

# A New Graph Theory based Routing Protocol for Wireless Sensor Networks

<sup>1</sup>B.Baranidharan, <sup>2</sup>B.Shanthi

SASTRA University, School of Computing, Thanjavur, India

<sup>1</sup>baranidharan@it.sastra.edu, <sup>2</sup>shanthi@cse.sastra.edu

## ***Abstract.***

Energy efficiency in wireless sensor network [WSN] is the highly sorted area for the researchers. Number of protocols has been suggested for energy efficient information gathering for sensor networks. These protocols come under two broad categories called tree based approach and clustering techniques. In these techniques clustering is more suitable for real time applications and has much more scalability factor when compared with its previous counterpart. This paper presents the importance and factors affecting the clustering. Also this paper surveyed the different clustering algorithms with its extensions till date and proposed the clustering technique using Minimum Spanning Tree [MST] and shortest path concept with its strength and limitations.

**Keywords:** Clustering, Energy efficiency, MST, shortest path.

## **1 Introduction**

The advancement in wireless technologies and miniaturized hardware has led to the development of the new area called pervasive computing. The base concept behind this pervasive computing is 'anywhere' and 'anytime' computing. Wireless sensor network is one of the pervasive networks which sense our environment through various parameters like heat, temperature, pressure, etc... For battle field surveillance in military applications, habitat monitoring, industrial applications wireless sensor network is an essential one nowadays. A wireless sensor network is built of thousands of sensor nodes. A sensor node has embedded low power processor, limited memory and battery. In WSN, devices are battery operated and un rechargeable, to meet out this challenges, an energy efficient operation of the WSN is the need of the hour for some critical applications like military surveillance, remote patient monitoring. For energy efficient data gathering in the sensor networks we have been following two approaches,

- Tree based approach
- Clustering based approach

When compared with tree based approaches clustering of sensor nodes have more advantages like scalability, avoiding redundant data, latency.

On studying the clustering algorithms proposed for ad hoc networks their main goal was node reach ability and stability. These protocols are not suitable for the WSN since coverage area is an

important factor in WSN. The section 2 in this paper gives a brief idea about importance and objectives of clustering suitable for WSN. The section 3 analyzes the existing clustering algorithms in WSN. A new energy efficient clustering scheme has been proposed in section 4 and in section 5 concludes with the future direction.

## 2 Importance and objectives of clustering

Clustering is the process of dividing the sensor nodes into groups based on some attributes. Generally based upon geographical location and remaining residual energy value clusters are formed. A Cluster Head would be selected for each cluster which is having more responsibilities than cluster members. Clustering sensor network and electing the cluster heads can be in distributed or in centralized way. In distributed mechanism each sensor nodes will broadcast its location and energy level to its one hop neighbors and the node which is having higher energy level and connectivity will be elected as cluster head. In centralized mechanism all the nodes have to transmit their location and residual energy to the base station and base station will form the new clusters with the cluster head and broadcast it to the nodes.

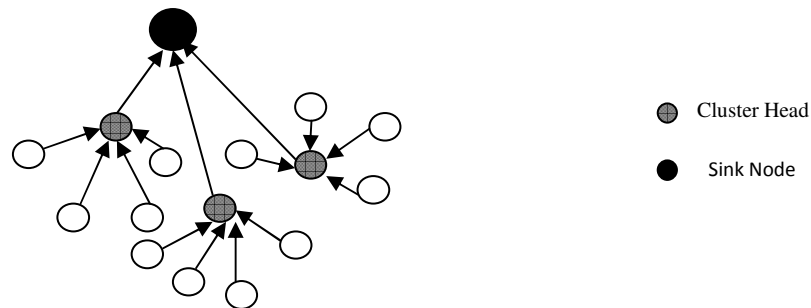


Fig 1 – Clustering Architectural Diagram

In case of homogenous sensor network the cluster heads are selected from the available nodes. The cluster head nodes are relieved from other sensor activities like monitoring the environment to conserve its energy for data aggregation and communication with base station. In case of heterogeneous sensor network, the node which is having more energy than other nodes is selected as cluster head or co-coordinators.

For ad hoc networks clustering is used for improving the reach ability of the nodes. But for wireless sensor networks clustering is an energy efficient scheme which concentrates on network longevity and better coverage. Also scalability is a major challenge in clustering for WSN. For different applications, objectives of the clustering also varies like load balancing, fault tolerance, improved coverage. Based upon the objective of the applications appropriate clustering algorithms are used.

## 3 Existing Clustering Algorithms

Since wireless communication always consumes higher percentage of energy than wired the number of message transmitted should be reduced. But when we reduce the number of message transmission in the network there may be the chances of reduced performance of the network. To make a tradeoff between performance and energy efficiency we are going for the clustering algorithms. Based upon its cluster formation we are having two categories of clusters,

- Geographical area based clusters.
- Clustering based on residual energy level.

### **3.1 Based on residual energy level:**

#### **3.1.1 LEACH:**

Low energy adaptive clustering hierarchy [13] is a distributed clustering protocol to distribute the energy consumption all over its network. Here, based on data collection, network is divided into Clusters and Cluster heads are elected randomly. The cluster head collects the information from the nodes which are coming under its cluster. The phases involved in each round in the LEACH protocol as follows,

*Advertisement phase:* This is the starting phase in LEACH protocol. The eligible cluster head nodes will be sending a request to its nearby nodes to join in its cluster. The non-CH node will be joining with the cluster head which offers higher Received Signal Strength (RSS).

*Cluster set-up phase:* In this step the nodes with its new cluster head form a new cluster.

*Schedule creation:* After cluster set-up phase, the cluster head have to generate a TDMA scheme and pass it to its cluster members to intimate them when they have to send their data to it.

*Data transmission:* The data sensed by the individual sensors will be forwarded to its cluster head during its TDMA time interval.

Here in the LEACH protocol multi cluster interference problem was solved by using unique CDMA codes for each cluster.

It helps to prevent energy drain for the same sensor nodes which has been elected as the cluster leader, using randomization for each time cluster head would be changed. The cluster head nodes collect data from its cluster members and aggregate it. Finally each cluster head will be forwarding the aggregated data to the base station. When compared with LEACH, it have shown a better improved lifetime, in terms of number of data gathering rounds.

LEACH-Centralized [13] works in the same way as LEACH. It follows the centralized mechanism. All the nodes have to transmit their current location and residual energy to the base station. Then the base station forms the new cluster with a cluster head for each of it. The newly formed clusters with its cluster head IDs is transmitted to the nodes. If the nodes receive the message with its own ID as cluster ID it assumes the cluster head role. The steady state phase is same to both LEACH and LEACH-C.

LEACH-F [13] is another variant of LEACH protocol. The cluster formed in the setup phase is fixed. The energy wasted due to new cluster formation in each data collection round is reduced by maintain fixed clusters. But the major drawback in this scheme is the newly arriving nodes cannot be included in the fixed clusters.

### 3.1.2 HEED:

Though the LEACH protocol is much more energy efficient when compared with its predecessors like Direct Transmission (DT), the main drawbacks in LEACH is the random selection of cluster head. In the worst case the CH nodes may not be evenly distributed among the nodes and it will have its effect on the data gathering rounds. To avoid the random selection of CHs a new algorithm called HEED [12] was developed which selects the CHs based on both residual energy level and communication cost. The HEED protocol works in three subsequent phases,

*Initialization phase:* During this phase the initial CHs nodes percentage will be given to the nodes. It is represented by the variable  $C_{prob}$ . Each sensor node compute its probability to become CH by the formula,  $CH_{prob} = C_{prob} * E_{residual}/E_{max}$  where  $E_{residual}$  to residual energy level of the concerned node,  $E_{max}$  corresponds to maximum battery energy. Since HEED supports heterogeneous sensor nodes  $E_{max}$  may vary for different nodes according to its functionality and capacity.

*Repetition phase:* Until the CH node was found with the least transmission cost, this phase was iterated. If the node cannot find the appropriate CH, then the concerned node itself was selected as the CH.

*Finalization phase:* The selection of CH is finalized here. The tentative CH now becomes the final CH node.

### 3.1.3 DECA:

DECA is an acronym for **D**istributed **E**fficient **C**lustering **A**pproach [9]. DECA differs from HEED in deciding and arriving at the score computation. The phases involved in DECA operations are:

*Start Clustering:* In the initial phase all the nodes will compute its score with the help of the function  $score = w_1E + w_2C + w_3I$ . E refers to residual energy, C to node connectivity, and I to node identifier and 'w' to weight which is equal to unity. After some delay the score value will be given to the neighboring nodes with the node ID and cluster ID if the computed score is of a higher value.

*Receive Clustering Message:* When the node is receiving the score value more than its own value and if it is not attached to any cluster it accepts the sender node as its CH.

*Actual announcement:* After the previous phase, the new nodes with the already existing nodes from some other previous cluster which are intended to form a new cluster with a new head, the CHs ID, cluster ID and score value would be broadcasted.

*Finalize Clustering:* In this last step the CH nodes with its Cluster Members forms the new clusters.

### 3.1.4 TEEN:

In TEEN [18], Threshold sensitive Energy Efficient sensor Network protocols two values called Hard Threshold and Soft Threshold are broadcasted to its members by the respective cluster heads. Also they are another variable SV, Sensed Value for this scheme. The Hard Threshold value is the sensed attribute value and Soft Threshold value is a minor change in Hard Threshold value. When a cluster member senses a value beyond its Hard Threshold value it sends the data to its cluster head and stores a copy of it in sensed value (SV) variable. Next time the data is transmitted only when it senses a data greater than Hard Threshold and differs by an amount equal to or greater than Soft Threshold value. The last sent data value is stored in Sensed Value variable replacing previous value. The major advantage in this algorithm is it reduces the number of transmissions between cluster members and cluster heads. The limitations in this algorithm were the nodes does not sent its data to its CH when threshold values does not reaches them, not suitable for applications requiring periodical updates from nodes, also there is the possibility of data collision when two nodes sends the data at same time since TDMA is not a suitable one for the time critical applications.

### 3.1.5 APTEEN:

Adaptive Threshold sensitive Energy Efficient Network protocol [15] combines both proactive and reactive strategies followed by LEACH & TEEN respectively. Even though Hard Threshold and Soft Threshold values are used in this algorithm, by using TDMA schedule for each nodes and Count Time ( $T_c$ ) periodical data collections from the nodes is achieved. The Count Time is the maximum time difference between two successive data transmission from the node. It offers the flexibility to the users that they can set both  $T_c$  and Threshold values. But the complexity of the algorithm increases due to the above factors.

### 3.1.6 MOCA:

Multi-hopping Overlapping Clustering Algorithm (MOCA) [10] is used to improve the inter cluster communication. It differs from the previous clustering algorithms in which MOCA clusters were overlapped with each other. The nodes falling under two adjacent clusters acts as the relay node for Cluster Head communication.

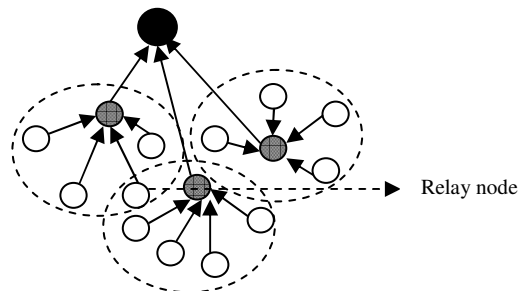


Fig 2 – MOCA – Overlapping Clusters

### **3.1.7 EECPL:**

In order to prevent rapid depletion of energy in cluster head nodes, EECPL [4] algorithm uses one node as cluster head and cluster sender for each cluster. Generally, EECPL algorithm follows ring topology within each cluster and each node will get data from its previous node, fuses it with its own data and transmit it to the next node in the ring. The cluster sender nodes are responsible for transmitting the aggregated data to the base station. As like LEACH data gathering was done in two phases.

*Setup phase:* Based on the remaining energy level of each nodes and their geographical location cluster heads and cluster senders would be selected. Then the cluster head nodes create TDMA schedule for its cluster members and distribute it. The cluster sender nodes take care of sending the aggregated data to the base station.

*Steady state phase:* Initially cluster sender will send its sensed data to the neighboring node and each node is responsible for aggregating the received data with its sensed data and transmits the aggregated data to its neighbor node. When the cluster sender receives aggregated data it transmits it to the base station.

### **3.1.8 ADRP:**

ADRP, Adaptive Decentralized Reclustering Protocol [1] follows the centralized approach for cluster formation by collecting the remaining level and geographical location from the sensor nodes. It reduces the energy wastage due to cluster formation for each round by electing the next eligible cluster head for each cluster. It works in two stages they are,

*Initial Phase:* Initial phase is again divided into three sub stages as follows, In partition stage each sensor node have to send its current location and remaining energy level to the base station. Using this information base station is dividing the network into clusters with appropriate cluster head. In selection stage the next eligible cluster heads would be selected based on the predefined threshold value. At last in the advertisement stage the cluster head ID and next eligible cluster heads are transmitted to each node.

*Cycle phase:* In cycle phase also ADRP works in three stages. In Schedule stage the cluster heads creates the TDMA schedule for each of its cluster members. Next in the transmission stage the data are gathered from the cluster members, aggregated at cluster head and transmitted to the base station from there. Finally in recluster stage, the cluster members switch to their new cluster head in the next cluster head sequence.

Since ADRP follows centralized approach, each time during the new cluster formation the sensor nodes have to send its current location and remaining energy level to the base station. The nodes which are distantly located to the base station would rapidly deplete its energy compared with other nodes.

### **3.1.9 GESC:**

GESC, Geodesic sensor clustering [3] is also one of the distributed clustered algorithms for wireless sensor networks. Here in this scheme the Node Significance NI value is used to select

the cluster head. The cluster heads is called as articulation points, that these nodes are having shortest path to its neighboring nodes. The energy wastage due to inter cluster communication is reduced using GESC.

### **3.2 Geographical area based clustering:**

In Geographical area based clustering approach the nodes coming under particular geographical area are combined to form a cluster. Some of the protocols coming under this category are GAF, SPAN, and PANEL.

#### **3.2.1 GAF:**

The geographical regions in which the sensor nodes are fixed are divided into equal sized grids [16]. The nodes coming under the particular geographical range will be associated with a particular grid. The communication cost of nodes coming under the same grid will be same. During the routing decision any one node from the particular grid will wake up and takes part in routing and all other nodes in the same grid will go to the sleeping state to avoid unnecessary energy depletion.

#### **3.2.2 SPAN:**

SPAN [14] is similar to GAF protocol. In SPAN for every node if its neighbors are not directly or indirectly connected with each other it is elected as the coordinator node. And in any situation if the neighboring nodes can be connected without it assistance it can withdrawn from its coordinator responsibilities. The nodes nearby the coordinator node will be in sleeping state unless when it has to send or receive any data. All the routing activities will be carried over by the coordinator nodes. The major advantage of SPAN is that the nodes can integrates with 802.11 power saving mode easily. The limitation in this method is the coordinator node will drain its energy quickly since it follows geographical based clustering.

#### **3.2.3 PANEL:**

PANEL, Position based aggregator node election [7] is also one of the distributed clustering algorithm. Here, in this protocol the sensor deployed area is divided into fixed equal sized rectangular area. For each data gathering round, the sensor nodes in each clusters computes the reference points for its cluster. The node coming near to the reference point is selected as the aggregator node for that round. The nodes learn the shortest path to the aggregator node at the end of aggregator election procedure which is useful for intra cluster communication. For inter cluster communication any position based routing protocols can be used.

#### **3.2.4 VGA:**

In Virtual Grid Architecture Routing [19] the clusters are fixed and equal sized. Two kinds of aggregations are done here. The aggregations done at Cluster Head level are known as Local Aggregation and further aggregations done at special nodes called as Master Aggregations.

## 4 Proposed Algorithm

To improve the efficiency in clustering protocol, this paper suggests few modifications in the existing clustering techniques. On analyzing the existing clustering algorithms it can be categorized as follows,

- Partitioning
- Hierarchical
- Graph theoretic
- Density based
- Grid based

Every clustering method has its own strengths and limitations. The proper hard and fast rules are not needed in this network topology. Because according to the application and size of data user can define the methodology.

Here this paper tailors the methods to suit the limited energy of the network and also it considers the scalability factor. MST, Minimum Spanning Tree approach is followed in our proposed scheme for Cluster and Super Cluster formation. In the previously mentioned algorithms the energy wastage in data transmission from distance cluster head node to the sink node is not considered. In our proposed scheme the energy wastage in distance CH node transmission to the sink node is reduced by having multi hop communication between cluster head node to the sink node and having Super cluster head nodes, which aggregates the information from different cluster heads and transmits it to the sink node. The Proposed algorithm has three phases:

1. Cluster Formation
2. Cluster head selection.
3. Data transmission using shortest path.

### Cluster Formation:

Unlike previous algorithms, cluster formation precedes before cluster head selection. This is based on Minimum Spanning Tree (MST) concept. A tree is a connected graph without cycles. The spanning tree is 'minimal' when the total length of the edges is the minimum necessary to connect all the vertices in the graph. MST may be constructed using kruskal's (or) Prim's algorithm. But in our proposed algorithm MST is used in the initial cluster formation phase and in super cluster formation phase. Here cluster formation is done in the distributed fashion. Some of the pre assumptions for this new method are,

- All the participating nodes are aware of its position through GPS technique.
- The nodes are assumed to be static during the data gathering phase.
- The wireless sensor network is assumed to be a homogeneous network. All the nodes are having same computational power, storage capacity and communication radius.

The nodes calculate its communication cost with all other nodes by using the expression,



$$E_{TX} = E_{elec} + \epsilon_{amp}kd^{\alpha}$$

Where,  $E_{TX}$  = Transmission energy,  $k$  = number of bits,  $d$  = distance between nodes,  $\alpha$  = path loss component (2 to 5) and  $E_{elec}$  = Electronics energy.

The values for the above parameters are,  $E_{elec} = 50$  nJ/bit and  $\epsilon_{amp} = 100$  pJ/bit/m<sup>2</sup>. The distance between the two nodes is calculated by using Euclidean distance  $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ , where these  $x$  and  $y$  points are known through GPS techniques.

The Base Station will form a Minimum Spanning Tree to connect all the nodes through the minimum path cost. After the formation of the MST, the edges having high costs are cut. If the application needs 'n' clusters for its energy efficient working, then 'n-1' edges should be cut to form clusters.

### **Cluster Head Selection:**

In the newly formed clusters, the node with the highest energy level is selected as the cluster head and the next higher energy level node is selected as the next CH node. To maintain the stability within the clusters, next CH nodes were selected. Once the cluster head are selected, it generates the TDMA schedule for its cluster members and broadcasts to its cluster members.

### **Data Transmission using shortest path:**

In order to reduce further energy wastage due to data transmission between the long distanced Cluster head and sink node, multi-hop data transmission takes place. The data from the nearby cluster heads to the sink node will be directly transmitted to the sink node whereas the data from the distanced cluster head will be transmitted through the shortest multi-hop path. The steps involved in transmission as follows,

- Preparing proximity matrix using distance metric; distance between CH and sink and between CH's.
- Constructing MST to form a super cluster.
- Find shortest path between each CH to sink.
- Find the predominant node [node in maximum number of path]
- Select that node as super cluster head node and aggregation takes place at this node.
- Forward aggregated information to the sink node.

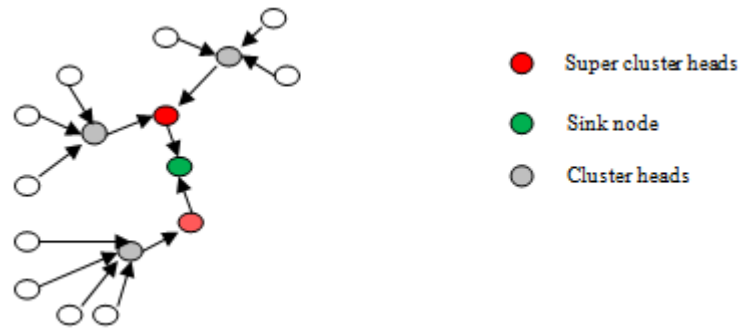


Fig 3 – Architectural diagram of proposed scheme

Strengths of the proposed algorithms are,

- Scalability is better when compared with the previous methods.
- To avoid fault tolerance, node leader is selected.
- All nodes are taken into account by MST technique.
- Aggregation reduces the redundancy.

Limitations:

- Cluster formation is based on certain parameters like path cost and distance; On changing the cluster formation parameters, cluster efficiency may be improved
- Cluster efficiency also changes.
- Mobility is not considered.
- This algorithm based on the assumption that the network is static.

## 5. Conclusions

In this paper, the detailed study about existing clustering algorithms in the wireless sensor networks have been presented and a new algorithm which focuses on energy efficient data transmission between the cluster heads and the sink node is proposed. Using MST mechanism we proposed to find clusters and shortest path for the data transmission. Our future work concentrates on the following,

- Taking Node Importance (NI) which describes the node connectivity as a basic parameter for Cluster Head selection.
- Incorporating Genetic Algorithms (GA) for the energy efficient Cluster Formation phase.
- Modifying the Data Transmission phase to make it suitable for Time critical applications.

## References

- [1] Fuad Bajaber, Irfan Awan.: Adaptive decentralized re-clustering protocol for wireless sensor networks, *Journal of computer and Systems sciences*, doi:10.1016/j.jcss.2010.01.007.
- [2] Yi-hua zhu, Wan-deng wu, Jian pan, Yi-ping tang.: An energy efficient data gathering algorithm to prolong lifetime of wireless sensor networks, *Computer Communications* 33, pp. 639-647 (2010).
- [3] Fuad Bajaber, Irfan Awan.: Energy efficient clustering protocol to enhance lifetime of wireless sensor network, *Journal of Ambient Intelligence and Human Computing* 1, pp. 239-248 (2010).
- [4] N.Dimokas, D.Katsaros, Y.Manolopoulos.: Energy-efficient distributed clustering in wireless sensor networks, *Journal of Parallel and Distributed Computing* 70, pp. 371-383 (2010).
- [5] Jamal N. Al-Karaki, Raza Ul-Mustafa, Ahmed E. Kamal, "Data Aggregation in Wireless Sensor Networks - Exact and Approximate Algorithms", *Proceedings of IEEE Workshop on High Performance Switching and Routing (HPSR)*, USA.
- [6] F.Bajaber, I.lawan.: Centralized dynamic clustering for wireless sensor network. In: *International Conference on Advanced Information Networking and Applications Workshops*, pp. 193-198 (2009).
- [7] L. Buttyan, P. Schaffer.: PANEL: Position-based Aggregator Node Election in Wireless Sensor Networks, In: *Proceedings of the IEEE International Conference on Mobile Ad hoc and Sensor Systems*, MASS, pp. 1 – 9 (2007).
- [8] Ameer Ahmed Abbasi, Mohamed Younis.: A Survey on Clustering Algorithms for Wireless Sensor Networks, *Computer Communications* 30, pp. 2826-2841 (2007).
- [9] Miao Yu, Jason H.Li and Renato Levy.: Mobility Resistant Clustering in Multi-Hop Wireless Networks, *Journal of Networks*, Vol.1, No.1, pp. 12-19(2006).
- [10] Youssef, M. Younis, M. Youssef, A. Agrawala.: Distributed formation of overlapping multi-hop clusters in wireless sensor networks, In: *Proceedings of the 49th Annual IEEE Global Communication Conference (Globecom'06)*, San Francisco, CA, pp. 1-6 (2006).
- [11] Kemal Akkaya, Mohamed Younis.: A Survey on routing protocols for Wireless Sensor Networks", *Ad Hoc Networks* 3, pp. 325-349 (2005).
- [12] O.Younis, S. Fahmy.: HEED: A Hybrid, Energy-Efficient, Distributed clustering approach for Ad Hoc sensor networks, In: *IEEE Transactions on Mobile Computing* 3 (4), pp. 366–379 (2004).
- [13] W.B. Heinzelman, A.P. Chandrakasan, H.Balakrishnan.: Application specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communication*, Vol 1, Issue 4, pp. 660-670 (2002).
- [14] B. Chen, K. Jamieson, H. Balakrishnan, R. Morris.: SPAN: An energy efficient coordination algorithm for topology maintenance in ad hoc networks, *ACM/Kluwer Wireless Networks* 8 (5), pp. 481\_494 (2002).

- [15] Manjeshwar A, Agarwal D (2002).: APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks, In: Proceedings of the 16th international parallel and distributed processing symposium, pp. 195-202 (2002).
- [16] Y. Xu, J. Heidemann, D. Estrin.: Geography-informed Energy Conservation for Ad-hoc Routing. In: Proceedings of the Seventh Annual ACM/IEEE International Conference on Mobile Computing and Networking, pp. 70-84 (2001).
- [7] Y. Yu, D. Estrin, and R. Govindan.: Geographical and Energy-Aware Routing: A Recursive Data Dissemination Protocol for Wireless Sensor Networks, UCLA Computer Science Department Technical Report, UCLA-CSD TR-01-0023 (2001).
- [18] Manjeshwar, D.P. Agarwal.: TEEN: a routing protocol for enhanced efficiency in wireless sensor networks, In: 15th International Parallel and Distributed Processing Symposium (IPDPS'01), vol. 3, pp. 2009-2015 2001.