DESIGN AND IMPLEMENTATION OF METADATA MODEL FOR SENSOR DATA STREAM

Yang Koo Lee*, Young Jin Jung*, Keun Ho Ryu*, Kwang Deuk Kim**

*Database/Bioinformatics Laboratory, Chungbuk National University, Korea
*{ leeyangkoo, yjjeong, khryu }@dblab.chungbuk.ac.kr

** Korea Institute of Energy Research (KIER)
**kdkim@kier.re.kr

ABSTRACT:

In WSN(Wireless Sensor Network) environment, a large amount of sensors, which are small and heterogeneous, generates data stream successively in physical space. These sensors are composed of measured data and metadata. Metadata includes various features such as location, sampling time, measurement unit, and their types. Until now, wireless sensors have been managed with individual specification, not the explicit standardization of metadata, so it is difficult to collect and communicate between heterogeneous sensors. To solve this problem, OGC(Open Geospatial Consortium) has proposed a SensorML(Sensor Model Language) which can manage metadata of heterogeneous sensors with unique format. In this paper, we introduce a metadata model using SensorML specification to manage various sensors, which are distributed in a wide scope. In addition, we implement the metadata management module applied to the sensor data stream management system. We provide many functions, namely generating metadata file, registering and storing them according to definition of SensorML.

KEY WORDS: SensorML, Sensor Model Language, Wireless Sensor Network, Metadata, Data Stream

1. INTRODUCTION

Wireless sensor networks (WSN) consist of an amount of sensors that collect and communicate data stream continuously in the physical world [1, 2]. Modern hardware technologies make it possible to gather data by using cheap and small sensor devices (e.g., smart dust and RFIDs) and data are then transferred in the manner of a wireless communication. These sensors collect data about natural phenomenon such as temperature, light, sound, and pressure and then transmit them to server in real-time. They are mainly used in geophysical monitoring, highway congestion monitoring, movement tracking, and medical monitoring [3, 4].

Unfortunately, wireless sensors have serious limitations to the use of resources such as battery lifetime and small storage. Therefore, sensors mainly use their resources for simple task like collecting data, and other data managed in server to reduce resource overhead. The data managed in server means the metadata that represents static attributes of sensors such as sensors' locations and sensor's types.

Until now, wireless sensors have been managed with individual specification, not the explicit standardization of metadata, so it is difficult to collect and communicate between heterogeneous sensors. To solve this problem, OGC(Open Geospatial Consortium) has proposed a SensorML(Sensor Model Language) which can manage metadata of heterogeneous sensors with unique format[5].

The metadata defined from SensorML is composed of Identification, Classification, limitation, function, feature, connection, and so on. These aim to express functional model for sensor not information about hardware.

In this paper, we introduce a metadata model using SensorML specification to manage various sensors, which are distributed in a wide scope. In addition, we implement the metadata management module applied to the sensor data stream management system. We provide many functions, namely generating metadata file, registering and storing them according to definition of SensorML.

The rest of this paper is organized as follows. In section 2, we introduce the concepts of sensor WEB and sensorML. Section 3 presents the metadata model for sensor data. Section 4 describes the results of our implementation. Conclusions and future work are given in section 5.

2. RELATED WORK

2.1 Sensor WEB

The Sensor Web defined by SWE(Sensor Web Enablement) in OGC(Open Geospatial Consortium) aims to apply sensors and detecting tools, which are accessible internet through the WWW without any environmental interference, and develops a standards foundation for plug-and-play web-based sensor networks [5].

The Sensor Web visions are to grasp the measured value easily in real time through web and provide

standardized web services which control sensor data and their measured values [6, 7]. For these visions, SWE use the schema based on XML(eXtensible Markup Language) to describe features of sensors, locations and their interface on the web according to an unified format.

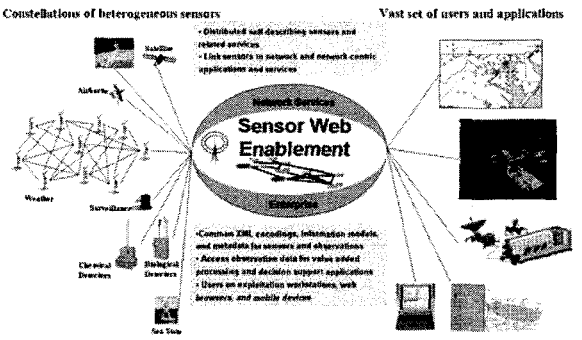

Figure 1. A view of sensor WEBs

Figure 1 illustrates the concept of sensor Webs. Sensor Webs as shown in the figure transmit sensor data collected from various sensors via network and provide them to services in application area such as flood measurement, air pollution monitoring, damage detection of the road and bridge, and so on.

2.2 SensorML

SensorML(Sensor Model Language) is the XML-based language developed by Earth observation scientists to overcome issues which are difficult to match sensor data with geographical information efficiently and combine with various sensors. sensorML was developed by Mike Botts with the support of GMTT(Global Mapping Task Team) affiliated CEOS(international Committee for Earth Observing Satellites) in April, 1998[7].

SensorML provides a model and XML schema to define the measured features from sensors using XML-based language for that describe attributes of dynamic sensors such as geometric, dynamic and radiometric attributes. SensorML supports both remote sensors and in-situ sensors, which are classified according to platform types as shown in Table 1.

Table 1. The classification of sensorML

Sensor \ Platform		Platform Type	
		Static	Dynamic
Sensor Type	Remote	Speed measurement	Satellite weather observation, Airplane weather observation
	in-situ	Chemical smell detection, Thermometer, gravity Measurement	Ozone detection, GPS, Weather detection

SensorML provides the information model and format to be able to detect between sensors and measured values which are registered on web. The schema to support this model is composed of identifiedAs, documentConstrainedBy, attachedTo, hasCRS, locatedUsing, measures, operatedBy, describedBy, documentedBy.

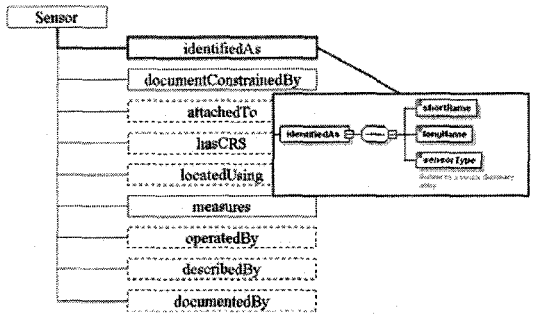

Figure 2. SensorML specification

Figure 2 shows the structure of schema to present on sensorML. Each element presents an attribute describing a feature of sensor. Elements bounded within solid lines in the figure (identifiedAs, documentConstrainedBy, measures) are described basically for sensorML and the rest of parts drawn in dotted line are optional parts possibly omitted.

3. DESIGN OF THE METADATA

3.1 Structure of the Metadata Management Module

Data stream collected from sensors is in the form of raw data. So the server has to conduct a transformation process on raw data stream to be able to manage in system. The transformation is performed for the metadata's attributes, which define data specification, such as sensors' features and functions. Therefore, data stream should refer to metadata while data stream arrive. We design the metadata management module along with the cooperation of sensor data and metadata.

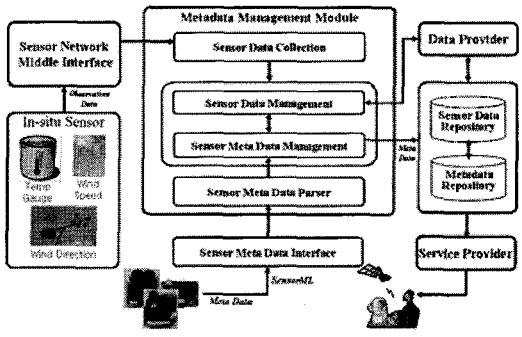

Figure 3. Structure of modules

Figure 3 depicts the whole structure of the proposed metadata management module. As shown in the Figure 3, the data measured from heterogeneous sensors are transmitted to Sensor Data Collection module via Sensor

Network Middle Interface. The measured values are transformed into a data format that can be stored and managed via Sensor Data Management module.

Besides that, metadata of sensors are converted into an XML format file through Sensor Meta Data Interface. Next, metadata are analyzed by Sensor Meta Data Parsing module, and then are stored in metadata repository through Sensor Meta Data Management module. The functions of each module included in the sensor data abstraction model are described as follows.

- **Sensor Data Collection**: This module collects sensor data transmitted from sensors and classifies them into attributes of collected data like sensorID, measuredValue, measuredTime, and so on.
- **Sensor Meta Data Interface**: This module makes out document file according to structure of XML schema suggested by sensorML to present metadata of in-situ sensor.
- **Sensor Meta Data Parser**: This module carries out the analysis of tasks to extract elements and attribute values from metadata file composed in the XML format.
- **Sensor Data Management**: This module performs the tasks of storing the collected sensor data from Sensor Data Collection in database.
- **Sensor Meta Data Management**: This module has the task of storing the elements and attributes extracted from Sensor Meta Data Parser in database.

3.2 Metadata File Format

Metadata schema is designed based on each element and their attributes according to the defined rules from SensorML. Basic information is composed of Identification elements including sensorID, modelNumber, sensorType, and so on. Other elements are defined by the detailed functions to refer the Identification.

```
<identifiedAs>
    <Identifier type="longName">Davis Anemometer 7911</Identifier>
</identifiedAs>
<identifiedAs>
    <Identifier type="modelNumber">7911</Identifier>
</identifiedAs>
<classifiedAs>
    <Classifier type="sensorType"
            codeSpace="http://vast.uah.edu/SensorML/sensorDictionary.xml">
        anemometer
    </Classifier>
</classifiedAs>
<classifiedAs>
    <Classifier type="application"
            codeSpace="http://vast.uah.edu/SensorML/applicationDictionary.xml">
        weather
    </Classifier>
</classifiedAs>
<classifiedAs>
    <Classifier type="phenomenon"
            codeSpace="http://vast.uah.edu/SensorML/phenomenonDictionary.xml">
        wind speed
    </Classifier>
</classifiedAs>
<classifiedAs>
    <Classifier type="phenomenon"
            codeSpace="http://vast.uah.edu/SensorML/phenomenonDictionary.xml">
        wind direction
    </Classifier>
</classifiedAs>
```

Figure 4. An example of SenserML file

Figure 4 shows an example of information about identifiedAs attributes among elements of sensorML. It is seen that identifiedAs is composed of metadata attributes longName, modelNumber, sensorType, and so on. The remaining elements described in document file are documentConstrainedBy, measures. In these elements, the element of documentConstrainedBy describes valid time of document, security degree, limit, and so on. And measures describe measurement unit, sampling interval, and so on.

3.3 Metadata Schema

Each element in SensorML file is parsed by metadata management module and stored in metadata repository.

measures
- description : String - survivableRange : Double - relativeAccuracy : Double - resolution : String - samplePeriod : Double

describedBy
- ResponsibleParty_id : String - individualName : String - organisationName : String - phone : String - email : String

identifiedAs
- shortName : String - longName : String - modelNumber : Integer

documentConstrainedBy
- startTime : String - stopTime : String

classifiedAs
- sensorType : String - application : String - phenomenon : String

locatedUsing
- latitude : Double - longitude : Double - altitude : Double - values : String

documentedBy
- Description : String

hasCRS
- CRS : String

Figure 5. SensorML schema

Figure 5 is an illustration of the entire schema defined by sensorML. The metadata schema shown in figure includes the units of SensorML element, and each attribute included in elements is defined as class.

4. IMPLEMENTATION

The metadata management module is implemented in JAVA language, version of jdk 1.4, and MYSQL 1.5. Initially, this module is executed by extraction step that loads XML document created by XML document according to SensorML format, separates elements and attributes.

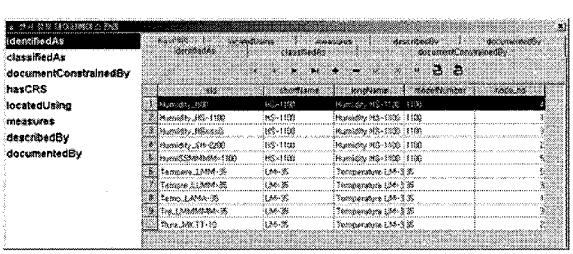

Figure 6. The metadata table

Figure 6 is a user interface to manage metadata extracted from SensorML to manage the used database. This module creates information about metadata, provides modification and deletion function.

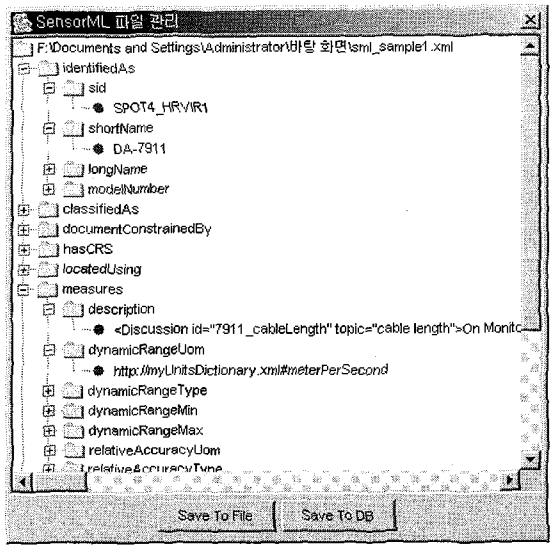

Figure 7. The metadata tree parsed from SensorML

Figure 7 displays a metadata tree as parsed result from sensorML. The parsed metadata is kept in main memory. And sensor data stream searches attributes of metadata for transformation, conducts transformation according to attributes corresponding sensorID.

5. CONCLUSION

Sensor data is composed of measured data and metadata. Metadata includes various features such as location, sampling time, measurement unit, and their types.

In this paper, we introduced a metadata model using SensorML specification to manage various sensors, which are distributed in a wide scope. In addition, we showed our implementation for the metadata management module applied to the sensor data stream management system.

Currently, we are implementing a system for sensor data stream management. In this system, we are considering metadata management and applying it to our system. We are also developing a framework which can provide appropriate service to various sensor data stream.

REFERENCES

[1] Golab, L., Özsu, M. T.: Issues in Data Stream Management. In: SIGMOD Record 32(2) (2003) 5-14

[2] Babcock, B., Babu, S., Datar, M., Motwani, R., Widom, J.: Models and Issues in Data Stream Systems. In: Proceedings of PODS (2002) 1-16

[3] Yao, Y., Gehrke, J.: Query Processing for Sensor Networks. In: Proceedings of CIDR (2003) 233-244

[4] Arasu, A., Cherniack, M., Galvez, E., Maier, D., Maskey, A. S., Ryvkina, E., Stonebraker, M., Tibbets, R.: Linear Road: A Stream Data Management Benchmark. In: Proceedings of VLDB (2004) 480-491

[5] M. Botts, "Sensor Web Enablement," http://www.opengeospatial.org/press/?page=papers&view=SensorWebWhPpr030512, November 1, 2005.

[6] M. Botts, "Sensor Model Language," http://vast.nsstc.uah.edu/SensorML/SensorML_04-019_1.0_beta.doc, November 1, 2004.

[7] M, Botts, "SensorML version 0.6," http://vast.nsstc.uah.edu/SensorML/, November 1, 2002.

5.1 Acknowledgements (optional)

This work was supported by the Korea Institute of Energy Research(KIER) and the RRC program of MOCIE