Querying Sensor Networks: Techniques, Evaluation, and New Directions

Sayed Ahmed, M. Rasit Eskicioglu, and Sajid Hussain
Department of Computer Science, University of Manitoba, Winnipeg, MB, R3T 2N2, Canada
{sayed, rasit, sajid}@cs.umanitoba.ca

***Abstract*— Advances in sensor nodes hardware, which comprises sensors, embedded processors, and communication components, have made the large-scale deployment of sensor networks a reality. Various sensor network applications ranging from monitoring to military require sensor nodes to collect data over a continuous time period. The placement, management, and processing of the sensor data necessitates an effective data storage, management and query processing policy. This poster attempts to identify the key query processing techniques in the sensor networks. Design goals and challenges of the query processing techniques are identified. The techniques are evaluated in terms of efficiency, scalability, applicability, and reliability. The evaluation of the techniques is guided by the distinctive query processing features supported by both types of sensor networks, conventional and wide area. Moreover, we argued for the integration of conventional and wide area sensor networks and addressed the integration issues and design goals. In particular, a query processing architecture is proposed to meet the emerging needs of the sensor networks. The architecture addressed the requirements for different layers of the integrated sensor network components such as base stations, sensor nodes, and the workstations in wide area. Additionally, the future research directions for the query processing are outlined.**

Proposed Architecture: To reduce communication and power consumption data aggregation and filtering capability in each sensor node and base station. At base stations proxy components to map between XPATH and SQL, or a traditional query into a continuous/aggregate query for the inter-operability of both parts. A query proxy layer at each sensor node to help the query plan and query optimization generated at the base stations. A query parser/optimizer at base stations and workstations in wide area network that will perform semantic analysis, normalization, optimization of the queries posed at the base stations. In addition, the workstations in wide area sensor networks will use XML document, a fragmentation strategy also a DNS like system to generate an efficient query plan. Security enabled services at each work station in wide area sensor databases also in the base stations along with a simple security module at each sensor node. Caching, partitioning and site finding components at the workstations in wide area to reduce query response time. Catalog Management, Meta Data Management at base stations, and sensor nodes would support them. Data structures, and architectural supports (multiple gateways) to support multiple queries and share the common intermediate results among the queries. TCP is not suitable for end to end connection for conventional sensor networks as sensor nodes are data centric, attribute based and address-less. An appropriate transport layer protocol is needed at sensor nodes. Base stations can work as bridges between transport protocols for wide (TCP) and conventional sensor networks. Energy efficient, data centric routing algorithms are required.

Aggregation Support	Security Proxy	Application Layer for Sensor Nodes
Data Structure for Multiple Queries	Query Proxy	
XPATH/SQL Converter	Security Proxy	Application Layer For Base
Query Parser, Planner, Optimizer	Power Control	
Meta Data Management	Catalog Mgmt.	
Aggregation/ACQP Support		
Continuous/Streaming Query Support		
Query Parser	Caching Comp.	Application Layer For Wide Area
Data Partitioning Logic	Site Finding	
Application Dependent Security		
Wide Area Network to sensor node communication support		Transport Layer
Energy Efficient Data Centric Routing		Network Layer
Protocols for Catalog Mgmt.		

Fig. 1. Proposed Query Processing Architecture

Future Directions: Conventional sensor networks use SQL like declarative languages while wide area networks use XPATH/XSLT. A proxy layer will work in the base stations to map between XML and SQL. We propose a common query language in both parts to remove the proxy layer. Besides, for very changing sensor data we propose XML and XPATH queries in both parts. Sensor network applications treat sensor data as stream or continuous. Besides, base station databases are hybrid where some static data is kept and continuously updated data are coming. Queries are sometimes continuous and long-lived. Supporting aggregate queries is also important. Continuous queries must also be supported in wide area. Hence, a generalized DBMS is required that can handle diversified types of applications. Since the sensor nodes support limited storage, the code size should be small. Alternatively, the base node can load appropriate modules to the sensor nodes depending on the application and might use swapping/paging in the base station hard drive, which requires much power consumption and power efficient algorithms. Both architectural support, and query language support also efficient optimization, and query processing plan are needed to pose continuous queries over wide area sensor networks. Sophisticated caching (view based semantic caching) of query results are needed to reduce query response time. Sophisticated algorithms for evicting of cached data are required. Different data mining approaches over historical sensor data will find historical trends to provide better answers to user queries. None of the current querying techniques for sensor networks are concerned with database security. We propose to introduce a security layer in application layer of each sensor node and the gateway nodes to prevent unauthorized users from querying. Besides, energy efficient encryption techniques in query request and reply might help.