Data Source Management using weight table in u-GIS DSMS

Sang Ki Kim[*]   Sung Ha Baek[**]   Dong Wook Lee[**]
Warn Il Chung[***]   Gyoung Bae Kim[****]   Hae Young Bae[*****]

Abstract The emergences of GeoSensor and researches about GIS have promoted many researches of u-GIS. The disaster application coupled in the u-GIS can apply to monitor accident area and to prevent spread of accident. The application needs the u-GIS DSMS technique to acquire, to process GeoSensor data and to integrate them with GIS data. The u-GIS DSMS must process big and large-volume data stream such as spatial data and multimedia data. Due to the feature of the data stream, in u-GIS DSMS, query processing can be delayed. Moreover, as increasing the input rate of data in the area generating events, the network traffic is increased. To solve this problem, in this paper we describe TRIGGER ACTION clause in CQ on the u-GIS DSMS environment and proposes data source management. Data source weight table controls GES information and incoming data rate. It controls incoming data rate as increasing weight at GES of disaster area. Consequently, it can contribute query processing rate and accuracy

Keywords : u-GIS, Disaster, Data Source, Data Stream Management System

1. Introduction

As the emergence of the GeoSensor, the research of the u-GIS is progressed. The GeoSensor can obtain information such as temperature, humidity, spatial information and multimedia information. The u-GIS requires real-time processing technique about the GeoSensor data.[1, 2, 3]

Data Stream is continuous and real-time data to be input to system. The DSMS(Data Stream Manage-ment System) manages and processes data stream. Recently, the growth of the DSMS requires u-GIS DSMS that process Spatial-temporal Information and Multimedia Information. The u-GIS DSMS consist of GES(GeoEdge Server) and GSS(GeoStream Server). GES coupled with GSN (GeoSensor Network) collects data in GeoSensor and manages it. The GSS processes data stream inputted from GES locally

distributed. Also, the GSS samples and analyzes information by registering CQ.[4, 5, 6, 7]

The u-GIS DSMS can be applied to a disaster monitoring application. The disaster monitoring application real-timely observes a specified area and prevents the accident and spreading disaster such as a fire and an earthquake. To support them, the GSS receive real-time data stream from sensor measuring temperature and surveillance cameras detecting current state when the disaster was caused.[8, 9]

For example of a disaster monitoring application, each GeoSensor periodically and continuously acquires information such as temperature and humidity in order to prevent and to detect a fire of school and a robbery case in the library. Suppose a case which installs the surveillance camera in dangerous area and observes it. At this time, in a lecture room of a building in the school, a sensor near it detects high temperature and

realizes it is continuously increasing. In this case, the sensor management system requires the following analysis.

Foremost, in order to correctly analysis the disaster area, the system increases sensing speed of the area detected with high temperature. Next, the system provides the state of disaster area as multimedia data to control surveillance cameras at the area in order to watch the damage in the area. Finally, the system monitors neighborhood area using other cameras in order to notice the reason of accident and prevent spreading the accident.

For the situation, the u-GIS DSMS needs many following require in the disaster monitoring application.

- TRIGGER ACTION has to be appended into CQ (Continuous Query). For example, the GSS analyzes the sensed data from GeoSensor and detects the accident area. And then, the GSS executes the surveillance camera in accident area by TRIGGER ACTION that obtains the multimedia data of the accident area. So, accident multimedia data from the surveillance camera should be obtained in the accident area.

- GSS incoming data is transmitted from locally distributed GES. If the GSS receives from locally distributed all of GES that transmits the same incoming rate, when occurs the accident, to process multimedia data and spatial data occur delay of query processing in the GSS, and network traffic increase.

Until a recent the DSMS is conducted research about load shedding and CQ optimization. But, incoming input adaptor of the GSS is not thoroughgoing enough Data Source Management so that solve delay of query processing and network traffic increase. Also, the CQ of the DSMS is not state of TRIGGER ACTION, so occurred accident is not thoroughgoing enough rapidly measure.

In this paper, TRIGGER ACTION states clearly DSMS and introduces processing method. Data Source Weight Table suggests technique which the GSS reduces delay of query processing and prevents network traffic. TRIGGER ACTION of the CQ in the u-GIS DSMS observes predefined threshold that is environment temperature from GeoSensor. If over the threshold, accident area related the GES controls surveillance camera and obtains multimedia data. Data Source Management uses Data Source Weight Table that remains the GSS. Data Source Weight Table manages the GES that connected the GSS and controls incoming rate of the Data Source in the accident area and far from the accident area. The Query divides the GES so that the GSS prevents delay of query processing. Then filtering data is transmitted into the GSS by the GES. Because of reduces data volume, the GSS will reduce delay of query processing and could more than exactly analyzed data.

The remainder of this paper is organized as follows. Section 2 describes the related work. Section 3 describes background in this paper. Section 4 describes executed TRIGGER ACTION and Data Source Weight Table for Data Source Management. Section 5 shows the performance evaluation using simulation tool. Section 6 concludes this paper and mentions future work.

2. Related Work

2.1 DSMS(Data Stream Management System)

The DSMS is processed the Data Stream which it is transmitted continuously sensing data from the GeoSensor. The data stream has some features that arrives order of time and continuous and infinite.

Until a recent researched representative DSMS model is STREAM[10] Aurora[11], COUGAR[12], and TelegraphCQ[13]. STREAM registers adaptive queries in main server, and variety of data is received from a lot of GES. Stream Server is cashing data until received new data. The drawback of cash model can't divide query to the GES because of query registers main server. Thus, STREAM increases network traffic. According to incoming data rate can't prevent delay of query processing.

Aurora uses the load shedding technique which the method discards data or use approximate technique because occurs memory overflow when a volume data receive from GES. Because of load shedding technique is not difficulty guarantee of accurate analysis. COUGAR perform pre-aggregation with the GES of middle area. But the drawback has not flexibility about query register because the technique registers previously query. TelegraphCQ uses resource sharing to each other queries. Resource Sharing need scheduling about query processing when the accident is caused because of registering many queries about

the same stream.

2.2 Trigger Action

Trigger Action is an operator which is automatically processed by DBMS when the event defined in DBMS is caused. Trigger Action process actions related with trigger pre-defined if the condition is true in case the event is caused.

In u-GIS DSMS disaster monitoring applications observe accident area from GeoSensor and need the definition of trigger action to acquire multimedia data for controlling surveillance camera in accident area.

3. Background

Fig. 1 shows monitoring application of DSMS for disaster detection in u-GIS. The GSS connecting GES (GeoEdge Server) receive data stream from data source as shown as below figure. GES is connected with GSN (GeoSensor Network) connecting sensor and surveillance cameras. The dispersed GES is connected GSS which receive sensing data of GeoSensor and video data of surveillance cameras. The GSS shows present situation of processed queries through GeoSensor Monitoring Tool. The following Fig. 2 shows structure of GSS.
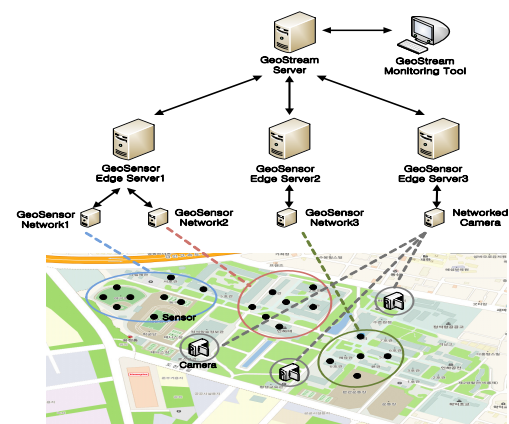

Fig. 1. GSS is Connected with GES Structure

The data source of GES, surveillance cameras and database which store geometric information are inserted into GesStream Server through input adaptor. Data Stream is transferred through Database, Web, and Monitor after those which process CQ processor are sent into Output Stream.
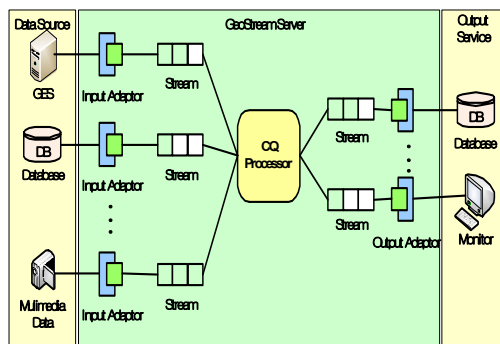

Fig. 2. GeoStream Server Processing

4. Data Source Management in GSS

In this chapter, we present the Data Source Management at the GSS in case the system is added with the TRIGGER ACTION at the CQ or an accident occurs.

4.1 TRIGGER ACTION in continuous query

The TRIGGER ACTION is the operation which performs automatically in case the event expressed at query is executed. Then, Disaster Monitoring Application of a u-GIS DSMS registers a query by adding the TRIGGER ACTION. If the condition which generates the TRIGGER ACTION of query registered is satisfied, the application generates an output stream for the TRIGGER ACTION. Generated the output stream is connected with a service expressed at the TRIGGER ACTION.

```
SELECT ts, locationX, locationY,
        AVG(IS1.temperature) as Indoor
FROM    InputStream1 IS1, InputStream2 IS2
WHERE IS1.GESID = IS2.nodeID
WINDOW RANGE 5ms SLIDE 3ms TYPE ts
INTO OutputStream
TRIGGER ACTION GeoPoint(NC1, locationX, location Y)
AFTER Indoor > 40
```

Fig. 3 describes a example of the TRIGGER ACTION in CQ. The execution of the TRIGGER ACTION describes execution of conditions in AFTER clause and BEFORE clause. Because data stream have infinite length, it use sliding window operation that process certain part of data. Commonly, sliding win-

dow uses aggregation operation in data stream. Aggregation operation in CQ needs rename for execution of TRIGGER ACTION. The aggregate rename is described the condition of AFTER clause or BEFORE clause that perform TRIGGER ACTION in CQ.
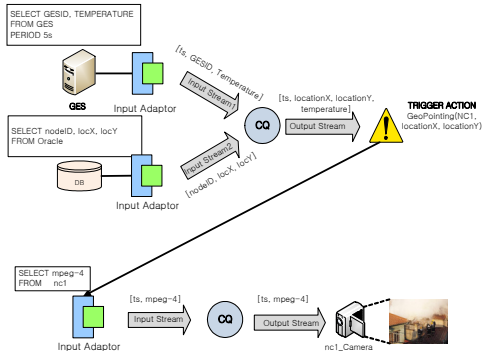

Fig. 3. Example of TRIGGER ACTION

After it receive by the GES server having Geo-Sensor data and a DBMS which spatial data were stored, an input adaptor conveys data stream to a CQ. Data Stream conveyed are processed and analyzed through a CQ. If average of the temperature about data analyzing data stream is above 40℃, the application executes automatically the TRIGGER ACTION. And then the TRIGGER ACTION is connected with a surveillance camera. The surveillance camera through the TRIGGER ACTION transmits image data such as type of data stream through an input adaptor and displays accident state.

4.2 Data Source Weight Management Table

The GSS receives a data source through an input adaptor. For example, data are transmitted from data source to the GSS. There are types of data such as environment information data(included with temperature or humidity) which are generated by the Geo-Sensor, spatial information data which are stored at DBMS, and image data which filmed by a surveillance camera. Generally, after the weight of the GES is granted similarly from each data source, the data are transmitted to the GSS.

In case an accident occurs, however, the data generated at accident area must need fast and accurate analysis in order to prevent secondary accident. According to this case, the GSS increases the weight of area which an accident occurred. Then data which occur at an accident area should be

entered more than data of other areas. However network traffic is increased because data granted with the weight are fast entered. And delay processing a query occurred for processing an image data received by a surveillance camera. To solve the problems, the GSS grants the weight of data source which divides the event level according to generating dangerous rate. The GSS need the Data Source Weight Table to manage the weights.

The Data Source Weight Table was organized from <GESID, DSType, EventLevel, Weight>. The GESID manages unique ID of the GES connected with an input adaptor. The DSType is type of data source(such as the GeoSensor data, a Database data, Multimedia data) which is received from the GES. The EventLevel shows dangerous rate from 1 level until n level. The GSS apply differently the weight according to dangerous rate and adjusts incoming rate of data source.

For example, we assume that the EventLevel was set up by 5 level if basis level is 3. Each data source was granted by 3 level of the weight. The GSS was connected with 5 of the GES. First, GESID from 1 to 3 are a GeoSensor data transmitted at the GES. Second, GESID 4 is a spatial data transmitted by a spatial database. Third, GESID 5 is an image data transmitted by a surveillance camera.
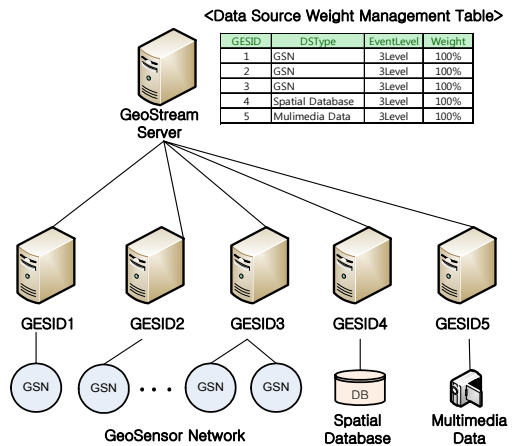

Fig. 4. Data Source Weight Management Table

If the temperature of GESID 1 is increased above standard, the GSS senses a danger of area with GESID 1. In this case, the GES of the event occurrence area is increased by 5 level of the EventLevel, and the method set up the weight from 100% to 200% in order to increase incoming rate of

data. In order to receive data of related building, the method increased GESID 4 connected with spatial database, received building data of GESID 1 sending snapshot query, increased EventLevel of GESID 5 to 4 for receiving related multimedia, and received multimedia data to control surveillance camera. Conversely, the GES without relation with an area which an accident was generated turn down load of the system which lowers the EventLevel. And then the GSS analyzes accurately data of area which an accident was generated.

Table 1. Event Change Table according to accident occurrence

GESID	DSType	EventLevel	Weight
1	GSN	5Level	200%
2	GSN	2Level	50%
3	GSN	2Level	50%
4	Spatial Database	4Level	150%
5	Multimedia Data	4Level	150%

4.3 Distributed Query Processing to GES in GSS

When some area has accident, the data inflow of accident area is increased. In addition, because the video data of surveillance cameras and the spatial data are increased, the network traffic has increased. In order to reduce work load of query processing, the GSS register parts of query to GES. Fig. 5 shows distributed query processing to GES.


```
SELECT   ts, locationX, locationY, AVG(IS1.temperature) as
Indoor
FROM   InputStream1 IS1, InputStream2 IS2
WHERE IS1.GESID = IS2.nodeID
WINDOW RANGE 5ms SLIDE 3ms TYPE ts
INTO OutputStream
TRIGGER ACTION GeoPoint(NC1, locationX, location Y)
AFTER Indoor > 40
```

INSERT INTO GESID1
FROM GESID4

GESID4   Map Data   GESID1   GeoStream Server
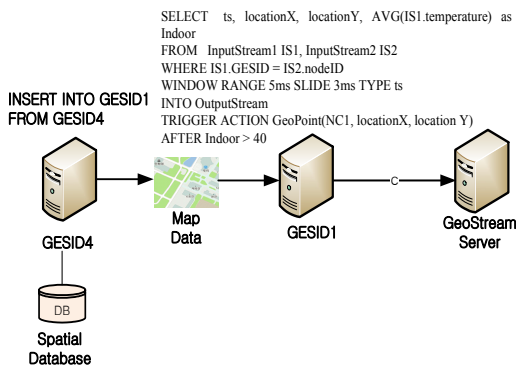
DB
Spatial Database

Fig. 5. Distributed Query Processing to GES

When the event is caused, the method sets the weight per each data source. The incoming rate of data source has fixed by weight, and the query of

GSS distribute to GES. In order to distribute many of the queries in GSS, the queries which related information of accident areas register to GES. Therefore, the data of processed query is inserted to GSS. Because the data size is reduced, the spatial data of accident areas and the video data of surveillance cameras can process rapidly. Moreover, because the query processing rate is increased, it can obtain correct data of related accident.

5. Performance

5.1 Experiment Environment

The system for performance evaluation was progressed in HP ProLiant DL380 G5 Server with 2GB memory and Intel Xeon 1.86GHz CPU, and the algorithm was implemented with C++ in Microsoft Visual Studio 2005. The data set is generated from GES simulator. GSS is connected with 10 numbers of GES, and the 6 GES send temperature data come from GeoSensor. 2 GES send spatial DBMS information storing spatial data and 2 remaining GES send multimedia data of camera. Every experiment compares Aurora, STREAM and using data source weight table managing data source and load balancing between GES and GSS.

5.2 Accuracy of Query Processing

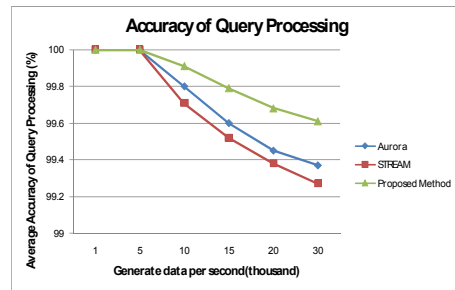Next experiment analyses accuracy of query processing in the accident area.

Fig. 6. Accuracy of Query Processing

Because the Aurora used a load shedding method, it dropped data which are generated by overflow at main memory. So, accuracy of query processing is fallen about accident area. As the STREAM uses data cashing method until receiving new data, it has lower accuracy than the Aurora. However, we could check lower accuracy than the proposed method.

5.3 Query Processing Time

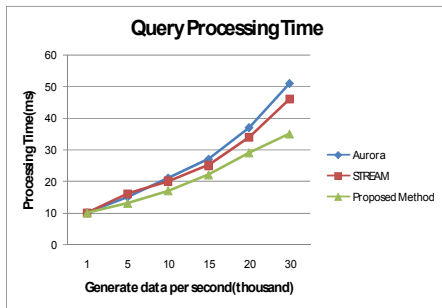Next experiment compares query processing rate of accident area data.

Fig. 7. Query Processing Time

Since the proposed method reduced the weight of the data excepted accident area, it is faster data processing with the accident area. Aurora and Stream is received with the data which are granted by the same weight from whole GES. Therefore, delay of query processing is appeared for the accident area data.

6. Conclusion and Future Work

In this paper, we propose the method which expresses TRIGGER ACTION at the CQ in u-GIS DSMS and which manages efficiently the GES. The TRIGGER ACTION operation could be fast decided the condition of an accident area, the Data Source Weight Table granted differently the weight for each GES according to dangerous level of an accident occurrence. The proposed method improves the accuracy and the query processing rate, and the network traffic of server reduce. In the future work, the management method of data source when simultaneous accidence has occurred is needed.

References

[1] Choong-Ho Lee, Kyoung-Hwan An, Moon-Su Lee, Ju-Wan Kim, "Trends of u-GIS Spatial Information Technology," Trend of Electronics and Telecommunications, ETRI, 2007.

[2] Suang Ha Baek, Dong Wook Lee, Sang Hun Eo, Weon Il Chung, Hae Young Bae, "Adaptive Memory Management Method based on Utilization Ratio to Process Continuous Query", Journal of KSISS Jun, 2009.

[3] Suang Ha Baek, Dong Wook Lee, Gyoung Bae Kim, Weon Il Chung, Hae Young Bae, "Load Shedding Method based on Grid Hash to Improve Accuracy of Spatial Sliding Window Aggregate Queries", Journal of KSISS Jun, 2009.

[4] B. Babcock, S. Babu, M. Datar, R. Motwani, and J.Widow, "Model and Issues in Data Stream System", Proc. of ACM SIGMOD-SIGACT-SIGART Symposium on PODS, 2002, pp1-16.

[5] A. Jain., E. Y. Chang, Y. Wang., "Adaptive Stream Resource Management Using Kalman Filters", Proc of SIGMOD, 2004.

[6] A. Arasu, S. Babu, and J. Widom. "The CQL Continuous query Language : Semantic Foundations and Query Execution," Stanford University Technical Report, 2003.

[7] A. Arasu, B. Babcock, S. Babu, J. McAlister, J. Widom, "Characterizing memory requirements for queries over continuous data stream" In Proc of the ACM SIGMOD, June 2002.

[8] C. Olston, J. Jiang, and J. Widom. "Adaptive filters for continuous queries over distributed data streams" In Proc of the ACM SIGMOD. June 2003.

[9] R. Motwani, J. Widom, A. Arasu, B. Babcock, S. Babu, M. Datar, G. Manku, and C. Olston, J. Rosenstein, and R. Varma, "Query Processing, Resource Management, and Approximation in a Data Stream Management System," In Proc of CIDR, 2003.

[10] A. Arasu, B.Babcock, S. Babu, M. Datar, K. Ito, R. Motwani, I. Nishizawa, U. Srivastava, D. Thomas, R.Varma, and J. Widom, "STREAM : The stanford stream data manager", Proc of IEEE Data Engineering Bulletin, Vol.26, 2003, pp19-26.

[11] D.J. Abadi, D. Carney, U. Cerintemel, M. Cherniack, C. Convey, S. Lee, M. Stonebraker, N. Tatbul, and S. Zdonik, "Aurora: a new model and architecture for data stream management", The VLDB Journal, 2003.

[12] Y. Yao, J. Gehrke. "Query processing for sensor networks" Proc of the CIDR Conf, January 2003.

[13] S. Chandrasekaran, "TelegraphCQ : Continuous dtaflow processing for an uncertain world" Proc of the CIDR Conf, January 2003.

Sang Ki, Kim

2008 Dept. of Computer Education, Seowon University(B.S.)
2008~Present Master Student, Dept. of Computer Science and Information Engineering, Inha University
Research Interests : Spatial Database, Data Stream, Sensor Network

Sung Ha Baek

2005 Dept. of Computer Science and Information Engineering, Inha University (B.S.)
2007 Dept. of Computer Science and Information Engineering, Inha University (M.S.)
2007~Present Doctoral Student, Dept. of Computer Science and Information Engineering, Inha University
Research Interests : Data Stream, Cluster, Location-based Service

Dong Wook Lee

2003 Dept. of Computer Science and Engineering, Sang Ji University (B.S.)
2005 Dept. of Computer Science and Information Engineering, Inha University (M.S.)
2005~Present Doctoral Sutdent, Dept. of Computer Science and Information Engineering, Inha University
Research Interests : Spatial Data Warehouse, Spatial Information Management, Spatial DBMS in Ubiquitous Environment

Warn Il Chung

1998 Dept. of Computer Science and Engineering, Inha University(B.S.)
2004 Dept. of Computer Science and Information Engineering, Inha University (Ph. D.)
2004~2006 Senior Researcher, ETRI
2007~Present Instructor, Dept. of Information Security Engineering, Hoseo University
Research Interests : Database, Data Stream, Moving Object

Gyoung Bae Kim

1992 Dept. of Computer Science and Engineering, Inha University(B.S.)
1994 Dept. of Computer Science and Engineering, Inha University(M.S.)
2000 Dept. of Computer Science and Engineering, Inha University(Ph. D.)
2000~2004 Senior Researcher, ETRI
2004~Present Assistant Professor, Dept. of Computer Education, Seowon University
Research Interests : Moving Object Database, Storage System

Hae Young Bae

1974 Dept. of Applied Physics, Inha University(B.S.)
1978 Dept. of Computer Science and Engineering, Yonsei Univercity(M.S.)
1989 Dept. of Computer Science and Engineering, Soongsil University(Ph.D.)
1985 Guest Professor, Univ. of Houston
1992~1994 Director, Electronic Calculations Office of Inha University
1982~Present Professor, Dept. of Computer Science and Information Engineering, Inha University
1999~Present Director, Intelligent GIS Research Center
2000~Present Emeritus Professor, Chongqing University of Posts and Telecommunication, China
2004~2006 Director, Graduate School of Information and Communication, Inha University
2006~2009 Director, Graduate School, Inha University
Research Interests : Distributed Database, Spatial Database, Geometric Information System, Multimedia Databases