

# A Survey On Data Aggregation Techniques to Improve the Energy Efficiency in Wireless Sensor Networks

Snehal S Lonare<sup>1</sup>, A S Hiwale<sup>2</sup>

<sup>1</sup>Department of Information technology, MIT college of Engineering, University of Pune,India

<sup>2</sup>Professor, HOD, Department of Information technology, MIT college of Engineering, University of Pune,India

**Abstract:** Data aggregation method in a wireless sensor network (WSN) is used for sending local representative data to the sink node which is based on the spatial correlation of sampled data. The problem occurs in recent spatial correlation models in sensor node's data are not appropriate for measuring the correlation in a complex environment. The representative data are inaccurate when compared with real data. The data density with correlation degree is necessary to resolve the problem. The spatial correlation is measured that measures the correlation between a sensor node's data and its neighboring sensor node's data. The data aggregation is necessary for minimizing power utilization during the data transmission in wireless sensor network and multilevel network clustering method is used for reducing power consumption. Based on a data density correlation degree (DDCD) clustering method representative data.

**Keywords:** Data aggregation, clustering methods, energy efficiency, wireless sensor network.

## 1. Introduction

Wireless sensor network (WSN) is a collection of devoted autonomous sensor nodes that observed physical or environmental conditions differently, such as pollution levels, humidity, temperature,

sound, wind direction, pressure, etc. To cooperatively deliver data through the network to a main destination. WSNs were initially designed to facilitate military applications such as battlefield surveillance; but its usages have since been drastically extended to monitoring machineries, industrial processes, health, and controls. The WSN consist of "nodes" which may be differ in numbers from a few to several hundreds or even thousands, where every node is in connection with at least one sensor (or sometimes several). The sensor node is equipped with a tiny processor, a small battery, a radio transceiver antenna, and situate of transducers that used to gather information. They describe the variations in the environment of the sensor node [3]. Topology for WSNs can vary from a simple star or mesh network to an advanced multi-hop wireless mesh network.

The propagation of data takes place between the hops of the network with the help of routing or flooding techniques. To attain high energy saving and long network lifetime, nodes can grouped into clusters and can be arranged in a hierarchical manner. In the working process Clustering algorithm includes partitioning of sensor nodes into several clusters, and each cluster selects one sensor node as cluster-head (CH) [2]. In clustered sensor networks, the nodes do not transmit their collected data to the sink or base station (BS), but to designated cluster heads (CH) which aggregate the data packets and send directly or via multi-hop communication to the BS [3], [11]. In-network data aggregation reduces the amount of data transmission and communication; hence the energy consumed, especially in large WSNs. The main idea is to combine partial results at intermediate nodes during data transmission. [1].

## 2. Literature Survey

In [4], Genetically derived and secured cluster based data aggregation (GSDSA) in WSN the CHs are chosen by considering the node connectivity, which interns act as a DAG. The genetic algorithm is used for cluster based data aggregation. When the cluster member wants to transmit data to the aggregator, a data encryption technique is utilized. The confidentiality to the data packet (DP) is offered by crypto module (CyM), thus ensuring the authenticity and integrity of the sensed data.

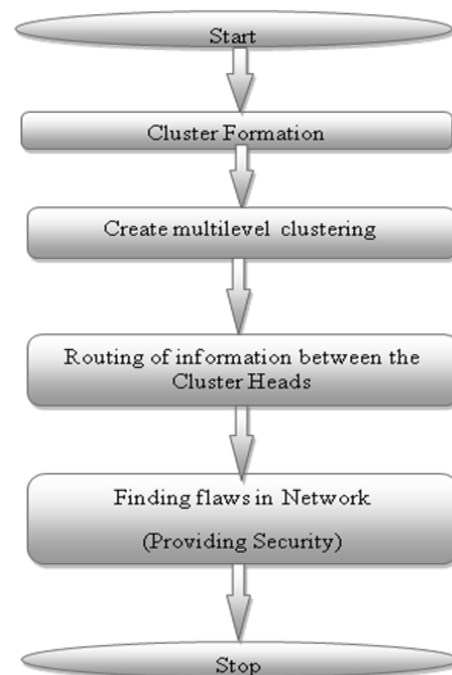


Figure 2: cluster based data aggregation

In [5], Tiny Aggregation (TAG), a generic aggregation service for ad hoc networks of Tiny OS motes. They classify aggregates according to four properties that are particularly

important to sensor networks. The first is duplicate sensitivity. Duplicate sensitive aggregates will change when a duplicate reading is reported while duplicate insensitive aggregates are unaffected by duplicate readings from a single device. Second, excellent aggregates return one or more representative values from the set of all values; summarized aggregates compute some property over all values. Third, monotonic aggregates have the property that when two partial state records,  $s_1$  and  $s_2$ , are combined via  $f$ , the resulting state record  $s'$  will be greater or smaller than the complete sample space. The fourth is related to the amount of state required for each partial state record. For example, a partial AVERAGE record, and a partial COUNT record constitutes only a single value.

In [6], they have proposed the DOC protocol, which can be select a set of disjoint potential clusters that maximize the global compression gain of Slepian-Wolf coding. In a cluster hierarchy constructed by the DOC protocol, they have proven that there exists an optimization algorithm that can find an optimal rate allocation within each cluster to minimize the intracluster communication cost and have presented an intracluster coding protocol to locally perform Slepian-Wolf coding within a single cluster.

In [7], K-Centers Mean-shift Reverse Mean-shift Clustering (KCMRM) protocol, cluster heads have both mobility and higher initial energy. This allows CHs to handle the heavy burden of receiving and transmitting data without running out of energy beforehand, thereby extending network's lifetime and preserving coverage of a monitoring area. However, cluster heads having high initial energy are more expensive than immobile sensor nodes with lower initial energy. KCMRM can reduce the number of empty clusters when the sensors and CHs are arbitrarily deployed. It calculates the mean-shift of sensor nodes and the reverse mean-shift of cluster heads to iteratively move cluster heads in accordance with sensor nodes' density and away from cluster heads' density.

In [8], they proposed an energy efficient algorithm using node ranking in electing cluster heads. The difference between this algorithm and other algorithms is that this algorithm uses a more competent mechanism to select cluster heads. This is accomplished by calculating the distance, the current energy level of nodes and scheming the number of rounds that they can be clustered heads for, to maximize the network lifetime and decrease unnecessary communication overheads used for electing new cluster heads. In this algorithm, nodes are ranked based on their current energy level ( $E_n$ ) and their positions ( $d_n$ ) in reference to the BS. This ranking is used for choosing cluster heads which are also ranked into levels based on their position, Euclidean distance, from the BS. Therefore, each node is assigned a rank  $R_n(E_n, D_n)$  reflecting its candidacy for being elected as a cluster head.

In [9], they have introduced data density correlation degree and the data density correlation degree (DDCD) clustering method. With the DDCD clustering method, the sensor nodes having high correlation are distributed in the same cluster, allowing more accurate aggregated data to be obtained in cluster-based data aggregation networks

produced by the DDCD clustering method. Also, the amount of data conveyed to the sink node can be minimized. The WSN is modelled by undirected graph  $G = (V, E)$ . Where  $V$  is the sensor node set consisting of all sensor nodes in the WSN,  $E$  is the edge set consisting of all links in the WSN. The antenna of sensor node  $I$  ( $I \in V$ ) aim an omnidirectional antenna, with a communication radius of  $\alpha$  ( $i$ ). In cluster-based data aggregation networks, the data transmission process is that every cluster head sends aggregated data obtained from its member nodes to the sink node by one hop or multi-hops.

### 3. Analysis

Papers are reviewed and analyzed related to clustering and aggregation mechanisms in WSN. Those techniques are reviewed, are different from each other, some are based on density of deployed sensor nodes, while some are based on node connectivity, energy, position, etc. In TAG approach, the most of the significant computation time is utilized for data aggregation purpose. Also the number of messages to be exchange is high. Data aggregation is performed on the information present in the database in the form of queries. This extends time complexity.

In GDSDA, the CHs are chosen based on the node connectivity, then the genetic algorithm is used. When the cluster member wants to transmit data to the aggregator, a data encryption technique is utilized. Here nodes connectivity does not guarantee the largest possible energy node will be elected as cluster head and hence may lead to node exhaustion.

KCMRM selects CHs based on high mobility and high initial energy. However, evaluating high mobility is expensive than finding immobility in WSN. This algorithm focuses on removing empty clusters. In node ranking approach, nodes are ranked by calculating the distance, the current energy level of nodes and scheming the number of rounds that they can be clustered heads for. This algorithm highly enhances the network lifetime by choosing highest energy node.

In DDCD clustering method, clusters are formed based on their level of correlation, the sensor nodes having high correlation are distributed in the same cluster. This allows aggregated data to be obtained more accurately.

### 4. Conclusion and Future Work

The comparison of various clustering and aggregation techniques have been done and presented in this paper. The direct further work on the DDCD clustering method is developing a method which could confirm the parameters adapted to the real sampled data, especially the data threshold has major effect on clustering performance. In the data transmitting process, the energy of sensor nodes should be considered to construct an energy balanced networks. Hence, this will be researched in our future work as well. Based on the analyzed parameters, data density correlation degree and node ranking algorithms are energy efficient for immobile nodes in Wireless Sensor Network. Implementing data

density correlation degree, clustering and electing a cluster head based on node rank would increase the network lifetime.

### References:

- [1] Sankardas Roy, Mauro Conti, Sanjeev Setia, Sushil Jajodia, "Secure Data Aggregation in Wireless Sensor Networks: Filtering out the Attacker's Impact," *IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY*, VOL. 9, NO. 4, APRIL 2014
- [2] Jinhui Yuan; Hong Chen, "The optimized clustering technique based on spatial-correlation in wireless sensor networks," *Information, Computing and Telecommunication*, 2009. YC-ICT '09. IEEE Youth Conference on, vol., no. , pp. 411, 414, 20-21 Sept. 2009
- [3] Deepak Kumar Yadav, Mayank Kumar Sharma, "Novel Multi Level Clustering Approach to Reduce Energy Consumption during Clustering Process in Wireless Sensor Networks," *International Journal of Scientific and Research Publications*, Volume 3, Issue 8, August 2013 | ISSN 2250-3153
- [4] Lathes Bhasker, "Genetically derived secure cluster-based data aggregation in wireless sensor networks," *IET Inf. Secure.* , 2014, Vol. 8, Iss. 1, pp. 1–7
- [5] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TAG: A Tiny AGgregation service for ad-hoc sensor networks," *ACM SIGOPS Operating Syst. Rev.* , vol. 36, no. 1, pp. 131–146, 2002
- [6] Jun Zheng, Pu Wang, Cheng Li, "Distributed Data Aggregation Using Slepian–Wolf Coding in Cluster-Based Wireless Sensor Networks," *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY*, VOL. 59, NO. 5, JUNE 2010
- [7] Qing Yan Xie; Yizong Cheng, "K-Centers Mean-shift Reverse Mean-shift clustering algorithm over heterogeneous wireless sensor networks," *Wireless Telecommunications Symposium (WTS)*, 2014, vol., no., pp. 1,6, 9-11 April 2014
- [8] Alnuaimi, Mariam; Shuaib, Khaled; Alnuaimi, Klaihem; Abed-Hafez, Mohammed, "Clustering in Wireless Sensor networks Based on Node Ranking," *Wireless Communications and Mobile Computing Conference (IWCMC)*, 2014 International , vol., no., pp.488,493, 4-8 Aug. 2014
- [9] Fei Yuan, Yiju Zhan, and Yonghua Wang, "Data Density Correlation Degree Clustering Method for Data Aggregation in WSN," *IEEE SENSORS JOURNAL*, VOL. 14, NO. 4, APRIL 2014
- [10] J. N. Al-Karaki, R. UL-Mustafa, and A. E. Kamal, "Data aggregation and routing in wireless sensor networks: Optimal and heuristic algorithms," *Comput. Netw.*, vol. 53, no. 7, pp. 945–960, 2009.
- [11] J. Yuan and H. Chen, "The optimized clustering technique based on spatial-correlation in wireless sensor networks," *in Proc. IEEE Youth Conf. Inf., Comput. Telecommun. YC-ICT*, Sep. 2009, pp. 411–414.
- [12] A. Seetharam, A. Acharya, A. Bhattacharyya & M. K. Naskar, (2008) "An Energy Efficient Data Gathering Protocol for Wireless Sensor Networks", *Journal of Applied Computer Science*, Vol.2, No. 1. International

- Journal of Distributed and Parallel Systems (IJDPSS)  
Vol.2, No.2, March 2011
- [13] Ahmad Khadem Zadeh, Ali Hosseinalipour & Shahram Babaie, (2010) "New Clustering protocol for increasing Wireless Sensor Networks Lifetime" *International Journal of Computer and Network Security(IJCNS)* , Vol. 2, No. 1.