

# The Technology and Implementation of Automatic Identification of Sensors Based on Semantic IoT

Dufang Fan <sup>a)</sup>, Xiushan Zhang <sup>b)</sup> and Jian Huang <sup>c)</sup>

*Naval University of Engineering, Wuhan 430033, China*

<sup>a)</sup>596344273@qq.com, <sup>b)</sup>xiu3zh@139.com, <sup>c)</sup>275970675@qq.com

**Abstract.** With the development of semantic Internet of Things, there are many smart applications and services springing up in different domains. However, the gap between semantic representations and various sensor data formats used in IoT devices raises doubts and challenges for us. This paper aims to convert general data formats into unified SenML (Sensor Markup Language) and improves the algorithm about transforming SenML data into RDF (Resource Description Framework), a standardized semantic model. We exploited the knowledge-based system and designed the Internet of Things(IoT) structure of “smart campus” with LoRa technology, one of Low Power Wide Area Network(LPWAN). Such a proposition and application facilitate intelligent functions in IoT, mainly including semantic interoperability and scalability. In the end, an IoT system is presented to illustrate the usability of the whole approach and designed the structure of database storing RDF data.

**Key words:** SenML, RDF, LoRa, The Internet of Things.

## INTRODUCTION

The Internet of Things(IoT), also called the Internet of Everything or the Industrial Internet, is a new technology paradigm envisioned as a global network of machines and devices capable of interacting with each other [1]. At the beginning, the system structure of IOT mainly includes the types: one is based on the structure of sensor networks, the second is RFID based on EPCG Globle architecture, and the last is based on M2M(machine-to-machine) architecture. The Internet of things can also be divided into four layers: the perception recognition layer, the network construction layer, the management service layer and the comprehensive application layer.

In spite of its widely-spreading emergence, IoT is still in its infant stage and has a huge potential for research in variety of issues like standards, scalability, heterogeneity of different devices, common service description language, domain specific service discovery, integration with existing IT systems, and so forth [2].

The challenge of the Web of Things is how to build a semantic Web of Things (SWoT) in order to ensure a common understanding. The SWoT is, on the one hand, the fusion of the trends of the IoT for moving towards the web technologies with protocols such as CoAP, REST architecture and the Web of Things concept; on the other hand, it is the evolution of the web with the semantic web technologies [3]. Semantic technology means that different resources provide a unified way of data description to make machine recognizable and understandable, so that the computer can display and disseminate information to the user. Semantic interoperability refers to the access and parsing of mutual data among different resources based on semantics.

In SWoT, another significant thing should be considered that information of heterogeneous IoT devices in sensor network can be converted into uniform data format, and then data will be transformed into message processing engine in IoT to deal with the data.

The primary contribution of this paper is that the utilization of RDF (Resource Description Framework) describes various sensor data and through LPWA technology-LoRa data can be transmitted by a node is typically received by multiple gateways. The second contribution is the introduction to how SenML (Sensor Markup

LanguageSensor Markup Language) language is transferred to RDF. In this paper, the fusion of the two key technologies is finally applied to the Information integration system of classroom based on Intelligent Campus.

The rest of the paper is organized as follows. Section 2 describes the related research status at home and abroad. Section 3 is basic concepts and techniques involved and the description of transforming SenML to RDF. Section 4 tells the design and implementation of the Information integration system of classroom based on Intelligent Campus, and Section 5 gives the conclusion and discusses the future work.

## RELATED WORK

In view of the disunity of data representation and transmission standards or the inability of single data resources to meet the requirements of the application, the relevant research institutions have proposed that the data of sensor networks should be published as linked sensor data. At present, some foreign scholars have done comprehensive research on link sensor data, but there are few related studies in China.

Heiko Muller and others have built a cloud platform of semantic sensor called SERAW [4], and it is used to add some semantic information to sensor data and to find information related to sensor data. The sensor ontology of the platform is built on the basis of the universal sensor ontology. It can label the sensor data in JOSN format, and the ontology also has some specific concepts, which can improve the adaptability of the system. The platform also has the function of Linked Open Data Cloud (LOD Cloud) [5], and it runs through the construction of some functional functions to find the relevant information. But it has a certain limitation, and the versatility is not strong. Furthermore, Cuan LinNa invented Link sensor data publishing system [6]. It makes the application not only understand and utilize sensor network data, but also obtain more related resources based on the resource description framework (RDF) link between linked sensor data sets.

Despite the significant number of existing tools, incorporating data from multiple sources and different formats into the Linked Open Data cloud remains complicated. Another solution to the question is the conversion form various sensor data formats into semantic representations. Anastasia Dimou and others proposed how to map such heterogeneous sources into RDF in an integrated and interoperable fashion and introduced The RDF Mapping language(RML) [7].

## THE PROPOSED SCHEME

This section briefly reviews conceptions of RDF and LoRA, which are essential components of query processing. Besides, an algorithm is introduced about the description of transforming SenML to RDF.

### RDF

RDF is a W3C recommendation which has gained popularity as a mean of expressing and exchanging semantic metadata [8]. RDF is designed as a general language and method providing a description of the information data of Web resources, and it makes various data readable and understandable by computer applications. RDF is the data model of the relationship between network resource object and other objects, and it have simple semantics, following the syntax, structure of XML and available language writing of XML. The design of RDF is based on the ideology: the described resources have some attributes, each of which has specific attribute values. Resources refer to all the objects that can be identified by URI on the web. the attribute values of resources can be resources or literal values. If the attribute value is a resource, the corresponding attribute can be understood as the relationship between the two resources; if the attribute value is a literal type, the attribute is a description of the resource property. The description of resources is a statement of the attributes of resources and attribute value. We call them the statement. The basic structure of statements is a three tuple (resource-attribute-attribute value), also called the three tuples of "subject-predicate-object". A statement is used to describe the value of the property that a resource has. Such as:

(Pride and Prejudice) ([http://vip. book. sina. com. cn/book/index\\_37590. html](http://vip.book.sina.com.cn/book/index_37590.html)), its author is Austen, Jane. The following three tuples can be used to describe:

([http://vip. book. sina. com. cn/book/index\\_37590. html](http://vip. book. sina. com. cn/book/index_37590. html) foaf: author "Austen, Jane".)

the subject of the statement: [http://vip. book. sina. com. cn/book/index\\_37590. html](http://vip. book. sina. com. cn/book/index_37590. html); the predicate of the statement: foaf: author; the object of the statement: Austen, Jane

Among them, foaf is a prefix and represents a namespace that is [http://xmlns. com/foaf/0. 1/](http://xmlns.com/foaf/0.1/).

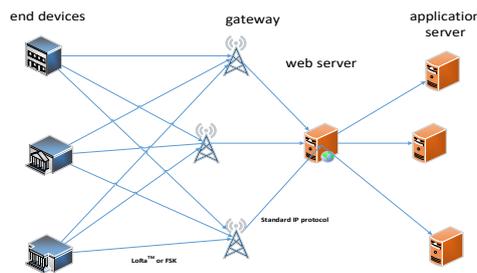
Defining and executing a mapping with rml requires the user to provide a valid and well-formatted input dataset to be mapped, and the mapping definition (mapping document) will be executed to generate the data's representation using the RDF data model (output dataset) [8].

## LoRa

New transceiver technologies have emerged which enable power efficient communication over very long distances. Examples of such Low-Power Wide-Area Network (LPWAN) technologies are LoRa, Sigfox and Weightless. LoRa (Long Range) is a proprietary spread spectrum modulation technique by Semtech. It is a derivative of Chirp Spread Spectrum (CSS). The LoRa physical layer may be used with any MAC layer; however, LoRaWAN is the currently proposed MAC which operates a network in a simple star topology [9].

Nodes transmit directly to a gateway which is powered and connected to a backbone infrastructure. Gateways are powerful devices with powerful radios capable to receive and decode multiple concurrent transmissions. As shown in Figure 1, all nodes are bidirectional communication, but uplink communication is dominant. The gateway communicates with the server through standard IP protocol, and the terminal device is connected to one or more gateways by single-hop LoRa™ or FSK.

LPWA has solved the problem of long distance and low power connection in the Internet of things. As the representative of LPWA, LoRa technology has the advantages of low power consumption and wide coverage. It is very suitable for the long-distance application scenarios needing battery powered in IOT. At present, LoRa has been widely applied to various IOT industry. It is believed that LoRa will certainly be prosperous in the development of future IOT.



**FIG. 1.** Network architecture of LoRa

## Transforming SenML Data to RDF

SenML is a sensor markup language that aims to simplify gathering data from different devices across the network. It simply is JSON containing named events together with an associated value and unit. The design goal is to be able to send simple sensor measurements in small packets on networks from large numbers of constrained devices.

SenML provides a simple model for retrieving data from sensors and controlling actuators. Transforming SenML to RDF facilitates intelligent functions in IoT, including reasoning over sensor data and semantic interoperability among devices.

The transformation allows connecting SenML-enabled IoT sensors to knowledge-based systems with minimal code changes. Sensors can utilize SenML without any additional computation to prepare the data and a simple parsing component can be employed at a knowledge-based system for transforming SenML into RDF. This component can be physically deployed in a gateway or a server machine of IoT systems. An algorithm for implementing such a component has the following steps [10].

Table 1 presents the mechanism to utilize URIs for assigning unambiguous identifiers to SenML elements. This table shows the mapping from SenML elements, their shorthands in JSON, and their corresponding types when transformed into RDF.

Following algorithm is specific.

Step 1. Converts SenML elements to their corresponding uniquely identified elements, usually URIs and characters. If a prefix or base name is defined in the SEMML document, it should be associated with the element name. This ensures that resources, attributes, types and values are given their complete representation.

Step 2. Restructure the SENML document to the RDF triples. RDF containers, RDF sets, etc. are introduced when needed.

Step 3. Serialize the RDF triples, and the results can be XML, N3 and JSON-LD. If required, define the XML namespaces and shorthands.

**TABLE 1.** Transforming SenML Elements to RDF

SenML Elements	JSON Shorthands	Types in RDF
Base Name	bn	URI(Subject)
Base Time	bt	xsd:dateTime
Base Units	bu	Xsd:int
Version	ver	Xsd:int
Measurement or Parameters	e	- (RDF Triples)
Name	n	URI
Units	u	Xsd:string
Value	v	Xsd:float
String Value	sv	Xsd:string
Boolean Value	bv	Xsd:boolean
Value Sum	s	Xsd:float
Time	t	Xsd:dateTime
Update Time	ut	Xsd:dateTime
Resource Type	rt	URI(rdf:type)

The algorithm is suitable for SENML in JSON and XML formats. When SenML has EXI format, it is necessary to convert data from EXI to XML. Some existing libraries can be used to perform such transformations in a simple way.

Figure 2 shows an example of the SEMML data produced by a micro environmental sensor used in the classroom IOT application, which shows the temperature and humidity in the environment around the sensor. The device has device ID TEMSALSENSORNO011 (with "bn") and environmental node of resource type (with "rt"). In order to solve the potential contradiction in the global level IOT system, the prefix element ("PR": "http://oto. y/o\*") is used in the namespace of the sensor. Similar to the resource type, if the IOT sensor and the IOT application use data to default this type of optimization, there is no need to transmit prefix symbols in each SENML description. Figure 2 shows the form of RDF/XML that can be converted to SEMML data. It should be noted that the measured values are in the XML mode data type, and the units of the measured values are defined separately. Figure 3 shows the corresponding RDF diagram of the RDF/XML described in Figure 4. It is common for an SEMML description to be converted into multiple RDF statements.

In this article, we do not study how to convert RDF statements into SENML, although information can be sent to IOT devices in this way. RDF is more expressive than SEMML, so when all the capabilities of RDF are taken into account, understanding information is too complex for many resource constrained IOT devices.

```

{"e": [
  {"n": "temperature", "u": "Cel", "v": 25}
  {"n": "humidity", "v": 58%}
  "bn": "http://iot.fi/o#temsalsensor009/"
  "rt": "EnvironSensingNode"
  "pr": "http://iot.fi/o#"
}]

```

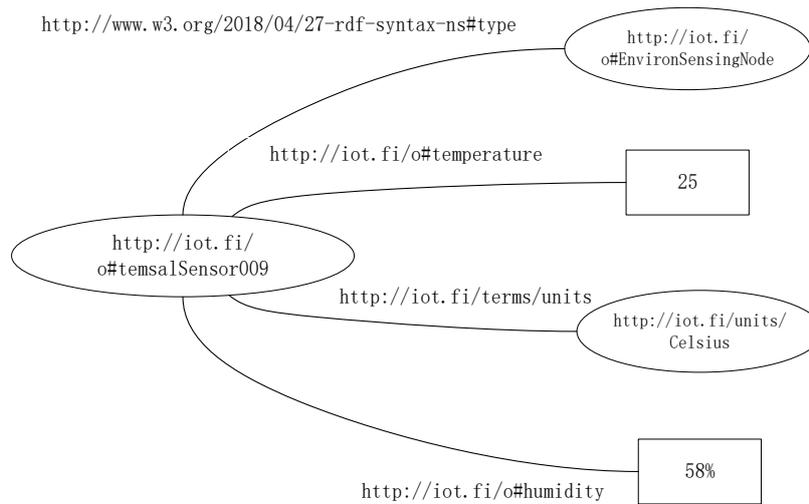
**FIG.2.** An Example of SenML Data Produced by an Environmental Sensor

```

<rdf:RDF xml:base="http://iot.fi/o"
xmlns:i="http://iot.fi/o#"
xmlns:rdf="http://www.w3.org/2018/04/27-rdf-syntax-ns#"
xmlns:iotterms="http://iot.fi/terms/">
  <i:EnviromenSensingNode rdf:ID="temsalsensor009">
    <iotterms:units rdf:resource="http://iot.fi/units/Celsius"/>
    <i:humidity>58%</i:humidity>
  </i:EnvironSensingNode>
</rdf:rdf>

```

**FIG.3.** Corresponding RDF/XML representation the SenML data shown in Figure 1



**FIG.4.** RDF Graph of the Marine Sensor Data Example

## DESIGN AND IMPLEMENTATION

The so-called "smart campus" refers to the intelligent campus work, learning and life integrated environment based on the Internet of things. This integrated environment runs all kinds of application service systems as the carrier, and integrates teaching, scientific research, management and campus life fully. In recent years, with the rapid development of information technology and Internet of things technology, numerous colleges and universities are developing intelligent campus project, using information exchange and sharing technology, eliminating the phenomenon of isolated information, realizing information sharing; Campuses usually use RFID, two-dimensional code, video surveillance and other physical awareness technology to construct campus intelligent perception environment; LDAP access technology is also adopted to realize unified authentication of the identity of the school.

The basic function of the intelligent laboratory is to automatically control the switch of lighting, air conditioning, multimedia and other equipment in the laboratory, and control the power supply and anti-theft alarm of the experimental instrument.

### The Intelligent Gateway Layers

In the Fig. 5, there are two functions of the intelligent gateway layer. One is the equipment used to access the underlying intelligent sensing layer and provide different hardware interfaces and communication interfaces for different acquisition devices. This needs to know which hardware devices, hardware interfaces and communication interfaces are used in the system. Another function of the intelligent gateway layer is to provide parsing service for the LoRa transport protocol. The two functions of the intelligent gateway layer, one is the equipment used to access the underlying intelligent sensing layer and provide different hardware interfaces and communication interfaces for different acquisition devices. This needs to know which hardware devices, hardware interfaces and communication

interfaces are used in the system. Another function of the intelligent gateway layer is to provide parsing service for the LoRa transport protocol.

### **The Data Processing Layer**

In order to solve this problem, the data processing layer must be set up to unify the data format and the data structure because of the differences in the data format and the difference of data structure. The layer mainly has the knowledge-based system, functioning the algorithm of section 3.

In the data link layer, this agreement is mainly from the subordinative structure of half duplex communication mode. The handheld unit or other data terminal is the main station, and the multi-function electric energy meter is the slave station. Each multifunctional electric energy meter has its own address code. The establishment and release of communication links are controlled by information frames issued by the main station. Each frame consists of 7 domains (Table 2): frame initiator, slave address domain, control code, data domain length, data field, vertical check code of frame information and frame terminator. Each part is made up of a number of bytes.

The meter data follows the dlt645-2007 statute as specified by the national electricity meter industry, such as following Voltage data:

After removing the header byte, command code and packet length, it is the sensor data. The data packet contains physical address, network address, node type and so on. Key value is 54c569, from the 36th bit on. Hexadecimal needs to be converted to decimal. Through the analysis, the voltage value is 230V. This process is called data format conversion in the fig. 7.

After data format conversion, the sensor data is disposed by Data format unification. The process is mainly packaging and somatization of data including converting to SenML format and RDF format. Now we have got the data and formatted the data. In common sense, we can directly transfer the converted data to the application layer to satisfy our design but considering that it is unnecessary to deal with different data format in the application, we should do data format unification. In order to optimize this problem, we have made the following conventions. On the basis of table 1, the sensor is converted to SenML format and then converted to RDF format based on algorithm in the section 3.

Figure 6 shows a IOT architecture of the intelligent laboratory. Environmental sensors and other devices are deployed for monitoring the temperature and humidity of laboratory conditions. These devices are connected with observatory networks wirelessly and send measurements to an intelligent gateway in SenML format or other formats. This gateway fuses the data and forwards it to a knowledge-based system. The knowledge-based system located in Data processing layer transforms SenML measurements into RDF statements (with the algorithm presented in Section 3) and integrates these statements into a domain ontology for reasoning. Real time temperature alert or electricity alert will be sent to data base for restoring, and client application of the user, and then they can employ instruments when necessary. Moreover, data management and visualization tools are utilized to store and visualize the data. The data is transferred to the information processing engine of the Internet of things to process the data. The information processing engine of the Internet of things feedback the processing results to the IOT relationship management component, and realizes the function of the management, monitoring and real-time configuration of the Internet gateway through the component. Among them, the data bus of IOT provides the corresponding communication interface for the communication between other components and is the binder of other components.

The sensing device information of the sensing network is transformed into a unified data format, and the data is transferred to message processing engine of IOT for processing data. Take the electricity meter as an example, the "electric power industry standard of People's Republic of China" stipulates that Multi-function watt-hour meter communication protocol.

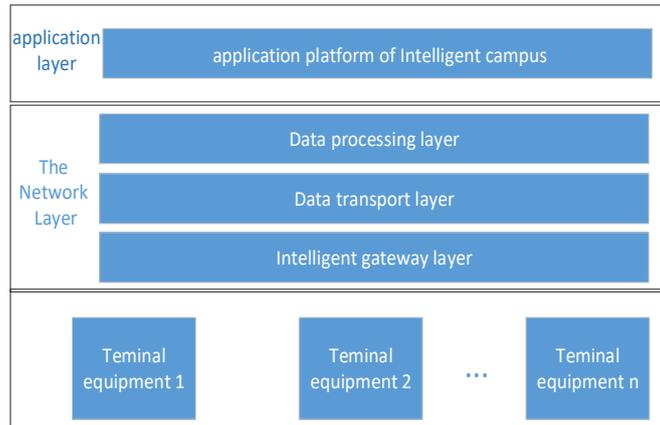


FIG. 5. Infrastructure of Physical network system

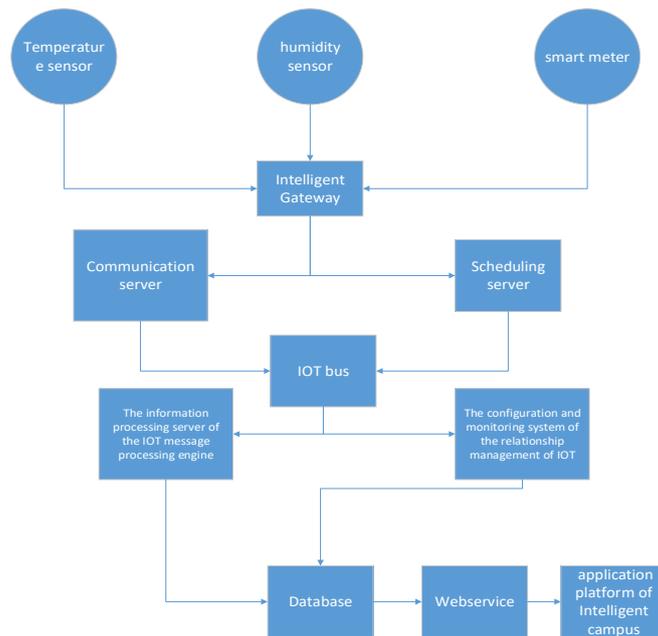


FIG. 6. An IOT Architecture of the Intelligent Laboratory

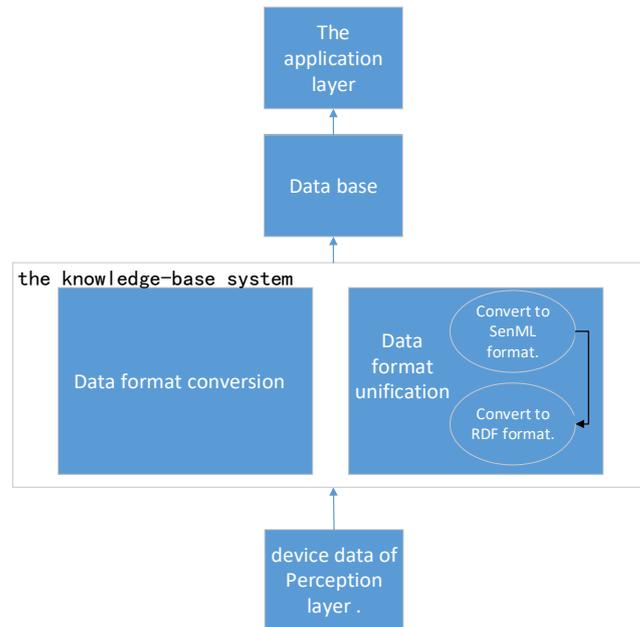


FIG. 7. Data processing module structure diagram.

TABLE 2. Frame Format

Illustration	Code
frame initiator	68H
address domain	A0
	A1
	A2
	A3
	A4
	A5
frame initiator	68H
control code	C
data domain length	L
data field	DATA
check code	CS
frame terminator	16H

### RDF Storage Solution

The basic structure of the vertical table database is to use a table to store all RDF assertions. Each record in the table corresponds to an RDF three tuple, which is called the "Statements". However, because the description of a

resource in RDF data generally contains multiple statements, or a certain attribute may be used to describe a number of different resources, resulting in the repeated storage of the same information, which greatly wastes the storage space. Nowadays, a large number of the RDF data in the Internet has higher demand for storage space. To solve this problem, the resource table (URI), the predicate table (Predicate) and the namespace table (Namespace) are introduced. The namespace table stores the prefix part of the predicate, and the resource table (URI) stores the URI reference.

The structure of the table is as Fig. 8.

Resource table (URI): an RDF resource (a stated subject or a resource-dependent object) that stores all URI references; because a resource in RDF may have multiple attributes, there are multiple statements to describe the resource and store all the resources in a single table; and only the foreign key is used in the statement table.

Namespace: describes the properties of RDF resources, usually composed of namespaces and local names.

Predicate: stores all attributes in the RDF file, which contains namespaces and local names. Namespace is foreign key.

Statement: it is the main table stored in RDF file, which stores all three tuples in RDF/RDFS.

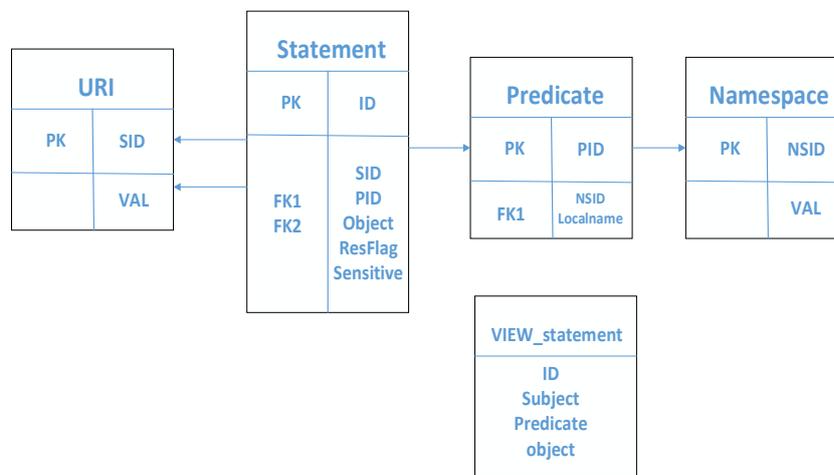


FIG. 8. Storage table structure of RDF

## CONCLUSION

In this paper, we introduced an approach for connecting SenML-enabled sensors to semantic knowledge-based systems and apply the algorithm to the project of “smart campus”. The details of this approach, design considerations and a real use case study integrated with LoRaWAN are introduced. In the end, we still sketched storage table structure of RDF in the database. By connecting a large amount of resource-constrained IoT sensors utilizing SenML to Web and Semantic Web, this approach can enable a large variety of intelligent IoT applications in different domains. The aim is that IoT applications take the benefit of semantics and Web technologies all the way, even though many objects are resource-constrained.

Even though SenML is a flexible data format, not all sensors adopted SenML language. Predecessors have not solved the problem. In this paper, we considered how the system recognizes different data format of sensors and how to build the whole “smart campus” system. This paper provided a decent example and solution for automatic identification of different sensor in IOT and promote the development of semantic IOT. The process can realize the feature of self-discovery, self-transmission, and no-configuration. Consistency, completeness is also included in the system.

In the future, a more complete algorithm and codes about conversion of different resource should be figured out. Next, the efficiency of conversion processing can be improved. A possible optimization is the use of execution plans that efficiently arrange the execution order depending on their dependencies. Finally, reliability is reflected more, and sub-maintenance should be considered, too

## REFERENCES

1. In Lee, Kyoochun Lee. The Internet of Things (IoT). Applications, investments, and challenges for enterprises, *Business Horizons*. Vol. 58(2015) Issue 4, p. 431-440.
2. Sohail Jabbar, Farhan Ullah, Shehzad Khalid, et al. Semantic Interoperability in Heterogeneous IoT Infrastructure for Healthcare. *Wireless Communications and Mobile Computing*. Vol. 2017, p. 10.
3. Antonio J. Jara, Alex C. Olivieri, Yann Bocchi, et al. Semantic Web of Things: an analysis of the application semantics for the IoT moving towards the IoT convergence. *International Journal of Web and Grid Services*. Vol. 10(2014) Issue 2-3.
4. Miiller H, Cabral L, Morshed A, et al. From RESTful to SPARQL: A Case Study on Generating Semantic Sensor Data. *Semantic Sensor Networks @International Semantic Web Conference*, 2013, p. 51-66.
5. Jentsch A. Linked open data cloud. *Linked Enterprise Data*. 2014, p. 209-219.
6. Cuan Linna, Shi Yiming, Li Guanyu, et al. Generation system of link sensor data generation system in semantic Internet of things. *Computer application*. Vol. 35(2015) No. 09, p. 2440-2446.
7. Dimou, A., Vander Sande, M., Colpaert, et al. RML: a generic language for integrated RDF mappings of heterogeneous data. In: *Proceedings of the 7th Workshop on Linked Data on the Web*. CEUR Workshop Proceedings, vol. 1184 (2014).
8. B. H. Song, B. Lee, K. T. Kim and H. Y. Youn. Enhanced query processing using weighted predicate tree in edge computing environment. *2017 IEEE Conference on Standards for Communications and Networking (CSCN)*. Helsinki, 2017, p. 48-53.
9. Bor, Martin, Vidler, et al. LoRa for the Internet of Things. In: *EWSN '16 Proceedings of the 2016 International Conference on Embedded Wireless Systems and Networks*. 2016, p. 361-366.
10. Xiang Su, Hao Zhang, Jukka Riekk, et al. Connecting IoT Sensors to Knowledge-based Systems by Transforming SenML to RDF. *Procedia Computer Science*. Vol. 32(2014), p. 215-222.