Towards an Ontology development for automated applications in Smart City environment of SmartME Project

Nidhi Kushwaha
Indian Institute of Information Technology, Allahabad
India
+91-9792881438
kkush6887@gmail.com

O P Vyas
International Institute of Information Technology, Naya Raipur
India
+91-9425501450
dropvyas@gmail.com

Carlo Puliafito
Dipartimento di Ingegneria Elettrica Elettronica e Informatica,
University of Catania, Italy
+39-3463315554
carlopulia@gmail.com

Ranjana Vyas
Indian Institute of Information Technology, Allahabad
India
+91-8677955160
ranjana@iiita.ac.in

ABSTRACT

Cities are becoming everyday smarter, with an increasing multitude of electronic nodes distributed throughout the territory. These range from rather simple ones, such as sensors/actuators, but also smartphones, to more complex ones, such as data centers and workstations. Citizens may have a central role in consuming, but also producing the data. The consequence of all this is that the amount of data collected is enormous and these data need to be properly processed in order to make the most of them. Unfortunately, there are several challenges in data processing, but exploiting the Semantic Web technologies and linking data among them is the right way to face them. This paper introduces the SmartME project developed by the University of Messina, Italy, and discusses the technologies and approached that cat be utilized to properly manage the collected data. In this paper, we are working towards the incorporation of semantic layer with the SmartME project of University of Messina. In this, our contribution is to built logic for maintaining sensors and their collected information query them in more meaningful way for getting accurate results. Also, we have presented a way of modifying a previously developed Ontology, SSN and customized it for our purpose. Other than manual entry of a new sensor, which mitigates the burden of manual entry. To build up this logic, we have exploited Jena API of Semantic Web in Java.

Keywords

Linked Open Sensor Data, Semantic Web; Sensor; IoT; Protégé; SmartME; Smart City.

ACM ISBN X-XXXXX-XX-X/XX/XX.

DOI: http://dx.doi.org/10.1145/00000.00000

1. EMERGENCE OF LINKED OPEN DATA

Linked Open Data provides an opportunity to harness the background data for the use of various applications such as Data Integration, Publishing, Information Retrieval, Data Miming and Feature Generation. It can be explored for the use of smart device due to its universal format and ease of data publishing and availability. These smart devices may be able to help their users to control HVAC (Heating, Ventilating, and Air Conditioning), window lighting control and so on. Already today, smartphone applications, such as CenceMe, exist that infer the activity of the person wearing phone, from sensor data and publish this on the Web [1]. Linked Open Sensor Data, which was produced by Kno.e.sis in 2011 (https://datahub.io/dataset/knoesis-linked-sensor-data), enable sensor data to be openly accessible on the Linked Open Data (LOD) Cloud. It contains expressive description of ~20000 weather stations in the United States, with more than 18,000 links to GeoNames places. Data are in the format of RDF descriptions of hurricane and blizzard observations in the United State. The data set includes observations within the entire United States during the time periods that several major storms were active including Hurricane Katrina, like Bill Bertha, Wilma, Charley, Gustav and a major blizzard in Nevada in 2002. These observation are generated by weather stations described in the MesoWest/Kno.e.sis Sensor Description Dataset introduced [2]. Currently, this dataset contains around one billion triples. On average, there are about five sensors per weather station measuring phenomena such as temperature, visibility, precipitation, pressure, wind speed, humidity, etc. In addition to location attributes such as latitude, longitude, and elevation, there are also links to locations in Geonames that are near each weather station. This sensors description dataset is now part of the Linked Open Data. This key work published the data in openly available knowledge base and also introduces the vocabulary of sensor network. This vocabulary was further exploited by many researches in recent years [3][4].

VALUETOOLS 2016, October 25-28, Taormina, Italy

DOI 10.4108/eai.25-10-2016.2266527

2. MOTIVATION & RELATED WORK

The cities in future will have billions of heterogeneous devices ranging from tiny communicating objects (e.g. actuators, sensor, tags), able to interact with the surrounding environments and remote system, to high-end nodes (e.g. data centers, workstations) capable of complex operations and to process a huge amount of information. They are also capable of delivering Intelligence on the go. The smartphone penetration in the urban population may be leveraged for its integration into the Smart City Technology framework as this will facilitate active and enabling the decision makers to feel "The pulse of your city in palm of your hand". In this futuristic scenario, citizen in Smart City are constantly connected with whatever surrounds them and they are formidable information consumers; their role is not limited to this though. In a study including Smart Bikes Project, it was highlighted by Nathalie and Herve et al [5] that citizens roaming around the city may be considered as mobile probes. It is proposed in their study that if properly instrumented, bikes going across the cities can constitute a dedicated crowd-sensing platform, monitoring the urban environment where and when it is most relevant. The useful data generation is this project was subjected to appropriate analytics for decision support purposes and suggests that smart phones can also be combined with the wireless sensing.

The wealth of data being generated by data sources needs to be appropriately processed using Data Mining techniques and Machine Learning approached. The complexity of a Smart City data analysis manifests due to a variety of issues as described in next subsection.

2.1 Challenges & Issues

1. The first challenging problem is that devices are not interoperable at any level with each other since most of the time technologies differ from one to another. For instance, in contemporary IoT (Internet of Things) applications multiple competing application level protocols such as CoAP (Constrained Application Protocol) (http://coap.technology/), MQTT (Message Queue Telemetry Transport) (http://mqtt.org/) and XMPP (Extensible Messaging and Presence Protocol) (http://xmpp.org/) are becoming popular. Each protocol possesses unique characteristics and messaging architecture helpful for different types of IoT (Internet of Things) applications. However, a smart IoT application architecture should be independent of messaging protocol standards, while also providing integration and translation between various popular messaging protocols. Thus, connecting and publishing the real time sensor data to LOD is a challenge;
2. Connecting sensors to the Internet and storing the data into Resource Description Framework (RDF). At the data level, devices do not use common terms or vocabulary to describe interoperable IoT data;
3. The current use of IoT is targeted to a single domain and most of the times the number of sensors are duplicated unnecessarily. For instance, temperature sensors in a building are primarily used for a Heating, Ventilating, and Air conditioning (HVAC) application; however, values produced by temperature sensors could be used in other applications, such as fire detection. The primary advantage of using common sensors in various applications is that it can reduce development, maintenance and deployment costs and promote device reusability. For this, automatic annotation of the new sensor which enters into an already deployed sensor network would play a significant role;
4. Mapping sensor's raw output into high level states. To achieve this requirement, a smart IoT system has to provide high level knowledge that can map sensors to real world entities and sensor's raw output to high level states;
5. Integrating the prediction model with SPARQL or exploiting periodic patterns from past state. For example, how to predict the probability of the heavy rainfall in June, 2017 will be approximately 0.8;
6. Multiple Data sources generate Structured, Semi-Structured and Un-structured Data;
7. Analysis and interpretation of cross thematic data (Urban Transport, Energy, Water etc);
8. The need for real-time processing of data streams for proactive information delivery and risks analysis.
9. To utilize LOD with semantic annotations for data analysis.

2.2 Research Possibilities

We can use the already built vocabulary of Linked Open Sensor Data. Furthermore, the newly entered sensor can be automatically annotated by matching the similarities of collected data on one side and location of the already deployed sensors on the other. Besides, Linked Open Sensor Data provides historical data regarding previous events that would help to predict future information. Smart IoT systems need to interconnect data produced by various sensors to understand the meaning of data in order to automatically take decisions or to provide suggestions. Thus, dealing with the interoperability of heterogeneous data is required. Data is stored in different files (e.g., CSV, Excel) and structured with different models (e.g., Ontology, schemas). To deal with heterogeneous data, Semantic Web technologies bring several benefits as they:

1. Unify data;
2. Link IoT data to external knowledge bases;
3. Explicitly add metadata (i.e. semantic enrichment/enhancement);
4. Deduce new knowledge.

Semantic Web technologies enable the interconnection with knowledge graphs. By interconnecting IoT data with such graphs (datasets and Ontologies used to structure the datasets), IoT systems are becoming smarter. A knowledge graphs that is widely available and useful is "Linked Open Data" (LOD). Major internet companies such as Facebook and Google are building private knowledge graphs based on Schema.org (agreement on common schemas to structure data) that is widely adopted by search engines, in addition to components of LOD such as DBpedia, Wikipedia and/or Wikidata. Knowledge graphs are built to get access to the information requested more easily and in an automatic way. Companies have chosen to use different representations for semantic data, which include unlabeled and labeled graphs, W3C ratified Semantic Web languages such as RDF and RDFS to explicitly describe the data, and OWL, a language to describe Ontologies/vocabularies [6]. Sensor's information can be collected on Web pages and parsed using Web crawler and converted into RDF version using Jena Semantic Web parser and converter [7]. Thus, we explore the IoT applications using Semantic Web technologies to semantically annotate sensor data, deduce new knowledge and combine IoT applicative domains.

3. REPRESENTING IoT ENTITIES AND THEIR ENVIRONMENT

In this paper we refer to the SmartME project of the University of Messina, Italy, according to which “smart cities may be regarded as complex ecosystems composed of heterogeneous interconnected things providing sensing and actuating facilities”, and we tried to give it our contribution. In the scenario depicted above, new services and applications can be envisioned like traffic monitoring, energy management, e-health, disaster management, etc. Though, to develop these applications, a powerful proliferation of mobile and IoT devices, including sensors/actuators, wearable, smart appliances and cyber physical system is needed. In the SmartME project, a data oriented approach is used, in which heterogeneous IoT devices send data to the Cloud and applications are built on top of standard Cloud facilities. Open source technologies and standards are involved in the SmartME project. Stack4Things[1], an OpenStack-based framework, has been developed by the University of Messina in order to "manage sensing and actuation resources, remotely controlling nodes as well as virtualizing their functions and creating network overlays among them. Thus, the aim is not that of using Cloud Computing simply to manage data, but to shift to Sensing-and-Actuation-as-a-Service (SAaaS) perspective in which to provide developers or final users with actual or virtualized sensing and actuation resources" [8]. Web Application Messaging Protocol (WAMP) has been used as an open standard WebSocket subprotocol for Publish/Subscribe and RPC communication patterns. At the edge of the system, the IoT nodes are Arduino YUN-like boards. Furthermore, for building IoT applications, Node.js has been used, as it is open source, lightweight and efficient. Data collected by the nodes are open and users can visualize them through Comprehensive Knowledge Archive Network (CKAN), open source data portal software. CKAN is built with Python on the backend and Javascript on the frontend and its database engine is PostgreSQL. It also offers a powerful API that allows third-part applications and services to be built around it [9]. Raw data captured by sensor networks in the SmartME project are required to be stored, thus a network level integration between that system and sensors is needed. To annotate raw data, explicit mappings are needed for a particular application: for example, a rain fall greater than 250mm in 24 hours means heavy rain fall. Besides, fusing the output of multiple sensors for meaningful information generation is also a challenging task to deal with. For the task of storing these raw data in a more meaningful way, their conversion will be done using the RDF (Resource Description Framework) and sensor networks vocabularies (SSN). This annotation provides the machine readability of the sensor data. For suitability and ease of access, the data will be stored inside a Cloud environment. To make this data open and available to all for future extensions, we have also proposed data to be published on LOD Cloud by following standards/rules of data publishing on LOD [9]. Among the possible applications, collected data can be utilized for event predictions (e.g. catastrophes). For this purpose, we will exploit some past as well as current data. These data are then analyzed for finding patterns in them, thus to help the municipality take the necessary actions as soon as possible. In the SmartME project, there are about six types of sensors (temperature, visibility, precipitation, pressure, wind speed and humidity), as shown in Figure 1. These sensors generate data and store them in Cloud environment, making them openly viewable by anyone. But, due to the lack of semantics, these sensors are not uniquely identified and queried, thus the information is of little interest for human beings. Our basic aim in this paper is to provide a semantic layer below the user interface.

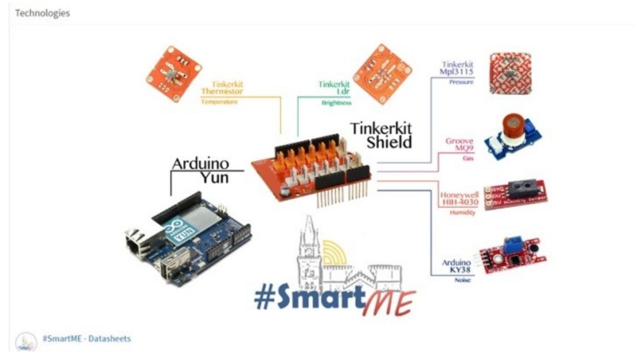

Figure 1. Sensor used in SmartME project

4. QUERYING WITH SPARQL FOR RETRIVAL

We have used the Protégé tool for modifying SSN Ontology which was previously developed by Kno.e.sis in 2011. Figure 2 shows the annotation (in Protégé) of all types of sensors present in SmartME, while Figure 3 shows a graphical view of a temperature sensor instance (Sendor_HMP50). A modification has been carried out in SSN Ontology, which made it suitable for specific queries as the ones shown in Figures 4-8. Before SPARQL querying, SSN Ontology has been modified using the Protégé tool as the Temperature sensor hierarchy shown in the Figure 3. Our modified Ontology is capable to answer various different and important questions that are needed to develop a suitable IoT environment. These questions can be triggered any time, and their output can be combined to present many fruitful results. For example: *Sensor_Temp1* is an instance of temperature sensor, located near *Location_01* and its measured minimum & maximum temperature is 8.7 & 19.67 Celsius, during day hours.

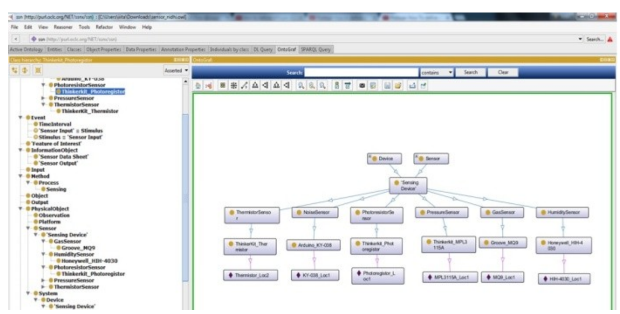

Figure 2. Annotate the type of sensors in protégé tool

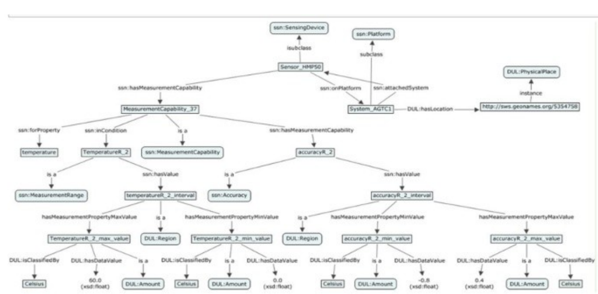

Figure 3. Graphical view of a temperature sensor instance

Q1: What type of sensor have been used

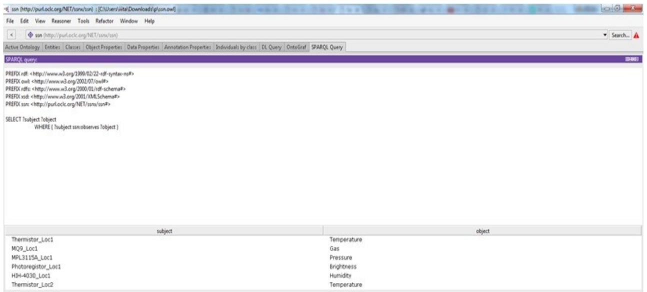

Figure 4. SPARQL query for different types of sensors

Q2: What is the type of specific sensor?

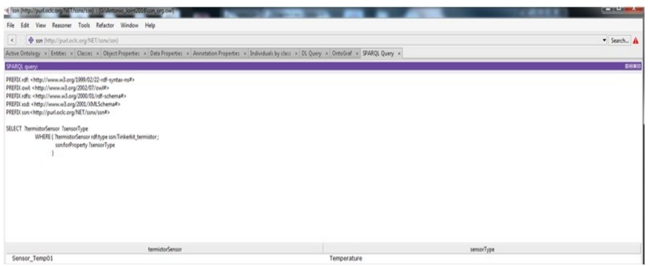

Figure 5. SPARQL query for displaying the type of a sensor instance

Q3: What is the location of the sensor?

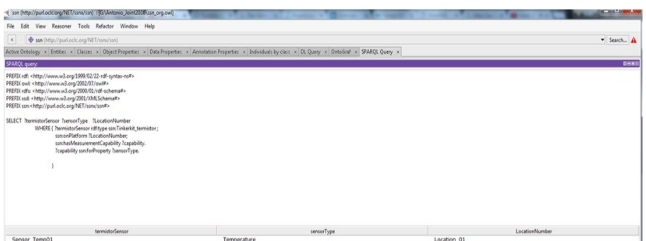

Figure 6. SPARQL query for displaying the location of a sensor instance

5. CONCLUDING REMARKS

Today, when we think about Smart Cities, we think about several applications such as traffic monitoring, energy management, e-health, disaster management, etc. For fast uptake of IoT devices, the key issues like semantic interoperability and cloud technologies have been taken into account in this project. Our main aim was to integrate a semantic layer with the SmartME project of the University of Messina for the betterment of several IoT applications. To do so, we have considered previously developed sensor Ontology, SSN, and according to the requirements of the SmartME project, we have customized this Ontology with Protégé tool. Furthermore, to check the applicability of the modified SSN ontology, we have fire various SPARQL queries on it.

Q4: What are the maximum and minimum temperatures measured by that sensor and when were they measured?

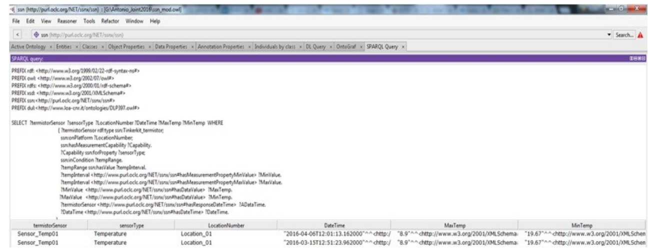

Figure 7. SPARQL query for displaying max. & min. temperature captured by a sensor instance

Q5: What is the unit of measured temperature?

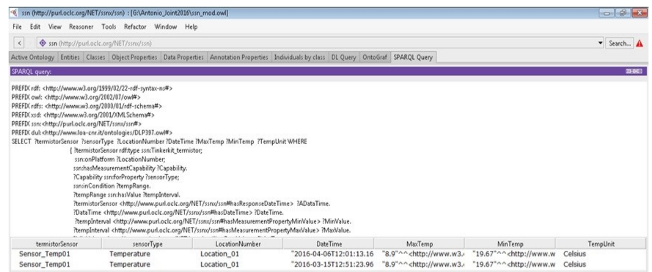

Figure 8. SPARQL query knowing the unit of temperature captured by a sensor instance

6. REFERENCES

[1] Pfisterer, D., Römer, K., Bimschas, D., Kleine, O., Mietz, R., Truong, C., Hasemann, H., Kröller, A., Pagel, M.,Hauswirth, M. and Karnstedt, M., 2011. SPITFIRE: toward a semantic web of things. *IEEE Communications Magazine,* 49, 40-48.

[2] Compton et. al. 2011. The SSN Ontology of the Semantic Sensor Networks Incubator Gr., *J. of Web Sem.*, 17, 25-32..

[3] Barnaghi, P. et al. 2012. Semantics for the Internet of Things: early progress and back to the future. *International J. on Semantic Web and Information Systems (IJSWIS)* 8, 1-21.

[4] Ploennigs, J., Schumann, A. and Lecue. F. 2014. Extending Semantic Sensor Networks for Automatically Tackling Smart Building Problems. *Frontiers in Artificial Intelligence and Applications*, 1211-1212.

[5] Mitton, N., Rivano, H. 2015. On the use of city bikes to make the city even smarter. *In Proceedings of SMARTCOMP workshop on Sensors and Smart Cities*, China. 3-8.

[6] Gyrard, A., Patel, P., Sheth, A. and Serrano, M. 2016. Building the Web of Knowledge with Smart IoT Applications, Kno.e.sis Central Technical Report.

[7] https://jena.apache.org/documentation/io/

[8] Giacobbe, M., Puliafito, C., Scarpa, M. 2016. The Big Bucket: An IoT Cloud Solution For Smart Waste Management in Smart Cities, In Proceedings of the 4th Int. W. on Cloud for IoT, Austria.

[9] https://www.w3.org/TR/ld-bp