International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

# A Survey on Enhanced Energy Efficient Data Gathering Scheme in Wireless Sensor Network

**R. Mehala[1], Dr. A. Balamurugan[2]**

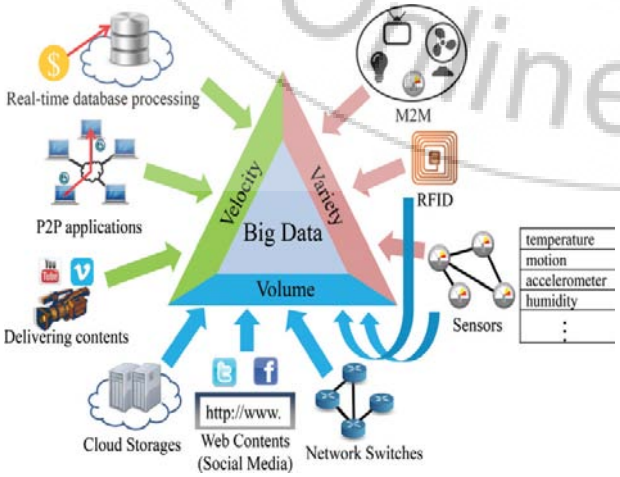
PG Scholar Professor, Department of Computer Science and Engineering, Sri Krishna College of Technology
Coimbatore, Tamilnadu, India

**Abstract:** ***A Wireless Sensor Network (WSN) is a collection of small sensor nodes with sensing, computation and communication capabilities. The sensor nodes are randomly deployed in distributed environment. The main goal of this paper is to form clusters of sensor nodes and to perform Data Aggregation to collect the data by maintaining the energy efficiency so that the network lifetime can be increased. Network lifetime is an important issue in an energy constrained sensor networks. The most widely used Expectation-maximization (EM) algorithm along with proposed Neighborhood density based clustering algorithm is used for clustering process and Directed Diffusion is used for data aggregation. One of the main problem in directed diffusion is the implementation of flooding diffusion which is used to forward interest and discover the routing map that reduces network lifetime and data latency through high energy consumption.***

**Keywords:** Expectation Maximization, Directed Diffusion, Gaussian mixture model

## 1. Introduction

With the recent technological advances in wireless communications, processor, memory, radio, low power, highly integrated digital electronics, and micro electro mechanical systems (MEMS), it becomes possible to significantly develop tiny and small size, low power, and low cost multi-functional sensor nodes. Big data comprises high volume, high velocity, and high variety information assets [6][7], which has to be gathered, stored and processed by using the available technologies Wireless sensor networks (WSNs) have been used for numerous applications and deployed in so many areas like vital signal monitoring in the homecare systems like analyzing patients health, ecology sensing which are widely used for monitoring wild-life, micro organisms, changes in the sea or lake water, soil texture after natural disasters like typhoon, tsunami, flood and soil erosion, monitoring climatic changes like temperature, structural monitoring like monitoring the conditions of a bridge after its construction, monitoring the historic buildings and Surveillance in Defence organizations including military surveillance, facility monitoring and environment monitoring[1,2,3]. Typically WSNs have a large number of sensor nodes with the ability to communicate among themselves and also to an external sink or a base-station.




In this paper, we are considering base station as static one where the mobile sink moves across each clusters to collect the sensed data. The sensors could be scattered randomly in environments such as a battlefield or deterministically placed at specified locations. The sensor nodes are grouped to form clusters based on clustering techniques. The sensors coordinate among themselves to form a communication network such as a single multi-hop network or a hierarchical organization with several clusters and cluster heads. The sensors periodically sense the data, process it and transmit it to the sink node. These sensor nodes can communicate within clusters and collaborate to accomplish a common task. The sensor nodes are energy constrained, therefore it is inefficient for all the sensor nodes to transmit the sensed data directly to the sink node. Data sensed by the sensor nodes which are nearer to each other is redundant. Hence, there is a need for a method which combines the data from different sensor nodes and reduces the number of packets to be transmitted to the sink node. Wireless sensor nodes require less power for processing the data than compared to transmitting data. This results in the saving of energy and increase in the network lifetime. Data aggregation usually involves the fusion of data from multiple sensors at intermediate nodes and transmission of the aggregated data to the mobile sink. Since sensor nodes may generate contain redundancy of data, packets from multiple nodes can be aggregated so that the number of transmissions is reduced. This can be accomplished by data aggregation techniques. The cluster head attempts to collect the critical data from the neighboring and intermediate sensor nodes and make it available to the mobile sink in an energy efficient manner with minimum data latency.

## 2. Clustering Based Data Aggregation in WSN

In this section, we first outline the clustering and data aggregation problem in WSN and considered network model in WSN. The overview of EM algorithm for clustering and directed diffusion for data gathering is discussed. Based on EM algorithm, we propose our clustering method and the procedure to gather data in WSN.

### A. Clustering Problem

When considering data gathering scheme in WSN using mobile sink, the biggest challenge is used to reduce the energy consumed by the sensor nodes. The energy depletion can be reduced by forming clusters of sensor nodes and then to perform data gathering. When forming clusters it should be ensure that no same node should be present in one or more clusters.1) What is the best algorithm for dividing nodes into clusters? The best way is to use clustering technique to group sensor nodes and after the cluster formation, cluster head has to be selected based upon some energy constraints. EM algorithm is powerful and well-known method to solve the clustering problem by iteratively calculate the sum of square of distance between every node and cluster centroid, we adopt EM algorithm over the 2-dimensional Gaussian mixture distribution.

### B. Data Aggregation problem

Data aggregation is a process of aggregating the sensor data using aggregation approaches. These approaches uses the sensor data from the sensor nodes and then aggregates the data by using some aggregation algorithms such as centralized approaches like LEACH(low energy adaptive clustering hierarchy),TAG(Tiny Aggregation) and Directed Diffusion[9] .The data aggregated is transfer to the mobile sink node by selecting the efficient path. The most popular data aggregation algorithms are cluster-based data aggregation algorithms, in which the nodes are grouped into clusters and each cluster consists of a cluster head (CH) and some members. Each member transmits data to its CH, then each CH aggregates the collected data and transmits the fused data to BS. The cluster- based WSNs have an inherent problem of unbalanced energy dissipation. Some nodes drain their energy faster than others and result in earlier failure of network.

### C. Data request flooding problem

The sink node sends data request message to invoke data transmission from sensor nodes when it arrives at the cluster centroids. The nodes that receive data request message send the data to the sink node and broadcast data request message to their neighboring nodes. The data request messages are broadcasted repeatedly until all nodes in the same group receives the message. Some nodes may receive data request message more than two times, but they send data and only at the first time of receiving the message. These broadcasts of data request message cause high energy consumption because the network will be flooded with redundant wireless communication. Thus, reducing data request transmission is also important for mobile sink in data gathering scheme.

### D. Overview of Expectation Maximization Algorithm

The EM algorithm is a classical clustering algorithm, where nodes are distributed according to Gaussian mixture distribution [8],

$$p(x) = \sum_{k=1}^{K} \left( \pi_k N\left(x|\mu_k, \sum_k\right) \right)$$

where K and $\pi_k$ indicate the total number of clusters and the mixing co-efficient of the $k^{th}$ cluster. The EM algorithm seeks to find the maximum likelihood estimation of the marginal likelihood by iteratively applying the E-step and M-step.

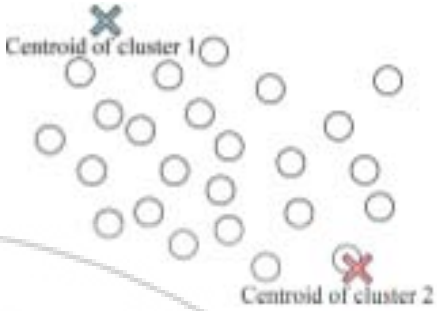



**Figure 1:** (a) First stage of EM algorithm

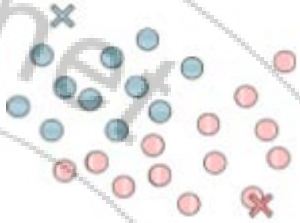


**Figure 1:** (b) Second stage of EM algorithm

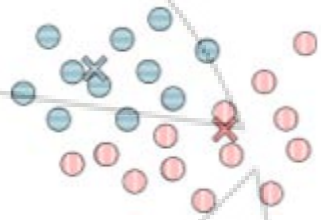


**Figure 1:** (c) Third stage of EM algorithm

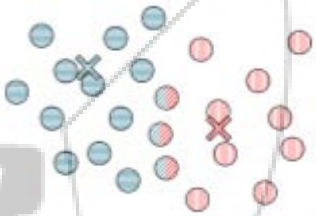


**Figure 1:** (d) Fourth stage of EM algorithm

Fig. 1(a) is the initial status of network. Every nodes do not belong to any cluster and centroids of clusters which represented by the cross are randomly decided. At the first step, shown in fig.1(b), EM algorithm calculates each node's degree of dependence that is referred to as responsibility. The responsibility value shows how much a node depends on a cluster. This responsibility value is calculated by the nodes location and centroid location of cluster. Normally, each node depends only on one cluster. However, it is possible for nodes to depend on more than one cluster so that those nodes will not focus their energy in a single cluster. After that EM algorithm calculates the responsibility value and forms centroid of the cluster by using node's location as shown in fig.1(c). The centroid of each cluster is calculated to minimize the distance between each node which belong to the cluster and the centroid.

## 3. Proposed Cluster-Based Data Aggregation System:

### A .Considered Network Model

The system consists of large number of sensor nodes. The sensor nodes are of equal energy at the initial phase. There

are two sinks in the proposed system one is the static sink node and other is mobile sink node[4].The static sink node is placed in the center of the network and in this system there are two mobile sinks. These mobile sinks travels around the cluster centroid to gather data. The cluster centroid gets data request message from the sensor nodes and forward the message to one of the mobile sink. When the mobile sink gets message from more than one cluster centroid it forwards to other mobile sink.

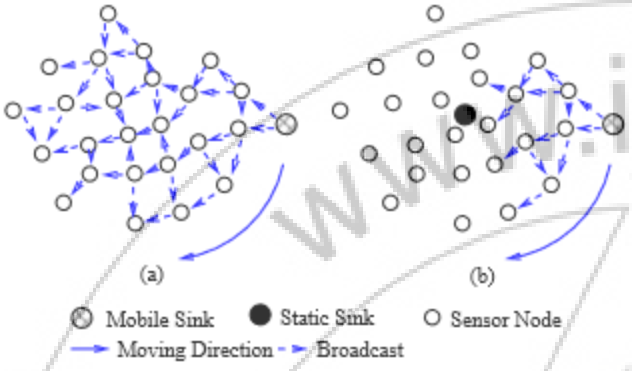



**Figure 2:** (a) The mobile sink uses network-wide broadcasting; (b) The mobile sink only broadcasts to a subset of nodes in Dual-Sink

**B. Clustering Phase**

The sensor nodes are deployed in a physical environment for monitoring and sensing the data. A large number of sensor nodes of equal energy and mobile sink of higher energy is deployed in the sensing area. At the initial stage, the nodes are scattered and cluster centroid is placed at random location by using EM algorithm [10].Then the clusters of nodes are formed by calculating the shortest distance between cluster centroid and senor nodes. The distance is calculated by taking one-hop neighbor from each sensor node to cluster centroids. The responsibility value for each node is measured. The responsibility value refers to the degree of dependence of sensor nodes. Some sensor nodes have lower degree and some nodes have higher degree value. Lower degree nodes are merged with higher degree nodes to form clusters by using Neighborhood density based clustering algorithm. The next step is to calculate the centroid location for each cluster. The centroid is calculated based on location information of cluster members. The centroid location place of each cluster is the data collection points in the network .Nodes are grouped based on fixed transmission distance from the centroid. Higher degree sensor nodes are also taken as to be member nodes. After calculating the distance from each node, the centroid location is calculated and grouped into the cluster.

**C. To find shortest path for Mobile sink**

To find the shortest path for data gathering path using a mobile sink, Travelling Salesperson (TSP) problem is used .The mobile sink patrols around the cluster centroids to collect the data from sensor nodes. The sum of the square of the distance between cluster centroid and nodes are reduced. When the transmission path distance reduces, the energy consumption of sensor nodes and mobile sink, data latency can be reduced.

**D. Data gathering using Mobile Sink**

The mobile sink remains idle till it get data request message from cluster centroids. Directed Diffusion using "Push Diffusion"[5] is used for data gathering. The sensor node sink identifies a set of attributes and propagates an interest message to the cluster centroids. Each node records the interests and establishes gradient, the state indicating the next hop direction for other nodes to report data of interest. When an interest arrives at a data producer, data are being forwarded to the sink along established gradients. The cluster centroid sends request message to the mobile sink once it reaches a fixed threshold value [11].It sends the sensor nodes message id along with the request to the mobile sink. Then the sink takes the shortest path to reach the particular centroid. The mobile sink remains idle till it receives the message request. The mobile sink collects the gathered data from each cluster centroid point and sends it to the base station.

## 4. Conclusion

This paper presents Clustering-based Data aggregation algorithm for Wireless Sensor Networks by using two mobile sinks, which improves network lifetime and reduces the transmission distance between each nodes. By using this proposed method, the sensed data are collected efficiently which minimizes data latency.

## References

[1] D. Culler, D. Estrin, and M. Srivastava, "Overview of Sensor Networks," IEEE Computer, August 2004.
[2] N. Xu, S. Rangwala, K. Chintalapudi, D. Ganesan, A. Broad, R. Govindan, and D. Estrin, "A Wireless Sensor Network for Structural Monitoring," Proceedings of the ACM Conference on Embedded Networked Sensor Systems, Baltimore, MD, November 2004.
[3] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, "Wireless Sensor Networks for Habitat Monitoring," WSNA'02, Atlanta,Georgia, September 2002.
[4] Xiaobing Wu , Nanjing Univ., Nanjing Guihai Chen (Aug. 2007),"Dual-Sink: Using Mobile and Static Sinks for Lifetime Improvement in Wireless Sensor Networks ",Computer Communications and Networks, ICCCN 2007, pp 1297 – 1302.
[5] C. Intanagonwiwat, R. Govindan and D. Estrin, "Directed Diffusion: A Scalable and robust communication paradigm for sensor networks," Proceedings of the Sixth Annual International Conference on Mobile Computing and Networking (MobiCOM '00)
[6] A. Divyakant, B. Philip, and et al., "Challenges and opportuni- ties with Big Data," 2012, a community white paper developed by leading researchers across the United States.[Online].Available:http://cra.org/ccc/docs/init/big data whitepaper.pdf.
[7] S. Sagiroglu and D. Sinanc, "Big data: A review," in International Conference on Collaboration Technologies and Systems (CTS), 2013.
[8] Nandini. S. Patil, Prof. P. R. Patil, "Data Aggregation in Wireless Sensor Network", IEEE International

Conference on Computational Intelligence and Computing Research, 2010.

[9] Xuxun Liu, "A Survey on Clustering Routing Protocols in Wireless Sensor Networks," Sensors 2012, 12, 11113-11153; doi:10.3390/s120811113, 9 August 2012

[10] Daisuke Takaishi, Hiroki Nishiyama, Nei Kato, and Ryu Miura, "On the Effect of Data Request Message Flooding in Dense Wireless Sensor Networks with a Mobile Sink," IEEE 78th Vehicular Technology Conference (VTC 2013 Fall), Las Vegas, Nevada, USA, Sep. 2013.

[11] Seino, W., Sakashita, S., Yoshihisa, T., Hara, T. and Nishio, S. (March 2012), "A Sensor Data Collection Method with a Mobile Sink for Communication Traffic Reduction by Delivering Predicted Values", pp 613 – 618.