



PERFECT STRATAGEM FOR QUERY PROCESSING IN WSN

Vandana Jindal

PG Dept of Computer Science, D.A.V. College
Bathinda, Punjab, India

Abstract: In a WSN, huge amount of data is generated by sensor nodes continuously. Data generated by nodes make a large distributed database. Information requirements for various applications can only be satisfied if the data available in the network is extracted to a suitable location and processed efficiently. However, query based systems are popular in database type environment. Wireless Sensor Networks find usage in a wide range of applications. It is possible due to various mechanisms such as adhoc routing protocols, in network data aggregation and signal processing etc. Performance of the network can be improved by tweaking these features. In order to evaluate performance of a network objectively a number of metrics have been defined. Some important ones are: lifetime, coverage, cost and ease of deployment, response time, temporal accuracy, security and effective sample rate. In this paper various metrics for WSN evaluation have been discussed. Many of these metrics are interrelated. Generally, performance in one metric may have to be decreased to achieve increase in some other metric.

Keywords: adhoc, Wireless Sensor Network (WSN), Base Station (BS), Multiple Query Optimization (MQO).

I. INTRODUCTION

Wireless Sensor Network (WSN) is a network composed of small devices called 'nodes' which are haphazardly distributed over some area of scrutiny. These nodes are 'sensors' which behave like a computer as they are capable of storing information gathered from the area under supervision, evaluate the data gathered and forward the same. The nodes in the network collect the data from the surroundings and forward the same through other nodes in the network, ultimately leading to the main location i.e. the base station. WSN is usually deployed in those areas where manual supervision is not possible like: battlefields, environmental observation, system's performance and industrial machine monitoring etc. Due to unfavorable environmental conditions these nodes are unsafe and thus become incapable of performing their destined tasks. As they are affordable they are easy to replace within the WSN and help in performing the required tasks with ease. Although the nodes are capable of aligning themselves within the network but the frequency with which they collapse influence the network terribly. Besides the nodes becoming useless due to the unfriendly environmental conditions, the depletion of restricted battery contributes to making the nodes purposeless. As the existence of the WSN is mainly due to the presence of the nodes within the network, the enhanced life of the nodes becomes desirous. Environmental conditions cannot be changed as these are to be monitored by the WSN, so node failure due to battery depletion is to be minimized.

Acquisition of new technologies has enhanced the working of the sensors. The researchers are not only bent upon bringing the refinement in the hardware but also at the programming level as well to furnish better results. The sensors deployed in the network for various applications accumulate the data and pass it on to the sink where it gets stored for further analysis. Various strategies such as application specific MAC and routing protocols, In-network processing, data aggregation techniques etc. have been employed to minimize communication. Query is the most widely accepted and well suited procedure of withdrawing the data. Writing queries in an optimal way at the Base Station (BS) is another approach to

minimize communication load. Modern WSN are so flexible that they are able to provide varied information simultaneously to multiple users having diversified data interests.

II. QUERY PROCESSING

In a WSN, huge amount of data is generated by sensor nodes continuously. Data generated by nodes make a large distributed database. Information requirements for various applications can only be satisfied if the data available in the network is extracted to a suitable location and processed efficiently. However, query based systems are popular in database type environment. For extraction of information from the data, queries are injected as in other databases. A good DBMS is a system which optimizes on resource consumption in response to a query. Query processing is an elementary part of any DBMS. The means by which we can obtain the best plan, used in implementing the data extraction request from database is "Query Processing".

Huge repository of data is available in the form of various databases. To make this data into information, it is essential that the user be able to extract precise and relevant data with minimum consumption of invaluable resources- time and energy in response to the query injected by the user. Query processing is an elementary part of any DBMS. Query is a tool to gather information. The word derives from the Latin "quaere" i.e., to ask or seek.

Choosing parameters from a menu is a method where the database puts forth the parameters in front of us to choose from amongst the given lot.

There is a way of telling a database exactly what you want it to do by speaking in a way that it is built from the ground up to understand. When it comes to working within existing data, whether it is to add, move, or delete data, Data Manipulation Language (DML) offers an effective way of doing it. DML allows a user to interact with the massive data. A computer programming language that requests and retrieves data from database and information systems by sending queries is the subcategory of the DML referred to as 'Query Language'. The

method employed for achieving the goal through a database request is known as “Query Processing”.

A. Query Types

Various query types available are depicted below:

- *Select queries:* These are used to retrieve data from one or more tables and display the result.
- *Parameter queries:* For creating on-the-fly queries which prompt the user for criteria at the time the query is run.
- *Crosstab queries:* These are used to summarize data from one field and group it in tabular form according to two criteria.
- *Action queries:* An **action query** is a **query** that makes changes to or moves many records in just one operation. Action queries may be subdivided into the following categories:
 - i) Delete queries – Employed for record removal from the tables.
 - ii) Update queries – Queries that wholly change a section of records in a table.
 - iii) Append queries – An Append query is often referred to as an Insert Query that adds records to a table.
 - iv) Make table queries – You can copy rows into a new **table** using a **Make Table query**, which is useful for creating subsets of data to work with or copying the contents of a **table** from one database to another. A **Make Table query** is similar to an Insert Results **query** but creates a new **table** to copy rows into.
- *SQL queries:* A query created using SQL, which is highly advanced query language.
- *Contextual queries:* The subclasses of contextual queries are:
 - i) *Search context* – It is a form of optimizing web-based search results based on the context supplied by the user. E.g. web indexes, bibliographic catalogs etc. In case of searching, the user may provide a single or group of clauses joined by the Boolean operator to the search engine.
 - ii) *Taxonomy* – It is concerned with the classification. A taxonomy γ is a tree depicting classes where each junction or vertex is a predestined class. Each class is marked by the tags along the line from the root to the corresponding node or the vertex.
 - iii) *Level-N category, ancestor category and sibling category:* This is for a category c in taxonomy γ , c is called a level- n category if the node at c is located at n th level of γ .

B. Approach for Query Processing

Activities involved in extracting data from the database are referred to as query processing. A two step approach involved in processing a query are: First, the query which is represented in the Structured Query Language (SQL) form is translated

into low level language form with the help of relational algebra [1] and the second step involves query execution.

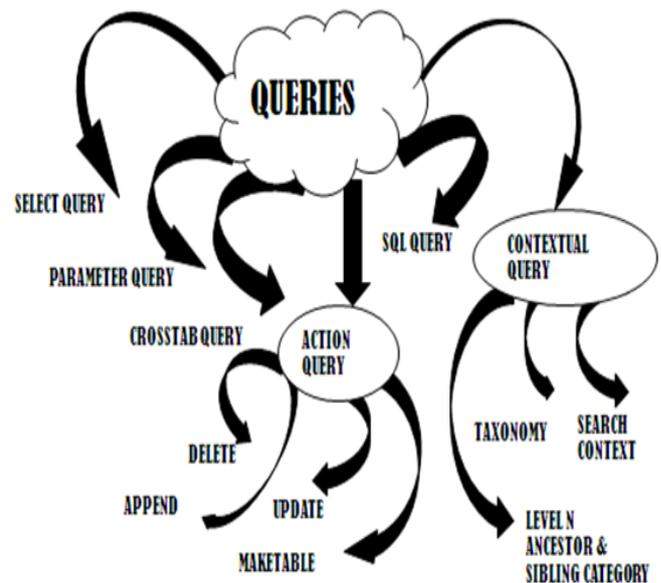


Figure 1: Query taxonomy

C. Relationship between Query processing and WSN

Recent studies have shown a great attention being focused on mobile computing, with intent of accessing data at any place, at any time. The devices so mentioned are not only capable of transmitting voice but are able to transmit the data as well. The main purpose of WSN is to collect data and aggregate the data. The presence of full-fledged computing hardware in the sensor node capable of performing multiple sensing functions like the temperature, light, humidity, etc., these sensor networks (which are composed of thousands of such nodes) may be thought of as a ‘distributed database’.

This database accumulates physical measurements, inside the memory available on the board about the environment and aids in providing answers to the queries injected by the users. Each sensor node produces data packets that are acquired by the sensing devices present in the node.

The nodes which are composed of micro processor, transceiver, battery, memory, sensing device and ADC enable the node to sense, compute and communicate the data. The nodes in the WSN generate the data which may be extracted by its name, which depicts similarity to the traditional view of relational databases.

D. Query Processing in WSN

Wherever databases exist, query processing takes place. i.e. data of interest is extracted from the given pool of data. However, the methods employed for data extraction from the conventional databases is altogether different from the ones employed in WSN because of the way the queries are represented in WSN, the goal of getting the data of interest and the restrictions of the nodes or sensors.

For convenience of users, WSN is viewed as a database [2-4]. This assumption makes it more comfortable for the usage along with the ease of modifiability and scalability resulting into better performance. All these factors enable WSN to

execute various tasks at the same time. Considering WSN to be a database approach, query processing in WSNs is executed as follows:

Initially, a user injects the query of his interest to the BS. The BS generates the query plan and communicates the query to the free nodes in the network. On receiving the query the nodes collect the required data (of the parameters specified) from the environment followed by data processing. Then the data is transmitted to the BS jointly with the other free nodes. On receiving the data final processing takes place at the base-station and finally the result is submitted to the user's query.

II. MULTIPLE QUERY OPTIMIZATIONS (MQO) IN WSN

Queries in WSN are continuous and we consider data as stream. Technically, an appropriate term given to it is "Window" i.e. data coming during a specific period of time as a unit. The query optimization methods may be categorized as: *Base station optimization* and *In-Network optimization*.

A. Base Station Optimization

This method helps in reduction in energy consumption in the network by reducing the communication messages as the number of injected queries in the network is reduced at the base station. This is made possible by drawing the conclusion if the latest query presented to the base station may be rephrased using the previously running queries. Attempts are made to run this freshly rephrased query at the base station employing the results of the already running queries.

B. In-Network Optimization

The sensor nodes collect the data for all the queries that have been injected into the WSNs. It is processed for deletion of redundancy etc. Resultant data is broadcasted thereby minimizing the number and quantity of radio messages thus making it bandwidth and energy efficient.

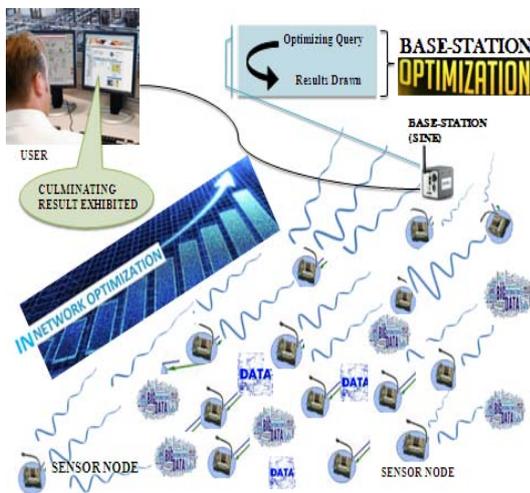


Figure 2: Stratagem for Query optimization

Sensor nodes being resource constrained devices, it is required that the queries be processed in an efficient way [5]. Therefore, it is necessary that all the nodes should be at the disposal for supplying continuous data streams of the data gathered to satisfy the query in hand. Various new proposals have come up regarding Two-Tier Multiple Query Optimization (TTMQO) leading to decreased number of radio transmissions and lowering the transmission time in the sensor network.

III. CONCLUSION

Generally, performance in one metric may have to be decreased to achieve increase in some other metric. All these metrics form a multidimensional space where every possible application is represented by a single point. The single point represents value of all these parameters which may be high in one metric and low in other as required by that specific application. A network will be able to serve the application if that point lies in the multidimensional space of performance metric. In other words a tradeoff is to be established.

IV. REFERENCES

- [1] David Gascón, Marcos Yarza, "Wireless Sensor Networks to Control Radiation Levels", Published in: Articles, Environment, Makers, Smart Metering April 19th, 2011.
- [2] Alberto Bielsa, Mark Boyd, "Wireless Sensor Networks to monitor food sustainability", Published in: Articles, Environment, February 28th, 2012.
- [3] E. Jovanov, et al., "A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation," Journal of Neuro Engineering and Rehabilitation, 2005, 2:6.
- [4] Martha Baer, "The Ultimate on-the-fly Network", On Newsstands Now, Issue 11.12 | December 2003.
- [5] <http://rp-www.cs.usyd.edu.au/~wsn/sunspots.shtml>