International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 4, July – August 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Data Aggregation in Wireless Sensor Networks: Issues, Challenges and Future Directions

[1]Dr. Nagendra Nath Giri, [2]Dr Anil Kumar R J, [3]Dr. Sumati Behera
[1]Maharani's Science college for women ,LB Road , Mysore ,Karnataka India
[2]Maharani's Science college for women,JLB Road , Mysuru, Karnataka.India
[3]Maharani's Science college for women,JLB Road , Mysore ,Karnataka , India,

Abstract:
In The large-scale deployment of wireless sensor ,Data should be aggregated in wireless sensor networks (WSN) to reserve energy and to avoid burst traffic in network. The regular sensor node generates data, based on its sensing mechanism, observation and then transmit sensed data packet to the base station (sink). Here the data can be transmitted directly to the base station located far away from the sensor nodes but requires more energy, hence it is better to transmit data by few nodes by collecting the data locally from neighboring nodes. These few nodes that collect and process data and later sent to base station are called aggregator nodes and the process is called aggregation and data dissemination in wireless sensor networks.

Keywords: — Wireless sensor network ,Sensor node ,Data aggregation, burst traffic,

I. INTRODUCTION

Wireless sensor networks are like ad-hoc network consisting of small light weighted wireless nodes called sensor nodes, deployed physically in the environment to be monitored and measures physical parameters such as sound, pressure, temperature, humidity, etc. These sensor nodes are deployed in large numbers and they collaborate to form an ad-hoc network capable of reporting to sink (base station). Such networks aroused for various applications like habitat monitoring, building monitoring, health monitoring, military surveillance, target tracking, etc. However these networks are highly resource constraint with respect to energy, computation, memory and limited communication capabilities. All sensor nodes in the wireless sensor network interact with each other or communicate hop by hop through intermediate sensor nodes.

II. DATA AGGREGATION

In typical wireless sensor networks, sensor nodes are usually resource-constrained and battery-limited. In order to save the resources and energy, data must be aggregated to avoid redundancy and amount of traffic in the network. There are extensive works on data aggregation schemes in sensor networks where the aim of data aggregation is to eliminate redundant data collection transmission that reduces traffic and power consumption, in turn enhances the lifetime of wireless sensor network. Data aggregation process also includes collecting, processing of secured data from other sensor nodes and forwarding to either for base stations or other sensor nodes to reduce transmission [01].

Normally WSN interacts with external world through base station also called querier with a proper planning. Data transmission between sensor nodes, aggregators and the querier consumes lot of energy. Figure 2.1(a) shows a typical example for data aggregation model and figure 2.1(b) a non data aggregation model in which sensor nodes 1,2,3,4, 5 and 6 are regular nodes that collect data packet and send them to the next higher level nodes 7 and 8 normally called as aggregator nodes that perform sensing and aggregation simultaneously. In figure 2.1(a) three data packets collected from 1, 2 and 3 nodes are aggregated at node 7 and only one data packet is transmitted to the base station. In figure 2.1(b) non data aggregation model three data packets forwarded from 4, 5 and 6 at node 8 are sent directly to the sink. This means that with the help of data aggregation process the number of data transmission are reduced before transmitting to the base station or sink.

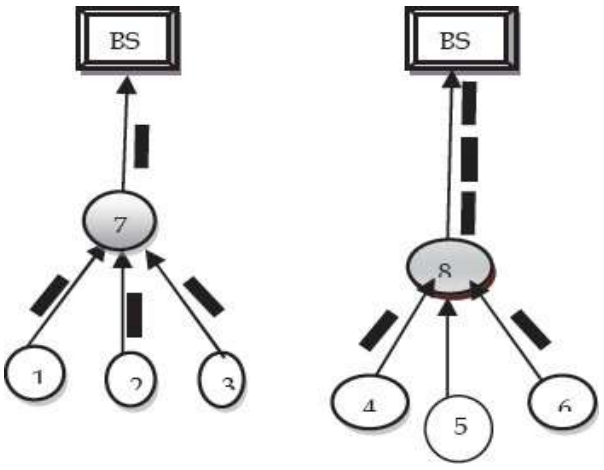

Figure 2.1 (a) Data aggregation model (b) Non data aggregation model

Usually, the wireless sensor network has three types of nodes. Simple regular sensor nodes, querier node and aggregator node. As mentioned in figure 2.1(a) aggregation model, regular sensor nodes sense data packet from the environment and send to the aggregator nodes. Querier nodes are those nodes that generate or issue the query without knowing how the results are generated, processed and returned by the sensor network. The aggregator nodes collects data from multiple sensor nodes of the network, aggregates the data packet using some aggregation functions, like sum, average, count, max, min and then forwards aggregated result to next aggregator node or the querier node that generated the query packet transmission and thus save the energy of the sensor node in the network. Hence, with the help of data aggregation we can enhance the lifetime of wireless sensor network and sink receives the data packets in an energy efficient manner with minimum data latency. Here the data latency is very important in many applications such as environment monitoring, health, monitoring, where the freshness of data is also an important factor.

The general data aggregation algorithm [02] works as shown in the below figure 2.2. The algorithm uses the sensor data from the sensor node and then aggregates the data by using some aggregation algorithms such as centralized approach, LEACH(low energy adaptive clustering hierarchy),TAG(Tiny Aggregation) etc. This aggregated data is transfer to the sink node by selecting the efficient path.

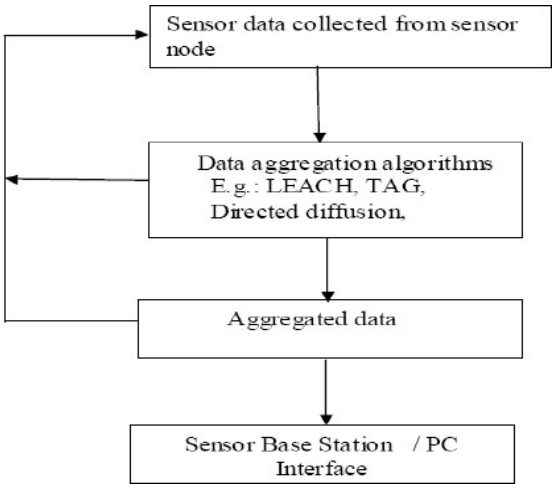

Fig 2.2 **General architecture of the data aggregation algorithm**

It is critical to develop energy-efficient data-aggregation algorithms so that network lifetime is enhanced. There are several factors which determine the energy efficiency of a sensor network, such as network architecture, the data aggregation mechanism, and the underlying routing protocol. Data aggregation is the technique that describes the processing method that is applied on the data received by a sensor node as well as how data is to be routed in the network by reducing energy consumption of the sensor nodes and also reduce the number of transmissions or length of the data packet.

Elena Fosolo et al in [03] describe "In-network aggregation is the exclusive process of collecting, routing information through a multi hop network and processing of data packet with the help of intermediate sensor nodes". The objective of this approach is to increase the life time of the network and also to reduce resource consumption. There are two types of approaches for In- network aggregation i.e., with size reduction and without size reduction. Firstly, the aggregation technique with size reduction is the process of combining and compressing the data received by a sensor node from its neighbors and is done in order to reduce the length of data packet to be sent towards the base station. For example, in some circumstance a node receives two data packets that are highly correlated in such scenarios it is inefficient to send both packets, instead the functions like MAX, AVG, and MIN can be applied and sent as a single data packet to base station. With the above approach the number of bits transmitted is reduced in the network and also saves a lot of energy. Secondly, In-network aggregation without size reduction involves processing of data packets received by different neighbors and forming into a single data packet but without processing the actual value of data. This process reduces the packet length with respect to multiple header information and consumes less transmission energy and in turn increases life time of the network. In total, data aggregation reduces redundancy of data, reduces the traffic and conserves energy of the sensor nodes. There are many other types of aggregation techniques are present some of them are listed below.

Centralized Approach: This is an address centric approach where each node sends data to a central node via the shortest possible route using a multi-hop wireless protocol. The sensor nodes simply send the data packets to a leader, which is the powerful node. The leader aggregates the data which can be queried. Each intermediate node has to send the data packets addressed to leader from the child nodes. So a large number of messages have to be transmitted for a query in the best case equal to the sum of external path lengths for each node.

Tree-Based Approach [04]: In the tree-based approach perform aggregation by constructing an aggregation tree, which could be a minimum spanning tree, rooted at sink and source nodes are considered as leaves. Each node has a parent node to forward its data. Flow of data starts from leaves nodes up to the

sink and therein the aggregation done by parent nodes.

Cluster-Based Approach[05]: In cluster-based approach, whole network is divided in to several clusters. Each cluster has a cluster-head which is selected among cluster members. Cluster heads do the role of aggregator which aggregate data received from cluster members locally and then transmit the result to sink.

III. CHALLENGES OF DATA AGGREGATION

The prominent challenges encountered at the time of data aggregation process in WSN are as follows:

Energy Consumption: In a resource constrained network environment, energy consumption determines the lifetime of the wireless sensor network and hence it becomes the major issue in wireless sensor network. Most of the sensor nodes use battery power as their energy source. The sensor network can be deployed in hazards conditions so it becomes difficult to change their batteries or provide the energy, here it is necessary to develop techniques that use the battery power efficiently. The energy consumption depends upon major operations of the sensor nodes that includes sensing, communication and data processing out of which communication requires more energy.

Data Gathering:. Data gathering is a task of systematic collection of sensed data from multiple sensor nodes to be eventually transmitted to the sink for processing without loss of information. Sensor nodes send their own sensed data packet and also forward packets forwarded by other sensor nodes and hence tends to consume more energy and form the energy holes near the sink.

Scheduling: It is also a important factor that determines the time period during which the sensor node will be in either sleep, active or in standby mode. Scheduling also plays important role for coverage and connectivity by knowing the status of the sensor nodes. By using proper scheduling we can conserve the energy.

Scalability: A WSN usually consists of thousands of sensor nodes, densely deployed in a region to be sensed and the number of sensor nodes in wireless sensor network depends on the requirements of the application. Scalability is defined as measure of a method's ability to maintain efficiency, as some network parameters increase to a large value, such as number of sensor nodes. Protocols must scale well with the number of nodes.

Reliability: - Reliability is defined as the probability that every pair of nodes can communicate with each other in a network. A node may fail due to various reasons. The failure of one node should not affect the overall performance of the wireless sensor network.

IV. PERFORMANCE MEASURES OF DATA AGGREGATION

The performance measures of data aggregation techniques are discussed below and these performance measures are highly dependent on the application requirements.

Energy Efficiency: It is defined as the ratio of the amount of energy consumed with respect to successfully delivered packet. By data-aggregation technique, we can increase the lifetime of the wireless sensor network. A data aggregation technique is energy efficient if it maximizes the life of the network. Network lifetime, data accuracy, and latency are some of the significant performance measures of data-aggregation algorithms. The definitions of these measures are highly dependent on the application requirements.

Network Lifetime: Network lifetime is measure with different yardstick based on application. The network lifetime is defined as the number of data aggregation rounds till the specified percentage of the total nodes dies and this percentage depend on the application. If we talk about some applications where in, simultaneous working of all the sensor nodes is a must, then the lifetime of the network is number of rounds until the first node dies .

Latency: Latency is evaluation of time delay experienced by the system, the length of the route that the packets have to take between sender (source node and sink).Basically delay is encountered in data transmission, routing and data aggregation.

Data accuracy: It is a ratio of total number of data packets received at the sink to the total number of packets generated at the source node.

To meet the above performance metrics, different types of data-aggregation protocols like network architecture based data-aggregation protocols, network-flow-based data-aggregation protocols and quality of service (QOS)-aware data-aggregation protocols are designed.

V. IMPACT OF DATA AGGREGATION IN WSN

In this section, we discuss the two main factors that affect the performance of data aggregation techniques in wireless sensor networks, such as energy saving and delay. With the help of this process we can save the energy in the network as discussed in the earlier section. Delay is the latency that is caused at the intermediate node for holding the data received from nearby sources in order to combine them with the data from sources that are farther away. Basically aggregation method is based on the position of the sources in the network, number of sources and the

network topology. By examining the position and number of sources the authors consider two models of the source placement. The event radius (ER) model and random source model [05]. The modeling of ER and random source model says that where the sources are clustered near to each other or located randomly, so that significant energy gains are possible with data aggregation. These gains are greatest when the number of sources is large, and when the sources are located relatively close to each other and far from base station. The modeling though, also seems to suggest that aggregation latency could be non negligible.

VI. CONCLUSION

In this work we have studied many important parts of data communication in sensor networks. From them the most important are query processing and data aggregation. We realized how communication in sensor networks is different from other wireless networks. Wireless sensor networks are energy constrained network. Since most of the energy consumed for transmitting and receiving data, the process of data aggregation becomes an important issue and optimization is needed. Efficient data aggregations not only provide energy conservation but also remove redundancy data and hence provide useful data only. We saw when the data from source node is send to sink through neighbors nodes in a multi-hop fashion by reducing transmission and receiving power, the energy consumption is low as compared to that of sending data directly to sink that is aggregation reduces the data transmission then the without aggregation.

VII. FUTURE SCOPE

Future work will focuses on developing of new routing algorithms to improve the processing time of critical data which is very important in real time systems, Node failure or unresponsiveness of node, Transmission of packets with network delay and Congestion of data in the aggregator etc.

VIII. REFERENCES

01 Gopal Pandurangan, V. S. Anil Kumar, Gu Ming, Yunhao Liu, Yingshu Li, Wireless Algorithms, Systems, and Applications, Springer, 30-Sep-2010 - Computers - 366 pages

02. Nadini S Patil And Prof P.R Patil, Data Aggregation in Wireless Sensor Network"2010 IEEE International Conference on Computational Intelligence and Computing Research

03 E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-Network Aggregation Techniques for Wireless Sensor Networks: A Survey", IEEE Wireless communication 2007.

04. M. Lee and V.W.S. Wong, "An Energy-aware Spanning Tree Algorithm for Data Aggreg ation in Wireless Sensor Networks," IEEE PacRrim 2005, Victoria, BC, Canada, Aug. 2005.

05. Ramesh Rajagopalan and Pramod K. Varshney, Syracure University "Data aggregation techniques in sensor network : A survey" "IEEE communication survey & tutorial.4th quarter 2006

06. K. Dasgupta, K. Kalpakis, and P. Namjoshi, "An Efficient Clustering-based Heuristic for Data Gathering and Aggregation in Sensor Networks", IEEE 2003.

07. S. Lindsey and C. Raghavendra, "PEGASIS: Powerefficient gathering in sensor information systems," inProceedings of IEEE AerospaceConference, vol. 3, Mar. 2002, pp. 1125–1130.

08. M. Lee, and S. Lee, "Data Dissemination for Wireless Sensor Networks", in Proceedings of the 10 th IEEE International Symposium on Object and Component- Oriented Real-Time Distributed Computing (ISORC'07)

09. A.Boulis, Saurabh Ganeriwal, Mani B.Srivastava,"Aggregation in Sensor Networks: An energy-accuracy trade-off", Elseiver, Ad Hoc Networks 1 (2003) 317-331.

10. Gunti Spandan, Archana Patel, C R Manjunath, Nagaraj GS, Data Aggregation Protocols in Wireless Sensor Networks, International Journal of Computational Engineering Research, Vol, 03, Issue.5, 2013

