Proposal for Knowledge Model Using RDF-based Service Control for Balancing Security and Privacy in Ubiquitous Sensor Networks

Makoto Sato*, Yoshimi Teshigawara**, and Ryoichi Sasaki***

*Graduate School of Advanced Science and Technology, Tokyo Denki University, Japan
**Cyber Security Laboratory, The Research Institute of Science and Technology, Tokyo Denki University, Japan

{sato_m, teshiga}@isl.im.dendai.ac.jp, sasaki@im.dendai.ac.jp

Abstract - In ubiquitous sensor networks, various sensors and tag readers automatically collect information in space and relevant information is acquired. Efficient utilization of the acquired information is important for providing high-quality services that meet the users' privacy requirements. We use RDF triples that represent spatial information at the granularity of the requested security levels. In earlier work, we created a knowledge model that considers privacy by representing user information hierarchically, and we verified its feasibility by a simulator that we developed. Then, we extended this knowledge model. In this paper, we discuss our newly proposed extended knowledge model and its applicability to various spaces. In addition, we evaluate the feasibility of the model by using a test simulator that we developed.

Keywords: Security and Privacy, Knowledge Model, RDF, Semantic Sensor Network Ontology, Sensor Network.

1 INTRODUCTION

In ubiquitous sensor networks, various sensors and tag readers automatically collect information in space and relevant information is acquired. It is expected that the amount of information in the sensor network space will further increase due to advances of these networks. It is also expected that personal information on users will be presented with various levels of granularity. For example, GPS can acquire rough location information, and cameras can acquire detailed location information. Efficient utilization of the acquired information is important for providing high-quality services that meet the users' privacy requirements. In this regard, it is possible to identify users' personal information by combining sensor information with user information that seems to be trivial by itself. Therefore, the risk of an indirect violation of privacy makes it difficult to provide high-quality services, because protecting the user's privacy means limiting the information obtained. Thus, it is necessary to consider privacy and security. Privacy requires protection from a third party and meeting the user's privacy requirements. Security requires that the network does not leak information to the outside. Terminals that use privacy information must preserve confidentiality by means such as encryption or anonymity protection, and must impose security measures to prevent privacy information leakage or tampering.

Our research objective is RDF-based service control for balancing security and privacy. In this study, privacy information is defined as information about a behavior of the user that the sensor collects. In this paper, we focus on privacy protection.

We have been developing a platform that integrates all the information in a space by using the Resource Description Framework (RDF). An RDF represents information about a resource (subject, predicate, object) in the form of an RDF triple [1]. An RDF triple is represented by a graph in Fig. 1. The RDF expresses the subject of the resources associated with the object through the predicate. By combining inference rules and a set of vocabulary, it is possible to connect different types of data and to make the aggregation of over the partial sums. RDF triples are represented with the granularity of any spatial information. Therefore, service control information or privacy information is represented flexibly. For this reason, using the information efficiently to provide a flexible service requires organizing the RDF triples of the control information and the service state information of the space required by each service.

On the other hand, protecting personal information requires collecting this information with restrictions and proper control [2]. In our previous study, we discussed only the use of restrictions, because collecting restrictions is outside our work scope. Thus, we define privacy protection as follows. Services are allowed to use only intended information on users. Sensors are not permitted to collect unintended information on users.

We previously proposed a knowledge model that can be applied to a platform using RDF-based practical services [3]. We created a knowledge model, which is a set of vocabulary required for expressing services provided by the RDF and analyzing the RDF obtained at that time. Because our knowledge model considers privacy by representing user information hierarchically, we were able to control user information by adding a function that reflected user requests [4]. In addition, we verified the feasibility of the knowledge model by developing a test simulator.

![Figure 1 Example of RDF Triple](image-url)

Makoto affiliation
located
University
Security Lab

ISSN1883-4566 © 2015 - Informatics Society and the authors. All rights reserved.
The knowledge model is represented by simple and common logic. The service provider benefits by verifying whether personal information is properly used when the service is under development. The user benefits by limiting personal information in accordance with the user's intention.

Our current work provides a level of service management for a particular space. That is, we extended the same service to a different service management system. In this paper, we discuss the applicability of our extended knowledge model to various services and various spaces, and we evaluate the feasibility of the extended model by using the test simulator.

2 RESEARCH BACKGROUND

2.1 Related Work

Various integrated management methods for sensor networks have been proposed [5]. Some of the studies represent sensor network information by using the RDF. Fujinami et al. represented a physical environment model by a location model and an object model using the RDF [6]. The location model is represented by relations between a unit space, such as a room and a building, and unit territories, such as an entrance and a kitchen. The object model is represented by object information, such as specification information and operating conditions. By using these models, developers can handle directly required information for a variety of applications. Held et al. represented user-specific information, such as user preferences, by using the RDF [7]. By evaluating context information and managing user profiles, the RDF allows for personalized, context-aware service mediation and content adaptation. Noguchi et al. managed sensor information by using the RDF to realize intelligent support systems in a room in a home [8]. In this case, the system needed a mechanism for automatically understanding information such as the sensor configurations of rooms. Therefore, they proposed an RDF sensor description to inclusively portray sensor information. It not only could describe the characteristics of the sensors, but could also easily realize an extension of the description in collaboration with other knowledge information, including new information. With these features, it allowed unified processing of sensor data. An example of the applied RDF description is the implementation of applications, such as component discovery in middleware. In our study, the service execution rules and the user requirements are centrally managed in the same way as the sensor information. Therefore, our model is expected to provide both high-quality service and protection of privacy.

Some research has discussed access control on the Web. Sacco et al. proposed the Privacy Preference Ontology that enables fine-grained access control [10]. This ontology has a vocabulary for defining fine-grained privacy preferences for RDF data. This ontology restricts a resource, a particular triple and a group of triples. By using this ontology, access control to privacy information is restricted by the properties that a requester must satisfy. Carminati et al. proposed an access control framework for social networks by specifying privacy rules using SWRL (Semantic Web Rule Language) [9].

Additionally, user/resource relations were modeled by using RDF/OWL (Web Ontology Language). Because the Web contains a lot of privacy information, access control is effective as a method of privacy protection. Similarly, privacy protection using the RDF in a sensor network was studied. Jagtap et al. investigated privacy protection by using the RDF [11]. They proposed a model for representing the user's environment, position, and activities. An important element of their study was the use of collaborative information among sharing devices, which share and integrate knowledge about the contexts of the collaborative information. Therefore, mechanisms for privacy and security were required. They used the RDF to specify a high-level declarative policy describing the settings for sharing user information.

Our study presents a framework to provide users with an appropriate level of privacy for a mobile device and to protect the personal information gathered, including personal information that can be inferred from other information. Our study assumes an environment where the mobile devices are owned by individuals, and sensors, such as camera sensors and positioning sensors, are placed in each location. Therefore, our model is expected to protect privacy while providing a variety of services.

2.2 Development of Platform

As described in Section 1, we have been studying a method to integrate all the information in a space by describing sensor information, user information, and service states for using the RDF [12]. We aim to control services in the sensor network space by using RDF triples to provide services and information corresponding to the users' requests. Furthermore, to provide a high-quality service and to protect the privacy information of the user by reflecting the user requirements into the usage rights of RDF triples, we have been developing a platform that uses the appropriate information that satisfies the users' requirements. Figure 2 shows an overview of the functions of this platform. The functions are to generate RDF triples from the spatial
information acquired from sensors and to select provided services based on the RDF triples. These processes are carried out in "the RDF triple generation rule management unit" and "the service execution rule management unit". In the RDF triple generation rule management unit, RDF triples are generated from the acquired sensor information based on the RDF triples from the RDF triple generation rules. Here, the generating rule for the RDF triples is managed as a set of rules, or triggers, for generating new and already generated RDF triples. Table 1 shows an example of a generated RDF triple rule. The service execution rule management unit selects the services that can be provided by checking the RDF triples passed from the service control unit to the service execution rules. In addition, the service providing service execution rules, the RDF triples that trigger the service, and the service execution instruction are managed as a single set of rules. Table 2 shows an example of a service execution rule.

### 2.3 Creation of Knowledge Model

As described in Section 2.2, spatial information is represented by an RDF. We define the vocabulary and the relations of spatial information as a knowledge model expressed by the RDF. To create a knowledge model that can provide a service, it must be created after stipulating the service requirements envisioned. However, the service that runs on this platform is not yet defined. Therefore, an effective approach is to create a primary knowledge model first and then extend it gradually.

The primary knowledge model is created with a clear description of the technical issues for practical use, while considering and evaluating the services as a prototype. Specifically, the resources required for the services are assumed. Next, the state transition of the resources is expressed by an RDF graph (a set of RDF triples). Then, the resources within the RDF graph are classified into sets of the same type. A knowledge model is created to represent the relation between sets. For example, the primary knowledge model is applied to a service of the same type. A new service concept is introduced when one is lacking. Thus, by extending the knowledge model, a more general knowledge model is created.

### 2.4 Development of Simulator

No real system has been developed to provide services by using the knowledge model created in Section 2.3. Because the platform includes ambiguous parts, such as the storage method of the service execution rules, we cannot clearly verify the feasibility of the knowledge model. Therefore, we developed a simulator to apply the knowledge model [13].

In the simulator, we developed several functions, such as input of RDF triples, introduction of new RDF triples by inferring, reflection of user requests, and selection of executable services. Jena was used for development of the simulator [14]. Jena provides a framework for processing RDFs, and an inference engine. Graphviz was used to visualize the RDF graph [15]. We demonstrated the operation of each function and verified the feasibility of a service control by the knowledge model [4].

One of the beneficial features of the simulator is a function for reflecting user requests in order to limit the information used in the sensor network space. In Fig. 3, which is intended for a university. In this figure, an ellipse represents a resource, and an arrow expresses a predicate. A feature of this knowledge model is that the domain corresponds to a subject, and the range corresponds to an object, shown as (predicate_property, domain, domain_name, (predicate_property, range, range_name). This RDF triple expresses a resource that is the subject of the relation and the object of the relation. Thus, when RDF triples are added to the RDF graph, the inference is that resources belong to a classification with a focus on the predicate [3]. In addition, by using a hierarchical representation of the affiliation information of the user, it becomes possible to restrict the use of privacy information [4]. We examined the flexibility of the service execution rules by an experiment using this knowledge model in a simulator [3].
2.5 The Need for the Sensor Concept and Collecting Restrictions

The main resources in the sensor network space can be divided into space, user and service. Space is divided into "environment" and "sensor". "Environment" is a place for providing services. For example, the environment is stations, a university or a home. "Sensor" obtains the spatial information. In the previous study, we also focused on the service control using the RDF, but it was limited to only "environment" in spatial information. We considered service control information as information directly related to the service. We did not discuss the sensor at that time. Therefore, as a next step, it is necessary to incorporate the concept of the sensor to create a more general knowledge model.

In addition to introducing the concept of the sensor, it is also necessary to consider again the restrictions on collecting information, as discussed in our previous study [4]. To meet the requirements of more users, the RDF triples acquired from a sensor only use information allowed by the user. In addition, RDF triples are not applied except for those uses. The results of the study are as follows.

3 EXTENDED KNOWLEDGE MODEL

3.1 Assumed Service

The assumed service is entry management in a university campus according to the affiliation information of a user. For example, one service is unlocking the entrance door if the user is enrolled in the affiliated faculty. We make the following assumptions. The service manager is able to attach the affiliation information for the user (faculty), and the user is able to specify the affiliation information for which the service is available. We considered the "environment" as a cafeteria, three buildings, five rooms in each building and the main gate to the university campus. Figure 4 illustrates the relation between entry permission and user affiliation. A user must belong to the university in order to pass through the main gate. Similarly, the faculty must belong to the university in order to enter a building. The faculty must also belong to the corresponding department in order to enter a classroom. For example, users can enter room A1110 if they have the affiliation information of faculty. They can enter the cafeteria if they have the affiliation information of the university. Each entry has a keycard system in conjunction with the entry permit information outputted from the service, and the RDF triples are assumed to have been inputted into the system in advance by the service administrator.

As described in Section 2.3, our RDF-based service extracts the information required for the service provision from the assumed service. Service execution rules are created by analyzing the information to trigger the service execution from the service contents. The service is then provided when the affiliation information for the user is inputted and room entry is allowed. For example, an RDF triple indicates whether a user can pass through the main gate of the university (university, entrypermit, maingate). In another example, the trigger for room entry is represented by the RDF triples (user, affiliation, Department of Information Media), (user, permit, Room A1110). Table 3 shows the service execution rule for room entry.
3.2 Privacy of Extended Knowledge Model

We created an extended knowledge model based on the assumed service described in Section 3.1 [4]. Specifically, the model contains the affiliation information for the user and the model of the university's sensor network space corresponding to "environment".

For the "sensor" that obtains spatial information, we used the Semantic Sensor Network Ontology (SSN) proposed by the W3C [16]. This ontology describes sensors, observations, and related concepts. For example, Sensor, Sensor Output, Sensor Input, and Device are basic resources of the sensor [17]. Therefore, we consider that these resources are sufficient as a primary model of the sensor.

It is necessary to incorporate the collecting restrictions, as described in Section 2.5. The collecting restrictions can be realized by sensors that are not permitted to collect unintended information on users. For example, a camera sensor can acquire position information, but a user who does not want to be recorded might feel that he or she wants to stop this camera sensor. If sensors are permitted to simply reflect the user's request, all sensors will be stopped. Thus, the services are likely to be provided frequently. Therefore, in terms of efficiency, we decided to stop the sensor only when all users in the space do not wish to collect the information. For this reason, to indicate the availability of the sensor, we added a predicate "hasAvailability". The predicate is the relation between the user and the sensor. In addition, a word with the prefix "ssn." indicates that it is a vocabulary of the SSN. Figure 5 shows the extended knowledge model based on this information.

Moreover, we needed to introduce new inference rules on the collecting restrictions. The collecting restrictions are realized by the following formulas using SWRL. If the user does not have the "hasAvailability" predicate, the simulator does not use the affiliation information of the user (Formula 1). If the RDF triple (user, hasAvailability, sensor) is not added, the sensor stops working (Formula 2).

\[\text{Affiliation}(?user, ?affiliationof) \land
\text{noValue}(?user, \text{ssn}.\text{hasAvailability}) \rightarrow \text{Permit}(?user, \text{ssn}.\text{noValue})\]  

\[\text{noValue}((\text{ssn}.\text{hasAvailability}), \text{down}) \rightarrow \text{Permit}(\text{Sensor}, \text{down})\]

4 SIMULATION EXPERIMENTS

This section describes our simulation experiments. The experiments were executed to verify that the extended knowledge model is able to protect privacy information by using the user affiliation information.

4.1 Experimental Environment

The environment is the same as that described in Section 3.1. We made the following assumptions for the experiment. Each user has a smartcard. The physical location of the university is the sensor network space. The physical location can be uniquely identified by the geographical coordinates of latitude and longitude. All locations in the university have the names of identified strings assigned by public authorities. Sensors are installed in the vicinity of the door or the gate for each location. Sensors used in this space are a camera sensor and a smartcard reader. The camera sensor obtains name information for the users present in the space. The smartcard reader obtains affiliation information for the users. The individual can then be authenticated by comparing the information in the smartcard and the information acquired by the camera sensor. The acquired sensor information is converted to RDF triples automatically. The difference from the previous experiment [4] is the sensor information and increased number of relations. The relations of entry permission and affiliation information for the user are shown in Table 4.

We assumed the following two scenarios in the experiments. The difference between the two scenarios is that sensor information input is added only in scenario A (Steps 2 and 3). Then, service execution rules and the knowledge model are assumed to be stored in the database in advance. The user is assumed to be a student of Faculty of Engineering at this university in the Department of Electronic Engineering.

Scenario A:

1) Input the initial state of the assumed space. Specifically, the service administrator enters RDF triples indicating the building permit and the space information into the simulator.
2) Enter the user requirements. The user selects the available user information and the available sensors and enters them into the simulator. All sensor and user information is supposed to be available at this time.
3) The acquired sensor information, such as affiliation information, is inputted in the simulator.
4) Generate new RDF triples to perform inference processing by using the input information.
5) Determine whether the entry management service can be performed by using an RDF graph.
6) Suggest the possible entry locations.

Scenario B:

A similar scenario is carried out, but this one does not use the available sensors in Step 2 above.

Figure 5: A part of the extended knowledge model

Experimental Environment

Figure 6: A part of the knowledge model

We evaluated the feasibility and flexibility of the extended knowledge model. In addition, we evaluated the feasibility of collecting restrictions to meet the user requirements.

4.2 Experimental Results

Here we present the results of the experiments described in Section 4.1, where the scenarios were executed by the simulator. Figure 6 illustrates a part of the result of Step 1. It can be seen that the smartcard and the camera sensor were associated with the SSN.

Figure 7 illustrates the result of Step 4 in scenario A. From this figure, it can be seen that by inference processing, many RDF triples were automatically generated. Figure 8 shows a list of the rooms where the user is permitted to enter.

Figure 9 shows the results of applying the inference rule collecting restrictions on the user request inputted in Step 2 in scenario B. The available information was eliminated because the user does not have the "hasAvailability" predicate. Figure 10 illustrates the result of Step 4. As compared with Fig. 7, Fig. 10 shows that the RDF graph is divided into two groups. This is because the affiliation information for the user is not bound to the entry permission information. Figure 11 shows a list of the rooms where the user is permitted to enter.

Table 4: Some of the inputted RDF triples

<table>
<thead>
<tr>
<th>Subject</th>
<th>Predicate</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
<td>department</td>
<td>Engineering</td>
</tr>
<tr>
<td>University</td>
<td>department</td>
<td>Future Science</td>
</tr>
<tr>
<td>Future Science</td>
<td>faculty</td>
<td>Information Media</td>
</tr>
<tr>
<td>A</td>
<td>room</td>
<td>A1101</td>
</tr>
<tr>
<td>University</td>
<td>entrypermit</td>
<td>Maingate</td>
</tr>
<tr>
<td>Engineering</td>
<td>entrypermit</td>
<td>A</td>
</tr>
<tr>
<td>Information Media</td>
<td>entrypermit</td>
<td>A1110</td>
</tr>
</tbody>
</table>
4.3 Discussion

The RDF graph in the upper part of Fig. 9 shows the user requirement. The RDF graph in the middle part of Fig. 9 can be derived from the inference rule in Formula 1 and the RDF graph in the upper part of Fig. 9. One example is (Makoto, affiliation, Electronic Engineering). The "Makoto" subject does not have the "hasAvailability" predicate. Therefore, (Electronic Engineering, use permit, no) is generated. The RDF graph in the lower part of Fig. 9 shows that the user’s affiliation information was eliminated. For this reason, collecting restrictions was executed in accordance with the user by using the extended knowledge model. Moreover, the derived RDF graph shows the user affiliation information cannot be used. Therefore, it was considered that the restriction of privacy information that satisfies the users’ requirements was fulfilled. However, the resource indicating the user name remained in the RDF graph of the lower part of Fig. 9. Therefore, it seems that this privacy information needs to be removed.

In Scenario A, all user information is supposed to be available. As compared with Fig. 4 and the user’s affiliation information (University: TDU, department: Engineering, faculty: Electronic Engineering), all the permitted rooms are found. As a result, the entry management service lists all the permitted rooms, as shown in Fig. 8. In Scenario B, all...
sensors are not supposed to be available. Thus, the permitted room does not exist. Figure 11 shows that the entry management service is not provided. For this reason, entry into permitted rooms was properly listed in the newly defined space. In the two scenarios, we confirmed that the provision of services can be automatically executed. This result shows the feasibility of the sensor network space by using the knowledge model in a variety of spaces.

The results of this study show a possible resolution to the security issue in privacy protection.

5 CONCLUSION

The purpose of this study was to confirm whether an RDF-based service implementation method keeps balance of service provisions that efficiently employ state information and privacy protection at the same time in a ubiquitous sensor network. In this paper, we extended the knowledge model by introducing the concept of a sensor by Semantic Sensor Network Ontology. In addition, we expressed the collecting restrictions by adding inference rules. We also verified the feasibility of the extended knowledge model and collecting of restrictions by experiments on the simulator that we developed. As a result, we found a possible resolution to the issue of balancing security and privacy.

REFERENCES


(Received November 20, 2014)
(Revised April 23, 2015)
Makoto Sato received his B.E. and M.E. degrees from Soka University in 2011 and 2013, respectively. He is currently doing his Ph.D. project at Tokyo Denki University. His research interests include sensor network system, ubiquitous computing and privacy issue.

Dr. Yoshimi Teshigawara is currently a senior researcher of Department of Information Systems and Multimedia Design, School of Science and Technology for Future Life as well as Cyber Security Laboratory, the Research Institute of Science and Technology at Tokyo Denki University since 2013. He began his professional career in 1970 at NEC Corporation, engaged in the design and developments of network architecture and computer systems via satellite. He worked for Soka University from 1995 to 2013, and served Dean of Faculty of Engineering and Graduate School of Engineering. His current interests are network security, e-learning, ubiquitous computing. Dr. Teshigawara received his PhD from Tokyo Institute of Technology, Japan, in 1970.

Ryoichi Sasaki is a professor of Dept. of Information Systems and Multi Media, School of Science and Technology for Future Life, Tokyo Denki University. He received his B.S. Degree in health science and Ph.D. Degree in system engineering, both from the University of Tokyo in 1971 and 1981, respectively. From April of 1971 to March of 2001, he was engaged in the research and research management on systems safety, network management and information security at Systems Development Laboratory of Hitachi Ltd. From April of 2001, he is a professor of Tokyo Denki University, and engaged in the research and education on information security. Now, he is also an advisor of Information Security in Cabinet Secretariat for Government of Japan, and a visiting professor of National Institute of Informatics, Japan.