Environment Observation System based on Semantics in the Internet of Things

Liang Hu, Jingyan Jiang, Jin Zhou, and Kuo Zhao*
College of Computer Science and Technology, Jilin University, Changchun 130012, China
Email: hul@mail.jlu.edu.cn, zhaokuo@jlu.edu.cn
*Corresponding author

Liang Chen
College of Software, Jilin University, Changchun 130012, China

Huimin Lu
College of Software, Changchun University of Technology, Changchun 130012, China

Abstract—The Internet of Things is rapidly developing in recent years. A number of devices connect with the Internet. Hence, the interoperability, which is the access and interpretation of unambiguous data, is strongly needed by distributed and heterogeneous devices. The semantics promotes the interoperability in the Internet of Things using ontology to provide precise definition of concepts and relations. In this paper, we demonstrate the importance of the semantics in three aspects: firstly, the semantics means that the machines could understand and respond to the human command. Secondly, the semantics is mainly reflected in the ontology of the physical world. Thirdly, the semantics is important for interoperability, data integration and reasoning. Then, we introduce a semantic approach to construct an environment observation system in the Internet of Things. The environment observation system contains three major components, the first is the ontology of the environment observation system, the second is the semantic map of the environment observation system, and the last is exposing the observation data. Finally, the system publishes the environment observation data on the Web successfully. The environment observation system with semantics provides a better service in the Internet of Things.

Index Terms—Internet of Things; Semantics; Ontology; Semantic Sensor Network; Environment Observation System

I. INTRODUCTION

The Internet of Things (IoT) is an expansion of the current Internet, which refers to interconnecting “Things” has attracted lots of interest in recent years. It is the potentialities of the IoT that makes it possible for the development of a huge number of applications [1], such as transportation and logistics domain, healthcare domain, smart environment (home, office, plant) domain. Reference [2] proposed a smart hospital model using IoT technology in healthcare domain. Reference [3] proposed an intelligent traffic monitoring system based on IoT technology in transportation and logistics domain. In smart environment domain, reference [4] puts forward a smart home system based on the Internet of Things. The essential requirements of IoT are the interoperability, automation and data analytics, in which case, the semantics is a significant perspective and solution.

With the help of the semantic technology, we could effectively help promote the interoperability of the IoT. So, a number of works has been developed.

The ontology is a modeling tool for describing knowledge in a semantic way. The ontologies in IoT have a great improvement. The Sensor Web Enablement (SWE) of OGC defines a suite of interfaces and services to connect heterogeneous sensors. It contains several specifications: Observations & Measurement (O&M) defines standard models and XML Schema for encoding observations and measurements, which are archived and real-time; Sensor Model Language (SensorML) defines standard models and XML Schema for describing sensors systems and processes refer to sensor observations; Sensor Observations Service (SOS) is a standard web service interface for a client to request, filter, and retrieve; Transducer Model Language (TransducerML or TML), Sensor Planning Service (SPS), Sensor Alert Service (SAS) and Web Notification Services (WNS) are also part and parcel of SWE. Recently, an ontology called SSN for describing sensors and sensor network resources has been developed by the W3C Semantic Sensor Networks Incubator Group. The SSN ontology provides a better support for abstraction, categorization, and reasoning offered by semantic technologies [5]. Since the SSN ontology defines high level schema of deployments, systems, the operating restrictions, processes, devices and data, it answers the need for interoperating with other ontologies.

Linked data plays an important role in effective reasoning, processing mechanisms and better interoperability. It is based on the HTTP and URIs and enables not only the people but also the machines to explore data on the Web. The resources are identified by URI and described in RDF, URIs link to other URIs, so that data from different sources are connected [6].
linked open data (LOD) integrates numerous distributed
data in form of RDF across the Web [7]. Many
applications use linked data principles to annotate IoT
data: Kno. e. sis linked sensor data describes the weather
data using linked data principles and links the Geonames
for searching by geographical location [8]. Sense2Web
provides a GUI to describe IoT data [9]. Page et al.
propose an API to publish data of observatory using
linked data principles and REST [10].

In this paper, we just focus on the environment
observation, which is in a higher level of the Internet of
Things. The lower level is the part of data sensing and
data collection. We deployed many sensors in different
rooms of our lab, and these sensors sensed the
environment and collected real data (e. g. temperature,
humidity and illumination) into a relational database.
In the higher level, we introduce an approach of semantics
to publish the real sensor data on the Web, it contains
three parts: ontology, semantic mapping and exposing.
The ontology is a data abstract of the existing sensor data,
the real things are mapped to concept and properties.
In our environment observation system, the nodes and some
environmental indicators (e. g. temperature) are described
as classes with properties. The goal of semantic mapping
is to map the relational database schema to RDF or
ontology using D2R mapping language, in this process,
the data integration can be applied for context and
situation awareness. Semantic interpretation is a crucial
issue for exploring the semantic descriptions. In this work,
the D2R server is applied to implement the publishment.

The rest of this paper is structured as follows. In
section 2, we demonstrate the importance of the
semantics in three aspects. In section 3, we introduce
the architecture of the environment observation system in
the Internet of Things. In section 4, we general evaluate
the environment observation system. Section 5 is a summary
of this paper.

II. SEMANTICS IN THE INTERNET OF THINGS

The semantics makes a great contribution in many
the FOAF vocabulary for social network domain.
video retrieval approach based on semantics in the multi-
media domain. The semantics is important for the Internet
of Things as well. In order to shed light on the semantics
in the Internet of Things, three issues should be discussed.

Firstly, we will explain the meaning of the semantics.
The semantics is the interpretation of an expression. M.
Uschold divided semantics into four levels: implicit,
explicit and informal, explicit and formal for human
processing, and explicit and formal for machine
processing. The implicit semantics is a kind of shared
human consensus such as meaning tags. The explicit
and informal semantics is an informally expressed notation
or language which is mainly for humans. For example,
the meaning terms in the Dublin core. Explicit and formal
for human processing semantics is a formal documentation
just for humans. For example, the modal logic for the
semantic definition of ontological categories. The explicit
and formal for machine processing semantics is the kind
of semantics which is used at runtime for automated
inference for machines. In Semantic Web, the semantics
means that the machines could understand and respond to
the human command. It is explicit and formal for
machine processing as M. Uschold explained in 2003
[14]. For example, the machines understand the
“temperature” by meaning instead of a character string.
And then the machines will have a reaction relate to
“temperature”, for instance, the machines will give a
command or instruction to a device to open the air
condition. In this process, semantically structured
information (Explicit and formal for human processing
semantics) should change into the machines processing
semantics level for automated inference.

Secondly, we will set forth the formal description of
the semantics. The “understanding” ability of machine
requires the information which is semantically structured
derived from the Internet of Things. Hence, the IoT data
should be modeled as ontology. The ontology is an
explicit, formal and shared conceptual model which is
structured using standard ontology language such as
OWL and RDF, and it provides the precise definition of
concepts and relations. The precise definition is the key
of knowledge representation, information sharing and
semantic reasoning. Therefore, the semantics is mainly
reflected in the ontology. In the Internet of Things, the
scenario could be described as ontology. The nodes and
observation attributes are the classes and relations.

Thirdly, we will underline importance of the semantics
in the Internet of Things. First, the semantics is important
for interoperability. The Internet of Things requires a
stunning number of devices connecting with Internet. For
this reason, the interoperability, which is the access and
interpretation of unambiguous data, is intensively needed
by distributed and heterogeneous devices and is provided
by the unambiguous data representation. The
semantically structured data precisely answers the need to
unambiguity. Therefore, the semantics is essential or even
indispensable in interoperability. Second, the semantics
plays a vital role in data integration. The goal of the
Internet of Things is to sense the environment and react
intelligently. The IoT data can combine with other
domain information and integrate to infer new
information. Consequently, the unambiguous data
representation is an approach for heterogeneous data
integration and semantic descriptions. Third, the
semantics is a key foundation in reasoning. The reasoning
enables the inference from existing rules and information.
The semantically structured data is the premise of the
reasoning. In this way, the Internet of Things could react
intelligently.

III. THE ARCHITECTURE OF THE ENVIRONMENT
OBSErvATION SYSTEM

This section introduces the architecture of the
environment observation system. The environment
observation system is based on semantics, using standard
ontology languages such as RDF or OWL to represent the
data which is stored in the RDB (relational database),

2722 JOURNAL OF NETWORKS, VOL. 8, NO. 12, DECEMBER 2013

© 2013 ACADEMY PUBLISHER
And this system is collecting real world data (i.e., humidity and temperature) through the sensors deployed in the Internet of Things platform. Consequently, this system aims to publish the real observational environment data in a semantic way and integrate more information in specific scenario.

The environment observation system comprises three major components as illustrated in Fig. 1, the first one is the ontology of the environment observation system, the second one is the semantic map of the environment observation system, and the last one is exposing the observation data. And the system is in the top level of Internet of Things platform. Obviously, the initial data is provided by sensors and stored in the database in the lower level.

The ontology of the environment observation system describes the observation data in sensor network domain, and exploits the Semantic Sensor Network (SSN) ontology which is the standard ontology for describing sensors and sensor network by the W3C Semantic Sensor Networks Incubator Group. In this paper, we just focus on the observation data of the sensors so that only a subset of the SSN ontology concepts is applied to the system. The ontology is modeled through the actual IoT scene to describe the environment of two rooms in our labs.

The semantic map of the environment observation system is the core component of this system. The RDFS vocabularies and OWL ontologies could be published on the Web, thus, the observation data which stored in the relational database (i.e., MySQL) could be mapped into the RDFS vocabularies and OWL ontologies. In order to describe the relation between a relational database schema and RDFS vocabularies or OWL ontologies, the D2RQ platform is applied in the system. The observation table is mapped into an observation class, and the columns in the table are mapped into the properties. The classes and properties provide precise descriptions of the environment observation scene and the semantization of data. In addition, the semantic map could provide more information by integrating other domain knowledge such as the location. The data in relational database is flexibly reassembled and linked as well.

In order to illustrate the data flow of the environment observation system, a data flow diagram is shown in Fig. 2. The data transformation process is mainly divided into three stages. First, the raw sensory data is perceived by the sensors, and collected by a relational database directly. The data is structured in the form of database table without any semantics. Then the data is described as a sensory ontology based on the content of database table. All the observation items are described as classes, and each class has several properties. The mapping file is created on the basis of the ontology and database schema using D2R mapping language. Finally, the data from relational database is mapped into a virtual RDF graphs via a mapping file. It is worth noting that a dotted line points to virtual RDF graphs from RDB. In fact, the transformation from RDB to virtual RDF graph is direct under the guidance of mapping file. Since the mapping file guides how the instances of virtual RDF graph are created, it is crucial in the process. On the right of the figure, a "knowledge hierarchy" corresponding to the three transformation stages are shown. The data layer is a lower layer refers to numberous raw data collected by the IoT devices. The information layer is in the middle, the interoperability is achieved by structuring with semantics from the raw data. The knowledge layer provide machine-readable and machine-comprehensible data, which could be browsed and searched using web browsers and SPARQL easily.

The exposing the observation data process publishes data on the Web. In order to access the content of the database as linked data over the Web, the D2R server is an appropriate choice. D2R server provides a linked data view, a HTML view and a SPARQL Protocol endpoint over the database by HTTP protocol [15]. Hence, the linked data could be browsed and searched using web browsers and SPARQL easily.
A. The Ontology of the Environment Observation System

In general, ontology is a specification of a conceptualization [16], it is a modeling tool for describing knowledge in a semantic way. In this paper, the ontology is applied to modeling the concept of the environment observation system based on the actual IoT scene and take advantage of the SSN ontology.

The observation data is obtained from the sensors which are deployed in our lab, and is stored in the database. The database schema is shown in Fig. 3.

Idsh11_5w is the primary key, it is an ID of all these data. Nodeid is the ID of the nodes which are placed in the lab. The parent records the node ID which represents routing path of the data. Board_id is applied to manage sensor boards by recording the ID of sensor boards which is attached to the nodes. The airHumidity is the real air humidity of the room, and measured by the relative humidity. The airTemperature is the real temperature of the room, and its unit is centigrade. The light is the real illumination intensity of the room, its unit of measurement is lux. The time is the date and time, and the pattern demonstrates like this “yyyy-mm-dd hh: mm: ss”.

The ontology of the system is proposed on the base of database schema as illustrated in Fig. 4. The ellipse means classes and the arrows means properties. Explicitly, the ontology provides more information, the measuring unit and the room ID that deployed nodes is the additional information. The data is coordinated by the location information and the time information which makes the observation model more concrete and precise. Also the ontology reassembles the data in a hierarchical way. The properties are shown in the Table I.

B. The Semantic Map of the Environment Observation System

The semantic map plays a very important role in the system. It maps the schema of a relational database to RDF and defines a virtual RDF graph which contains information from the database. D2RQ platform provides a mapping language to accomplish the map process. The main goal of the mapping language is to map a RDF from the RDB without changing the existing database schema. The complexity is reduced greatly by joining the SQL statement into the mapping process. The mapping process is mainly divided into four steps as follows:

1. A record set is selected for each class by SQL statement.
2. Grouping the record set according to the specific ClassMap.
3. Creating the instances of each class and assigning the URI.
4. Datatype and property bridges are applied to each instance property.

A virtual RDF graph is created after the four steps above. The virtual RDF structure for the system is shown in Fig. 5.

The database is mapped to RDF terms through d2rq: ClassMap and d2rq: PropertyBridges, shown on the right. The class map is the most important concept. The class map represents a class or a group of similar classes of the ontology. A class map specifies how URIs are generated for the instances of the classes [17]. In this semantic map, the map: observation_sht11_5w_ClassMap is the major class map, and all the observation items are included in it. In order to expand the observation model, we structure additional class map, nodeIdCls_classMap provides more information of the nodes, such as serial numbers, board IDs and the locations. It is noteworthy that the nodeid property of observation_sht11_5w_ClassMap doesn’t have the concrete value, instead, it references another d2rq: ClassMap which creates the instances which are used as the values of nodeIdCls_classMap by d2rq:
referenceClassMap. It is just like the concept of the foreign key. Along the same lines, airHumidityCls_classMap, airTemperatureCls_classMap and lightCls_classMap are structured for the same purpose.

Each class map has many property bridges, which specifies how the properties of an instance are created [17]. Therefore, property bridges map a database column to an RDF property. The d2rq: column points the properties with literal values which are the table column names in the relational database. For example, d2rq: column "observation_sht11_5w.time" means the value of map: time_PropertyBridge is observation_sht11_5w.time. The d2rq: constantValue points that the property has the same constant value on all the instances of the class map. For example, d2rq: constantValue eos: roomTwo is the constant value of the map: room_two_PropertyBridge, which implies that the location for some nodes. Taking advantage of the d2rq: constantValue, we could provide some constant value inexistent in the database. The d2rq: referenceClassMap is also a property of d2rq: PropertyBridge and we have discussed it before.
C. Exposing the Observation Data

In order to publish the RDF that we have created before, the D2R server is applied. D2R Server is a tool for publishing the content of relational databases on the Semantic Web. D2R Server uses a D2RQ mapping to map relational database schemas into RDF, and allows the RDF data to be browsed and searched on the Semantic Web. To start D2R Server, we need the command as follows: ”d2r-server -p 8080 mappingName.ttl”, then open the URL “http://localhost:8080/swaprql/” in a web browser. A query interface is shown in the web just as the Fig 6. Then we could write the query into the text of the SPARQL section and push the ”Go!” button to obtain the results of the query in the SPARQL results section. The classes and properties of the system is shown in the Fig. 7 and Fig. 8. An example of a query for air temperature resources is shown in the Fig. 9. This example searches all the useful information about the air temperature of No. 997 through the data links such as time, unit, value and observed nodeid. The result is shown in the Fig. 10. The information of air temperature of No. 997 is centigrade. The underlying sentence is made by connecting the subject-verb-object. For example, the unit of air temperature of No. 997 is centigrade. The underlying sentence demonstrates the characteristics of semantic.

WHERE {
  { <http://localhost:8080/resource/airTemperatureCls/997> ?property ?hasValue }
  UNION
  { ?isValueOf "http://localhost:8080/resource/airTemperatureCls/997";  
    FILTER (?isValueOf != rdf:type) .
    FILTER (?property != subject-verb-object.)
  } 
  ORDER BY (!BOUND(?hasValue)) ?property ?hasValue ?isValueOf
}

IV. Evaluation

The Internet of Things is rapidly developing in recent years. Many applications based on the Internet of Things are developed in various domains. Reference [4] puts forward a smart home system based on the Internet of Things. The monitoring system has three levels, the sensing layer, network layer and application layer. The data is collected in the sensing layer, transmitted in the networks layer and processed in application layer. Obviously, the whole system is without the semantics.

In this paper, we propose an environment observation system architecture which exposed the observation data on the Semantic Web. Besides the three layers mentioned above, environment observation system has a semantics layer. Hence, the observation data which is provided by the sensors could be browsed and searched in a semantic way. The semantics is derived from the given ontology model. The ontology is modeled through the actual IoT scene to describe the environment of two rooms in our labs. The ontology provides precise definition of concepts and relations which reflects semantics. However, the smart home system could not be searched in the semantic way and different devices do not share the same concepts which lead to the lower interoperability.

The D2R mapping language perfectly maps the relational database schemas to the RDF. The mapping process mixes the data from the database and integrates more information as specific scenario. However, in smart home system, the data is structured in relation database without semantics. Hence, the integration of the data is only implemented in application layer and the degree of integration is lower.

In general, to publish environment observation system in a semantic way is a non-trivial task. In the Internet of Things, it is not enough to network the devices, but the semanticization of data, which promote data understanding and semantic reasoning, so that it makes the Internet of Things smarter.
V. CONCLUSIONS AND FUTURE WORK

In this paper, we first illustrate the importance of the semantics in the Internet of Things by discussing three issues: What is the semantics? Where is semantics in the Internet of Things? Why is semantics important? Then, we introduce a semantic approach to the Internet of Things through constructing an environment observation system. The environment observation system comprises three major components, the first is the ontology of the environment observation system, the second is the semantic map of the environment observation system, and the last is exposing the observation data. The system publishes the environment observation data on the Semantic Web successfully. In the future, we could add our semantic environment observation data in the Internet of Things into the Linked Open Data (LOD) for sharing and interoperability.

ACKNOWLEDGEMENT

This work was supported in part by the National High Technology Research and Development Program of China under Grant No. 2011AA010101, the National Natural Science Foundation of China under Grant No. 61103197, 61073009 and 61240029, the Science and Technology Key Project of Jilin Province under Grant No. 2011ZDGG0007, the Youth Foundation of Jilin Province under Grant No. 201101035, and the Fundamental Research Funds for the Central Universities of China under Grant NO. 200903179, the China Postdoctoral Science Foundation under Grant No. 2011MS00611, he 2011 Industrial Technology Research and Development Special Project of Jilin Province under Grant No. 20110069, the 2012 National College Students’ Innovative Training Program of China.

REFERENCES


