International Journal of Information and Computation Technology.
ISSN 0974-2239 Volume 3, Number 3 (2013), pp. 167-174

http://www. irphouse.com /ijict.htm

Data Aggregation in Wireless Sensor Network: A Survey

Mousam Dagar[1] and Shilpa Mahajan[2]

[1, 2]Department of CSE/IT, ITM University, Gurgaon, India.

Abstract

Wireless sensor networks (WSNs) consist of sensor nodes. These networks have huge application in habitat monitoring, disaster management, security and military, etc. Wireless sensor nodes are very small in size and have limited processing capability very low battery power. This restriction of low battery power makes the sensor network prone to failure. Data aggregation is very crucial technique in wireless sensor networks. With the help of data aggregation we reduce the energy consumption by eliminating redundancy. In this paper we discuss about data aggregation and its various energy-efficient technique used for data aggregation in WSN.

Keywords: Data Aggregation, Wireless sensor network, security.

1. Introduction

Data aggregation is the process of collecting and aggregating the useful data. Data aggregation is considered as one of the fundamental processing procedures for saving the energy. In WSN data aggregation is an effective way to save the limited resources. The main goal of data aggregation algorithms is to gather and aggregate data in an energy efficient manner so that network lifetime is enhanced.. Wireless sensor networks have limited computational power and limited memory and battery power, this leads to increased complexity for application developers and often results in applications that are closely coupled with network protocols. In this paper, a data aggregation framework on wireless sensor networks is presented and a survey on various energy-efficient algorithm for data aggregation. The framework works as a middleware for aggregating data measured by number of nodes within a network**.**

2. Data Aggregation: An Overview

The data aggregation is a technique used to solve the implosion and overlap problems in data centric routing. Data coming from multiple sensor nodes are aggregated as if they are about the same attribute of the phenomenon when they reach the same routing node on the way back to the sink. Data aggregation is a widely used technique in wireless sensor networks. The security issues, data confidentiality and integrity, in data aggregation become vital when the sensor network is deployed in a hostile environment. Data aggregation is a process of aggregating the sensor data using aggregation approaches. The general data aggregation algorithm works as shown in the below figure1.

Figure 1 illustrates that data aggregation is the process of aggregating the sensor data using aggregation approaches. Then the algorithm uses the sensor data from the sensor nodes 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.

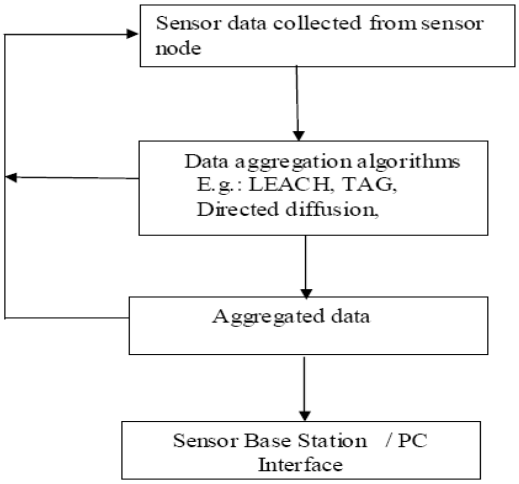

Figure 1: Architecture of data aggregation.

3. Data Aggregation Based Network

3.1 Flat Networks:

In flat networks, each sensor node plays the same role and is equipped with approximately the same battery power. In such networks, data aggregation is accomplished by data centric routing where the sink usually transmits a query message to the sensors, for example, via flooding and sensors which have data matching the query send response messages back to the sink. The choice of a particular communication protocol depends on the specific application at hand.

3.1.1 Diffusion:
Directed diffusion (DD) [2] may be a popular information aggregation paradigm for wireless device networks. it's a data-centric and application aware paradigm, within the sense that every one information generated by sensor nodes is called by attribute-value pairs. Such a scheme combines the information coming back from totally different sources en-route to the sink by eliminating redundancy and minimizing the amount of transmissions. during this means, it saves the energy consumption and will increase the network lifespan of WSNs. during this theme usually base station broadcast the message to the interested supply node. subsequently every node receives interest. These interests outline the attribute worth like name of object. every node get the interest will cache it for later use. because the interest is broadcasted by the network hop by hop, gradient square measure setups to draw information satisfying the query toward the requesting node. A gradient may be a reply link to the nearer from that the interest was received

3.2.2 SPIN:
The sensor protocol for data via negotiation[3] The staring node that has new data advertises the data to the close nodes within the network using the meta data. A close node that is interested in this type of information sends asking to the leader node for data. The leader node responds and send data to the sinks every node has a resource managing capability to keeps track of its energy usage within the sensing element network. every node polls its resources like battery power before data transmission. SPIN is also well-suited for environments with mobile sensors, since the forwarding decisions are based on native neighbourhood data.

3.2 Hierarichical Networks:
In the hierarchical network, In which data aggregation data has to be done at special nodes, with the help of these special node we can reduce the number of number of data packet transmitted to the sink. So with this network improves the energy efficiency of the whole network. Various type hierarchical data-aggregation protocols as follows

3.2.1 Cluster-Based Networks for data aggregation:
These Wireless sensor network is resource constraint that’s why sensor cannot directly transmit data to the base station. In which all regular sensors can send data packet to a cluster head (local aggregator) which aggregates data packet from all the regular sensors in its cluster and sends the concise digest to the base station. With the help of the scheme we save the energy of the sensors. LEACH [4]: Low energy adaptive clustering has been proposed to organise a sensor network into a set of clusters so that the energy consumption can be event distributed among all the sensor nodes.

3.2.2 Chain –Based Networks for Data Aggregation
In which each sensor sends data to the closer neighbour. Power- Efficient Data-Gathering Protocol for Sensor Information Systems (PEGASIS) is type of chain based data aggregation. In PEGASIS [5], all sensors are structured into a linear chain for data aggregation. The nodes can form a chain by employing a greedy algorithm or the sink

can decide the chain in a centralized manner. In the Greedy chain formation assumes that all sensors have inclusive knowledge of the network. The farthest node from the sink initiates chain formation and, at each step, the closest neighbour of a node is selected as its successor in the chain. In each data-gathering round, a node receives data packet from one of its neighbours, aggregates the data with its own, and sends the aggregates data packet to its other neighbour along the chain. Eventually, the leader node in the are similar to cluster head sends the aggregated data to the base station. Figure below shows the chain based data-aggregation procedure in PEGASIS.

3.2.3 Tree Based Networks for Data Aggregation

In which all node are organized in form of tree means hierarchical, with then help of intermediate node we can perform data aggregation process and data transmit leaf node root node. Tree based data aggregation is suitable for applications which involve in-network data aggregation. An example application is radiation-level monitoring in a nuclear plant where the maximum value provides the most useful information for the safety of the plant. One of the main aspects of tree-based networks is the construction of an energy efficient data-aggregation tree.

4. Working Principle of Data Aggregation

The working of WSN proposed architecture model illustrated in Figure 2 below that starts working by choosing selecting of nodes and divided into clusters. These clusters can satisfy the intended parameter requirements and conditions. The parameters like RSSI, TTL, MRIC, bandwidth, battery consumption are accustomed verify the amount of nodes that will be considered in a cluster. thereafter a cluster head (CH) is selected among nodes lies within the each cluster. CH are going to be responsible for administration of all different nodes inside several cluster and collecting the data} from the nodes within the cluster and transferring the information to the neighbouring cluster head for more information exchange and updation . The newly arrived nodes will be assigned as cluster head if the global cost of arrived node is minimum , otherwise other cluster nodes are going to be given opportunity to participate and global cost is once more recalculated. thereafter the data aggregation approach is presumed as the collection of data and numerous queries from the user end are checked and transformed into low level schemes by a query processor. All data collected and aggregated is stored at a storage location in database server. Finally at last the data is aggregated by data cube approach[6] and every one the aggregated data are going to be transfer to the base station for further use.

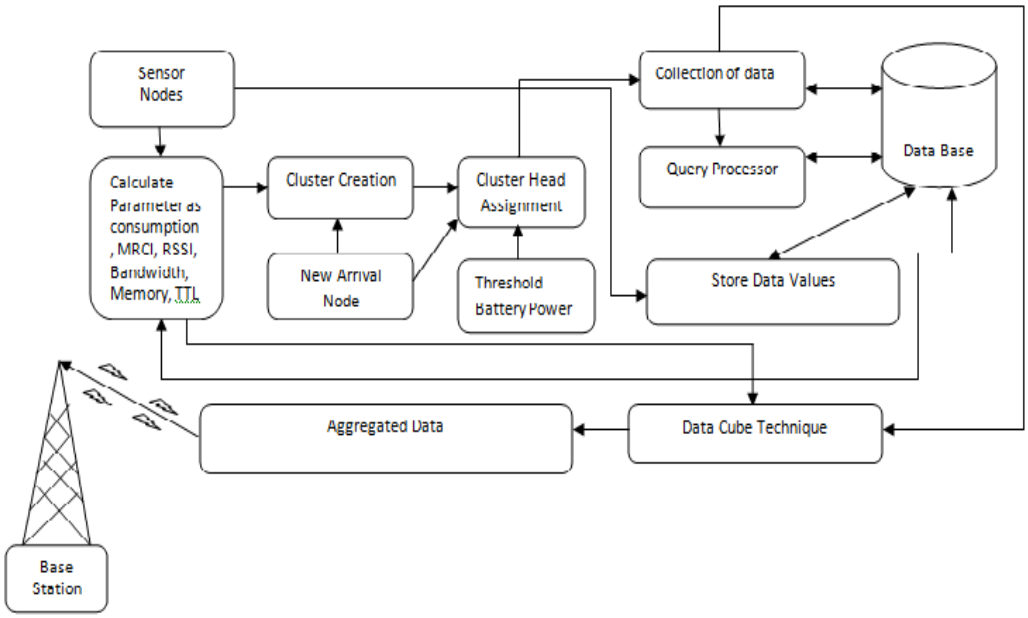

Figure 2: Architecture of Data Collection and aggregation for WSN.

5. Different Energy-efficient Techniques in Data Aggregation

5.1. Grid-Based Architecture

It is energy–efficient data storage scheme[7] in which Snake-like Energy Efficient Scheduling is given in the network is divided into 2 dimentional logical grids where the number of sensors in a grid is N. This works on Active and sleep mode procedure i.e if one sensor is active at one time slot then the other is in sleep mode at that time of slot.

	E	F	G
A	1	2	3
B	6	5	4
C	7	8	9
D	12	11	10

Figure 3: The time slots are assigned in a snake-like direction into T with 4*3.

When the active sensor changes from one row to another row in figure 4, i.e., from row A to row B, there exists one column sensor in active mode with two sequential time slots, i.e., column G. Hence, when two active sensors receive the query packet, exactly one of two sensors can continue to stay in active mode in the next time slot. This mechanism can guarantee that the query packet is preserved in one node of the grid to continue performing the query task.

5.2. Temporal Correlation Based Data Aggregation Scheme

In this scheme the author uses the ARIMA model[8] also called Box-Jenkins model is a widely used forecast model for univariate time series. Data aggregation in this scheme the ordinary sensor node collects sensed value from environment. If the periodical update time is up, it will save the sensed value into the historical data queue and send the sensed value to the aggregator. Otherwise, it will calculate the forecast value using ARIMA model and compare the sensed value with the forecast value. If the difference between them is less than the predefined error threshold, the sensor will store the forecast value into the historical data queue. Otherwise, it will store the sensed value into the historical data queue and send the sensed value to the aggregator at the same time. The periodical update time is a preset and tunable scheme parameter which is used to periodically collect real sensed value and avoid cumulative error in continuous forecasts. The aggregator listens on the wireless channel to retrieve sensed values from ordinary sensor node and store them into the historical data queue. If the aggregator does not receive any data from sensor node after a predefined periodical data collect time, it means the difference between the sensed value and forecast value is within an acceptable range. Then the aggregator will calculate the forecast value using ARIMA model with historical data. The periodical data collect time should be selected carefully to ensure it is enough to deliver the message from sensor to the aggregator.

5.3 Steiner Tree

In this technique the author[9] describe the wireless sensor network communication model as an undirected graph G <V, E>, where the node set V contains all nodes that have been aware within the region, the distance between nodes in the graph are the weights E associate with the vertices, the communication distance of the node is R, introduces a Steiner tree a weighted undirected graph G is given, and the demand Steiner tree T, and find the shortest path from the root of T to other nodes.

5.4 Polynomial Regression Based Secure Data Aggregation

PRDA[10] in this the author suggests that it is a protocol in which sensor nodes represents their sensed data as polynomial functions. This paper proposes a Polynomial Regression based secure Data aggregation protocol, called PRDA, to preserve the privacy of the data being aggregated. PRDA is an additive data aggregation protocol and achieves data privacy by employing polynomial regression on sensor data series. The novel idea behind PRDA protocol is to perform data aggregation using polynomial coefficients that represent sensor data.

5.5 Clustering Based Lifetime Maximizing Aggregation Tree

CLMAT[11] is in which we create aggregation tree which aim to reduce energy consumption, minimizing the distance traversed and minimizing the cost in terms of energy consumption. In CLMAT the node having maximum available energy is used as parent node/ aggregator node. We concluded with the best possible aggregation tree minimizing energy utilization, minimizing cost and hence maximizing network

lifetime. Hence by achieving the above mentioned parameters the obtained aggregation tree proves to be the best for enhancing the network lifetime.

6. Security Issues in Data Aggregation

Data aggregation in Wireless sensor Network refers to exploit the sensed data from the sensors to the gateway node. data aggregation plays a significant role in Wireless sensor Networks since the aggregation schemes followed here involve in reducing the amount of power consumed throughout data transmission between the sensor nodes. within the data aggregation of WSN, 2 security requirements, confidentiality and integrity, ought to be consummated. Specifically, the fundamental security issue is data confidentiality, that protects the sensitive transmitted data from passive attacks, such as eavesdropping. data confidentiality is especially very important in a hostile environment, where the wireless channel is at risk of eavesdropping. though there are many methods provided by cryptography, the difficult encryption and decryption operations, like modular multiplications of large numbers in public key primarily based cryptosystems, will assign the sensor's power quickly [8]. the other security issue is data integrity, that prevents the compromised source nodes or aggregator nodes from considerably altering the final aggregation value [9]. sensor nodes are easy to be compromised because they lack expensive tampering-resistant hardware, and even that tampering-resistant hardware may not continually be reliable. A compromised node will modify, forge or discard messages.

7. Conclusion

In this paper we have studied the data communication in sensor networks i.e data aggregation and 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. When the data from source node is send to sink through neighbours nodes in a multihop 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. In this paper, we have proposed an energy-efficient techniques for data aggregation in wireless sensor networks. Our scheme integrates energy-efficient and data storage mechanisms. This survey paper shows that these techniques not only reduces power consumption but also prolongs the lifetime of a network.

8. Future Scope

The Future work will focuses on developing a new routing algorithms for routing the data from the source to the sink. Our approach should confront with the difficulties of topology construction, data routing, loss tolerance by including several optimization

techniques that further decrease message costs and improve tolerance to failure and loss. Later, we will simulating our developed technique and compare it with some protocols to prove its efficiency.

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] Kiran Maraiya, Kamal Kant, Nitin Gupta "Architectural Based Data Aggregation Techniques in Wireless Sensor Network: A Comparative Study", International Journal on Computer Science and Engineering (IJCSE), Vol. 3 No. 3 Mar 2011

[3] Vaibhav Pandey, Amarjeet Kaur and Narottam Chand "A review on data aggregation techniques in wireless sensor network", Journal of Electronic and Electrical Engineering, ISSN: 0976–8106 & E-ISSN: 0976–8114, Vol. 1, Issue 2, 2010

[4] Wendi Rabiner Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan "Energy-Efficient Communication Protocol for Wireless Microsensor Networks", Proceedings of the 33rd Hawaii International Conference on System Sciences IEEE, 2000.

[5] Stephanie Lmdsey and Cauligi S. Raghavendra "PEGASIS: Power-Efficient GAthering in Sensor Information Systems", IEEE 2002.

[6] Sumit Chaudhary, Neha Singh, Avinav Pathak and A.K Vatsa, "Energy Efficient Techniques for Data aggregation and collection in WSN" , International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol.2, No.4,2012

[7] Wen- Hwa Liao , Hung-Chun Yang , "An Energy- Efficient Data Storage Scheme in Wireless Sensor Networks", Network Operations And Management Symposium(NOMS), IEEE 2012

[8] Guorui Li, Ying Wang "efficient Data aggregation scheme leveraging time series prediction in WSN" International Journal of Machine learning and Computing Vol.1. No.4. October 2011.

[9] Shen Xueli , Wu Wenjum ,"The Research Of Data Aggregation In Wireless Sensor Networks", International Forum Of Information And Technology, IEEE Computer Society, 2010.

[10] Suat Ozdemir, Yang Xiao,"Polynimial Regression Based Secure Data Aggregation For Wireless Sensor Network", IEEE Communication Society, 2011

[11] Jingsha He, Mingxin Yang , Xuguang Sun, "Research On Data Aggregation in Wireless Sensor Networks Based on Clustering Method", International Technology and Applications,IEEE Conference , 2011.