

Middleware Mediated Semantic Sensor Networks

Cristian Lai^(✉) and Antonio Pintus

CRS4 Center for Advanced Studies, Research and Development in Sardinia,
Ed. 1 Loc. Piscina Manna, 09010 Pula, CA, Italy
{cristian.lai,antonio.pintus}@crs4.it

Abstract. This paper investigates how the Internet of Things (IoT) can take advantage of Semantic Technologies when combined with a prototyping middleware. Although Sensor Networks within the IoT offer enormous potential, the surprising variability in terms of communication and interoperability among Things is a challenging problem. This paper proposes to move from the classic meaning of Sensor Networks to the concept of Semantic Sensor Networks. The proposed methodology uses Semantic Web technologies, based on machine-interpretable representation formalisms, for combining Sensor Networks and the Paraimpu middleware. We propose to prototype and deploy a Semantic Sensor Network, focusing on the semantic aspects, with the tangible advantage to delegate to Paraimpu the low-level network operations. As a result we obtain a shared RDF Knowledge Base useful for improving integration and communication between different networks.

Keywords: Internet of Things · Sensor network · Semantic Sensor Network · Ontology · RDF · Knowledge Base

1 Introduction

The Internet of Things (IoT) is the network of physical objects that contain embedded technology to communicate and sense or interact with their internal states or the external environment (Gartner Newsroom¹). The IoT will grow to 26 billion units installed in 2020 representing an almost 30-fold increase from 0.9 billion in 2009. Nowadays it is possible to utilize wide networks with multiple elements (sensors, actuators) or in general *Things*. *Things* can be considered active participants in information processes and their heterogeneous nature makes interoperability among them a challenging issue. One specific issue is how to consider and manage networks' elements, namely devices and information obtained from raw data. This paper shows that semantic technologies, based on machine interpretable representation formalisms, appear as a promise for describing devices and data, sharing and integrating information together with other intelligent processing techniques. Data collected from sensors has to be exchanged among *Things* and other users on the Internet. Data originating from

¹ <http://www.gartner.com/newsroom/id/2636073>.

a device or a human being can be combined with other data to create different abstractions of the physical environment.

Semantics can support this integration and allows movement towards the concept of Semantic Sensor Networks (SSN). Our methodology proposes to encode sensor descriptions and observation data with Semantic Web technologies and to embed them directly within devices. Such descriptions compound a collection of semantic resources. Resources are linked to each other as well as to other types of virtual and/or real world objects through semantic links following the Linked Open Data principles². Moreover, we propose to collect sensor data by using the cloud-based Paraimpu platform. The platform is used as a middleware and, thanks to its web API, freed us from handling all the low-level network operations, data collection and real-time data transformation/annotation.

This paper describes a dynamic and scalable SSN system supporting the following features:

- Geographically distributed sensors, grouped in stations, are able to form ad hoc networking topologies; not only sending measurements data but also announce themselves and their presence in the network;
- Raw data sent by sensors is semantically annotated and enriched on-the-fly by Paraimpu as it flows through the system, using Semantic Web formats;
- Resulting data and descriptions, semantically annotated, compose a shared RDF Knowledge Base useful to improve integration and communication processes between different networks.

The remainder of this paper is organized as follows: Sect. 2 briefly introduces related work; Sect. 3 introduces the proposed methodology; Sect. 4 illustrates the annotation and transformation processes; Sect. 5 introduces the Paraimpu middleware; Sect. 6 presents a case study; and Sect. 7 provides conclusions.

2 Related Work

In recent years, several middleware-like solutions for the IoT emerged to simplify the overall adoption of the IoT by companies and final users. This adoption process includes building new applications connecting and managing smart *Things*, sensors and devices and gathering data produced by them. Many existing platforms propose quite similar solutions and tools: API, web dashboards, data collection, while a large sub-set of them are simply cloud-based services. Available platforms such as Xively³, Evrythng⁴, Carriots⁵, Paraimpu⁶ and many others try to offer these services and tools to build and maintain new IoT applications. *Mineraud et al.* [6] report a gap analysis of today's IoT landscape and about available platforms, in particular aspects concerning a broad adoption of

² <http://linkeddata.org>.

³ <https://xively.com>.

⁴ <https://www.evrythng.com>.

⁵ <https://www.carriots.com>.

⁶ <https://www.paraimpu.com>.

the IoT. In that analysis, some of the emerged gaps concerning data are nonuniform data sharing formats and processing data streams efficiently and handling different formats and models. *Mineraud et al.* [6] recommend to address these gaps by aligning data models and semantic indexing to ease the uniform processing of data sharing and fusion. *Patni et al.* [7] proposed a framework for Linked Open Data (LOD) about sensor data collected from the physical environment, introducing a publishing method for LOD using RDF and SPARQL. An interesting part of this framework is the use of a sensor dataset description for LOD. Other interesting works include the knowledge representation for integrated IoT semantic services [12], but these works did not consider dynamic, semantic annotation and description of sensors and data in real time. As also remarked in another relevant work [10], in IoT Sensor Networks, devices are dynamically added to sensor networks in real-time. To address this issue the authors [10] propose an integrated platform, which includes a web-based authoring tool and a service-based platform. This work tries to address the problems of a dynamic semantic expression of IoT resources/devices in real time and a semantic IoT data repository.

We did not focus on building a new semantic-enabled middleware platform from scratch. We decided to use an existing IoT platform in order to address our needs. After exploring many platforms, sorting by features, flexibility and data transformation tools, we decided to base our work on the Paraimpu platform. This middleware has been chosen mainly for its user-friendly workspace, for the simple but effective APIs and for its powerful embedded Javascript-based rule engine. This latter tool allows us to transform, in real-time, annotated data snippets coming from sensors in semantic descriptions.

3 Methodology

The open issues that our work tries to address can be summarized in two questions. How can sensors, stations and in general, smart, connected things, be instructed to provide semantic annotated data about their presence announcements in a SSN and related measurements? How can an existing IoT middleware platform be used to enrich data to transform it into Linked Open Data entities? Thus, our methodology addresses two objectives. The **first objective** focuses on sensors and observations data annotations, encoded through Semantic Web technologies. *Semantic annotations* are embedded within sensing stations based on Arduino microcontrollers. Sensing stations are basically devices that provide a small piece of information about their features; they are able to announce themselves in the network and declare what they observe. This information is minimized, in order to relieve sensors from details which keeps information representation very small. We consider some devices in order to easily lead the modeling and design of the basic set of sensing capabilities. We identify a set of basic components called *Things*, that can be shared within a sensor network and then used jointly (e.g., temperature and humidity sensor).

The **second objective** provides enriched *semantic descriptions* thanks to the idea of Linked Sensor Data [1–3,5,11], which facilitates in publishing and

using enriched sensor data with the help of the Linked Open Data principles. Resources are enriched and associated with each other as well as with other types of virtual and/or real world objects through semantic links. The second step is based on the Paraimpu middleware and is automated starting from the annotations provided by devices. Devices are directly connected to Paraimpu through its web API. For this second objective, we divert our attention to the Paraimpu rule engine. We identify the rules necessary to transform semantic annotations to semantic descriptions. This approach allows us to create a distributed RDF Knowledge Base integrating the various annotations first embedded within sensors and then transformed into enriched descriptions through the Paraimpu middleware. We focus our attention on sensor network components such as *stations*, *sensors* and *observations*.

4 Semantic Sensor Networks

In our vision, a SSN is the set of devices, semantic meta-modeling formalisms and RDF Knowledge Bases. To address our purposes we use state-of-art ontologies while giving due consideration to the specific concepts and properties. As a result, we build the SSN Knowledge Base containing all the semantic descriptions. To implement our methodology, we apply the proposed technique to real-world devices. We generate semantic annotations, the so called *snippets*. Snippets are plain JSON objects and are embedded directly within devices. Snippets provide information concerning the sensing providers and the kind of sensing measures. We address two different cases: *(i) sensing features*; *(ii) sensing data*. In case of *sensing features*, a snippet is composed of an individual and an observed property. The individual identifies the specific sensor, while the observed property refers to the measure. Individuals become mainly from state-of-art ontologies, SSN [4] and CF⁷. In many cases we need to define our specific individuals. As a common convention we use specific namespaces inherited from state-of-art ontologies, whereas the base namespace is defined as “:”.

5 The Paraimpu Middleware

In this work, a dynamic and scalable SSN system is based on the Paraimpu platform [8,9]. Platform main tool is the workspace. The workspace allows us to quickly prototype and deploy the whole system, focusing on the semantic aspects of research and delegating all the low-level, network operations to Paraimpu. The Paraimpu platform simplifies the connection, and the resulting collection of data, of an heterogeneous range of devices, applications and general data sources. By its API, almost any smart device able to communicate through the HTTP protocol can be connected in few simple steps, including Arduino boards sporting sensors and actuators to mobile applications, social networks and lighting systems. After the smart devices (from now we will refer to them as

⁷ <http://www.w3.org/2005/Incubator/ssn/ssnx/cf/cf-property>.

Things) are connected to the platform they acquire a virtual peer on the visual Paraimpu workspace which allows them to be easily linked (inter-connected) with other Things. Adopting some abstraction paradigms, all things in Paraimpu can be considered as belonging to two main categories: *Sensors*, or any Thing capable of measuring and producing data, physical or not (e.g., temperature or humidity sensors, social networks or external API and web services); *Actuators*, or any Thing able to receive data and perform an action, such as a physical, connected device or another web service or API. On Paraimpu, a *connection* is a logic data flow between two things. When two things are interconnected, it is possible to write and set a variable number of rules for each separate connection. Rules can apply filters or transformations on data. The powerful, integrated, rule-based Javascript engine inside the platform allows the system to transform data coming from a heterogeneous range of Things in real-time such as raw measurements and packets, or structured Linked Data, in compliance with Semantic Web standards. Thanks to Paraimpu APIs it is then possible to retrieve transformed/semantic annotated data to be stored in a dedicated Knowledge Base.

6 Case Study

Raw data sources used in the experimentation includes several physical sensors deployed at CRS4 buildings along with others data sources, like weather streams and traffic information, queried in near real-time using the open-data provided by the Cagliari Municipality Government and released under the IODLv2.0 License. This section reports some details of the overall case study, including the deployment of the system.

Referring to Fig. 1, the Paraimpu platform acts as a middleware playing a very central role in the general architecture. Not only does it expose all the required API to let Things gather or push data, but also it provides the facilities to transform and annotate semantically data on-the-fly as it flows from them. Our SSN deployment using Paraimpu middleware is shown in Fig. 1. The numbered main modules/steps are as follows: (1) Physical things have been built using Arduino Yun programmable boards. The boards were equipped with temperature and luminosity sensors and connected to the Internet through WiFi. They were deployed inside the CRS4 HQ buildings and the software running on them allowed to announce their presence on the network sending data payloads to Paraimpu (i.e., temperature and luminosity values). The communication happens through the Paraimpu REST API which accepts sending data in JSON format. Other Things pushing heterogeneous data to the middleware included: city weather data, gathered through OpenWeatherMap⁸ and traffic data measurements at several locations in the city of Cagliari, through their Open Data providers⁹. (2) All data sources are connected, managed and interconnected with ease by the facilities provided by the Paraimpu workspace. Incoming data, about network announcements and observations, are represented inside the Paraimpu

⁸ <http://openweathermap.org/>.

⁹ http://www.comune.cagliari.it/portale/it/openservice_info.page.

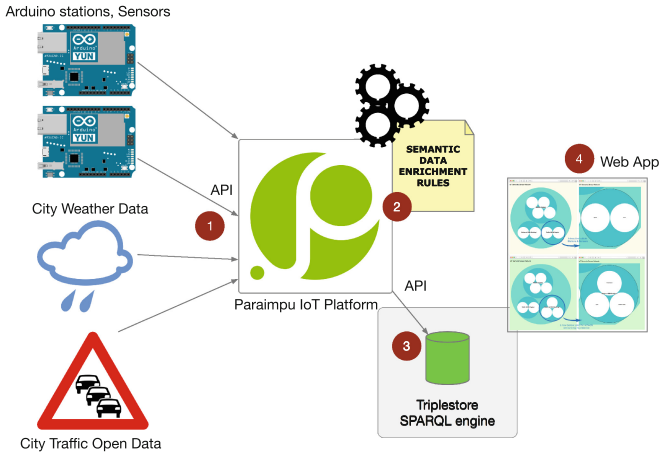


Fig. 1. The SSN deployment using the Paraimpu IoT platform

workspace as *Sensors* (see Sect. 5). Each of these sensors are connected to a corresponding *Actuator*. The *Registry* is a special *Actuator* that receives all the data concerning SSN network announcements. Thus, all Things are connected and all the connections are properly configured by *rules*. Rules are basically regular JavaScript expressions. They instruct the system about how to filter or transform data as it flows in near real-time from one Thing to another. We took advantage of this mechanism and wrote several semantic data enrichment rules to transform raw data, coming from direct measurements, to Linked Data formats to be sent to the actuators. (3) Every *Actuator* registered in the Paraimpu platform can be reached by a REST API endpoint. Using APIs, we wrote an external service to gather semantic annotated data from the platform and to store it in a Triplestore. (4) As soon as data is stored in the Triplestore, a query-based event triggering is fired and a web-page automatically reflects the current status of the SSN, including sensor stations presence and related measurements. The web page shows an interactive bubble view of the SSN which updates in real-time through the websocket technology.

In this case study, the sensor sends an announcement annotated snippet to Paraimpu. The following example shows the snippet related to a temperature sensor:

```
{ "whois": "sensor_Temperature_1",
  "observes": "cf-property:air_temperature" }
```

The snippet is composed of two key/value pairs. The *whois* key identifies the sensor, while the *observes* value addresses the observed property.

Paraimpu transforms the snippet into a semantic description and sends it (through the JSON-LD notation) to a connected *Actuator* representing our *Registry*. Data transformation/annotation is applied in real-time by a defined Paraimpu Rule (see Fig. 2). The semantic descriptions stored in the Triplestore

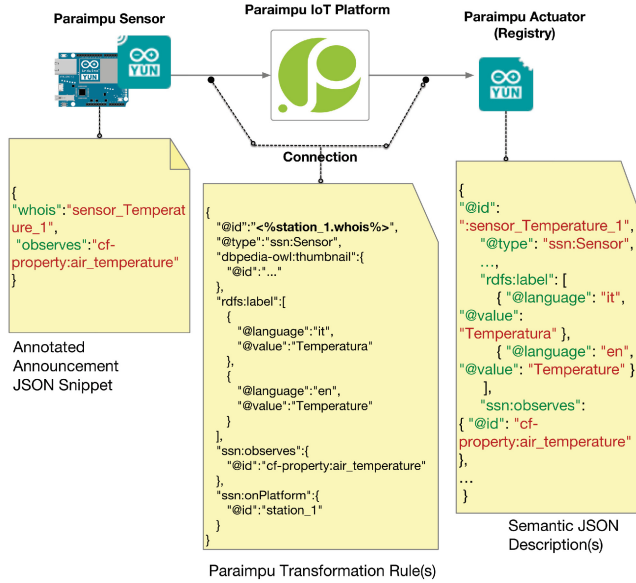


Fig. 2. Deployment of the SSN through Paraimpu sensor, actuator and connection abstractions, in case of a sensor announcement in the network.

compound a wide LOD knowledge base that allows to discover the interlinked information of the SSN.

Actually, a Rule can be seen as a generic data template, in which every snippet surrounded by the specific `<%%>` tag is dynamically executed as JavaScript code. Thus, the snippet can refer to data coming from a sensor to reduce the template to the sought real data.

7 Conclusion

In Sensor Networks, *Things* are active participants in information processes. In this paper we used semantic technologies for providing unambiguous data modelling, describing *Things*, sharing and integrating information. We discussed how to benefit from Semantic Web technologies in SSN. In our approach, devices are active parts of heterogeneous networks, able to announce themselves and to provide sensing data. Data originating from devices will be part of the world of Linked Sensor Data, through distributed RDF Knowledge Bases. To move from devices to Knowledge Bases we adopted the Paraimpu platform. The current Paraimpu connection model doesn't allow to group sensors together, thus the defined rules must be replicated in each new connection, event though they respond to similar aims. Future works will investigate how to avoid duplication of rules in connections sharing the same goal. In our vision, future networks will be able to exploit novel approaches of sensor analysis that infers semantic properties

such as the type of observed property, using the raw sensor observations as input. In a SSN, the network, its sensors and the resulting data can be organized, managed and queried through high-level specifications.

References

1. Barnaghi, P., Presser, M.: Publishing linked sensor data. In: The 3rd International Workshop on Semantic Sensor Networks 2010 (SSN 2010) in Conjunction with the 9th International Semantic Web Conference (ISWC 2010) (2010)
2. Barnaghi, P., Sheth, A., Henson, C.: From data to actionable knowledge: big data challenges in the web of things. *IEEE Intell. Syst.* **28**(6), 6–11 (2013)
3. Botts, M., Percivall, G., Reed, C., Davidson, J.: OGC Sensor Web Enablement. The Open Geospatial Consortium (2008)
4. Compton, M., et al.: The SSN ontology of the W3C Semantic Sensor Network incubator group. *Web Semant. Sci. Serv. Agents World Wide Web* **17**, 25–32 (2012)
5. Kebler, C., Janowicz, K.: Linking sensor data - why, to what, and how? In: Taylor, K., Ayyagari, A., Roure, D.D. (eds.) *Proceedings of the 3rd International Workshop on Semantic Sensor Networks, SSN*, Shanghai, China, 7 November, CEUR Workshop Proceedings, CEUR-WS.org, vol. 668 (2010)
6. Mineraud, J., Mazhelis, O., Su, X., Tarkoma, S.: A gap analysis of Internet-of-Things platforms, *Computer Communications*, Special issue on the Internet of Things: Research challenges and Solutions, March 2016
7. Patni, H., Henson, C., Sheth, A.: Linked sensor data. In: *Proceedings of the 2010 International Symposium on Collaborative Technologies and Systems (CTS)*, Chicago, IL, USA, pp. 362–370 (2010)
8. Paraimpu, a web platform for the Internet of Things. <https://www.paraimpu.com>
9. Pintus, A., Carboni, D., Piras, A.: Paraimpu: a platform for a social web of things. In: *Proceedings of the 21st International Conference Companion on World Wide Web*, pp. 401–404. ACM, Lyon (2012)
10. Ryu, M., Kim, J., Yun, J.: Integrated semantics service platform for the Internet of Things: a case study of a smart office. *Sensors (Basel, Switzerland)* **15**(1), 2137–2160 (2015)
11. Sheth, A., Henson, C., Sahoo, S.: Semantic sensor web. *IEEE Inter. Comput.* **12**(4), 78–83 (2008)
12. Wang, W., De, S., Cassar, G., Moessner, K.: Knowledge representation in the Internet of Things: semantic modelling and its applications. *Automatika J. Control Meas. Electron. Comput. Commun.* **54**, 388–400 (2013)