Cluster Formation and Cluster Head Selection for TEEN Routing Protocol

Prakash Dhatti[#1] and Pradeepkumar M P[*2]

[#] PG Student, Dept. ISE PESIT, Bengaluru, India

[*] PG Student, Dept. ISE PESIT, Bengaluru, India

***Abstract*— The Wireless Sensor Network (WSN) has become an interesting field of research of the 21st century. This has brought about developing low cost, low-power and multi-function sensor nodes. The network life for wireless sensor network plays an important role in survivability. Energy efficiency is one of the critical concerns for wireless sensor networks. Sensor nodes are strictly constrained in terms of storage, board energy and processing capacity. For these reasons, many new protocols have been proposed for the purpose of data routing in sensor networks. These protocols can be classified into three main categories: data-centric, location-based and hierarchical. TEEN is type of hierarchical routing protocol, which is used to operate as environment friendly in sensor network. In this paper we are mentioning an energy efficient cluster formation and cluster head selection algorithms for TEEN routing protocol. All the simulations of the proposed idea will be simulated on Berkeley's NS2 network simulator and the performance of the proposed scheme has been evaluated.**

***Index Terms*— Routing Protocol; Cluster Formation; Cluster Head Selection; Wireless Sensor Network Protocol.**

I. Introduction

Sensor networks present unique opportunities for a broad spectrum of applications such as industrial automation, situation awareness, tactical surveillance for military applications, environmental monitoring, chemical/biological detection etc. Micro sensor systems enable the reliable monitoring and control of a variety of applications. WSN can be viewed as a network consisting of hundreds or thousands of wireless sensor nodes which collect the information from their surrounding environment and send their sensed data to Base Station or sink node. Each sensor node is battery powered and equipped with:

1. Integrated sensors
2. Data processing capabilities
3. Short-range radio communications

The ideal wireless sensor is networked and scalable, fault tolerance, consume very little power, smart and software programmable, efficient, capable of fast data acquisition, reliable and accurate over long term, cost little to purchase and required no real maintenance [1]. An ideal sensor network should have the following additional features:

So, the routing protocol should be fault-tolerant in such a dynamic environment. The traditional routing protocols defined for wireless ad hoc networks [2] [10] are not well suited due to the following reasons:

1. Sensor networks are "data centric" i.e., unlike traditional networks where data is requested from a specific node, data is requested based on certain attributes such as, which area has temperature > 50F?

2. The requirements of the network change with the application and so, it is application-specific [4]. For example, in some applications the sensor nodes are fixed and not mobile, while others need data based only on one attribute (i.e., attribute is fixed in this network).

3. Adjacent nodes may have similar data. So, rather than sending data separately from each node to the requesting node, it is desirable to aggregate similar data and send it.

4. In traditional wired and wireless networks, each node is given a unique id, used for routing. This cannot be effectively used in sensor networks. This is because, these networks being data centric, routing to and from specific nodes is not required. Also, the large number of nodes in the network implies large ids [3], which might be substantially larger than the actual data being transmitted.

Attribute based addressing is typically employed in sensor networks. The attribute based addresses are composed of a series of attribute-value pairs which specify certain physical parameters to be sensed. For example, an attribute address may be (temperature > 60F). So, all nodes which sense a temperature greater than 60F should respond with their location.

Location awareness is another important issue. Since most data collection is based on location, it is desirable that the nodes know their position whenever needed.

II. OVERVIEW

In this section, we provide a brief overview of some related research work and different types of routing.

Intanagonwiwat et. al [8] have introduced a data dissemination paradigm called directed diffusion for sensor networks. It is a data-centric paradigm and its application to query dissemination and processing has been demonstrated in this work.

Estrin et. al [4] discuss a hierarchical clustering method with emphasis on localized behavior and the need for asymmetric communication and energy conservation in sensor networks.

A cluster based routing protocol (CBRP) has been proposed by Jiang et. al in [10] for mobile ad-hoc networks. It divides the network nodes into a number of overlapping or disjoint two-hop-diameter clusters in a distributed manner. However, this protocol is not suitable for energy constrained sensor networks in this form.

Heinzelman et. al [5] introduce a hierarchical clustering algorithm for sensor networks, called LEACH.

A WSN can have network structure based or protocol operation based routing protocol. Routing protocols in WSNs might differ depending on the application (Protocol-Operation-based) and network architecture (Network-Structure-based).Depending on protocol operation WSN can be classified into,

A. *Multipath-based routing*

It uses multiple paths rather than a single path in order to enhance network performance. For instance the fault tolerance can be increased by maintaining multiple paths between the source and destination at the expense of increased energy consumption and traffic generation.

B. *Query-based routing*

The destination nodes propagate a query for data from a node through the network. A node with this data sends the data that matches the query back to the node that initiated it.

C. *Negotiation-based routing*

This negotiation based routing is done in order to eliminate redundant data transmissions. In this communication decisions are also made based on the resources available in the network scenario.

D. *QOS-based routing*

When delivering process of data in ongoing with the help of this routing, balances the network in between energy consumption and data quality through certain QOS metrics such as delay, energy or bandwidth.

E. *Coherent-based routing*

The entity of local data processing on the nodes is being distinguished between the coherent (minimum processing) and the non-coherent (full processing) routing protocols.

F. *Flat-based routing*

In this routing protocol each node plays the same role and sensor nodes collaborate to perform the sensing task.

G. *Hierarchical-based routing*

In this type of routing, the nodes having the higher-energy are used to process and send the information, while the nodes having the low-energy are used to perform the sensing in the proximity of the target. The process of creation of clusters and assigning special tasks to cluster heads can efficiently increase the overall system scalability, lifetime, and energy efficiency. Hierarchical routing is an efficient way to lower the energy consumption within a cluster with the help of performing data aggregation and fusion within the different clusters in order to decrease the number of transmitted messages to the sink node.

H. *Location-based routing*

In this type of protocol sensor nodes are addressed by means of their locations. The distance between neighboring nodes can be estimated on the basis of incoming signal strengths from the source nodes. Relative coordinates of neighboring nodes can be obtained by exchanging such information between neighbors or by communicating with a satellite using GPS. To save energy, some location-based schemes also suggest that nodes should go to sleep if there is no activity to perform in a definite time.

III. CLASSIFICATION OF SENSOR NETWORKS

Here, we present a simple classification of sensor networks on the basis of their mode of functioning and the type of target application.

A. *Proactive Networks*

The nodes in this network periodically switch on their sensors and transmitters, sense the environment and transmit the data of interest. Thus, they provide a snapshot of the relevant parameters at regular intervals. They are well suited for applications requiring periodic data monitoring.

B. *Reactive Networks*

The nodes react immediately to sudden and drastic changes in the value of a sensed attribute. As such, they are well suited for time critical applications.

IV. SENSOR NETWORK MODEL

The sensor network must have network model so consider a model which is well suited for these networks. It consists of a base station (BS), away from the nodes, through which the end user can access data from the sensor network. All the nodes in the network are homogeneous and begin with the same initial energy. The BS however has a constant power supply and so, has no energy constraints. It can transmit with high power to all the nodes. Thus, there is no need for routing from the BS to any specific node. However, the nodes cannot always reply to the BS directly due to their power constraints, resulting in asymmetric communication.

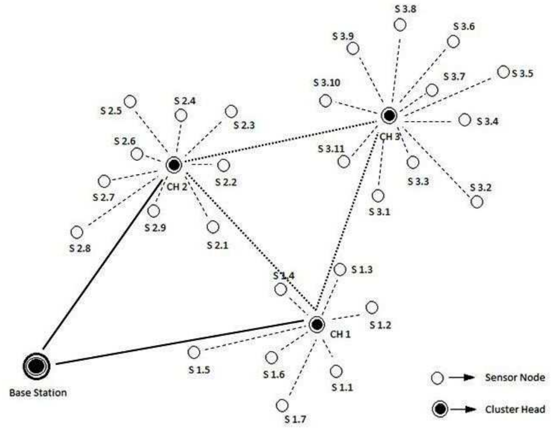

Fig.1 Hierarchical clustering

This model uses a hierarchical clustering scheme. Consider the network structure as shown in *Fig.1*. Each cluster has a cluster head which collects data from its cluster members, aggregates it and sends it to the BS or an upper level cluster head. For example, nodes S 3.1, S 3.2, S 3.3, ………. S 3.11

forms a cluster with node CH 3 as the cluster head. Similarly there exist other cluster heads such as CH 1, CH 2 etc.

These cluster-heads, in turn, form a cluster with their nodes. So, node CH 1 becomes a second level cluster head too. This pattern is repeated to form a hierarchy of clusters with the uppermost level cluster nodes reporting directly to the BS. The BS forms the root of this hierarchy and supervises the entire network. The main features of such architecture are:

All the nodes need to transmit only to their immediate cluster-head, thus saving energy. Only the cluster head needs to perform additional computations on the data. So, energy is again conserved. Cluster-heads at increasing levels in the hierarchy need to transmit data over correspondingly larger distances. Combined with the extra computations they perform, they end up consuming energy faster than the other nodes.

V. TEEN ROUTING PROTOCOL

In Wireless Sensor Networks the consumption of energy is one of the most important issues. TEEN routing protocol is found to be energy efficient than other protocols. By the use of a clustering technique they minimize the consumption of energy greatly in collecting and disseminating data. TEEN routing protocol minimizes energy consumption by dividing nodes into clusters. In each cluster, higher energy nodes can be used to process and send the information while low energy nodes can be used to perform the sensing in the proximity of the target. This means that creation of clusters and assigning special tasks to cluster heads can greatly contribute to overall system scalability, lifetime, and energy efficiency.

The cluster head broadcasts two thresholds namely hard and soft thresholds to the nodes after the clusters are formed. Hard threshold is the minimum threshold used to trigger a sensor node to switch on its transmitter and transmit the cluster head. Thus, the hard threshold will perform transmission only when the sensed attribute is in the required range and reduces the number of transmissions. On the other hand, in 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. The time line for TEEN is as shown in fig. Thereafter the node transmits the sensed value if one of the following conditions satisfied:

- Sensed value > hard threshold (HT).
- Sensed value ~ hard threshold >= soft threshold

The hard and soft threshold values can be adjusted to control the number of packet transmissions.

A. Routing challenges and design issues in WSNs

There are different parameters are there which provides a very challenging criterion in routing for WSN and they are as follows

- Node deployment in the sensor network.
- Energy Consumption in the network should occur without losing of accuracy of the network.
- Scalability of the network.
- Network Dynamics.
- Transmission Media should be fault tolerance.
- Coverage area of different sensor nodes in the network.
- Data Aggregation process within the clusters in the network.

B. Cluster Head Selection Algorithm

Each sensor node n generates a random number such that 0< random < 1 and compares it to a pre-defined threshold T (n). If random < T (n), the sensor node becomes cluster-head in that round, otherwise it is cluster member.

$$T(n) = \frac{p}{1-P\left(r \bmod \frac{1}{p}\right)} \quad n \epsilon G \ldots\ldots\ldots 1$$

In this formula, p is the percentage of cluster heads over all nodes in the network, i.e., the probability that a node is selected as a cluster head; r the number of rounds of selection and G is the set of nodes that are not selected in round 1/p. As we can see here, the selection of cluster heads is totally randomly.

- Protocol doesn’t take residual energy of each node into consideration for the selection of cluster head node as each node has equal probability of becoming cluster head. If low-energy node is being selected as cluster head node, then the network fails soon due to high energy consumption causes adverse to energy balancing among the network. This results data loss and lower in survival time of the network.
- The random selection algorithm of cluster head nodes causes problem of imbalance in energy load. Distance factor is not considered in cluster formation due to which sometimes very big clusters and very class clusters exist at the same time in the network. More the distance between cluster head node and base station, more the energy consumption of that node.
- Cluster head nodes perform data aggregation and send processed data to the base station in single-hop due to which cluster head nodes deplete their energy too fast as compared to normal nodes. Also if a cluster head node fails, the whole nodes linked to it will deplete their energy too.

C. K-means Clustering Algorithm

Clustering is the process of partitioning or grouping a given set of patterns into disjoint clusters. This is done such that patterns in the same cluster are alike and patterns belonging to two different clusters are different. Clustering has been a widely studied problem in a variety of application domains.

The k-means method has been shown to be effective in producing good clustering results for many practical applications as shown in *Fig.2*. However, a direct algorithm of k-means method requires time proportional to the product of number of patterns and number of clusters per iteration. This is computationally very expensive especially for large datasets.

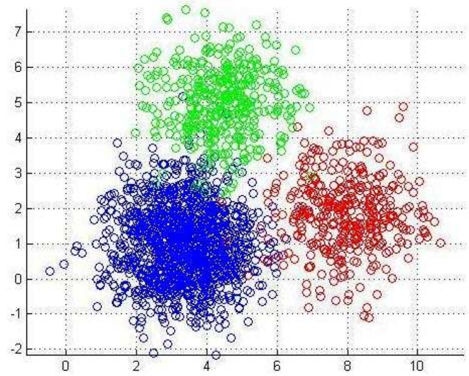

Fig.2 K-means for Clusters

The number of clusters k is assumed to be fixed in k-means clustering. Let the k prototypes be initialized to one of the n input patterns and X = {x1,x2,x3,……..,xn} and Y = {y1,y2,y3…….. yn} be the set of data points and V = {v1,v2,…….,vc} be the set of centers.

1) Select *'c'* cluster centers of cluster heads.

2) Calculate the distance between each data point and cluster centers.

3) Assign the data point to the cluster center whose distance from the cluster center is minimum of all the cluster centers.

4) Recalculate the new cluster center using:

$$Vi = \frac{1}{ci} \sum_{k=1}^{ci} xi\ yi \quad \ldots\ldots\ldots\ldots\ldots\ 2$$

Where, 'c_i' represents the number of data points in i^{th} cluster.

5) Recalculate the distance between each data point and new obtained cluster centers.

The number of iterations required can vary on the number of patterns, number of clusters, and the input data distribution. Thus, a direct implementation of the k-means method can be computationally very intensive. This is especially true for typical data mining applications with large number of pattern vectors.

VI. SIMULATION

In our experiment, we have used NS-2(Network Simulator 2) software tool under Linux environment and simulated in an area of 500x500 units.

TABLE.I TECHNICAL SPECIFICATION OF SIMULATED ENVIRONMENT.

Parameter	Details
Channel/WirelessChannel	channel type
Propagation/TwoRayGround	radio-propagation model
Phy/WirelessPhy	network interface type
Mac/802_11	MAC type
Queue/DropTail/PriQueue	interface queue type
LL	link layer type
Antenna/OmniAntenna	antenna model
50	max packet in ifq
30	number of nodes
1	number of base station
3	number of cluster heads
500	X dimension of topography
500	Y dimension of topography
EnergyModel	Energy set up

The simulation has been performed on a network of 30 nodes and a fixed base station. The nodes are placed randomly in the network. All the nodes start with an initial energy of 0J. Cluster formation is done using k-means algorithm as shown in *Fig.3*.

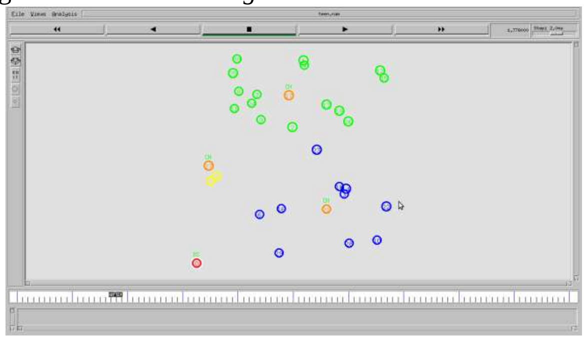

Fig.3 Cluster Formation and Cluster Head Selection

VII. RESULTS

We have executed several trails of the protocol and the readings from these trials were plotted using x-graph of network simulator.

The energy model represents the energy level of nodes in the network. In simulation, the variable energy represents the energy level in a node at any specified time. The value of initial Energy is passed as an input argument. A node loses a particular amount of energy for every packet transmitted and every packet received. As a result, the value of initial Energy in a node gets decreased. The energy consumption level of a node at any time of the simulation can be determined by finding the difference between the current energy value and initial Energy value. Energy consumed is almost linear line as shown in *Fig.4*, as time increases the energy consumed by the nodes also increases, more packet sending to the base station will result in more energy consumption, the graph shows that energy is moderately used which increase the network lifetime and efficiency of the protocol.

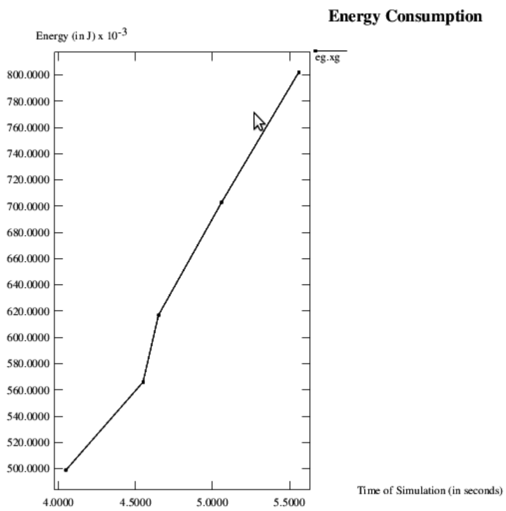

Fig.4 Energy Consumption

Throughput is the rate of production or the rate at which something can be processed. Network throughput is the rate of successful message delivery over a communication channel. The data these messages belong to may be delivered over a physical or logical link or it can pass through a certain network node. From the graph shown in *Fig.5*, throughput is varying at start of simulation and attains to a constant value for throughout the network life time. Every packet sent may not be delivered to the target node but an efficient routing mechanism makes the packet to reach destination and the packets which are lost are retransmitted using the link layer acknowledgement status.

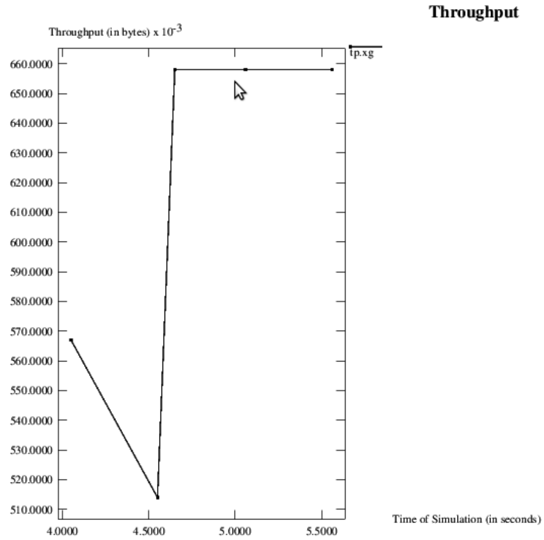

Fig.5 Throughput in the Network

End-to-end Delay is the average time taken by a data packet to arrive in the destination. It also includes the delay caused by route discovery process and the queue in data packet transmission. Only the data packets that successfully delivered to destinations that counted. From *Fig.6* we can see that Delay is high as simulation starts, and goes on decreases as time of simulation increases. The lower value of end to end delay means the better performance of the protocol and also good network performance. Delay may be caused by number of reasons in sensor network, but it can act as differentiating parameter for the real time applications. The network which has more delay in the packet transformation may fails to update the real time, which will results in miss-matching of the situations in time critical enents.

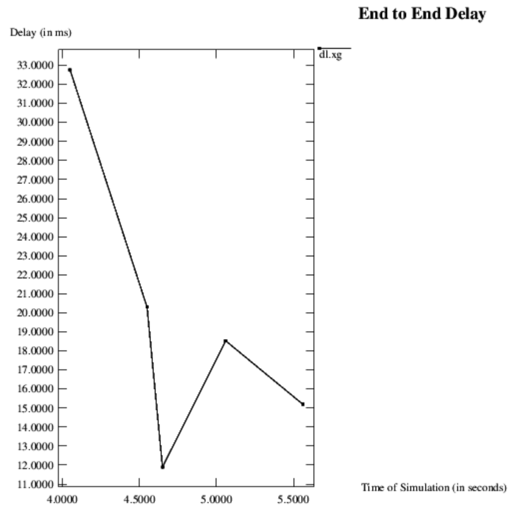

Fig.6 End to End Delay

Packet delivery ratio is the ratio of the number of delivered data packet to the destination. This illustrates the level of delivered data to the destination. Packet delivery ratio is less initially, as the time of simulation increases the PDR also increase, the greater value of packet delivery ratio the better performance of the protocol. This achieved only when one must have concentrated on the sent packets, received packets and the lost packets. If the number of lost packets required to be reduced then the constant bit rate must be maintained between the nodes, so that which will gives good PDR in the network.

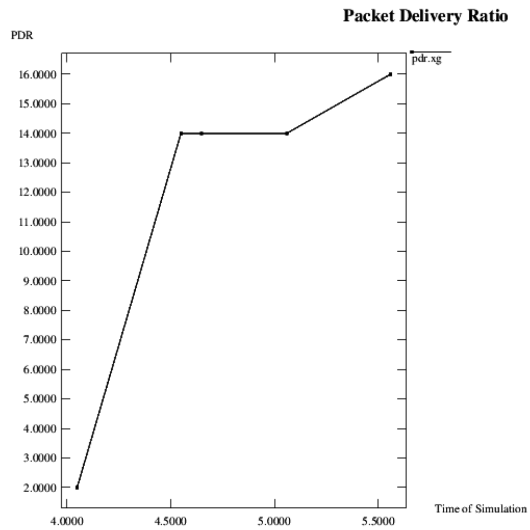

Fig.7 Packet Delivery Ratio

VIII. CONCLUSION

In this paper we have proposed a method for cluster formation and cluster head selection for protocol TEEN. The approach towards creating energy efficient clusters is effective compared to the others, the parameters analysis like packet delivery ratio and energy consumption proves the same. TEEN is well suited for time critical applications and is also quite efficient in terms of energy consumption and response time. It also allows the user to control the energy consumption and accuracy to suit the application.

REFERENCES

[1] Arati Manjeshwar and Dharma P. Agrawal, "TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Network". IEEE International Conference on Mobile Computing and Networking 2001.

[2] J. Broch, D. Maltz, D. Johnson, Y. Hu, and J. Jetcheva. "A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols". In Proceedings of the Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking(MOBICOM), ACM, October 1998.

[3] J. Elson and D. Estrin. "An Address-Free Architecture for Dynamic Sensor Networks". Technical Report 00-724, Computer Science Department, USC, January 2000.

[4] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar. "Next Century Challenges: Scalable Coordination inWireless Networks". In Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM), pages 263–270, 1999.

[5] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan. "uAMPS ns Code Extensions". http://wwwmtl.mit.edu/research/icsystems/uamps/leach.

[6] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan. "Energy-Efficient Communication Protocols for Wireless Microsensor Networks". In Proceedings of Hawaiian International Conference on Systems Science, January 2000.

[7] W. B. Heinzelman. "Application-Specific Protocol Architectures for Wireless Networks". PhD thesis, Massachusetts Institute of Technology, June 2000.

[8] C. Intanagonwiwat, R. Govindan, and D. Estrin. "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks ". In Proceedings of the 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking(MOBICOM), pages 56–67, August2000.

[9] Rajesh Patel, Sunil Pariyani and Vijay Ukani, "Energy and Throughput Analysis of Hierarchical Routing Protocol (LEACH) for Wireless Sensor Network". International Journal of Computer Applications (0975 – 8887) Volume 20– No.4, April 2011.

[10] M. Jiang, J. Li, and Y. C. Tay. "Cluster Based Routing Protocol". Internet Draft, 1999.