finally delivering to base station. At the same time optimized energy consumption and balancing the load for the region is required for guaranteed lifespan because heavily loaded nodes die earlier. The network lifetime comes to end when any node collapses its energy, the best way to satisfy the WSN's user or application demands for a pre-configured network lifetime is to distribute energy consumption adequately among sensor nodes, therefore the task of being cluster head has been rotated among sensors in each epoch to distribute the energy consumption across the network in the proposed algorithm. Hence, energy efficient unequal clustering has been proposed here for balancing the network load, it should be must be installed on both sink (CH) and general nodes in the network. The operation of algorithm is divided in three states. Primarily in the setup state, all the nodes are deployed in the region and divide the region into homogeneous unequal clusters. There after steady state phase begins with cluster head selection followed by collection of information locally, data fusion processing and further transmit the aggregated dynamically compressed data towards the global sink for further processing in last phase.

All the nodes of each cluster communicate to each other by broadcasting the join message with node_id and potential value (P_value) to enrol nodes in a cluster. The P_value of each node contained crisp value using fuzzy system of Residual energy (RE), Localization (LO) and Signal strength (SS), if it is more than T% than node will only be able to participate in the competition of cluster head selection.

Definitions of terminology:

Residual Energy (RE) of sensor node must be higher than threshold to participate in election of CH. If

its energy is less than threshold than the node will be declared as active node till exhausted but it will not participate in the election. Under this standard, the size of the battery capacity in the current node is in focus.

The value of **Localization (LO)** of sensor nodes is used to determine the accurate physical position which classifies the nodes based on how central the node is to the cluster, so ideal cluster head will be selected with minimum value of LO. Let sensor nodes be randomly distributed in a topology at (x1,y1), (x2,y2), (x3,y3),....(xn,yn) and distances between node i and other nodes be:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

The average distance of node (Dave) for ith node among n distributed nodes is:

$$(D_{ave})_i = \frac{\sum_{j=1}^n d_{ij}}{n-1} \quad \text{for} \quad \forall j; j \neq i \text{ and } LO_grade = \left[\frac{D_{ave}}{(D_{ave})_{max}} \right]^{-1}$$

The signal strength (SS) of the node is critical, which is depends on transmission efficiency, transmission loss, noise level of surrounding and directivity index of antenna of the node. A node with maximum *signal strength* should be selected as a cluster head, which is measured using following formula:

$$SS = TE - TL - NS + DI$$

where *SE*=Transmitted Energy, *TL*=Transmission Loss, *NS*=Noise Level of the surrounding and *DI*=Directivity index of antenna.

BinInv:

Initially all sensor nodes deployed in R_region
Each node contains its ID, value of RE, LO_grade and expected SS.

Divide the R_region into levels Lj and

Area of levels = r+Δj

Where, r = Area of level &

i=M to 1 and j= 1 to M

Set L1 = Hotspot Region

Cluster formation starts at each level using inverted binary tree concept.

Number of cluster at level i = 2j

Total number of clusters in R_region = 2j + 2j2M.

Fig. 1: Algorithm for cluster formation.

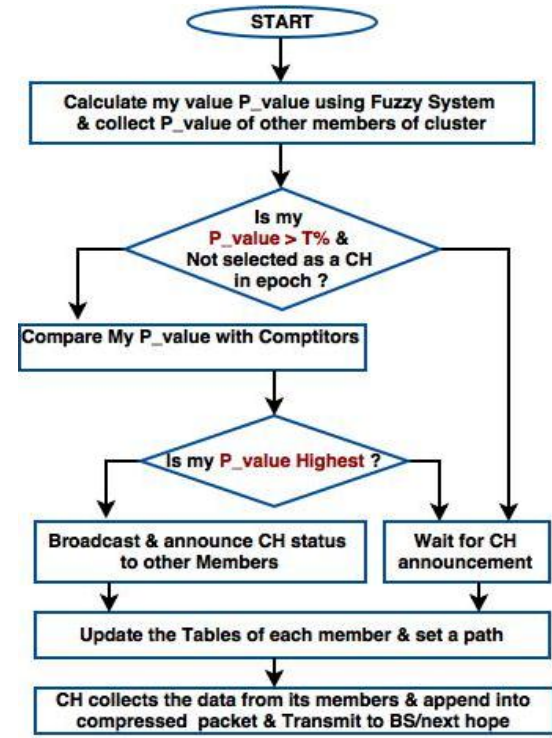


Fig. 2: Flowchart of steady state phase.

3.1. Setup state: Cluster formation

The CHs near to BS positioned in hotspot region exhaust early because they transmit large amount of data in multi-hop routing. Hence the cluster formation starts here by dividing region (R_region) into unequal clusters using inverted binary tree concept, in which area of the clusters decreases but number of cluster increases when moving towards the base station to balance the load. First the region is divided in levels and then form clusters from furthest level LM by using above BinInv algorithm and ends at BS which

is a root of binary tree, which minimizes the number of dead nodes to overcome the problem of hotspot region.

3.2. Steady state: Cluster head selection process

Instead of classical linear controllers or mathematical models, we introduce a fuzzy logic-based controller in this article to select CH due to its simplicity, clarity and suitability with WSN applications. Fuzzy systems are very useful in situations involving a highly complex system whose behaviours are not well understood and as well as where an approximate, but fast, solution is required. A flowchart of steady state phase of the algorithm is shown in figure 2. Each participant will broadcast their node_id and P_value to other competitors in the steady state phase of the algorithm and the node with largest P_value will be selected as a CH for prescheduled period. If there is tie then CH with least ID is chosen among tied ones. Any eligible node becomes CH once in an epoch and rest round it will remain out of race. The inference engine performs different calculations in fuzzy logic-based systems to select optimum CH. In this paper, the Mamdani inference engine works with the input of a Mamdani fuzzy logic system, that is a crisp value. Thus, the RE, LO_grade & SS input fuzzy values from the fuzzification stage apply to all fuzzy rules by the inference engine. The linguistic variables of residual energy, localization and signal strength are divided into three levels: low, Average and high. Following control output derived from the combination of input, output membership functions and fuzzy rules are still a fuzzy element. To make fuzzy output accessible to real applications, a defuzzification process is required.

P_value is the ability of node to become cluster head, the linguistic variables of P_value is divided in seven levels: Impossible (I), Very Unlikely (VU), Unlikely (U), Even Chance (C), Likely (L), Very Likely (VL), and Possible (W). The node with output membership function 'W' will be selected as CH for the current epoch, nodes with 'C', L, 'VL', 'W' marked as participant nodes whereas the nodes 'VU' & 'U' will declared as active nodes and remaining nodes will marked as dead nodes. Active nodes will never participate in election process in upcoming epoch to save the energy, whereas 'VU' nodes are not having sufficient energy level so they handover their data to nearby participant node to transmit their data towards CH.

Rules evaluation:

Knowledge base of FIS contains the definitions for each of the fuzzy sets and maintains a system operator AND or OR to express fuzzy rules mapping with a rule reliability matrix. Inference units with fuzzifiers and antifuzzifiers calculate real output values from the real input values. The fuzzifier takes input as a fuzzy set, the inference unit match with it in the rules stored in the knowledge base. Then, inference units calculate the strength of each rule and aggregate of a fuzzy set encompasses arrange of output values and value is defuzzified in order to resolve a single output value from the set using Centroid method.

Table 1. Membership Functions Table

RULES	Antecedents			Consequents
	SS	Lo grade	RE	P_value
R (1)	High	Low	Low	I
R (2)	High	Low	Average	C
R (3)	High	Low	High	VL
R (4)	High	Average	Low	U
R (5)	High	Average	Average	C
R (6)	High	Average	High	VL
R (7)	High	High	Low	U
R (8)	High	High	Average	L
R (9)	High	High	High	VL
R (10)	Average	Low	Low	I
R (11)	Average	Low	Average	U
R (12)	Average	Low	High	C
R (13)	Average	Average	Low	VU
R (14)	Average	Average	Average	L
R (15)	Average	Average	High	VL
R (16)	Average	High	Low	VU
R (17)	Average	High	Average	L
R (18)	Average	High	High	VL

R (19)	Low	Low	Low	I
R (20)	Low	Low	Average	VU
R (21)	Low	Low	High	U
R (22)	Low	Average	Low	I
R (23)	Low	Average	Average	U
R (24)	Low	Average	High	C
R (25)	Low	High	Low	VU
R (26)	Low	High	Average	L
R (27)	Low	High	High	VL

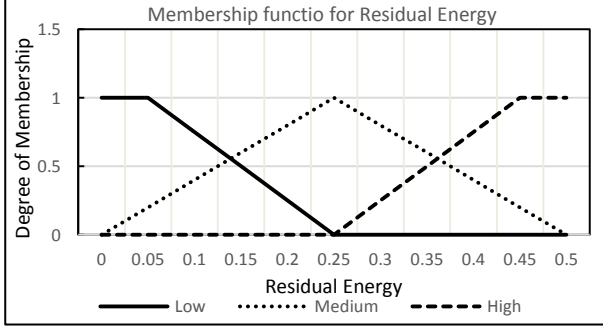


Fig. 3: Fuzzy sets for residual energy.

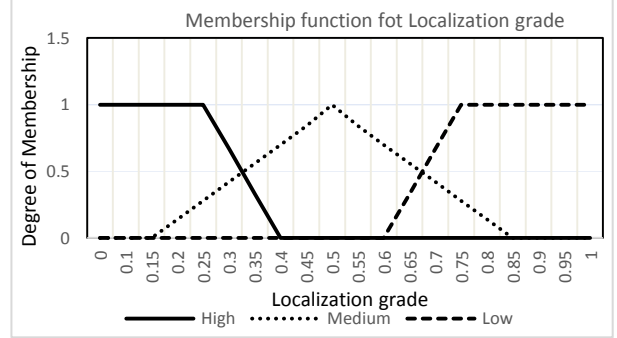


Fig. 4: Fuzzy sets for localization grade.

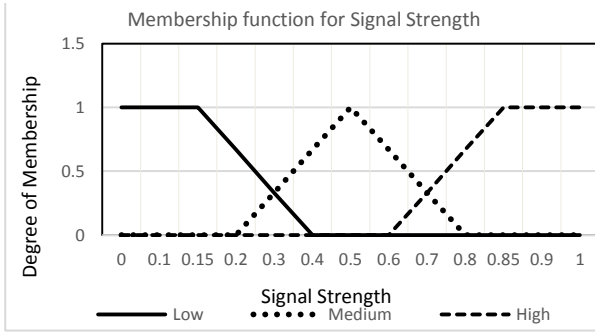


Fig. 5: Fuzzy sets for signal strength.

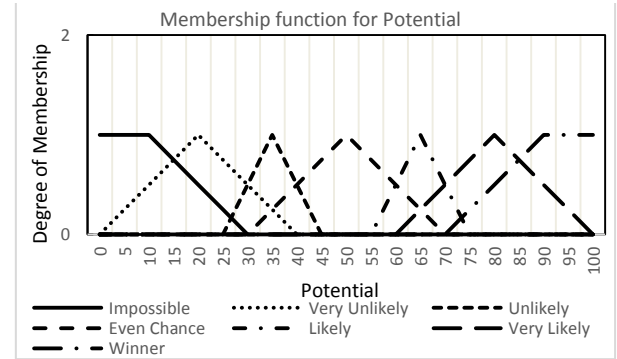


Fig. 6: Fuzzy sets for potential (P_value).

Triangle and trapezoidal membership functions are used since their degree can be easily determined. Triangle membership functions are used to represent the fuzzy input sets Average and trapezoid membership functions to represent low and high fuzzy sets. The fuzzy inference rules (if-then) defined as binary fuzzy relations between input and output. The Input Function uses three membership functions to show the degrees of input variables and generate 27 possible combinations for Rule base. Table 1 shows the rule base defined using “if-then” rules with “and” operators among input variables. The membership functions developed in MATLAB for the network scenario in this paper are represented in Figures 3 to 6.

3.3. Final State: Transmission of aggregated data to BS

The shortest path to forward the aggregated data to base station from host CH is selected on the basis of link quality indicator (LQI), which depends on signal strength. Every selected CH communicates with Base Station/route node and vice versa and established a multi hope tree. Transmission loss of the path will be higher if CHs of the path far away from each other and it will also increase with increasing number of hopes, so transmission loss is subtracted from signal strength. Therefore base station estimates the signal strength of all the paths hope by hope and link with maximum LQI ($\sum_{i=0}^n SS_i$) will be selected as a transmission path in proposed algorithm.

In addition to above, our proposed algorithm also save the energy of the transmitter (CH) by transmitting the compressed data to BS hope using algorithm [11]. The data collected by intermediate nodes of the path is appended with the received data dynamically from previous hope in compressed form without

decompression followed by compression. This approach reduces unnecessary decompression followed by compression at each intermediate node to save energy and time leg.

4. Simulation and Results

The proposed network scenario is designed and implemented using Network Simulator and the Fuzzy Interference system used for CH selection is implemented using MATLAB. In our 200-node network, there are specifically three categories of sensors (excluding the base station) – general sensing nodes, assistant node and the cluster head node. All the nodes are homogeneous but are assigned different tasks to function. Several measurement metrics were collected from our propose simulator to evaluate the performance. The nodes are deployed randomly in the area of 100×100 m², and the BS is located at the centre. The simulation is performed with similar initial energy (0.5J) of all the sensor nodes and with random generated different initial energy (0.2-0.6J). The other parameters including energy consumption model is same as discussed in paper [4], and whereas DAE= 5nJ/bit/signal, REE= 50nJ/bit, RAE= 10 pJ /bit/m². The energy variation of Hotspot Zone with different number of communications (Rounds), Number of nodes at Hotspot Zone = 125 BI=Before Index, AI=After Index and for different rounds. Simulation results are shown in Figure-7 & Figure-8. The network life time is depends on alive nodes, alive nodes are active nodes and energy of these nodes not exhausted its energy yet after sufficient simulation time. It can be seen from result as shown in Figure-7 that the number of nodes alive with same initial energy in the proposed method is always higher than the LEACH-C & EAUCF algorithms in noisy environment. The result also shows that the fuzzy based node selection technique increases packet delivery ratio, decreases the link failure, lowers error rate and increases throughput performance.

Figure-8 shows the total remaining energy of a WSN with simulation time increasing for the three approaches. This validates that the proposed BUCAN method actually consumes less energy than the LEACH-C and EAUCF algorithms. LEACH-C algorithm presents the higher energy dissipation per simulation time, which is consequence of the problems already cited about the CH election way of the algorithm and lack of discriminatory treatment of the energetic discrepancies of the network nodes.

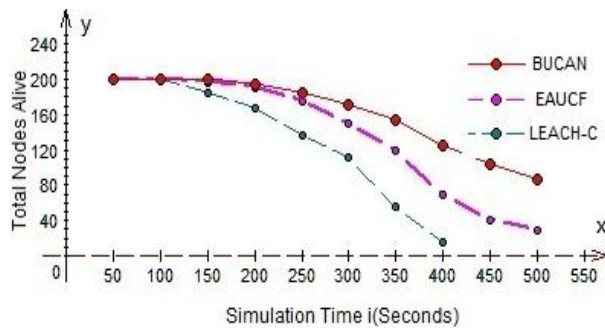


Fig. 7: Sensor network life time.

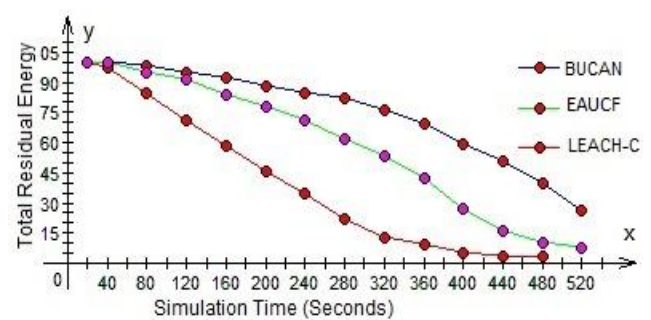


Fig. 8: Average energy consumption in the network.

5. Conclusion

Ultimate objective behind all the clustering protocols is to prolong the network lifetime by saving energy and enhance the network performance. Our simulated results shows that proposed fuzzy Based balanced unequal Clustering algorithm BUCAN offers great benefits and optimize the cluster head selection process, which adjusts the cluster-head radius values considering residual energy and localization and signal strength parameters of the sensor nodes, it take care of noise of surrounding as well transmission loss to calculate the signal strength. Our proposed mechanism also solves this hotspot problem by reducing the size of clusters closer to the base station using binary tree concept to reduce the intra-cluster as well as inter-cluster work load by balancing the energy consumption.

The nodes which are not in the radio coverage of cluster heads then these nodes will be considered as a dead nodes, our results show that alive nodes ratio and average residual energy of the network improved and clusters lying on the hotspot region near to sink incur the balanced energy dissipation. Proposed protocol

efficiently minimizes the selection process time as well as time taken in re-election, which results in reduction of resource utilization by using the fuzzy based decision making technique for cluster head selection to discover the optimal and reliable path. Another great advantage that contributes to the results obtained in this paper to forward the dynamically compressed data to BS through optimized route which have maximum signal strength.

6. References

- [1] Macedo D.F., Santos A.L., Nogueira J.M., Pujolle G. Fuzzy-based load self-configuration in mobile P2P services. *Computer. Network* 2011; 55:1834–1848. doi: 10.1016/j.comnet.2011.01.017.
- [2] W.R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, Energy-efficient Communication Protocol for Wireless Microsensor Networks., in *IEEE Computer Society Proceedings of the Thirty Third Hawaii International Conference on System Sciences (HICSS '00)*, Washington, DC, USA, Jan. 2000, vol. 8, pp. 8020.
- [3] Jong-Myoung Kim, Seon-Ho Park, Young-Ju Han and Tai-Myoung Chung, “CHEF: Cluster Head Election mechanism using Fuzzy logic in Wireless Sensor Networks”, *Advanced Communication Technology, ICACT'2008. 10th International Conference*, vol.1, 2008.
- [4] Q. Liang and L. Wang, “Event detection in wireless sensor networks using fuzzy logic system”, *Proc. Of IEEE International Conference in Computational Intelligence for Homeland Security and Personal Safety*, 2005; pp. 52-55.
- [5] Heinzelman WB, Chandrakasan AP, Balakrishnan H., “An Application-Specific Protocol Architecture for Wireless Microsensor Networks,” *IEEE Transactions on Wireless Communications*, 2002; 1(4): 660—670.
- [6] Indranil Gupta, Denis Riordan and Srinivas Sampalli, “Cluster-head Election using Fuzzy Logic for Wireless Sensor Networks”, *The 3rd Annual Communication Networks and Services Research Conference (CNSR'05)*, 2005.
- [7] T. Sharma, B. Kumar, “F-MCHEL: Fuzzy Based Master Cluster Head Election Leach Protocol in Wireless Sensor Network,” *International Journal of Computer Science and Telecommunications*, Oct-2012, Vol. 3, Issue 10, pp. 08-13.
- [8] Hakan Bagci, “Fuzzy Unequal Clustering in wireless sensor networks,” *Master’s thesis, Middle East Technical University*, January 2010.
- [9] Park J.K., Hong S.J., Kim K.H., Kang T.H., Lee W.Y. “A lifetime-guaranteed routing scheme in wireless sensor networks,” *World Academic Science Eng. Technology*, 2010; 65:946–994.
- [10] Boyinbode O., Le H., Takizawa M., “A survey on clustering algorithms for wireless sensor networks,” *International Space-Based Situated Computer*, 2011; 1:130–136.
- [11] R. Gupta, A. Gupta, S. Agarwal, “A Novel Approach of Data Compression for Dynamic Data”, *3rd IEEE Intl Conf. on System of Systems Engg; California; USA*, June 2-4, 2008.
- [12] R. Singh I. Gupta, A. K. Daniel, “Position Based Energy-Efficient Clustering Protocol under Noisy Environment for sensor Networks using Fuzzy Logic Technique” *Science and Information Conference 2014 (IEEE Int’t Conf.)*, London (U.K.), during 27-29 August 2014.