International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) e-ISSN: 2395-5325
Volume 2, Issue 1 (January '2016)

A STUDY ON CLUSTERING BASED ROUTING PROTOCOLS IN WIRELSS SENSOR NETWORKS

K.Ranjitha,
M.Phil Research Scholar,
Department of Computer Science and Applications,
Vivekanandha College of Arts and Science College for Women (Autonomous),
Elayampalayam,Namakkal, Tamilnadu.

M.Jothilakshmi,
Assistant Professor,
Department of Computer Science and Applications,
Vivekanandha College of Arts and Science College for Women (Autonomous),
Elayampalayam,Namakkal, Tamilnadu.

Abstract: The wireless sensor networks (WSN) consists of several autonomous sensor nodes with sensing, processing and wireless communication capabilities. Clustering the sensor nodes is an effective technique to power management and data dissemination routing protocols have been specifically designed for WSNs where energy consumption is an essential design issue. In this paper we discussed about clustered based routing protocols because clustering increases scalability of the network, balances energy consumption among the nodes in the network, and reduces the amount of data that is actually transmitted to the base station due to the aggregation process.

Keywords: Wireless Sensor Networks; Routing Protocols, Clustering; Energy Consumption.

I. INTRODUCTION

Wireless Sensor networks (WSN) is a large network which is consist of huge number of sensor nodes and these nodes are directly interacting with their environment by sensing the physical parameters such as temperature, humidity, etc[1]. All the sensor nodes send or receive data to/from a fixed wired station called base station (BS). The base station usually serves as a gateway to some other network. WSNs have a comprehensive range of applications in this field including [6, 9, 10]; environmental applications, military applications, home security, etc. The unique characteristics of WSN in terms of limited energy and computation constraints, limited transmission range, no public identification, and the denser level of node deployment [2] make it distinguishable from traditional wireless networks and useful for variety of applications. Since it could be deployed in harsh environment, WSN are used in environmental monitoring such as hazard or disaster monitoring, military application such as battlefield monitoring or object protection, health care applications, home automation, and moreover in human everyday-life. Due to the complex nature of WSN and in order to prolong the lifetime of the sensors, efficient routing algorithms must be employed for data gathering [3]. There are two approaches for data gathering: Flat and Hierarchical approaches [4]. In a flat network, the base station flood a query to all nodes in the network and only the nodes that matches the query, will reply through multi hop path via its first hop neighbors.

II.CLUSTER BASED MODEL:

Clustering is an efficient approach that has been implemented in many communication protocols in wireless environment. In Cluster Based Model [5] network is grouped into different clusters .Each cluster is composed of one cluster head (CH) and cluster member nodes. "The respective CH gets the sensed data from cluster member nodes, aggregates the sensed information and then sends it to the Base Station.(BS).[5]." Cluster members do not communicate directly with the sink node. Thus minimizes the energy consumption and number of messages communicated to base station. "Ultimate result of clustering the sensor nodes is prolonged network lifetime.[38]". Figure 1 presents the cluster based model.

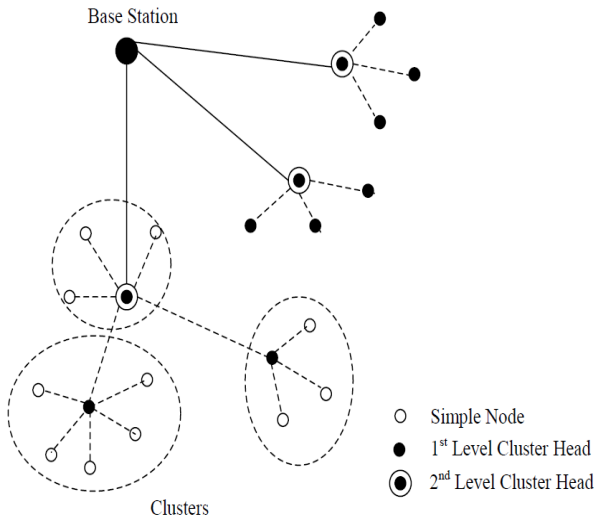

Figure 1: Cluster based model

Limitation: In cluster based model, formation of clusters, selection of CH of each node and formation of schedule is an overhead. Besides this the CHs receive the data, aggregate the

data and transmit the data, so required energy is much more of each CH.

Sensor Node: It is the core component of wireless sensor network. It has the capability of sensing, processing, routing, etc.

Base Station: Base station is considered as a main data collection node for the entire sensor network. It is the bridge (via communication link) between the sensor network and the end user.

Cluster Head: The Cluster head (CH) is considered as a leader for that specific cluster. And it is responsible for different activities carried out in the cluster, such as data aggregation, data transmission to base station, scheduling in the cluster, etc.

Cluster: It is the organizational unit of the network, created to simplify the communication in the sensor network.

Advantages of Clustering:

- Transmit aggregated data to the data sink
- Useful Energy consumption
- Scalability for large number of nodes
- Reduces communication overhead
- Efficient use of resources in WSNs

Disadvantages of clustering:

Dynamic nature (how different parameters of cluster formation is to be determined) of cluster is a difficult task in cluster based touting protocols in WSN.

III. CLUSTER BASED ROUTING PROTOCOLS

Based on the Clustering approach various routing protocols have been implemented in wireless networks. The cluster based protocols are LEACH, PEGASIS, HEED, TEEN and APTEEN

a) LEACH

LEACH [6] (Low-Energy Adaptive Clustering Hierarchy) is a self-organizing adaptive clustering-based protocol that utilizes randomized rotation of local cluster base stations (cluster-heads) to evenly distribute the energy load among the sensors in the network. Data aggregation reduces amount of information to be sent to base station; large reduction in energy dissipation as computation is much cheaper than communication. It can achieve as much as a factor of 8 in reduction in energy dissipation compared with conventional routing protocols. The conventional protocols likes' direct transmission, minimum-transmission energy MTE, multi-hop routing and static clustering may not be optimal for static sensor networks.

Problem Definition: The reason we need network protocol such as LEACH is due to the fact that a node in the network is no longer useful when its battery dies. This protocol allows us to space out the lifespan of the nodes, allowing it to do only the minimum work it needs to transmit data.

Protocol Description: In LEACH [7] the nodes organize themselves into local clusters, with one node acting as the local base station or *cluster-head*. If the cluster-heads were chosen a priori and fixed throughout the system lifetime, as in conventional clustering algorithms, it is easy to see that the unlucky sensors chosen to be cluster-heads would die quickly, ending the useful lifetime of all nodes belonging to those clusters. Thus LEACH includes randomized rotation of the high-energy cluster-head position such that it rotates among the various sensors in order to not drain the battery of a single sensor. In addition, LEACH performs local data fusion to "compress" the
amount of data being sent from the clusters to the base station, further reducing energy dissipation and enhancing system lifetime.

Key Features:

- Self-organizing adaptive Clustering Protocol.
- Randomized rotation of the high energy cluster head.
- Sensors elect themselves to choose a cluster head.
- Each node selects it's own cluster.
- Local data fusion at local BS- compresses data to send, reduce the transmit power.
- Within each cluster nodes access channel using the TDMA(MAC) scheme.
- Each local BS selects CDMA code randomly to communicate within the cluster.

Operation:

The operation of LEACH is broken up into *rounds*, where each round begins with a set-up phase, when the clusters are organized, followed by a steady-state phase, when data transfers to the base station occur. In order to minimize overhead, the steady-state phase is long compared to the set-up phase.

Set-up Phase: The set-up phase consists of the following phases:

A. Advertisement Phase
B. Cluster Set-up Phase

Steady Phase:

A. Schedule Creation Phase
B. Data Transmission Phase

Advertisement Phase:

Node n chooses a number between 0 and 1, if the number is less than the threshold, then n becomes a cluster head.

$$T(n)= \begin{cases} P / 1- P*(r \bmod 1/P) & \text{if } n \varepsilon G \\ 0 & \text{otherwise} \end{cases}$$

Notation:

n is the threshold number.
P is the desired percentage of the Cluster Heads
r is the current round
G is the set of nodes that have not Cluster Head in the last 1/P rounds.

- **E**lected node broadcasts the advertisement using CSMA MAC protocol.
- All elected nodes transmits with same energy.

- Based on the received advertisement signal strength the non-cluster –head node chooses its cluster.
- The node chooses the cluster-head whose advertisement has the highest received power .Since the amount of transmit energy needed to communicate to this cluster head is minimal.

Cluster Set-up Phase:

A. Each node informs the cluster-head node about its choice.

B. Each node uses the CSMA MAC protocol to transmit this information.

C. Cluster heads must keep receivers on.

Data Transmission:

A. Nodes send with minimal required energy to the local cluster head, and turn off their radio when not sending.

B. The local cluster heads receive data, compress it into a single signal.

C. At the end of the round, start from the beginning.

D. To avoid interference between clusters, CDMA codes are picked at random by local cluster heads broadcasted to their cluster nodes.

Advantages:

• Distributed, no global knowledge required
• Energy saving due to aggregation by CHs

Disadvantages:

• LEACH assumes all nodes can transmit with enough power to reach BS if necessary (e.g., elected as CHs)
• Each node should support both TDMA & CDMA
• High level negotiation.
• Only data providing new info is transmitted to BS

Algorithm :

Operation of LEACH protocol is broken into rounds. Each round begins with set up phase where cluster formation and cluster heads selection is performed which is followed by steady state phase where data transfer is done. Steady state phase duration is long enough than set up phase to overcome the overhead of cluster formation and other corresponding tasks.[1].

B) PEGASIS:

The PEGASIS [8] (Power –Efficient Gathering in Sensor Information System**)** a near optimal chain –based protocol aims to provide improvements to the LEACH protocol. PEGASIS aims to address the overhead caused by the cluster formation in LEACH by constructing chains of nodes instead of clusters. The chain construction is performed according to a greedy algorithm, where nodes select their closest neighbors as next hops in the chain. It is assumed that the nodes have a global knowledge of the network and the chain construction starts from the nodes that are farthest from the sink. As a result of chain operation, instead of maintaining cluster formation and membership, each node only keeps track of its previous and next neighbor in the chain.

The basic idea of the protocol is that in order to extend the network lifetime, nodes only need to communicate with their neighbors and they take turns into communicating with the base-station. When the round of all nodes communicating with the base-station ends, a new round will start and so on. This reduces the power required to transmit the data per round as the power draining is spread uniformly over all nodes.

Objectives:

- Increase the lifetime of each node by using the collaborative techniques and as a result the network lifetime will be increased.
- Allow only local coordination between nodes that are close together so that the bandwidth consumed in communication is reduced.

Strengths to measure the distance to all neighboring nodes and then adjust the signal strength so that only one node can be heard. .The in PEGASIS will consists of those nodes that are closest to each other and form a path to the base station. The aggregated form of the data will be sent to the base-station by any node in the chain and the nodes in the chain will take turns in sending to the base station. Communication in the chain is performed sequentially such that each node within a chain aggregates data from its neighbor until all the data are aggregated at one of the nodes, i.e. chain leader. The chain leader controls the communication order by passing a token among the nodes. .The chain construction is performed in a greedy fashion. The simulation result showed that PEGASIS is able to increase the lifetime of the network twice as much the lifetime of the network under the LEACH protocol. Such performance gain is achieved through the elimination of the overhead caused by the dynamic cluster formation in LEACH and through decreasing the number of transmission and reception by using the data aggregation.

Problems:

Although the clustering overhead is avoided, PEGASIS still requires dynamic topology adjustment since a sensor node needs to know about energy status about its neighbors in-order to know where to route its data. Such topology adjustment can introduced significant overhead especially for highly utilized networks. Moreover, PEGASIS assumes that each sensor node can be able to communicate with the base station directly. In practical cases, nodes use multihop communication to reach the base station. Also, PEGASIS assumes that all nodes maintain a complete database about the location of all other nodes in the network. In addition, PEGASIS assumes that all sensor nodes have the same level of energy and they are likely to die at the same time. PEGASIS also introduces excessive delay for distant node on the chain. Finally single leader may be a bottleneck. The following figure describes the operation of the protocol.

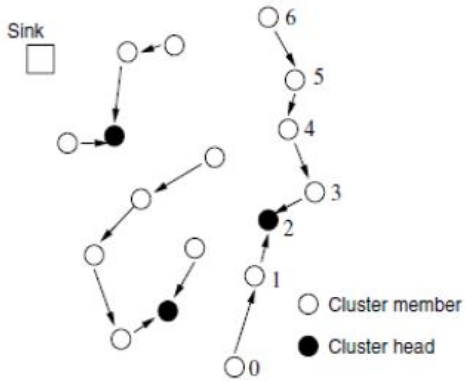

Figure 2 : Chain Structure of PEGASIS

Node 0 transmits its data to node 1, which aggregates these data with its own to create a packet of the same length. This packet is transmitted to node 2. Once node 2 receives the packet from node 1, it passes the token to the other end of the chain, i.e. mode 6. Information from nodes 6,5,4,and 3 also aggregated and sent to node 2 in the same fashion. Upon receiving the aggregated information in the chain , node 2 uses a single hop communication to transmit the data to the sink.

Advantages of PEGASIS

- Compare to LEACH transmitting distance for most of the node reduces in PEGASIS.
- Messages received by each head node are at most 2 in PEGASIS is less compared to LEACH.
- Since each node gets selected once, energy dissipation is balanced among sensor nodes

Drawbacks of the PEGASIS protocol

- When a head node is selected, there is no consideration how far the BS is located from the head node.
- When a head node is selected its energy level is not considered.
- Since there is only one node head, it may be the bottle neck of the network causing delay.
- Redundant transmission of data as only one head node is selected.

C) TEEN

TEEN [9] (Threshold Sensitive Energy Efficient Sensor Network**)** is a cluster based hierarchical routing protocol in WSN. This protocol is used for time-critical applications.

Assumptions:

1. The BS and the sensor nodes have same initial energy.
2. The BS can transmit data to all the nodes in the network directly.

In this protocol, nodes sense the medium continuously, but the data transmission is done frequently. The network consists of simple nodes, first level cluster heads and second level cluster heads. TEEN uses LEACH's strategy to form cluster. First level cluster heads are formed away from the BS and second level cluster heads are formed near the BS.

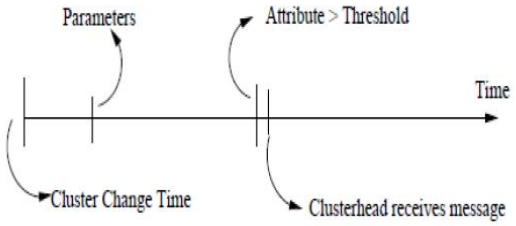

Figure 3: Operation of TEEN

A CH sends two types of data to its neighbors- one is the hard threshold(HT) and other is soft threshold(ST). In the hard threshold, the nodes transmit data if the sensed attribute is in the range of interest and thus it reduces the number of transmissions. Figure 3 shows the operation of TEEN protocol where if the attribute value(sensed data) > threshold value then only CH received the data otherwise not.

On the other hand, in the soft threshold mode, any small change in the value of the sensed attribute is transmitted. The nodes sense their environment continuously and store the sensed value for transmission. Thereafter the node transmits the sensed value if one of the following conditions satisfied.

a. Sensed Value > hard threshold(HT)

b. Sensed Value ~ hard threshold >= soft threshold (ST)

Advantages

(1) Based on the two thresholds, data transmission can be controlled commendably, *i.e.*, only the sensitive data we demand can be transmitted, so that it reduces the energy transmission consumption and improves the effectiveness and usefulness of the receiving data;

(2) TEEN is complement for reacting to large changes in the sensed attributes, which is suitable for reactive scenes and time critical applications

Drawbacks:

(1) It is not suitable for periodic reports applications since the user may not get any data at all if the values of the attributes may not reach the threshold .

(2) There exist wasted time-slots and a possibility that the BS may not be able to distinguish dead nodes from alive ones, because only when the data arrive at the hard threshold and has a variant higher than the soft threshold did the sensors report the data to the BS;

(3) If CHs are not in the communication range of each other the data may be lost, because information propagation is accomplished only by CHs. To overcome the drawbacks of the TEEN, a new protocol is proposed named APTEEN.

D) APTEEN:

The Adaptive Threshold sensitive Energy Efficient sensor Network [12](APTEEN) protocol is an extension to TEEN

and aims at both transmitting periodic data and reacting to time critical events. APTEEN, on the other hand, is a hybrid protocol that changes the periodicity or threshold values used in TEEN according to the requirement of users and the type of the application.

APTEEN is based on a query system which allows three types of queries: historical, on-time, and persistent which can be used in a hybrid network. Moreover, QoS requirements are introduced for the on-time queries and minimum delay is achieved by a TDMA schedule with a special time slot assignment manner.

In APTEEN, CHs broadcast the following four parameters:

(1) Attributes (A)—this is a set of physical parameters which the user is interested in obtaining data about.

(2) Thresholds—this parameter consists of the hard threshold (HT) and soft threshold (ST). HT is a particular value of an attribute beyond which a node can be triggered to transmit data. ST is a small change in the value of an attribute which can trigger a node to transmit data again;

(3) Schedule—this is a TDMA schedule, assigning a slot to each node;

(4) Count time (CT)—it is the maximum time period between two successive reports sent by a node. It can be a multiple of the TDMA schedule length and it accounts for the proactive component.

The distinctive feature of APTEEN is to switch between proactive and reactive modes to transmit data. All nodes sense the environment continuously, but only those nodes which sense a data value at or beyond the hard threshold permit transmitting. Once a node senses a value at or beyond the hard threshold, it transmits data. If a node does not send data for a time period equal to the count time, it must sense and transmit the data again. In APTEEN, each CH aggregates the data from the MNs within its cluster and transmits the aggregated data to the BS. During the process of data aggregation, it is assumed that the data received from the corresponding MNs are sufficiently correlated, thus it reduces a large amount of redundancy of the data to be transmitted to the BS. Moreover, a modified TDMA schedule is used to implement the hybrid network by assigning each node in the cluster a transmission slot. Additionally, APTEEN offers a lot of flexibility by allowing the user to set the CT interval and the threshold values for energy consumption can be controlled by changing the CT as well as the threshold values [11].

APTEEN supports quires like:

a. Historical analysis of past data values

b. A snapshot of the current network view.

c. Persistent monitoring of an event for a period of time.

E) HEED

HEED [14] (Hybrid Energy-Efficient Distributed clustering**)** was designed to select different cluster heads in a field according to the amount of energy that is distributed in relation to a neighboring node.

Primary Goals:

a. Prolonging network lifetime by distributing energy consumption

b. Terminating the clustering process within a constant number of iterations or steps.

c. Minimizing the control overhead

d. Producing well distributed cluster heads and compact clusters.

Hybrid Energy-Efficient Distributed clustering (HEED) , introduced by Younis and Fahmy, is a multi hop WSN clustering algorithm which brings an energy-efficient clustering routing with explicit consideration of energy. Different from LEACH in the manner of CH election, HEED does not select nodes as CHs randomly. The manner of cluster construction is performed based on the hybrid combination of two parameters. One parameter depends on the node's residual energy, and the other parameter is the intra-cluster communication cost. In HEED, elected CHs have relatively high average residual energy compared to MNs. Additionally, one of the main goals of HEED is to get an even-distributed CHs throughout the networks. Moreover, despite the phenomena that two nodes, within each other's communication range, become CHs together, but the probability of this phenomena is very small in HEED. In HEED, CHs are periodically elected based on two important parameters: residual energy and intra-cluster communication cost of the candidate nodes. Initially, in HEED, a percentage of CHs among all nodes, *C*prob, is set to assume that an optimal percentage cannot be computed a priori. The probability that a node becomes a CH is:

CHprob=Cprob *(Eresidual / Emax)

where *E*residual is the estimated current energy of the node, and *E*max is a reference maximum energy, which is typically identical for all nodes in the network. The value of *CH*prob, however, is not allowed to fall below a certain threshold that is selected to be inversely proportional to *E*max. Afterwards, each node goes through several iterations until it finds the CH. If it hears from no CH, the node elects itself to be a CH and sends an announcement message to its neighbors. Each node doubles its *CH*prob value

and goes to the next iteration until its *CH*prob reaches 1. Therefore, there are two types of status that a sensor node could announce to its neighbors: tentative status and final status. If its *CH*prob is less than 1, the node becomes a tentative CH and can change its status to a regular node at a later iteration if it finds a lower cost CH. If it's *CH*prob has reached 1, the node permanently becomes a CH. In HEED, every node elects the least communication cost CH in order to join it. On the other hand, CHs send the aggregated data to the BS in a multi-hop fashion rather than single-hop fashion of LEACH.

Advantages:

- It is a fully distributed clustering method that benefits from the use of the two important parameters for CH election;

- Low power levels of clusters promote an increase in spatial reuse while high power levels of clusters are required for inter-cluster communication. This provides uniform CH distribution across the network and load balancing;
- Communications in a multi-hop fashion between CHs and the BS promote more energy conservation and scalability in contrast with the single-hop fashion, *i.e.*, long-range communications directly from CHs to the sink, in the LEACH protocol .

Disadvantages:

- The use of tentative CHs that do not become final CHs leave some uncovered
- nodes. As per HEED implementation, these nodes are forced to become a CH and these forced CHs may be in range of other CHs or may not have any member associated with them. As a result, more CHs are generated than the expected number and this also accounts for unbalanced energy consumption in the network .
- Similar to LEACH, the performing of clustering in each round imposes significant overhead in the network. This overhead causes noticeable energy dissipation which results in decreasing the network lifetime;
- HEED suffers from a consequent overhead since it needs several iterations to for clusters. At each iteration, a lot of packets are broadcast.

IV.CONCLUSION

Routing is one of the main problems in WSNs and many solutions have been developed to address this problem. Ensuring efficient routing faces many challenges due to both wireless communication effects and the peculiarities of sensor networks. In this paper we have discussed concepts, advantage and disadvantage of cluster based routing protocols LEACH, PEGASIS, TEEN, APTEEN and HEED.

V.REFERENCES

[1] Yick, J.; Mukherjee, B.; Ghosal D.: Wireless sensor network survey. Computer Networks. Vol. 52, Issue 12, pp. 2292-2330 (2008).

[2] Chong, C.; Kumar, S.: Sensor networks: Evolution, opportunities, and challenges. Proceedings of the IEEE, vol. 91, no. 8, pp. 1247–1256 (2003).

[3] Yu, L.; Wei, L; Zhenhua, K.: Study on Energy Efficient Hierarchical Routing Protocols of Wireless Sensor Network. WASE International Conference on Information Engineering, ICIE'09. Vol. 1, 10-11 July 2009, pp. 325–328 (2009).

[4] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," Proceedings of the 33rd Hawaii International Conference on System Sciences (HICSS '00), January 2000.

[5] L.Cui, F. wang and H. Luo. "Network and Parallel Computing," Springer Berlin / HeidelbergLtd.14 Oct 2004.[38] J. Ibriq and I. Mahgoub. Cluster-based routing in wireless sensor networks: Issues and challenges. Department of computer science and Engineering, Florida Atlantic University.

[6]. W.Heinzelman, A.Chandrakasanand and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, Maui, HI, 2000, pp. 3005- 3014.

[7]. Akyildiz I. F., Su W, Sankarasubramaniam Y, Cayirci E, 2002"Wireless sensor network: a survey" Computer Networks, 393-422.

[8]. Lindsey S. and Raghavendra C., "PEGASIS: Power-Efficient Gathering in Sensor Information Systems, in Proceedings of IEEE Aerospace Conference, vol. 3, pp. 1125-1130, 2002.

[9]. Manjeshwar,A. Agrawal,"TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks,"in Parallel and Distributed Processing Symposium,IEEE Proceedings 15th International,Aug 2008.

[10]. S. Corson and J. Macker, "Routing Protocol Performance Issues and Evaluation Considerations," Naval Research Laboratory, Jan.1999.

[11]. S. Vijayanand, R.M. suresh, "AN OVERLOOK ON ROUTING TECHNIQUES IN WIRELESS SENSOR NETWORKS," IET-UK International Conference on Information and Communication Technology in Electrical Sciences, Dr. M.G.R. University, Chennai, Tamil Nadu, India, 2007, pp.557-998.

[12]. A. Manjeshwar and D. P. Agarwal, "APTEEN: A hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks," Parallel and Distributed Processing Symposium., Proceedings International, IPDPS 2002, pp. 195-202.

[13]. Neha Rathi, Jyoti Saraswat and Partha Pratim Bhattacharya, "A Review On Routing Protocols For Application In Wireless Sensor Networks", International Journal of Distributed and Parallel Systems (IJDPS) Vol.3, No.5, September 2012, pp : 39-58.

[14] Harneet Kour, Ajay K. Sharma "Hybrid Energy Efficient Distributed Protocol for Heterogeneous Wireless Sensor Network", International Journal of Computer Applications ,Vol.4, No.6,July2010, pp. 37- 41.