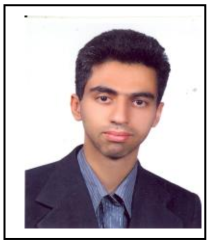
SHD: a New Sensor Data Storage



M.Gheisari[1], A.R.Bagheri[2]
[1]Damavand Branch, Islamic Azad University, Damavand, Iran;
[2]AmirKabir University of Technology, Tehran, Iran

0102-675

Mehdi Gheisari

Abstract

Sensor networks are dense wired or wireless networks for collecting and disseminating environmental data. They have some limited like energy that usually provide by battery and storages in order that we cannot save any generated data. The most energy consumer of energy is transmitting. Sensor networks generate immense amount of data. They sends collected data somewhere for storing to response users queries. In this paper we describe a new sensor data storage we called SHD, arrange sensors in hierarchical form that shows data in semantic model.

Keywords: sensor data storage, SHD, semantic, hierarchical

1. Introduction

Many sensor network applications that are related to pervasive computing, e.g., monitoring learning behavior of the children, senior care system, environment sensing, etc, generate a large amount of data continuously over a long period. Often, the large volumes of data have to be stored somewhere for future retrieval and data analysis. One of the biggest challenges in these applications is how to store and retrieve the collected data. We store sensor data's in a form that machines can collect and understand the data provided by the various types of sensors and networks. Section2 describes background studies, semantic Web technologies that include XML[1], RDF[2], SWE[3] and ontology. In Sections 3 we describe a novel sensor data storage based on SWE standards, we use a universal language to provide semantic data

[1] Extensible markup language
[2] Resource Description Framework
[3] Sensor Web Enablement

modeling for sensor networks. In Section 4 we say state-of-the-art of sensor data storage discussion. Section 5 concludes the paper and discusses the future Work.

2. Background

The OGC has recently established the Sensor Web Enablement Group in order to address problem of, Lack of standardization is the primary barrier to the realization of a progressive Sensor Web, by developing a suite of specifications related to sensors, sensor data models, and sensor web services.

2.1 Semantic web

Semantic Web is an extension to the current Web in which the meaningful relationships between resources is represented in machine process able formats. The main idea in the semantic Web is to provide well-defined and machine accessible representation of the resources and their relationships rather than simple links as they are offered by the link structure on the current Web (i.e. href links in HTML).

The World Wide Web Consortium (W3C) has defined different standards for representing the semantic Web data in machine accessible and process able formats .The primary technologies for the Semantic Web include the Extensible Markup Language (XML), Resource Description Framework (RDF), RDF Schema(RDF-S), and the Web Ontology Language (OWL).

2.2 Extensible Markup LANGUAGE (XML)

XML is actually a set of syntax rules for creating semantically rich markup languages in a particular domain. The fundamental construct in an XML document is the element. An element is simply a pair of matching start- and end-tags, and all the text that appears between them.

XML documents must have a single root element that encompasses all other elements in the document. Elements may have sub elements nested within them, to any level of nesting. Elements may also have attributes. The following example shows an Xml document.

```
<Account >
<account-number> A-101 </account-number>
<branch-name> Downtown </branch-name>
<balance> 500 </balance>
</Account>
```

This example generates an account with account number A-101, its branch name is Downtown, and amount of its balance is 500.

2.3 Resource Description Framework (RDF)

At the simplest level, the Resource Description Framework is an XML-based Language to describe resources. While XML documents attach Meta data to parts of a document, one use of RDF is to create Meta data about the document as a standalone entity. The Resource Description Framework (RDF) is a framework that allows data within a domain to be linked through named relationships. An RDF graph is encoded as a set of subject-predicate-object triples which resemble the subject, verb, and object of a sentence. The subject and object are nodes in the graph and the predicate is a directional named link between the subject and object. This simple triple structure turns out to be a natural way to describe a large majority of

the data processed by machines. A Universal Resource Identifier (URI), an address just like that used for Web pages, identifies each the subjects, verbs and objects.
Thus, anyone can define a new concept, or a new verb, by defining a URI for it on the Web.

2.4 Sensor Web Enablement (SWE)

The Open Geospatial Consortium recently established the Sensor Web Enablement as a suite of specifications related to sensors, sensor data models, and sensor web services that would enable sensors to be accessible and controllable via the Web.
The core suite of language and service interface specifications includes the following:
(1) Observations & Measurements (O&M) - Standard models and XML Schema for encoding observations and measurements from a sensor, both archived and real-time.
(2) Sensor Model Language (SensorML) - Standard models and XML Schema for describing sensors systems and processes; provides information needed for discovery of sensors, location of sensor observations, processing of Low-level sensor observations, and listing of task able properties.
(3) Transducer Model Language (TransducerML) – Standard models and XML Schema for describing transducers and supporting real-time streaming of data to and from sensor systems.
(4) Sensor Observations Service (SOS) - Standard web service interface for requesting, filtering, and retrieving observations and sensor system information. This is the intermediary between a client and an observation repository or near real-time sensor channel.
The following example shows a timestamp encoded in O&M and semantically annotated with RDFa [4].
The timestamp's semantic annotation describes an instance of time: Instant (here, time is the namespace for OWL-Time ontology):

```
<swe:component rdfa:about="time_1"
rdfa:instanceof="time:Instant">
<swe:Time rdfa:property="xs:date-time">
2010-0308T05:00:00
</swe:Time>
</swe:component>
```

This example generates two RDF triples. The first, time_1 rdf:type time:Instant, describes time_1 as an instance of time:Instant (subject is time_1, predicate is rdf:type, object is time:Instant). The second, time_1 xs: date-time "2010-03-08T05:00:00,"describes a data-type property of time_1 specifying the time as a literal value (subject is time_1, predicate is xs:date-time, object is "2008-03-08T05:00:00").

2.5 ONTOLOGY

Ontologies are typically defined as an abstract model of a domain of interest with a formal semantics in the sense that they constitute a logical theory. These models are supposed to represent a shared conceptualization of a domain as they are assumed to reflect the agreement of a certain community or group of people. In the simplest case, ontologies consist of a set of concepts or classes, which are relevant for the domain of interest, as well as a set of relations defined on these concepts. The general idea is that data and services are semantically

[4] Resource Description Framework in Attribute

described with respect to ontologies, which are formal specifications of a domain of interest, and can thus be shared and reused in a way such that the shared meaning specified by the ontology remains formally the same across different parties and applications. Ontologies are utilized by the semantic Web Applications to offer conceptualized representation of domains and to specify meaningful relationships between the resources. Ontologies provide a common and shared understanding of different domains.OWL is a language that is based on description logic and facilitates construction of ontologies for different domains. The OWL representation of data enables expression of semantics and meaningful relationships between resources and amongst different attributes of complex data.
The OWL data can be accessed by software agents for reasoning and inference purposes and to enable systems to derive additional knowledge from the represented data. There are common query languages such as SPARQL available for the OWL data, in other words the stored ontology can be accessed via SPARQL queries. There are also widely used software systems such as Jena and Sesame to deploy and manage the constructed ontologies.

2.6 XLINK

The XML Linking Language, or XLink, is an XML markup language used for creating hyperlinks in XML documents. XLink is a W3C specification that outlines methods of describing links between resources in XML documents, whether internal or external to the original document. XLink defines a set of attributes that may be added to elements of other XML namespaces. XLink provides two kinds of hyper linking for use in XML documents. Extended links are out of band hyperlinks that, in a link base document, can link resources over which the link editor has no control. Simple links offer similar functionality to HTML links, which are in band links.

3. Related Works

Russomanno discuss a broad sensor ontology which is called OntoSensor. OntoSensor primarily adapts parts of SensorML descriptions and uses extensions to the IEEE Suggested Upper Merged Ontology (SUMO) to describe sensor information and capabilities. The ontology is developed to support sensor information system applications in dynamic Sensor selection, reasoning and querying various types of sensor. Onto Sensor relies on deep knowledge models and provides extensive information about different aspects of the sensor nodes and devices. The ontology is represented in OWL format and the authors have discussed the advantages of the proposed approach compared to SensorML and XML based solutions. The main enhancement is providing self-descriptive meta-data for the transducer elements and embedded semantics in the descriptions which could be utilized in various sensor discoveries and reasoning applications. Although OntoSensor illustrates a semantic approach to sensor description and provides an extensive knowledge model, there is no distinctive data description model to facilitate interoperable data representation for sensors observation and measurement data.
A universal sensor observation and measurement data model in collaboration with a sensor specification model create semantic sensor network architecture. Semantic sensor network will utilize semantic Web technologies and reasoning mechanisms to interpret sensor data from physical devices that perform observations and measurements. This will support building automated sensor information processing mechanisms to extract additional knowledge from real-time or archived sensor data.

Ontology-based description of a service oriented sensor network is discussed in p.Barnaghi. The SWE and Geography Markup Language (GML) classes and properties in collaboration with SensorML, Suggested Upper Ontology (SUMO) and OntoSensor are used to develop ontology for sensor service description. The ontology consists of three main components Service Property, Location Property, and Physical Property.

Service Property explains what a service does and properties in the other two components describe the contextual and physical characteristics of the sensor nodes in wireless sensor network architecture. The ontology is represented in OWL form and some initial consistency checking and query results are provided to evaluate the validity of the proposed solution. The system, however, does not specify how complex sensor data will be described and interpreted in a sensor network application.

The proposed framework concentrates on building sensor description ontology for sensor discovery and description of sensor meta-data in a heterogeneous environment. Although sensor device and service description will contribute to build more autonomous sensor networks, providing an interoperable data description model would be also an essential requirement in architecture for semantically enabled sensor networks.

A high level design for a universal ontology which consists of extension plug-in ontologies, sensor data ontology and sensor hierarchy ontology is described in S.Meissner .The extension plug-in ontologies enable the developers to integrate domain specific ontologies into the main ontology. This describes the sensor network capabilities and provides relations between the domain concepts and the sensor functionalities. The sensor hierarchy ontology is a knowledge model for the sensors and actuators and other physical devices in the network. It describes the features and capabilities of the elements and contains meta-data related to devices such as measurement range, accuracy and calibration. The sensor data ontology describes the dynamic observational data for transducers. The ontology model describes the contextual data with respect to the spatio-temporal attributes. However the illustrated model does not specify the details of sensor data specification and relationships between various types of complex sensor data. The taxonomy provided for the sensor hierarchy ontology specifies a set of primary numerical attributes for common types of sensors. In a practical scenario, sensor data will include more complex data types and there will be a requirement for a universal structure to define the sensor data and emerging semantics.

Seth and Hanson, discuss the idea of a semantic sensor Web framework to provide Enhanced meanings to sensor data and to create situation awareness for the sensor networks. The semantics of sensor nodes is described within space and time dimensions, and it also includes thematic data. The spatial meta-data provides sensor location and data information in terms of a geographical reference system, location reference, or named locations. The main assumption is that although the sensor's location might be changing, its location can be determined relative to the moving object. The temporal meta-data refers to the time interval duration whose sensor data has been captured. Thematic meta-data provides descriptive information about the sensor node which can be derived by sensor data analysis, and utilizing tagging and textual descriptions. The sensor Web facilitates interoperable architecture for sensor networks and enables the application to process and interpret the contextual, observation and measurement data obtained from a sensor in a heterogeneous environment. The authors describe different scenarios for applying the semantic Web technologies and ontologies to the sensor networks. One of the main issues in the semantic sensor Web architecture is employing a unified data model which supports universal interoperability and semantic description for sensor data. The latter will enable construction of content and context aware sensor network applications.

Henson et al, describe a prototype application for the sensor Web by using annotated video data. The dataset contains Youtube videos annotated with SensorML and XLINK models with reference to time ontology. The authors discuss how utilizing the semantic leads to retrieve videos by specifying temporal concepts such as “within”, “contains”, or “Overlaps” during a time interval query submission. The proposed application demonstrates the main benefits of adding semantics to the sensor network and sensor data. The authors use keyword tagging and meta-data description to provide references to temporal concepts and domain ontologies. An extension to this idea could be seen as providing a universal meta-data structure with a broader scope to accommodate various sensor data types.

4. Sensor hierarchical data Storage

In this section, we introduce new sensor data storage.

At first, we arrange sensor nodes into some clusters. A sensor node in a cluster plays role of a cluster head, collect sensor data from sensors that relies on relevant cluster, then aggregate data and send them to sink for future querying. Sensors send their data in XML form. In figure 1 we see a snapshot of network view.

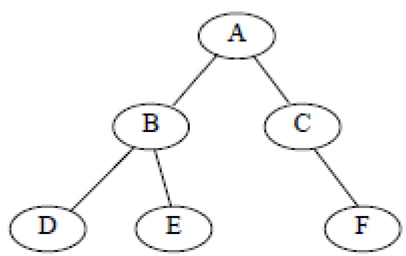


Fig 1. Network view

The network in figure 1, divide into two parts. Node B and C plays the role of cluster head in the network. Node B and node C aggregate received data. Then they send them to sink node, which in this example is Node A. We have done our simulation using j-sim sensor network simulator and protégé 2000 software. We also use LEACH algorithm that is a hierarchical protocol for clustering sensors.

In figure 2, we see amount of received data in different situation.

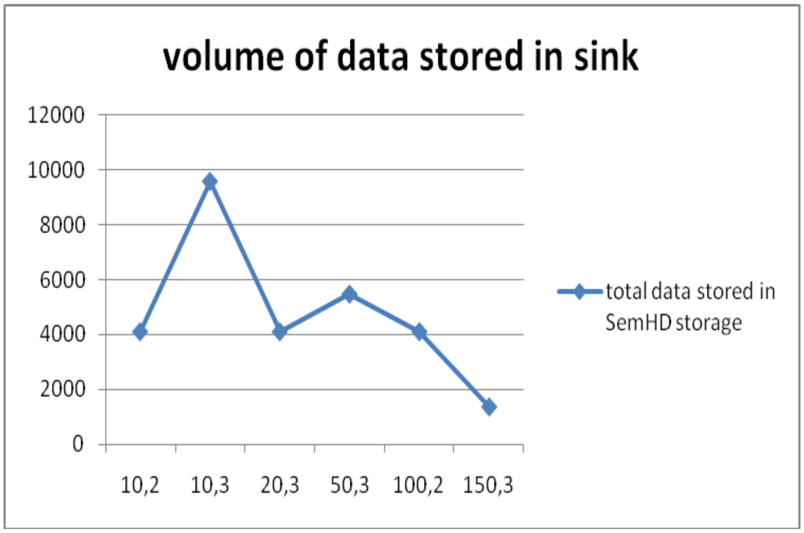


Fig 2. Amount of received data in sink node

Horizontal vertex shows an ordered pair; X, Y; X is the number of sensors in the sensor network and Y shows the number of clusters that the sensor network divides into.

As we can see, usually increasing sensors in sensor network result less amount of data received in Sink node. On of possible reason is aggregation of data because less data transmitted in network. However, in (10, 3), when we have 10 sensors that divides into 3 clusters; we have a trade-off in total amount of received data and number of clusters.

Figure 3 shows the lifetime and remaining energy of the sensor network in variety of situations.

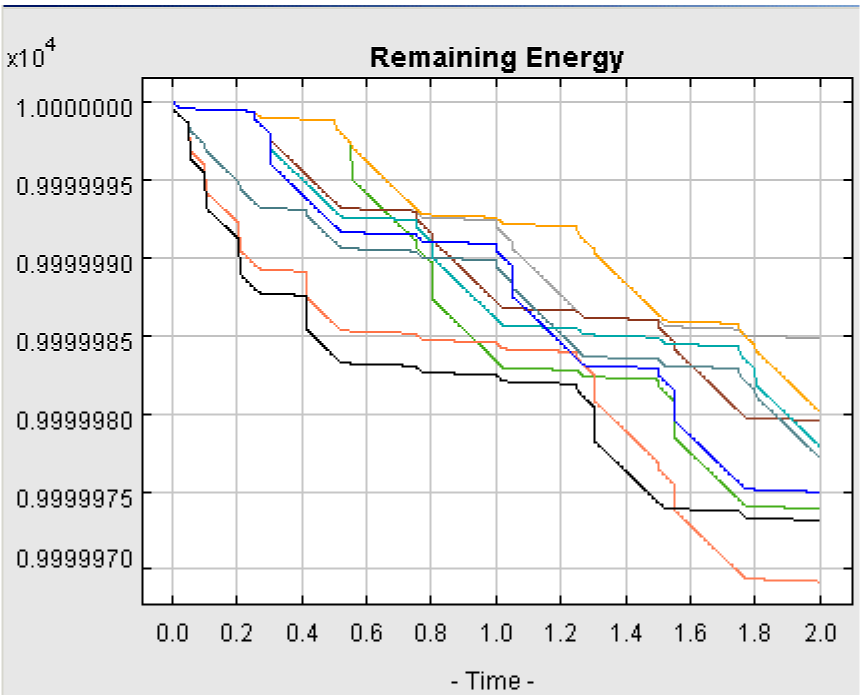


Fig 3 rate of remaining energy during times

As we can see, when a sensor plays the role of cluster head, the rate of energy consumption is increased because it done more processing like data aggregation, etcetera. The rate of consumption depends on amount of processing.

4. Conclusion and future Works

In this paper we introduced and formalized a new Hierarchical Sensor Data Storage that divide sensors into some clusters, the node in a cluster that collect sensor data; sensor data send their data in SWE form; named cluster head, then aggregate received sensor data, then send aggregated data into sink. Sink nodes collect data for further process like response more variety of queries, etc. For future work, we plan to explore a more reliable mechanism, in other words explore a new mechanism to deal with link failures between sensors in the network. Sending data more semantically will be also another step with corresponding evaluation.

Acknowledgements

We thanks from Iran telecommunications research center with the URL of www.itrc.ac.ir for their support.

References

S.Santini, D.Rauch "*Minos: A Generic Tool for Sensor Data Acquisition and Storage*". 19th International Conference on Scientific and Statistical Database Management IEEE, 2008

M.Lewis, D.Cameron, S.Xie, B.Arpinar*"Es3n: A Semantic Approach to Data Management in Sensor Networks*"Semantic Sensor network workshop, the 5th International Semantic Web Conference ISWC 2006, November 5-9, Athens, Georgia, USA 2006

P.Barnaghi, S.Meissner, M.Presser, and K.Moessner" *Sense and Sens'ability: Semantic Data Modelling for Sensor Networks*"Proceedings of ICT-MobileSummit 2009 conference

C. Henson, A. Sheth, P. Jain, and T. Rapoch, "*video on the semantic sensor web,*" W3C Video on the Web Workshop, 2007.http://www.w3.org/2007/08/video/papers.html.

G. Antoniou and F. van Harmelen, A Semantic Web Primer (Cooperative Information Systems). The MIT Press, April 2004.

A. Sheth, C. Henson, and S. Sahoo, "*Semantic sensor web*," Internet Computing, IEEE, vol. 12, pp. 78–83, July-Aug. 2008.

A. Sheth and M. Perry, "*Traveling the Semantic Web through Space, Time, and Theme,*"*IEEE Internet Computing*, vol. 12, no. 2, 2008, pp. 81–86.

Cory A. Henson, Josh K. Pschorr, Amit P. Sheth, and Krishnaprasad Thirunarayan," *SemSOS: Semantic Sensor Observation Service*"IEEE computer society 2009

C.P.Singh, O.P.Vyas, ManojKu.Tiwari," *A Survey of Simulation in Sensor Networks*", In Proceeding of CIMCA 2008, IAWTIC 2008, and ISE 2008

M.Gheisari,"*Evaluation of two known methods in energy parameter*"3rd National Conference on computer engineering and information technology, Iran, 17 February 2011

M.Gheisari,"*Sensor data storages*", 13rd electrical engineering, Iran, 2010