2010 IEEE International Conference on Computational Intelligence and Computing Research

Data Aggregation in Wireless Sensor Network

Nandini. S. Patil, Prof. P. R. Patil

B.V. Bhoomaraddi College of Engineering and Technology, Hubli-580031, India,

Visvesvaraiya Technological University Belgum-590014, India.

Nandinipatil5@gmail.com

Prakshapatil@bvb.edu

Abstract— **Sensor networks are collection of sensor nodes which co-operatively send sensed data to base station. As sensor nodes are battery driven, an efficient utilization of power is essential in order to use networks for long duration hence it is needed to reduce data traffic inside sensor networks, reduce amount of data that need to send to base station. 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 (WSN) offer an increasingly Sensor nodes need less power for processing as compared to transmitting data. It is preferable to do in network processing inside network and reduce packet size. One such approach is data aggregation which attractive method of data gathering in distributed system architectures and dynamic access via wireless connectivity. 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. The framework works as a middleware for aggregating data measured by a number of nodes within a network.**

The aim of the proposed work is to compare the performance of TAG in terms of energy efficiency in comparison with and without data aggregation in wireless sensor networks and to assess the suitability of the protocol in an environment where resources are limited.

I. INTRODUCTION

With advance in technology, sensor networks composed of small and cost effective sensing devices equipped with wireless radio transceiver for environment monitoring have become feasible. The key advantage of using these small devices to monitor the environment is that it does not require infrastructure such as electric mains for power supply and wired lines for Internet connections to collect data, nor need human interaction while deploying. These sensor nodes can monitor the environment by collecting information from their surroundings, and work cooperatively to send the data to a base station, or sink, for analysis.

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 (WSN) offer an increasingly attractive method of data gathering in distributed system architectures and dynamic access via wireless connectivity.

Clustering in WSN[6]: The process of grouping the sensor nodes in a densely deployed large-scale sensor network is known as clustering. The intelligent way to combine and compress the data belonging to a single cluster is known as data aggregation in cluster based environment. There are some issues involved with the process of clustering in a wireless sensor network. First issue is, how many clusters should be formed that could optimize some performance parameter. Second could be how many nodes should be taken in to a single cluster. Third important issue is the selection procedure of cluster-head in a cluster. Another issue is that user can put some more powerful nodes, in terms of energy, in the network which can act as a cluster-head and other simple node work as cluster-member only.

ISBN: 97881 8371 362 7

II. PROBLEM DEFINITION

Data aggregation protocols aims at eliminating redundant data transmission and thus improve the lifetime of energy constrained wireless sensor network. In wireless sensor network, data transmission took place in multi-hop fashion where each node forwards its data to the neighbor node which is nearer to sink. Since closely placed nodes may sense same data, above approach cannot be considered as energy efficient. An improvement over the above approach would be clustering where each node sends data to cluster-head (CH) and then cluster-head perform aggregation on the received raw data and then send it to sink. Performing aggregation function over cluster-head still causes significant energy wastage. In case of homogeneous sensor network cluster-head will soon die out and again re-clustering has to be done which again cause energy consumption.

III. DATA AGGREGATION: AN OVERVIEW

Data aggregation is a process of aggregating the sensor data using aggregation approaches. The general data aggregation algorithm works as shown in the below figure. 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.

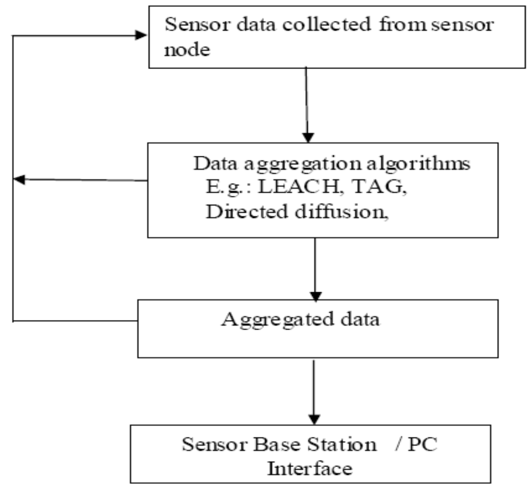


Fig 3.1: General architecture of the data aggregation algorithm

There are many 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 multihop 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.

In-Network Aggregation[7]: In-network aggregation is the global process of gathering and routing information through a multi-hop network, processing data at intermediate nodes with the objective of reducing resource consumption (in particular energy), thereby increasing network lifetime. There are two approaches for in-network aggregation: with size reduction and without size reduction. In-network aggregation with size reduction refers to the process of combining & compressing the data packets received by a node from its neighbors in order to reduce the packet length to be transmitted or forwarded towards sink. In-network aggregation without size reduction refers to the process merging data packets received from different neighbors in to a single data packet but without processing the value of data.

Tree-Based Approach[8]: 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[6]: 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.

IV. QUERY PROCESSING

1.Query Models

COUGAR approach [10] proposes a query layer to support aggregate queries. With the interface provided, the clients can issue queries without knowing how the results are generated, processed and returned by the sensor network to them. The query layer processes declarative queries and generate a cost effective query plan. They follow a database approach to design a query interface for sensor networks. The view of cost is different for sensor networks. The major factor under consideration is the communication cost, involving the cost of routing the queries and aggregating data over the sensor networks. TAG also proposes a query model for supporting aggregate queries.

TAG and COUGAR are tightly coupled with the underlying aggregation schemes. [11] Proposes a Query Agent that provides application independent query interface and an API support to map the user specified queries to lower level semantics corresponding to underlying routing and aggregating protocols. It supports different communication models - anycast, unicast, multicast and broadcast. Query agent will support a wide variety of routing and aggregation protocols selecting the best combination based on the type of the query.

2. Query Language in TinyDB

TinyDB's query language is based on SQL, and we will refer to it as TinySQL. Query Language in TinySQL supports selection, projection, determining sampling rate, group aggregation, user defined aggregation, event trigger, lifetime query, setting storing point and simple join [13].

3. Queries and Aggregates

The probable queries for the sensor networks can be categorized into:

1) Simple Queries

These are non aggregate queries.

Eg. "SELECT temperature FROM sensor WHERE node = z".

These are generally mapped into broadcast or point to point queries.

2) Complex Queries

They may contain sub queries.

Eg. "SELECT temperature FROM sensor WHERE room = (SELECT room WHERE floor = '3')"

3) Event Driven Queries

These are the continuous query that returns the values periodically at specified time intervals.

Eg: "SELECT AVG (temperature) FROM sensor where node = z"

The Grammar of TinySQL query language is as follows:

SELECT select-list
[FROM sensors]
WHERE predicate 294
[GROUP BY gb-list]
[TRIGGER ACTION command-name[(param)]]
[EPOCH DURATION time]

Where, select−list is the attribute list of the unlimited virtual relational table, which can include an aggregation function.

Predicate is the query condition. gb−list is an attributes list. command−name is a trigger operation. Param is the parameters of trigger. Time is the value of time. TRIGGERACTION is the subordinate clause which defines the trigger. It determines the operations executed when WHERE clause is satisfied. EPOCH DURATION defines the query cycle. The meaning of the other clauses is the same as SQL. Following is an example of a TinyDB query.

SELECT nodeid, AVG(light), AVG(temp)
FROM sensors
WHERE AVG(light)=100
GROUP BY nodeid
EPOCH DURATION 5min

The meaning of the query is detecting nodeid per five minutes in which the average light is equal to 100 and returning the nodeid and its average light and temperature. Currently, the functions of TinyDB are very limited. Some functions supported by SQL are not supported by TinyDB.

V. SIMULATION AND EXPERIMENTAL ANALYSIS

Simulation Tools: We have plenty of simulation tools or simulators for simulating wireless networks. The simulators which are most popular are TOSSIM, NS-2, OPNET, OMNet++, J-Sim, GlomoSim, and Qualnet and so on. TOSSIM is a discrete event simulator for TinyOS (TinyOS is a popular sensor network operating system) sensor networks. Instead of compiling a TinyOS application for a mote, users can compile it into the TOSSIM [20] framework, which runs on a PC. This allows users to debug, test, and analyze algorithms in a controlled and repeatable environment. As TOSSIM runs on a PC, users can examine their TinyOS code using debuggers and other development tools. TOSSIM's primary goal is to provide a high fidelity simulation of TinyOS applications. For this reason, it focuses on simulating TinyOS and its execution, rather than simulating the real world.

Simulation run

This simulation is run for the following with aggregation and clustering Query-1.

QUERY-1: SELECT AVG (light) FROM SENSORS GROUP BY NODEID % 2 SAMPLE PERIOD 2048

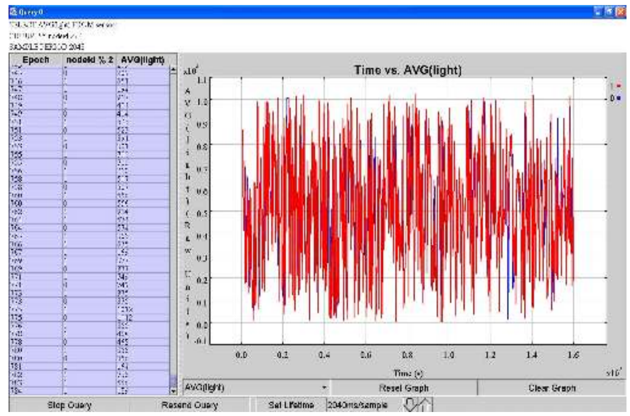


Fig:Result window for with aggregation and clustering

QUERY-2: SELECT MAX (temp), AVG (light) FROM SENSORS SAMPLE PERIOD 2048

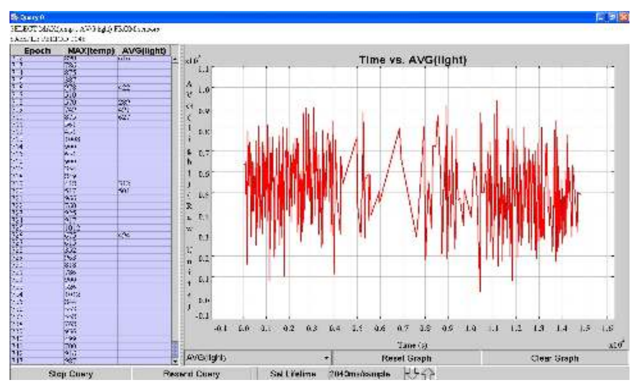


Fig:Result window for with aggregation and without clustering

QUERY-3: SELECT temp, light FROM SENSORS SAMPLE PERIOD 2048

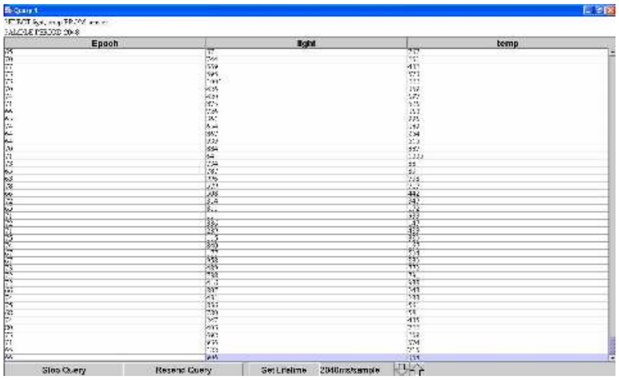


Fig:Result window for with out aggregation and clustering

Simulation results and comparison

With aggregation query

- SELECT MAX (temp), AVG (light) FROM SENSORS SAMPLE PERIOD 2048

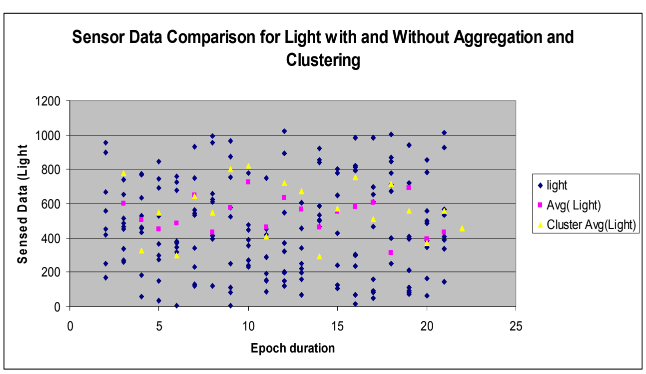
Without aggregation query

- SELECT light FROM sensors SAMPLE PERIOD 2048

With aggregation and with clustering query

- SELECT AVG(light) FROM SENSORS GROUP BY NODEID % 2 SAMPLE PERIOD 2048.

	Without Aggregation		With Aggregation		Cluster based Aggregation	
No of nodes	10	20	10	20	10	20
No of messages transmitted	895	687	87	72	266	122



VI. CONCLUSION

In this work we have studied the two most important parts of data communication in sensor networks- query processing, 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.

The simulation result shows that 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. We have showed how aggregate queries are efficiently executed in wireless sensor networks.

VII. FUTURE SCOPE

Future work will focuses on the using new different 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. In addition to implementing these techniques, we need to rethink some of these techniques to present more efficiency to network changes and external factors which could affect our approach such as node mobility, obstacles and other issues. In addition as future work, we could also extend our simulator to incorporate a 3D tree construction technique.

REFERENCES

1. S. Lindsey and C. Raghavendra, “PEGASIS: Power-efficient gathering in sensor information systems,” in

Proceedings of IEEE AerospaceConference, vol. 3, Mar. 2002, pp. 1125–1130.

2. 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).*
3. C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed Diffusion for Wireless Sensor Networking", *IEEE/ACM Transactions on Networking*, Vol. 11, no. 1, Feb 2003.
4. H. Cam, S. Ozdemir, P. Nair, and D.Muthuavina shiappan, "ESPDA: Energy-Efficient and Secure Pattern-based Data Aggregation for Wireless Sensor Networks", in *Proceedings of IEEE Sensor- The Second IEEE Conference on Sensors, Toronto, Canada*, Oct. 22-24, 2003, pp. 732-736.
5. Chalermek Intanagonwiwat, Ramesh Govindan, and Deborah Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks", (MobiCom 2000) pp 56-67.
6. K. Dasgupta, K. Kalpakis, and P. Namjoshi, "An Efficient Clustering-based Heuristic for Data Gathering and Aggregation in Sensor Networks", IEEE 2003.
7. E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-Network Aggregation Techniques for Wireless Sensor Networks: A Survey", IEEE Wireless communication 2007.
8. M. Lee and V.W.S. Wong, "An Energy-aware Spanning Tree Algorithm for Data Aggregation in Wireless Sensor Networks," *IEEE PacRrim 2005*, Victoria, BC, Canada, Aug. 2005.
9. M. Ding, X. Cheng, and G. Xue, "Aggregation tree construction in sensor networks," in *Proc. of IEEE VTC'03*, Vol. 4, Orlando, FL, Oct. 2003.
10. "The Design Space of Wireless Sensor Networks" by Kay R¨omer and Friedemann Mattern http://www.vs.inf.ethz.ch/publ/papers/wsn-designspace.pdf
11. http://en.wikipedia.org/wiki/Wireless_Sensor_Network
12. Wendi B. Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan, "Energy Efficient Communication Protocol for Wireless Microsensor Networks", (33rd Hawaii International Conference on System Sciences, 2000).
13. S. Nath *et al.*, "Synopsis Diffusion for Robust Aggregation in Sensor Networks," *ACM SenSys 2004*, Baltimore, MD, Nov. 2004.