A Survey on Data Aggregation Techniques in Wireless Sensor Networks
Shrikant A. Shinde1 Shailaja N. Uke2
1Student 2Assistant Professor
1,2Department of Information Technology
Smt. Kashibai Navale College of Engineering Pune, India

Abstract— Wireless sensor networks (WSNs) consist of sensor nodes and the sensor nodes are capable of collecting, sensing and gathering data from environment. These networks have huge application in disaster management, habitat monitoring, security and military, etc. Wireless sensor nodes are very small in size and very low battery power and have limited processing capability. We focus on data-aggregation in wireless sensor networks. Data aggregation is a very important technique in WSNs and helps in reducing the energy consumption by eliminating redundancy. The main goal of data aggregation is to collect and gather data in an energy efficient manner so that network lifetime is improved. In this paper we present a survey of data aggregation techniques and algorithms in wireless sensor networks. We compare and contrast different techniques as well as algorithms on the basis of performance measures such as lifetime, energy efficiency, latency, data accuracy and transmissions policies.

Key words: Data Aggregation, Energy Efficiency, Wireless Sensor Networks

I. INTRODUCTION
A Wireless Sensor Network (WSN) consists of spatially distributed autonomous sensor to monitor environmental or physical conditions, such as sound, temperature, pressure etc. and to cooperatively pass their data through the network to a destination. Wireless sensor network composed of self-organized wireless sensor nodes distributed in a monitored area and which collects, processes and transmits data acquired from the environment. The main goal of a WSN is accurately evaluating and reliably detecting the events in the monitored area with the collected data. If every sensor node transfers collected data to the sink node, the sensor nodes will consume much more energy.

The wireless sensor network is the collection of wireless sensors which communicates each other by sending and receiving data. These sensors are very small devices and are distributed in a large geographical area called as sensing field. It has sensing ability to sense some environmental conditions, processing ability and also has memory to store the data. The role of the sensors is to sense the environmental conditions, process the data and send that information to the base station [3]. There are different types of sensors such as seismic, visual, infrared and thermal sensors. These sensors monitor variety of conditions such as humidity, temperature, characteristic of object and their motion. These different types of sensors are used in number of applications like flood or fire detection, measuring temperature, smoke detectors, soil makeup, humidity in the air, pressure, vehicular movement, noise levels or mechanical stress levels. It can also be used in military, chemical processing and disaster, health, relief scenarios [8]. There are many issues related to the sensor network. For example- transmission policies, orderly transmission, flow and congestion control, power consumption, fairness, loss recovery, quality of service, energy efficiency, reliability. Congestion and transmission policies are major issue on which the rate of energy consumption depends [3].

Data aggregation is a very important technique in WSNs and helps in reducing the energy consumption by eliminating redundancy. The main goal of data aggregation is to collect and gather data in an energy efficient manner so that network lifetime is improved. Data aggregation is the process of collecting and aggregating the data. Data aggregation is one of the fundamental processing procedures for saving the energy. In WSN, data aggregation is an effective way to save the limited resources as well energy. The main goal of data aggregation is to gather, collect and aggregate the data in an energy efficient manner so that the network lifetime is enhanced [2]. Data aggregation improves the lifetime of sensor nodes by eliminating redundant data transmission. The data transmission uses a multi-hop fashion in which each node transfers data to the neighbor node nearer to the sink.

According to the level of sampled data in data aggregation strategy, data aggregation methods are divided into three classes namely data level aggregation, feature level aggregation and decision level aggregation [1].

The remainder of this paper is organized as follows: Section II gives an overview of related work already done on data aggregation. Section III includes result and comparison of transmission policies and aggregation techniques of existing system. Section IV concludes and gives future work for this survey.

II. RELATED WORK
Data aggregation improves the lifetime of sensor nodes by eliminating redundant data transmission. The data transmission uses a multi-hop fashion in which each node transfers data to the neighbor node nearer to the sink. According to the level of sampled data in data aggregation strategy, data aggregation methods are divided into three classes namely data level aggregation, feature level aggregation and decision level aggregation [1]. Data aggregation in a wireless sensor network refers to obtaining the sensed data from the sensors and sending to the central node for compression purpose. Aggregation minimizes the amount of network traffic which helps to minimize energy consumption on sensor nodes.

A. Data Aggregation Schemes
Two data aggregation schemes are proposed for nodes to exchange and compress the data [2].

The first scheme is Centralized Aggregation Scheme (CAS). In this scheme, a central node of a cluster collects data sensed by all the nodes in the cluster and then integrates and compresses the data and the compressed data is then distributed back to the nodes [2].

The other scheme is Distributed Aggregation Scheme (DAS), in which each node exchanges its data with
all other nodes in a cluster and then separately compresses the data [2].

1) CAS:
CAS works in three phases, namely gathering, compression and broadcasting and it is explained below.
- Gathering phase: The central node of the cluster is selected as the cluster head and certain amount of energy is utilized for this purpose [2].
- Compressing phase: The data sensed by many sensor nodes within a cluster are correlated and the degree of correlation in the data from different nodes is a function of the distance between them.
- Broadcasting phase: The central node contains the compressed data and sends the compressed and integrated data to other nodes [2].

2) DAS:
DAS has two phases namely gathering and compression and it is explained below.
- Gathering phase: Each node in a cluster exchanges its sensed data with all other nodes in that cluster by using many time slots to broadcast [2].
- Compressing phase: Each node separately collects and compresses the data gathered from the gathering phase and this compressed data is used in long-haul communication [2].

B. Transmission Policy Review:

1) Characteristic Distance for Transmission:
To obtain the optimal transmission distance at each hop that reduces the total energy usage along the path, a data link in a radio transmitter and a receiver separated by D meters are divided into K sub paths by introducing (K-1) relay nodes [3].

III. RESULT AND COMPARISON

On the basis of above survey we compare some mechanism used for the data aggregation. The Overview of various existing techniques and algorithms are described in following table. The techniques and mechanism used, advantages and disadvantages are included in the table.

<table>
<thead>
<tr>
<th>Paper Title</th>
<th>Mechanism</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data Density Correlation Degree Clustering Method for Data Aggregation in WSN.</td>
<td>Correlation degree is a spatial correlation measurement that measures the correlation between a sensor node’s data and its neighboring sensor nodes’ data.</td>
<td>The distance between representative data and the data center of its correlated data which are in its ε-neighborhood has significant influence on the value of DDCD.</td>
<td>A data density correlation degree (DDCD) was proposed to measure the spatial correlation of sampled data and try to resolve the drawbacks in existing spatial correlation Models.</td>
</tr>
<tr>
<td>2. Data Aggregation in Wireless Sensor Networks.</td>
<td>Data Aggregation is a technique which is combined with CMIMO to reduce the energy cost further by reducing the amount of data in transit.</td>
<td>To highlight the advantage of CMIMO-A, it has been compared with other techniques like CMIMO, DATAG, and the traditional SISO system.</td>
<td>The most prominent disadvantage of a WSN being its power constraint, many techniques has been introduced in literature to reduce the amount of battery power consumed by each node in the network.</td>
</tr>
<tr>
<td>3. Energy-Balanced Transmission Policies for Wireless Sensor Networks.</td>
<td>Decomposes the transmission distance of traditional multi-hop scheme into two parts: ring thickness and hop size, analyze the traffic and energy usage distribution among sensors and determine how energy usage varies and critical ring shifts with hop size.</td>
<td>In many applications, sensors could be preprogrammed before deployment with these optimal values, calculated in advance from predefined network parameters.</td>
<td>Relay traffic is higher no nodes closer to the sink than nodes near the boundary.</td>
</tr>
<tr>
<td>4. Distributed Data Aggregation Using Clustered Slepian-Wolf Coding in</td>
<td>Which aims at selecting a set of disjoint potential clusters to cover the whole network such that the global compression gain.</td>
<td>This can select a set of disjoint potential clusters that maximize the global compression gain of</td>
<td>In these studies, applying Slepian-Wolf coding globally is shown to be difficult and optimally constructing a cluster.</td>
</tr>
</tbody>
</table>
Wireless Sensor Networks.

5. Distributed Clustering-Based Aggregation Algorithm for Spatial Correlated Sensor Networks.

The evaluation shows that the resulting network achieved by our algorithm can provide environmental information at higher accuracy compared to other algorithms. The main advantage of time and message complexities of the algorithm, with the analysis of the size of the aggregated networks. Such an approach not only helps in optimizing the performance characteristics of the network, but also helps in avoiding information overload when monitoring and analyzing large data sets.

Table 1: Comparison table for Mechanisms

IV. CONCLUSION
We have presented a comprehensive survey of data-aggregation algorithms and techniques in wireless sensor networks. We compare and contrast different techniques as well as algorithms on the basis of performance measures such as lifetime, energy efficiency, latency, and data accuracy and transmissions policies. The survey mainly focus on optimizing important performance measures such as network lifetime, data latency, data accuracy, energy efficiency, transmissions policies, energy consumption and aggregation schemes. 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. As a future work, we wish to include simulation results for some of the analyzed works and present a comparative study based on the results.

REFERENCES