Efficient Sensor Stream Data Processing System to use Cache Technique for Ubiquitous Sensor Network Application Service

Jun-Yong Park and Ryum-Duck Oh
Department of Computer Science and Information Engineering,
Chung-Ju National University, Chungju-Si, South Korea

Abstract: Problem statement: For the purpose of supporting application services in the Ubiquitous Sensor Network (USN) environment, efficient processing of stream data that occur in the sensor nodes and relevant utilization technologies are being studied. **Approach:** The conversion of various sensor stream data that occur in the sensor nodes into sensor data suitable for use in the upper-level application system is required through sensing data monitoring by applying meta-database from the middleware on the server side. **Results:** In the earlier developed sensor data processing system, unnecessary search and operation were performed from the middleware on the server side as for sensor stream data that occurred in the same sensor node or in the similar area. **Conclusion/Recommendations:** In this study, sensing cache technique is used and efficient sensor data stream processing system is designed and built to reduce the redundant sensor data stream processing.

Key words: Ubiquitous Sensor Network (USN), data stream, dynamic metadata, sensor nodes, data processing, sensor stream data, cache technique, applying meta-database, wireless networks, S2CT technique

INTRODUCTION

Recently, with the development of Ubiquitous Sensor Network (USN) technology, necessary data for environmental monitoring, logistics surveillance, building management could be received through a wireless network from various small sensors. With the development of these technologies, various types of sensors were used in USN environments and sensor data generated from the nodes also changed very complicated. Metadata is defined and used in order to define the information of these sensor nodes. While such a large amount of metadata generated from the network, it is a cause to increase the data transmitted from sensor nodes. And because of the increased amount of data, data processing time of sensor nodes increased but the lifetime of sensor network reduced. In a sensor network system currently being built, static metadata are stored by entering data directly on the server middleware, dynamic metadata are stored to the database of server middleware by collecting on periodic basis. And when sensor stream data packet is sent, optimal sensor data is configured by obtaining the information needed for meta database where collected information are stored. But data collected from sensor network come in the form of repetitive stream, it access database each data stream processing. This type of database access is quite unnecessary in the nature of USN environment. The reason is because the events in the USN environment occur repeatedly in a particular space or node and because database is requesting the same metadata excessively when dealing with sensor data (Akyildiz *et al*., 2002; Madden *et al*., 2005).

This study proposed sensor stream data processing technique-Sensor Stream Cache Technique (S2CT) technique and S2CT system-which efficiently processed sensor data stream and excessively improved access to the database.

Related research: A study on the data stream is actively underway in the field of the current international database. Such interest recognized the need of efficiency of the DBMS for other areas in existing business areas through the development of the Internet. Such stream data processing technology is used to control real-time bank data management and provide intensive security services for the packet stream on the (Kawashima *et al*., 2006; Botts and Robin, 2007; Stonebraker *et al*., 2006) Internet. In addition, correlation analysis on URL filtering and different network stream has been used in more complex stream processing.

Corresponding Author: Jun-Yong Park, Department of Computer Science and Information Engineering,
Chung-Ju National University, Chungju-Si, South Korea

MATERIALS AND METHODS

Stream data processing technology has been used, as shown in (Fig. 1) in the environment which sensor data continuously flows through the wireless networks.
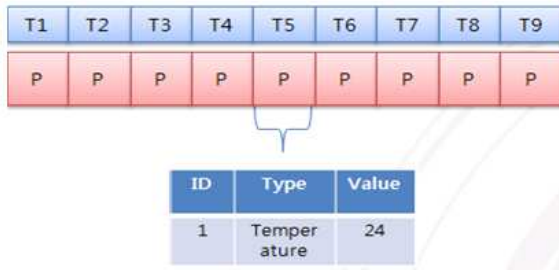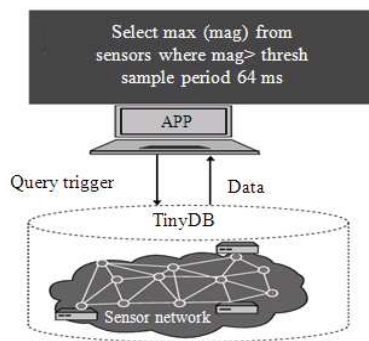

Fig. 1: Sensor stream data

Fig. 2: TinyDB

Fig. 3: Sensor data packet

Fig. 4: Sensor metadata

As the lifetime of the network remains in battery in the USN environment and life of the sensor network maintains for a long time as possible, efficient implementation of storage manager of sensor data stream or sensor stream data compression are being actively studied in the field of technology. But, the study on sensor stream data and optimization of metadata is very lacking (Golab and Ozsu, 2003).

TinyDB is a Ubiquitous sensor network system and a DB system for TinyOS environment which performed in MOTE sensor hardware developed in U.C Berkeley, USA. Considerations for the development of program in sensor networks include sensor's limited battery problems, loss of communications, impossible environment to maintain. Although programmers do not consider these things, TinyDB will start appearing as one of the measures that the system is responsible for. But, TinyDB regarded, unlike the concept of built-in DBMS where work independently under the operating system of TinyOS that is built into each sensor, the whole sensor network as an abstract concept and regarded the whole sensor network as a database. Figure 2 data that multiple sensors detect on the network are stored in one table named 'sensors' and work in a way that stored information transforms into Tiny SQL (Madden *et al*., 2005).

But, TinyDB has caused many problems when you use in reality because of unnecessary approach to the sensor network in practical use of USN environment and unnecessary data processing in sensor nodes.

Stream Data Cache Technique (S2CT):

Sensor data stream processing: USN environment sensor data stream processing refers to optimized information organization by using a combination of metadata and data packets when the packet data come from the sensor networks for an easier use in upper application. These sensor data stream processing is largely composed of three steps.

Step 1, as shown in (Fig. 3), come packets-which consists of header information such as the ID of sensor nodes and sensor data values. But the top application services cannot be used directly because it requires different information such as the location of sensor nodes in addition to sensor node's sensing data values.

Step 2, server middle ware, as shown in (Fig. 4), uses the ID of the node from the database which have metadata of sensors bring the metadata required for sensor nodes.

Step 3, the combination with the packet data incoming from sensor networks and the data obtained from the database will be optimized necessary information for upper application as shown (Fig. 5).

Sensor network ID	Sensor network node	Sensor type	Sensor value	Sensor network location	Sensor network admin	Sensor network protocol	Sensor node HWSpec
A2	A102	Temp	28	Upo	Chang nyeng	Zigbee	A

Fig. 5: Optimization data
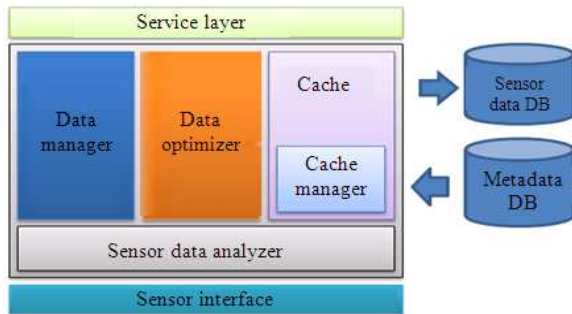
Fig. 6: System structure

Table 1: Experimental result (S2: S2CT, T: Traditional)

	1 h		2 h		3 h		4 h	
Node	S2	T	S2	T	S2	T	S2	T
DB access	87	226	111	325	118	400	131	448
Amount of packet	221		324		399		447	

And the optimized information will be provided to the upper application or if necessary will be stored in a database to accumulate sensor information.

RESULTS AND DISCUSSION

The S2CT technique improved the basic method of sensor stream processing. As for the processing order of this technique, if sensor stream data firstly transmitted from sensor network is delivered, index data for stream data processing is classified. Such classified data is used to access meta database and acquire data. Sensor data is acquired after using metadata. But since the sensor data is optimized, the data obtained from the database is stored in the cache. And in the S2CT technique, for the sensor data packet collected after, first you should figure out at a cache whether index information of sensor data packet and metadata index information are the same. And if necessary metadata in the cache is retrieved, cache uses metadata to optimize sensor data packet and metadata optimization. And metadata used should be located on the top of the cache so that the next incoming packet is to be retrieved for the first time. However, if the data does not exist, necessary metadata was obtained in the same way as the existing database and optimization proceeded.

S2CT sensor data steam processing system: The S2CT system consists of components such as Sensor Interface, Sensor Data Ana layer, Metadata DB, Sensor Database, Data Manager, Data Optimizer, Cache, Cache Manager as shown in (Fig. 6) in order to handle a variety of sensor data streams:

- Sensor Interface: provide a common interface for various sensor networks
- Sensor Data Analyzer: examine the error packet or null value for Sensor Data Packet coming from the sensor network and classify the index data required for search for metadata
- Cache/Cache Manager: Cache Manager provides cache information to data optimizer if there is cache information by comparing classified sensor index data and cache data, retrieves data from Metadata Database to store in Cache
- Data Optimizer: Use the information sent from the Cache and optimize the packet data of sensor networks
- Data Manager: The information configured from Data Optimizer provides data as Data Manger is an appropriate upper layer application

The S2CT system, unlike the other sensor stream data processing system, applies Cache techniques and processes the data faster than in other systems. And to verify the performance of the proposed S2CT system, simulation environment was configured and partial system was configured to measure performance.

S2CT system implementation and simulation: For critical cache capability simulation in S2CT system, TinyOS-based Telos platform model was used for sensor nodes and JDK 1.6 and MySQL were used to build software development platform. As for simulation method, the nodes were installed in a particular area and temperature change events were detected through the sensor nodes within the precinct. And depending on the events that occur, database access and recovery were measured for data packets that are sent. And on measurement, S2CT technique was used to measure the number of database access and unapplied database access number. Through this, it was tested to figure out how much the proposed technique can reduce the number of database access number.

As to simulation configuration environment, ten nodes were installed in area A to compare the number of times an event occurs and the number of packets and database access number in the case with S2CT technique applied or in the case without S2CT technique on hourly basis.

The simulation result shows a reduction in access to the database and operation, as shown in (Table 1), because a very high number of events occur on the same node, if sensor metadata is configured with cache, compared to the old data stream processing system.

As S2CT technique can reduce the cost consumed due to excessive access to database and provide high performance compared to the existing stream data processing system.

CONCLUSION

Currently, Water Quality Monitoring Network of Water Pure Control Office only provides monitoring and alarm or generate final elements of water pollution with integrated abnormalties of each item through the opinion of experts.

This study investigates various situations and data related to wetland environment water quality in order to understand user's intention and purpose within the area of wetland management and find the best results. Additionally, this promotes the convenience of the user search by organizing and analyzing a variety of related terms and proposes ontology model in the field of wetland environment, which is the basis of semantic situation detection of the relationship between words.

The proposed domain specific ontology context awareness model is not a deduction of fragmentary values through a simple query but, a provision with lexicons and terms to express and share the situational information as a means of describing concepts and correlations. This model support the share and reuse of situational knowledge in ubiquitous computing environment and creates lower-level of ontology appropriate for specific domain by using upper level ontology as a hierarchical situation ontology model. This provides the basis for the optimization of wetlands which is an important position of man and nature, systematic wetland management of wetland environment related data and data search of dictionary meaning and information.

For the future intelligent data retrieval and detection approach, more diverse subclass and relationship must be defined based on related water quality data collection from the Institute of Public Health and Environment. Further, by applying the proposed ontology model, in addition to information of eutrophication and algae phenomenon and alarm sending, more comprehensive management of wetland environment and continuous research on the methods of how to build and expand applicable to various domain will improve the completeness of system design of wetland environment context awareness.

REFERENCES

Akyildiz, I.F., W. Su, Y. Sankarasubramaniam and E. Cayirci, 2002. A survey on sensor networks. IEEE Commun. Mag., 40: 102-114. DOI: 10.1109/MCOM.2002.1024422

Botts, M. and A. Robin, 2007. OpenGIS sensor model language (sensorml) implementation specification. Open Geospatial Consortium Inc.

Golab, L. and M.T. Ozsu, 2003. Issues in data stream management. SIGMOD Record, 32: 5-14. DOI: 10.1145/776985.776986

Kawashima, H., Y. Hirota, S. Satake and M. Imai, 2006. Met: A real world oriented metadata management system for semantic sensor networks. Proceedings of the 3rd International Workshop on Data Management for Sensor Networks (DMSN'06), ACM New York, pp: 13-18. DOI: 10.1145/1315903.1315907

Madden, S.R., M.J. Franklin, J.M. Hellerstein and W. Hong, 2005. TinyDB: An acquisitional query processing system for sensor networks. ACM Trans. Database Syst., 30: 122-173. DOI: 10.1145/1061318.1061322

Stonebraker, M., U. Cetintemel and S. Zdonik, 2006. The eight rules of real-time stream processing. Mondo Visione.