Architecture for Automatic Semantic Annotation to Discover Knowledge from Heterogeneous Sensor Data

K. Raghava Rao [1] T. Ravi Kumar [2] M. Nagabhushnam Rao[3]

[1]*Department of M.C.A*
Swarna Bharathi Institute of Management,Khammam,
[2]*Dept. of CSE_, B.V.Raju Engg. College, Bhimavaram,*
[3]*Dept. of MCA, KL University, Vaddeswaram, Guntur, A.P. India*[123]

[1] trkumar5@gmail.com
[2]raghavarao1@yahoo.com
[3]mnraosir@gmail.com

***Abstract*—Sensor description represented through semantic technologies in heterogeneous sensor network is a better interoperability mechanism. Through semantic annotation, it is possible to provide context awareness information from sensor networks. This will useful to improve to extract knowledge from sensor data streams and discover new sensor capabilities. We propose an innovative architecture for a system be able to automatically semantic annotate sensors descriptions with semantic concepts. And it is enabling better analysis and processing from heterogeneous streams of data. The system we proposed, sufficient examples are mentioned and the improvements that semantic context brings in heterogeneous sensor networks.**

***Keywords*— Sensor Web, semantic annotations, heterogeneous sensors , context awareness.**

I. Introduction

Internet of Things relies on scalable networks, mobility of wirelessly connected objects and offering interoperability for heterogeneous and complex networks [1]. As per Open Geospatial Consortium (OGC) specification—A Web accessible sensor networks and archived sensor data that can be discovered and accessed using standard protocols and application program interfaces[2], one of the important characteristics of the Sensor Web is that its components share and use the information gathered [3]. Derived from Sensor Webs., the concept of Participatory Sensing as defined by the authors of [4], is —data collection and interpretation and it includes mainly mobile devices that can be used to build a sensor network for capturing and sharing local data. Searching and finding sensors published by different participants. These problems are caused by using different vocabularies in describing the sensors and by the large and increasing number of heterogeneous sensors. Therefore, extracting knowledge from sensor description for understanding the data that it sends can be difficult. The documentation provided in a natural language by the participants who publish their data is not enough for a machine to understand it. A solution for improving knowledge extraction from sensor data streams is to provide semantic context. Enriching sensor description with semantic concepts leads to the development of Automatic semantic annotation Web (ASW), increasing interoperability and enabling complex reasoning with the contextual knowledge resulted from the semantic concepts [5]. In this paper, we propose novel system architecture for automatic semantic annotating sensor descriptions with concepts from an ontology, in order to offer a common vocabulary and a representation model which will enable better sensor discovery and will provide reasoning capabilities. Afterwards, we demonstrate with illustrative examples how the semantic context can help in complex searching of sensors and how reasoning can be applied to discover new knowledge from the sensor descriptions. We also show the results of annotating descriptions of sensors from a real-world sensor web with semantic concepts and we discuss what improvements are required on the semantic level, without changing the design of the sensor web. The rest of the paper is organized as follows. In Section 2 we present related work, in section 3 Automatic semantic annotation for sensor data and discusses the technologies used in semantic annotation of sensor webs and describes the system architecture that we suggest for building the ASW. Section 3.1 is particular we focus on RDF's advantages over XML as a data modeling language for sensor data. Section 4 Knowledge Discovery with Sensor Data and section 5 about conclusions and future work.

II. Related Work

Previous works in the Sensor Web domain have proposed and discussed different possibilities of combining Sensor Web and semantic technologies [6][7][8][9]. Through illustrative examples they explained the advantages that semantics would bring, how the resulted ASW would ease the path on using sensor data in different studies and how it will enable better communication between parties involved in building and maintaining a heterogeneous Sensor Web. In this paper we try to apply similar scenarios on a real Sensor Web, from a collaborative environment of over 3700 sensor nodes. In [6], the authors discuss the design of the ASW, suggesting existing data on the semantic web to be used for annotations. They present some illustrative examples sensor data with description of sensor, exploiting the data already published. The assumption is that one is annotating sensor data directly with semantic concepts, which implies that all publishers will have to adopt a common ontology and annotate their sensor data with concepts from that ontology. The knowledge discovery on semantically annotated sensor data is given. Our work differs in the sense that we assume that the publisher can describe the sensor data using simple tags or natural language and, afterwards, we automatically annotate those descriptions with semantic concepts. The problem of geographical information retrieval is approached in [7]. The authors propose the use of semantic rules for adding additional processing capabilities for ontology represented in Web Ontology Language (OWL). These rules will overtake the lack of mathematical calculus of OWL and will enable context- aware geographical information retrieval.

III. AUTOMATIC SEMANTIC ANNOTATION FOR SENSOR DATA

Extending sensor webs with semantics implies finding a suitable representation of the domain knowledge in such a way as to enable interoperability and knowledge discovery mechanisms. One of the advantages that semantic technologies bring in knowledge representation are better scalability and interoperability, since adding or changing new information to a set of programs that use the same model resumes at changing the external model, while the design of those programs can remain the same, without the need of human involvement [10]. The complexity of ASW technology is derived both from the semantic and the sensor network point of view. Ontology used in knowledge representation play a key role in usefulness of combing semantics with sensor networks. Depending on how general is the knowledge represented by an ontology they can be categorized in domain ontology and upper ontology. The first category represents models of specific domains (e.g. sensors) and the particular meaning of concepts related to that domain, while the second category is used to model general concepts applicable on a large set of domain ontology. The authors of [7] are mentioning about an even more specified type of ontology, referred to as application ontology which ―specify the conceptualization that underlie specific application. Three of the existing sensor network ontology developed are presented in [11][12] and have a set of common concepts related to the taxonomy of different types of sensors, physical properties of sensor devices, data acquisition and sensed domain. However, the features of the sensed domain may vary depending on the application where the sensor network is used and further development of these set of concepts is required. Such ontology can be used for semantic annotation of sensor descriptions. Figure 1 presents the system architecture that we propose for building the ASW. We start form a sensor web composed of heterogeneous sensors which are described by their publishers. The sensor descriptions provide information about data streams, such as the type of measurements that the sensor performs (e.g. temperature, humidity, power consumption, etc.) or its physical location. Further, there are two assumptions on which we base the rest of the System:

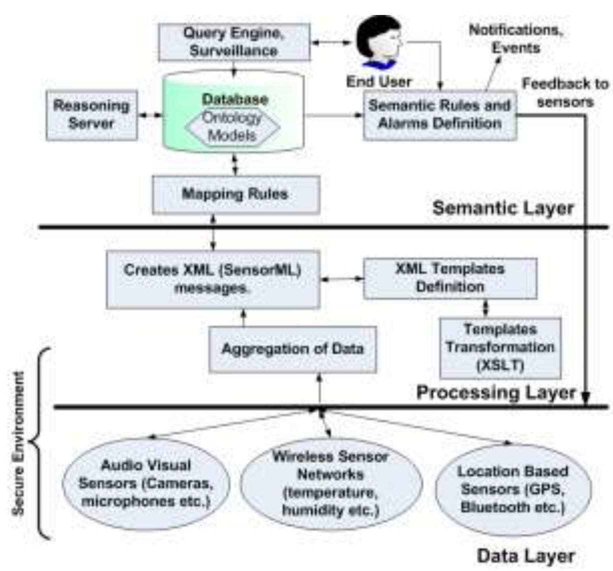

Fig. 1. System architecture building the ASW (automatic semantic annotation web)

the publishers provide at least a tag word for describing sensor measurements and/or location; the ontology concepts used for annotation are provided with a description, more exactly a string term is associated. The next component is represented by an ontology that contains the sensor web concepts needed for annotation. These concepts are used to automatically annotate the sensor descriptions, based on the string terms associated.

Apart from the ontology, there are also the logic rules that are applied on the ontology relationships and concepts, and together with the ontology they form a knowledge base. Applying the semantic concepts on the sensor web leads to the ASW which will provide more meaningful descriptions of sensors. The annotated sensor descriptions can be understood used for solving complex queries and deriving new information. A query formulated by a user will be solved by searching sensor data, which will first look into the knowledge base for the information needed for processing the query and then it will be able to search the sensors or data streams that are requested by that query, based on their annotations.

A. *Resource Description Framework*

A resource is any identifiable entity on the web. For example a web page is a resource identified by its URL. A description is any statement made about a resource comprising a resource, a property, and a value.[13] A property is itself resource that is used as the descriptive verb. A value is either another resource or literal text. For example, if a sensor 1 has a temperature reading of 54 F. The Statement representing this description would be: Since RDF <sensor1> <#has Temperature> "54" .in RDF format is a framework for describing and merging metadata, it is carefully designed to exhibit key characteristics. These characteristics include, but are not limited to: dependence, interoperability, scalability, and machine-readability. RDF resources must be independent, that is, anyone should be able to make statements about any resource. RDF must be deterministically parse reliably transfer. The most common format for exchanging RDF is RDF/XML, an XML serialization of RDF. RDF statements (triples) must scale to large numbers, since the entire web may describe resources in RDF. RDF statement and imagine how we would express the same data in XML.

```
<sensor>
<id>10</id>
<temperature>64</temperature>
</sensor>
 or maybe
<sensor id="10" temperature="64" />
```

Fig.2: Data representation using XML

IV. KNOWLEDGE DISCOVERY WITH SENSOR DATA

One of the most important arguments for semantically annotating sensor data is that of providing a support for performing reasoning on top of it. This will enable the possibility of applying logic rules through which new information can be inferred from the data available.

Example ontosensor description

```
<FLIR rdf:ID="FLIR_001">
<hasCapabilities>
<SensorCapabilities rdf:ID="FLIR_001_capabilities">
<supportedApplication
rdf:resource="#Fineresoultionimagery"/>
<supportedApplication
rdf:resource="#Daynightoperation"/>
<supportedApplication
rdf:resource="#Covert"/>
<performanceProperty>
<GenericProperty
rdf:ID="OntoSensor_Individual_337">
<Attr_Name rdf:datatype=
"http://www.w3.org/2001/XMLSchema#string">
Overscan ratio</Attr_Name>
<Attr_Value rdf:datatype=
"http://www.w3.org/2001/XMLSchema#string">
1</Attr_Value>
</GenericProperty>
</performanceProperty>
<performanceProperty>
<GenericProperty
rdf:ID="OntoSensor_Individual_339">
<Attr_Name rdf:datatype=
"http://www.w3.org/2001/XMLSchema#string">
Number of detectors</Attr_Name>
<Attr_Value rdf:datatype=
"http://www.w3.org/2001/XMLSchema#string">
180</Attr_Value>
</GenericProperty>
</performanceProperty>
</SensorCapabilities>
</hasCapabilities>
</FLIR>
```

Fig. 3 Sensor description using OntoSensor

Fig.2 displays an excerpt of the OWL file generated as the output of the instantiation. The excerpt shows the OWL constructs that capture the following knowledge: The resource *FLIR_001* is an instance of the OntoSensor FLIR class. This instance of a FLIR sensor is appropriate for obtaining fine resolution imagery, has day and night operability, and can be used for projects that require some concealment; hence the sensor also has a covert capability. In addition,*FLIR_001* also has an overscan ratio of 1 and is equipped with 180 detectors.

V. QUERYING THE KNOWLEDGE BASE

Querying knowledge from sensors

```
1 ?-sensor_capability('FLIR_001',
ListOfCapabilities).
ListOfCapabilities =
```

['Covert',
'Day/Night Operation',
'Fine Resolution Imagery']
2 ?-sensor_capability(SensorInstance,
'Fine Resolution Imagery').
SensorInstance = LASER_001 ;
SensorInstance = FLIR_001
3 ?-sensor_parameters('FLIR_001',
ListOfParameters).
ListOfParameters = [
'Entrance aperture diameter'='20 cm',
'Focal Length'='35 cm',
'Frame rate'='30 Hz',
'Horizontal DAS'='0.2 mrad',
'Horizontal FOV'='4 deg',
'Interlace'='2:1',
'Number of detectors'='180',
'Optics Transmission'='0.70',
'Overscan ratio'='1',
'Scan efficiency'='0.75',
'Vertical DAS'='0.2 mrad',
'Vertical FOV'='3 deg']

Fig.4: Example query result

Fig. 3 shows a representative query of the Onto Sensor knowledge base using SWI-Prolog's command line. Line one is a query that takes an instance of a FLIR sensor and determines its capabilities. The second query finds all sensors in the knowledge base that can capture fine resolution images. Line three shows a query that retrieves the parameters for a FLIR instance.

VI. Conclusions And Future Work

A solution for enriching sensor data streams for providing machine understandable meaning is represented by the semantic technologies. Semantic annotations can provide context for the sensor measurements and observations, transforming data streams from simple binary models into meaningful data, which can be used in further analysis. In this paper we proposed and discussed a novel system architecture that is able to automatically annotate, with semantic concepts, sensor description provided by publishers. We demonstrated through illustrative examples the advantages of applying reasoning mechanisms on semantically enriched sensor descriptions. However, the advantages of having access to common sense knowledge have been illustrated in the examples provided in this paper, namely extended knowledge can help in inferring geographical regions or creating more complex rules. In our future work we plan to provide more specific context for the concepts used in sensor annotation. This will enable more accurate annotation of the sensor description and will perform better in real-world scenarios. In addition, we are considering semantic solutions for sensor composition. The virtual sensor created through composition of sensors will introduce a level of abstractness that can enable better communication between people and sensor networks or between sensor networks themselves.

References

[1] Commission of the European Communities. 2009. Internet of Things — An action plan for Europe. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions (Brussels, June 18, 2009).

[2] Botts, M., Percivall, G., Reed, C., Davidson, J. 2007. OGC White Paper OGC® Sensor Web Enablement: Overview And High Level Architecture. White Paper. OpenGIS.

[3] Delin,K.A., Jackson S. 2001. The Sensor Web: A New Instrument Concept. Presented at SPIE's Symposium on Integrated Optics (San Jose, CA, January 20-26 2001).

[4] Goldman, J. Shilton, K., Burke, J., Estrin D., Hansen M. Ramanathan N., Reddy S., Samanta, V., Srivastava M. B., West, R. 2009. Participatoy Sensing: A Citizen-powered Approach to Illuminating the Patterns That Shape our World. Foresight & Governance Project, White Paper.

[5] Seth, A., Henson, C., Sahoo, S.S. 2008. Semantic Senor Web. Internet Computing, IEEE, pp 78-83.

[6] Wei, W., Barnaghi, P. 2009. Semantic Annotation and Reasoning for Sensor Data. 4th European Conference on Smart Sensing and Context, EuroSSC 2009, volume 5741 of Lecture Notes in Computer Science, pp. 67-76. Berlin, 2009. Springer.

[7] Keβler, C., Raubal, M., Wosniol, C. Semantic Rules for Context-Aware Geographical Information Retrieval. 4th

[8] Barnaghi, P., Meissner, S., Presser, M., Moessner, K. 2009. Sense and Sens'ability: Semantic Data Modelling for Sensor Networks. In Proceedings of ICT-Mobile Summit 2009.

[9] Janowicz, K., Schade, S., Bröring, A., Keßler, C., Maué, P., Stasch, C. 2010. Semantic Enablement for Spatial Data Infrastructures. Transactions in GIS (TGIS) volume 14, issue 2, 2010, pp. 111–129.

[10] Berners-Lee, T., Hendler, J., Lassila, O. 2001. The semantic web a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American, May 2001.

[11] Open Geospatial Consortium, "GML – A Markup Language for Geography," June 2004. Available at: http://opengis.net/gml/

[12] Z. Defne, A. Islam and M. Piasecki, "Ontology for Geography Markup Language (GML3.0) of Open GIS Consortium (OGC)," 2004. Available at:http://loki.cae.drexel.edu/~wbs /ontology/2004/09/ogc-gml.owl

[13] Benjamin Szekely, Elias Torres, A Semantic Data Collection Model for Sensor Network Applications Harvard University, {bszekely,torres}@fas.harvard