DATA AGGREGATION IN WIRELESS SENSOR NETWORKS: AN OVERVIEW

[1]S.Nanthini, [2] Dr.S.Nithya Kalyani, [3]Dr.S.Sudhakar
[1]Assistant Professor, [2]Associate Professor, [3]Associate Professor
[1]Department of Computer Science and Engineering, [2]Department of Computer Science and Engineering,
[3]Department of Computer Science and Engineering,
[1]Sri Ramakrishna College of Engineering, [2]KSR College of Engineering , [3]Easa College of Engineering and Technology, Tamil Nadu, India

***Abstract* –** Wireless Sensor nodes have restricted resources in terms of energy, bandwidth and computation capability. Energy efficiency is a key design goal in sensor network. As one of techniques to achieve energy efficiency, is data aggregation. It is the process of collecting and aggregating the useful data and it is one of the basic processing methods for saving energy. It has demonstrated its effectiveness in reducing traffic, easing congestion and decreasing the energy consumption. Conversely very few real world applications are designed and implemented in a running system.

***IndexTerms* - Data Aggregation, Wireless Sensor Network, Energy Efficiency, Bandwidth.**

1. INTRODUCTION

Data Aggregation techniques have been widely used in wireless sensor networks. Several basic forms of data aggregation methods are included here. They are i).The centre at the Nearest Source Method (CNS), where the destination aggregates the data from other nodes. ii).The Shortest Path Trees (SPT) method, where data aggregation happens at the intermediate nodes within a shortest path tree rooted at the sink; and iii).The greedy Incremental Trees (GIT) method, where an aggregation tree is constructed by connecting each destination sequentially to the existing tree via a shortest path.GIT assumes a complete knowledge of global topology information.

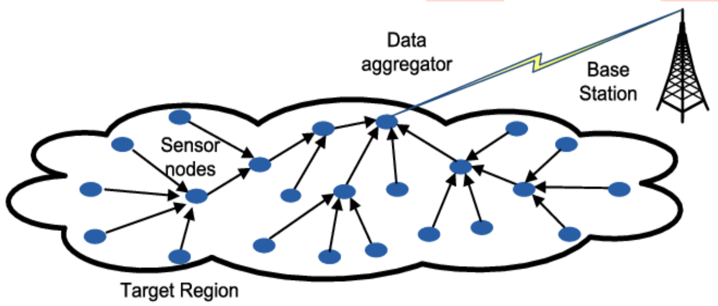

Figure 1. Data Aggregation Model

2. DATA AGGREGATION REQUIREMENTS

***2.1. Meet real-time constraints*:** Because the system deals with transient events, such as fast-moving targets, in the network, the sensor data needs to be processed in a timely manner. If the processing latency is too long, a target would move out the sensing range before the detection finishes.

2.2 *Deal with a large volume of inputs*: Endeavoring to accomplish reliable, timely, and quality sensing and tracking, a surveillance application often uses multiple sensors and samples them at very high rates.

3. AGGREGATION APPROACHES

The various approaches used by sink node for data aggregation is given below:

3.1. Node Level Data Aggregation

Each sensor node has four PIR sensors, one magnetometer with two axes, and one microphone. the node performs further aggregation after collecting confidence vectors from the sensors. it computes the averages of the sensors confidence vectors and form a single node-level confidence vector. Because the three types of sensors form their own confidence vectors per type ,the input of the node-level sensing and classification modules are three confidence vectors.

3.2. Group level Data Aggregation

Tracking of a target consists of two main parts; Position Estimation and Target representation. A simple way to tackle these problems would be to send the detection results and locations of individual nodes that detect targets to a centralized base station. Based on these received node positions, the base is able to estimate current positions of targets, and assign and maintain unique and consistent identifies for targets by running temporal and spatial correlation algorithms. However, such a centralized scheme is inefficient both in energy and latency.

3.3. Base level Data Aggregation

After receiving the reports from individual leader nodes, the base needs to aggregate the information further, an operation to serve three main objectives: Flow control, Filtering, Consolidated View. The long-haul link to the remote back-end could be the bottle neck of the system. A base is required to address the bandwidth mismatch between the sensor network output and the long-haul link capability. This is more likely to happen when the system tracks multiple objects simultaneously. A base needs to prevent sending the duplicate reports as well as the false alarms to the back-end. To filter out the false alarms, a system needs to correlate the spatiotemporal properties of consecutive reports. Since the base is the only place in the network that has the complete global knowledge of a tracked target, it is at a better position than in-network nodes to filter out the system-wide false alarms. End users is more interested in a consolidated view of the tracked targets instead of the individual sensing reports from the sensor network. Although the group-level aggregation does provide some persistent information such as object ID, it is not effective to keep a long trace history of a target through a sequence of hand-over operations among the leader nodes. To improve the efficiency, a base should be used to create a consolidated view instead.

4. DATA AGGREGATION STRATEGIES

Sensor nodes are resource constrained and possess limited battery. So, to avoid the usage of more resources and battery power, data sensed by sensor nodes must be aggregated and disseminated to other nodes. Data aggregation is the process of collecting data from different sensor nodes and combining it together by applying aggregate functions is known as data aggregation. Aggregation Strategies are used to enhance the network lifetime. The various aggregation strategies used in WSN are given below.

4.1. Continuous Packet Sensing and Dissemination (CPSD)

Since CPSD does not perform actual aggregation, it is also called as zero aggregation schemes. In this method, each node senses the data at fixed sensing intervals and the node immediately transmitting the received data to cluster head without storing in the buffer. CPSD is widely used in the situation where the fresh messages are required and the reception of data is required very urgent and any delay in receiving the data may lead to failure or performance degradation of the system. CPSD requires adequate energy to keep battery without dry.

4.2. Continuous Packet Collection and Dissemination (CPCD)

In CPCD, each node uses buffer to store the collected and sensed data. Buffer is a hardware unit and the storage area is allocated by the software used in sensor node. Each node has to wait until the buffer gets filled by data. The sensor node keeps on sensing data and tries to fill the buffer. Once the buffer is filled, the sensor node will start to disseminate data to other nodes. In CPCD, each data dissemination takes long interval since every node has to wait till the buffer is filled up. This scheme highly reduces network overhead and consumption of power.

4.3. Programmed Packet Collection and Dissemination (PPCD)

In PPCD, each sensor node senses the environment and collects data to store in the buffer. The sensor node will not disseminate the data immediately to other nodes. Instead, the node sets a dissemination time interval and waits until the time expires. If the dissemination interval time expires, then the sensor node will start to disseminate the buffered data to other nodes. If buffer overflow occurs before the dissemination interval, then the old packet will be replaced by the newly arrived data. So, increasing the dissemination time will highly reduce the regularity of packet transmission on the network. This scheme can be used in the situation where the data to be transmitted is not a critical case.

4.4. Programmed Packet Aggregation and Dissemination (PPAD)

In this scheme, each node senses the data and apply the required aggregate functions such as AVG,MIN, and MAX,STDDEV etc. on the sensed data. Each sensor node stores only the aggregated data rather than storing the sensed data. Finally, the buffer contains only one aggregated data and thereby saving memory.

5. DATA AGGREGATION ALGORITHMS

The objective of the data aggregation problem is to transmit the sensed data from each sensor node to BS. The goal of algorithms which implement data aggregation is to maximize the number of rounds of communication before the nodes die and the network become inoperable. This means minimum energy should be consumed and the transmission should occur with minimum delays, which are conflicting requirements. Hence, the energy x delay metric is used to compare algorithms, since this metric measures speedy and energy efficient data aggregation. A few algorithms that implement data aggregation are discussed in this section.

5.1. Direct Transmission

All sensor nodes transmit their data directly to the BS. This is extremely expensive in terms of energy consumed, since the BS may be very far away from some nodes. Also node must take turns while transmitting to the BS to avoid collision, so the media access delay is also large. Hence this scheme performs poorly with respect to the energy x delay metric.

5.2. Power Efficient Gathering For Sensor Information Systems

PEGASIS is a data aggregation protocol based on the assumption that all sensor nodes know the location of every other node, that is, the topology information is available to all nodes. Also any node has the required transmission range to reach the BS in one hop, when it is selected as a leader.

The Goals of PEGASIS are

- Minimize the distance over which each node transmits
- Minimize the broadcasting overhead
- Minimize the number of messages that need to be sent to the BS
- Distribute the energy consumption equally across all nodes

5.3. Binary Scheme

This is also a chain-based scheme like PEGASIS, which classifies nodes into different levels. All nodes which receive messages at one level rise to the next. The number of nodes is halved from one level to the next. This scheme is possible when nodes communicating using CDMA, so that transmission of each level can take place simultaneously.

5.4. Chain-Based Three-Level Scheme

For non-CDMA sensor nodes, a binary scheme is not applicable. The chain based three level scheme addresses this situation, where again a chain is constructed as in PEGASIS. The chain is divided into a number of groups to space out simultaneous transmission in order to minimize interference. Within a group, nodes transmit one at a time. One node out of each group aggregates data from all group members and rises to the next level. The index of this leader node is decided a priori. In the second level, all nodes are divided into two groups, and the third level consists of a message exchange between one node from each group of the second level. Finally, the leader transmits a single message to the BS.

5.5. Aggregated Packet Transmission (APT)

With aggregated packet transmission, nodes transmit a batch of packets in a single cycle. The benefit of this scheme is it provides shorter delay, throughput and energy efficiency tends to high.

6. PERFORMANCE MEASURE OF DATA AGGREGATION

Data aggregation consists of very important performance measures and the performance measures are highly dependent on the desired application.

Energy Efficiency: Every sensor nodes have spent some amount of energy while data gathering. Data aggregation scheme is highly energy efficient, while it maximizes the functionality of network.

Network Lifetime: Network lifetime is the time duration from the initial deployment to the instant when the network is considered nonfunctional.

Latency: It is defined as the data evaluation time delay experiences by the system, means the data are send by sensor nodes and the delivered data to be received by the base station. Normally delay involves during data transmission.

Communication Overhead: It is defined as the total numbers of packets are to be transferred or transmitted from one node to another is known as the *communication overhead*. It includes the overhead of routing process, routing table and packet preparation in a sensor node.

Data Accuracy: It is mainly used to evaluate the total number of reading (data) received at the base station to the total number of data generated.

7. IMPACT OF DATA AGGREGATION

In this section we discussed the following factors will affect the performance of data aggregation in wireless sensor network, such as saving energy and delay, fault-tolerance, security by surveying various protocols.

8. SECURITY ISSUES IN DATA AGGREGATION

Data aggregation exploits the sensed data from the sensors to the gateway node. It plays a vital role in WSN since the aggregation scheme followed and it involve reducing the power consumed during the data transmission between the sensor nodes. There are four important security requirements; Confidentiality, integrity, Authentication, Availability have to be consummated.

Confidentiality: The basic security issues is data confidentiality or privacy, it protect the sensitive transmitted data from passive attacks.

Data integrity: Data integrity prevents the compromised source node or aggregator nodes from considerably altering the final aggregation value. ***Authentication:*** Sensor requires authentication to find the malicious packets. To provide secure authentication sender and receiver sharing a secret key to compute the message authentication code for transmitted data.

Availability: Availability is more important than regular sensor nodes. It assures the survivability of network services against DOS (Denial-Of-Service) attacks. Wireless Sensor Network is deploying with high node redundancy to tolerate such availability losses.

9. CONCLUSION

In this paper we present the overview of data aggregation. Wireless sensor network is consisting of large number of sensor node and the sensor nodes are energy constraint. So, the life time of sensor nodes are limited. Various numbers of approaches and techniques are used to improve the life time of sensor nodes, in this paper we discuss the data aggregation is one of the most suitable techniques for improving the network life time. And also we discussed a security issue which is encountered in wireless sensor networks.

10. REFERENCES

[1] I.F.Akyildiz, W. Su, Y. Sankarasubramaniam, and E.Cayirci, “A Survey on Sensor Networks,” IEEE Communications Magazine, Vol. 40, No. 8, pp. 102-114, 2002.

[2] Stephanie Lmdsey and Cauligi S.Raghavendra “PEGASIS:Power Efficient Gathering in Sensor Information Systems”, IEEE 2002.

[3] Vaibhav Pandey, Amarjeet Karur and Narottam Chand “A review on data aggregation techniques in wireless sensor networks”, Journal of Electronic and Electrical Engineering, ISSN: 0976 -8106 & E-ISSN: 0976-8114, Vol. 3 No. 3 Mar 2011.

[4] S Sirsikar, S Anavatti,”Issues of Data Aggregation Methods in Wireless Sensor Network:A Survey”,Elsevier Procedia Computer Science 49,pp.194-201,2015.

[5] K. Akkaya, M. Demirbas and R.S. Aygun, “The Impact of Data Aggregation on the Performance of Wireless Sensor Networks”, Wiley Wireless Commun. Mobile Comput. (WCMC) J. 8 (2008) 171–193.

[6] Anita A.Gosavi and Sonali U.Nimbhorkar “Analysis of Secure Data Aggregation Mechanisms with the Impact of Collusion Attacks in Wireless Sensor Networks” Volume 6, Issue 2, February-2015.

[7] Ramesh Rajagopalan and Pramod K. Varshney, “Data aggregation techniques in sensor networks: A survey”, 2006

[8] Suat Ozdemir and Yang Xiao “Secure data aggregation in wireless sensor networks: A comprehensive overview” Computer Networks 53(2009) 2022–2037.