Data Model and Data Access Control Method on Service Platform for Smart Public Infrastructure

Yohei Kawada¹, Kojin Yano², Yoshihiro Mizuno¹ and Hirofumi Terada²

¹Information & Telecommunication Systems Company, Hitachi Ltd., Tokyo, Japan
²Yokohama Research Laboratory, Hitachi Ltd., Kanagawa, Japan

Keywords: Data Model, Data Access Control, Service Platform, Smart Public Infrastructure.

Abstract: In the smart city projects that will “smartize” urban infrastructures, a new access control technology is needed to offer appropriate consumer data to appropriate applications. In this paper, we analyze characteristics and problems of the data access in the service platform for smart public infrastructure and clarify requirements for data access control. Next, we propose a data model and a data access control method that satisfy those requirements. The data model includes access authorization that expresses the contracts between consumer and service provider. In the data access control method, data corresponding to the access authorization is filtered in RDBMS for the performance. Finally, we evaluate the proposed method by implementing a prototype and confirm that the requirements are satisfied.

1 INTRODUCTION

Recently some “smart grid” projects have included the widespread adoption of technologies such as renewable energy resources and electric vehicles (EVs). This has led to the concept of “smart city” attracting attention (Naphade et al., 2011). The roles for IT in “smartizing” public infrastructure are as follows.

- Data Acquisition
  Obtaining data from devices or equipment in real time. The data are about demanded and supplied amounts of electric varying from hour to hour.

- Data Management
  Collecting the data and providing them to users as a gateway service.

- Data Application
  Utilizing the data to forecast the demanded and supplied amounts of electric and shift the demand.

System architectures have been proposed that consist of a sensor layer, system layer, and service layer that respectively correspond to the above roles. Moreover, in these proposals, the system layer is realize by a service platform that collects the data from a wide variety of devices or equipment in the sensor layer and provides the data (transformed standardized formats) to a wide variety of applications in the service layer (Lee et al., 2011).

Data access control is considered a necessary function for the service platform (Naphade et al., 2011). In this case, data access control means providing data to only appropriate users on the basis of contracts between data owner and user. For example, in the electric power industry, it is suggested that the business model will change from one between consumers and suppliers (equal electric companies) to one among consumers, suppliers, and aggregators (NIST, 2010). The aggregators serve the consumers by visualizing their demanded amounts or control their appliance directly for saving the amount. The consumer means the consumer of electrical energy, which is the generator of data, and the application means a service provider for consumer and the user of data. Either manager or generator of data becomes the owner of data. The manager of data is an electric power company, and a generator of data is consumers. Figure 1 shows data flow in case that the owner of data is the consumer. As this figure shows, the consumers have the contract with the aggregators, so the service platform must grant the access to consumer’s data to only the aggregators that have the contract with their consumers. Moreover, the service platform must grant the access to only data permitted by their contract. The data that the service platform holds become huge, but the platform has to realize data management.
access control on the basis of individual contracts with realistic performance. This problem was not considered in the above papers.

In this paper, we describe the data access control in providing the data (about the history of consumer’s demanded amounts of electric or the states of their devices) to services like the aggregators. We assume that the access control of each service is handled on the platform. The access control of each user of the service is handled on service provider side.

The remainder of this paper is organized as follows. The next section describes the characteristics of data access and the requirements for data access control. Section 3 describes related work about data access control. Section 4 describes the proposed data model and implementation of data access control. Section 5 describes the evaluation our proposal from the aspect of the requirements. Section 6 concludes the paper.

2 CHARACTERSITICS AND PROBLEMS OF DATA ACCESS

2.1 Data Model and Data Providing Types to Applications

Figure 2 shows the system architecture that we assume in this paper. As this Figure shows, the platform collects the data about demanded/supplied amounts of electric from devices or equipment and provides the data to applications. Table 1 shows the example of the demanded-amount data on a relational database table. “Node ID”, “Class ID”, and “Timestamp” are based on CIM (Common Information Model) object in OpenADE specification (OpenADE, 2010). Node ID means the unique ID of devices, Class ID the kind of each Node ID device, and Timestamp the date and time when the following values are measured by sensor.

Table 1: Example of demanded-amount data.

<table>
<thead>
<tr>
<th>Node ID</th>
<th>Class ID</th>
<th>Timestamp</th>
<th>Instant consumption (kW)</th>
<th>Accumulative consumption (kWh)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device a-1</td>
<td>Smart Meter</td>
<td>2012/5/11 10:00:00</td>
<td>21</td>
<td>1343</td>
<td></td>
</tr>
<tr>
<td>Device b-2</td>
<td>Battery</td>
<td>2012/5/11 11:00:00</td>
<td>3</td>
<td>123</td>
<td>OFF</td>
</tr>
<tr>
<td>Device b-1</td>
<td>Smart Meter</td>
<td>2012/5/11 10:00:00</td>
<td>24</td>
<td>3441</td>
<td></td>
</tr>
<tr>
<td>Device b-2</td>
<td>Heat Pump</td>
<td>2012/5/12 12:00:00</td>
<td>4</td>
<td>63</td>
<td>ON</td>
</tr>
</tbody>
</table>

“Instant consumption”, “Accumulative consumption”, and “Power” are the values from the sensor attached on devices. Instant consumption means the instant electric power amount that each device consumes. Accumulative consumption means the accumulative electric power amount that each device consumes during a definite period. Power means whether each device is on or off. The kind of data in this example is the demanded amount on the consumer-side, but the platform can hold multiple kinds of data.

As to provide data to the applications, the following two types of data accesses are used (OpenADE, 2010).

- **Event Notification**
  The platform sends the data to defined application(s) when it receives the data from devices whose defined conditions correspond to any of the platforms data.

- **Batch Transfers**
  The platform sends the data to an application when it receives the request including search conditions from the application. The platform searches the data corresponding to the conditions from their database and replies to the result.

In next subsection, we will clarify the characteristics of the batch transfers data access.

2.2 Characteristics and Problems with Batch Transfers Data Access

There are various characteristics when the platform
provides data with batch transfers data access. We clarify two characteristics and problems below.

(A) The authorization is based on the individual contract between the application and the consumer.

In the model of Figure 1, the range of data to provide to service providers is based on contract between a consumer and the service provider. The following shows the example of the data the service provider’s applications make available.

- **Application A** (Adjustment of supply and demand)
  
  To coordinate the supply and demand between consumers, this service provider makes a contract with all consumers that states "Application A can refer to the data of the smart meter". For example, in Figure 3, this application can refer to only the data in the "Smart Meter" Class ID.

- **Application B** (Visualization of demanded amount)
  
  The devices targeted for visualization are prescribed in a contract between consumer and service provider. For example, in Figure 3, the service provider makes a contract with Consumer a. Thus, all devices Consumer a owns (Device a-1 and Device a-2) are the targets. Therefore, this application can refer to only the data where Node ID is Device a-1 or Device a-2 and Timestamp is within the contract period.

- **Application C** (Remote watch out for device user’s safety)
  
  This application confirms that the user of the devices is fine by checking the use history of the devices and notifies the requester such as the user’s families or primary care doctor. The devices targeted for checking are prescribed in a contract between consumer and service provider. For example, in Figure 3, the target user is Consumer b and the target devices are prescribed as Device b-2 in the contract. Therefore, this application can refer to only the data where Node ID is Device b-2 and Timestamp is within the contract period.

We define the rights to access data from given applications as data access authorizations. The data access authorizations are based contracts like the above. When a set of the application is assumed

\[ S = \{s_1, s_2, ..., s_i, ..., s_m\} \]

and a set of consumers is assumed

\[ O = \{o_1, o_2, ..., o_j, ..., o_n\} \]

a set of data access authorizations \( P \) is assumed as follows.

\[ P = \{s_1o_1s_2o_2..., s_ito_i, ..., smon, smon-1, smo_n\} \]

where \( m \) is the number of applications and \( n \) is the number of consumers. This means that the maximum number of authorizations accords with the number of combinations of application and consumers. The platform must define and hold this authorization based on individual contracts.

However in case that the owner of data corresponds to the data manager such as the electric power company, the application makes a contract with data manager. The platform must define and hold this authorization based on the contracts. It is necessary for the platform to define and hold the authorization based on both contract with the individual consumer and contract with the electric power company.

(B) A huge number of data is targeted for a judgment about whether application can access.

In the near future, the number of consumers and their devices is expected to grow along with urban population. In addition, data are expected to be collected from the devices more frequently so that the applications can grasp the state of the devices and consumers in detail. That is why the number of data (i.e., the number of rows in Table 1) per time to be provided to the application may increase. In this case, if the platform judges about whether application can access for every data, the time it takes for a judgment about data access authorization will increase when more of the data correspond to search conditions.

A result set of the rows corresponding to the search condition in which application \( s_1 \) requested is assumed as \( D = \{d_1, ..., d_k\} (k \text{ is a number of rows}) \). The time \( Time_i \) for judging whether application can access by every row is assumed as follows.

\[ Time_i = \frac{1}{T} \sum_{i=1}^{k} \text{time}_i (d_i, s_0) \propto k \]  

(2)

\( T \) is a number of threads for judging process, and \( \text{time}_i (d_i, s_0) \) is assumed as the time to judge whether row \( d_i \) corresponds to the data access authorization \( s_0 \) if \( o_i \) is an owner of data \( d_i \). The owner means the consumer who owns the device.
putting data \( d_j \) or not. If \( k \) is a large, then \( T_{m_i} \) is also large, so \( T_{m_i} \) is proportionate to \( k \).

3 RELATED WORK

This section describes related work about concept or method of data access control.

3.1 General Data Access Control

Identity Based Access Control (IBAC) performs access control based on the Access Control List (ACL) defined in every data. The ACL defines the applications that can access. This method is generally used in networks and operating systems.

If IBAC was applied to the platform in this paper, data access authorizations based on every contract would possibly be defined. In this case, ACL has to be defined by every consumer, device, or data row. It is necessary to judge whether the application that required data exists on the ACL for each data. As described in 2.2(B), when the number of data is huge, the processing time becomes long.

In Role Based Access Control (RBAC) (Sandhu et al., 1996), data access authorization is assigned to the application. Plural roles are able to be assigned to a user by relating many-to-many relationships between the user and role. In addition, a higher role can succeed to a lower role due to the hierarchical structure of the roles. These techniques make RBAC suitable to manage the authorizations of the user in an organization.

If RBAC was applied to the platform in this paper, RBAC could be applied when the contracts do not differ. On the other hand, when contracts differ, access control based on the individual contract is necessary. For example, in Figure 4, Application B can access the data from all devices of Consumer a but only the data from the smart meter of Consumer b. In the RBAC, different roles need to be assigned when authorization is different. This means that the stated advantage of the RBAC (its ability to simplify a definition by assigning authorization for a role) does not make sense.

Barker and Douglas introduced an example in which RBAC is implemented in an SQL database (Barker and Douglas, 2004). They showed that the performance of RBAC is reasonable in practice use but does not mention an implementation method or the results measuring RBAC for large data.

Attribute Based Access Control (ABAC) (Yuan and Tong, 2005); (Bertino et al., 2001); (Duri et al., 2004) is an access control method in which access authorization is based on the attribute of a data, application, or environment. It is suitable for open systems such as data dissemination systems on the Internet. In ABAC, the characteristic of the application is unidentified beforehand and an application is added to during use.

If ABAC is applied to the platform in this paper, ABAC may perform the access control based on contracts by regarding the contract between service provider and consumer as attributes. However, similarly, an implementation method and the attributes model of ABAC for large data have not been discussed.

3.2 Data Access Control for Smart Grid

As access control for smart grid systems, access control using XACML based on SOA is proposed (Jung et al., 2012). In this proposal, Policy Decision Point (PDP) acquires XACML policy, and judges whether the access to a resource is possible based on the policy. However, this proposal does not describe the details about the judgment processing on PDP.

In addition, a method to apply an attribute base encryption to the access control of the smart grid system is proposed (Ruj and Nayak, 2013).

4 PROPOSAL OF THE DATA ACCESS CONTROL METHOD

For the requirements mentioned in section 2, this section describes the proposed data model and data access control method on a service platform for smart public infrastructure. In the related work in section 3, access authorization and data from devices are managed separately. Therefore authority judgment processing is likely to become the bottleneck. In the proposal, we add the access authorization to a data model, which is composed on
the basis of the contract between the consumer and the service provider. This access authorization defines the condition of the data accessible by every application. The platform judges whether the access is admitted or refused by using this access authorization when it receives an access demand from the application.

4.1 Data Model with Access Authorization for Data Access Control

CIM defines the data model for the device configuration and the historical sensing data (e.g. demanded/supplied amounts of electric for devices), but does not define the data model for applications that utilize the data and access control. Therefore, as shown in Figure 5, we propose to extend the data model to add the access authorization. The access authorization is defined for every application, and the content is based on a contract between the service provider that operates the application and the consumer. For example, in the case of the data search demand from Application A, the platform searches for the data corresponding to the search conditions from Application A and judges whether the data corresponds to the access authorization of Application A or not. As a result, the platform returns only data corresponding to access authorization.

We assume that contracts have two kinds of content. One is common between consumers, and the other is may be different for each consumer. The former corresponds to the contents such as the kinds of data access (e.g. register, refer, update, delete), the kinds of data, and valid period for application. The latter corresponds to the contents such as contract period and targeted devices. Based on this supposition, the access authorization consists of two parts. One corresponds to former contents, called “application authority information”. The other corresponds to latter contents, called “access condition” and “relation information”. Figure 6 shows the constitution of access authorization.

Table 2 shows the example of application authority information. In this table, a record expresses an authority (a combination of one kind of data and one kind of access) that an application is given. This information includes Auth ID, which identifies the authority, Application ID, which means the targeted application, valid period, the time in which the application is in operation, and kind of access, and kind of data.

Access condition and relation information express the contract contents that may differ for each consumer. The access condition is related to the application authority information, meaning conditions of data in which application authority information becomes effective. In other words, if the data correspond to the contents defined in the access condition, the application can access data in a defined range in authority information. Table 3 shows an example access condition. In this table, a record expresses a condition (combination of item name of data and value) of targeted data. This includes Auth ID, which means related record in application authority information, item name of data that is targeted in the condition, condition type (discussed below), and value which, is the value at item name of data. Plural access conditions are able to be related to application authority information.
In this case, if there are plural conditions for the same item name of data, these conditions are OR conditions. If there are plural conditions for different item names of data, these conditions are AND conditions.

For example, in the case of Auth ID: “1” in Table 2, Application A can refer to the data of demanded amount of electric on the consumer-side. However, as defined in Table 3, Application A can refer to only the data where the value of Class ID equals “Smart Meter”.

In addition, the devices related to the applications are able to be assigned as the value in access conditions by using relation information. In this case, “relation” is assigned at “condition type”, and “relation type” is assigned at “value” in the access condition. Table 4 shows the example of the relation information. In this table, a record expresses a one-on-one relation. This information includes relation type, which means the kind of relation, Upper ID, Lower ID, start date, which is the date the relation started, and end date, which is the date the relation ends. The relation is valid from start date to end date. Upper ID and Lower ID are assigned Application ID, Node ID, or Consumer ID. In Table 4, Application B and Consumer a are related and the relation type is “service providing”, which means Application B provides a service to Consumer a. Start date correspond to the date on which the contract between Application B and Consumer a starts. Moreover, Consumer a is related to Device a-1 and Device a-2, and relation type is “own”. In the example in Table 2, application authority information in which Auth ID is “2” is for Application B. In Table 3, “relation” is assigned at condition type, and “service providing, own” is assigned at value. This means that Application B can refer to the data in which Node ID corresponds to the devices that Consumer a owns. Thus, Application B can refer to the data in which the Node ID is Device a-1 or Device a-2. Similarly, Application C can refer to the data in which the Node ID is Device b-1.

By using the relation information, the access authorization that reflected each contract content such as contract period can be defined.

4.2 Data Access Control Method

In the requirement at 2.2(B), the time for judging whether the access is allowed or refused must not drastically increase. Thus, in the proposed data access control method, a SQL based on an access authorization is generated and the data are filtered by the SQL performed at the RDBMS in the platform. If the filtering process is done at the RDBMS, the process can be speeded up by configuration of RDBMS (e.g. indexing, etc.).

Table 3: Example access condition.

<table>
<thead>
<tr>
<th>Auth ID</th>
<th>Item name of data</th>
<th>Condition type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Class ID</td>
<td>Smart Meter</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Node ID</td>
<td>relation</td>
<td>service providing, own</td>
</tr>
<tr>
<td>3</td>
<td>Node ID</td>
<td>relation</td>
<td>service providing, target</td>
</tr>
</tbody>
</table>

Table 4: Example of relation information.

<table>
<thead>
<tr>
<th>Relation type</th>
<th>Upper ID</th>
<th>Lower ID</th>
<th>Start date</th>
<th>End date</th>
</tr>
</thead>
<tbody>
<tr>
<td>service providing</td>
<td>Application B</td>
<td>Consumer a</td>
<td>2013/1/1</td>
<td></td>
</tr>
<tr>
<td>service providing</td>
<td>Application C</td>
<td>Consumer b</td>
<td>2013/1/1</td>
<td></td>
</tr>
<tr>
<td>own</td>
<td>Consumer a</td>
<td>Device a-1</td>
<td>2012/10/1</td>
<td></td>
</tr>
<tr>
<td>own</td>
<td>Consumer a</td>
<td>Device a-2</td>
<td>2012/11/1</td>
<td></td>
</tr>
<tr>
<td>target</td>
<td>Consumer b</td>
<td>Device b-2</td>
<td>2013/1/1</td>
<td></td>
</tr>
</tbody>
</table>

Figure 7 shows the flow chart of the proposed data access control method based on the above principle. The contents of processing are as follows:

1. When the platform receives the request from an application, the platform searches for the application authority information and related access conditions by Application ID. If the condition type in access conditions corresponds to “relation”, the relation information is also searched for by Application ID.

2. The platform generates WHERE-phrase of SQL based on application authority information and access condition obtained in (1). Figure 8 shows the contents of WHERE-phrase.

3. The platform generates WHERE-phrase of SQL based on search condition from the application.

4. The platform generates a SQL by combining the (2) and (3) WHERE-phrase and carries out the SQL.

5. The platform obtains the result of (4) and sends them to the application.

5 EVALUATION

This section describes an evaluation for the proposed data model and data access control method.

In the point of the data model, it is possible that the contract contents mentioned in 2.2(A) are defined in Tables 2, 3, and 4. However, the following limits at least.

- When designated data items in access conditions are the same, they become OR conditions. When
they are different, they become AND conditions. Therefore, a contract must not become the OR condition between different data items. At present, no application has been found that has a problem with the limit mentioned above.

Figure 7: Flow chart of proposed data access control method.

In the data access control method, a prototype system of the platform based on the proposed method was developed by Java. Table 5 lists the specifications of the data search function that the prototype provides for applications. An evaluation environment implementing the prototype system is shown at Figure 9 and Table 6. Application Client sends an XML message corresponding to the argument in Table 5 to Platform Server in HTTP Request. Platform Server sends select SQL via JDBC to DB Server and obtains the result. Platform Server composes the XML message corresponding to the return value of Table 5 and replies to Application Client as HTTP Response. After the demanded-amount data is registered with the DB Server and a data search is carried out, it is confirmed that only the data corresponding to the AND condition between the search condition and the access authorization were returned as a return value. These Servers secure availability by multiplexing servers and scalability by adding BES.

Next, the performance of the data search is measured by using the environment mentioned above. If the number of targeted consumer is

400,000, the number of the targeted devices per consumer is three, and if data are collected in half-hour periods, the data accumulated per day are 57,600,000. After registering data of this number

![Figure 7: Flow chart of proposed data access control method.](image7.png)

![Figure 8: Contents of WHERE-phrase.](image8.png)

![Table 5: Specification of data search function for applications.](table5.png)

![Table 6: Experiment Equipment Specifications.](table6.png)
with the DB Server, we experimented with five patterns of the number to correspond to search conditions: 1, 10, 100, 1,000, and 10,000. All data corresponding to the search condition corresponded to the access authorization that the application was given. As a comparison, we also experimented in the case in which access control was not valid. In this case, the platform processes only from (3) to (5) in Figure 7.

Table 7 shows the measurement results of processing time from the platform receiving the request to return the response. Figure 10 shows the increasing rate from the processing time of the case in which access control is not valid (Non-Access Control) to the processing time of the case in which it is (Access Control). As the results show, in case of Access Control, the processing time slightly increased in comparison with a case of Non-Access Control. This is because that Access Control needs the process to achieve access authorization incrementally. However, even in the cases of 1,000 and 10,000, the increasing rate of the time is less than 8%. Therefore, even if the number of data becomes huge, the processing time for judging the access allowed or refused does not largely increase. The cause that the case of 100 input data is the highest increasing rate is unclear, but characteristics of the RDBMS might be related. Meanwhile, in the case of equation (2) in 2.2(B), if it is assumed that the number of thread T=10 and the average of $t_{ij}(d_i,s_{ij})$ is 14.6msec (by the result that the number of data is 1), the processing time is 14,600 msec in the case of 10,000. In this case, the increasing rate is over 1,500%.

### Table 7: Experiment Result (unit: msec).

<table>
<thead>
<tr>
<th>Number of data corresponding to search condition</th>
<th>1</th>
<th>10</th>
<th>100</th>
<th>1,000</th>
<th>10,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access Control</td>
<td>132.8</td>
<td>134.7</td>
<td>195.7</td>
<td>286.0</td>
<td>980.6</td>
</tr>
<tr>
<td>Non-Access Control</td>
<td>118.2</td>
<td>119.4</td>
<td>136.1</td>
<td>265.1</td>
<td>908.8</td>
</tr>
<tr>
<td>Increasing Rate (%)</td>
<td>12.4</td>
<td>12.8</td>
<td>42.8</td>
<td>7.9</td>
<td>7.9</td>
</tr>
</tbody>
</table>

Figure 10: Experiment Results.

6 CONCLUSIONS

In this paper, we focused on a service platform that collects and manages the data collected from public infrastructure devices or equipment. We clarified requirements about the data access control when the platform provided the data to the service providers such as the aggregator. Next, we proposed a data model and a data access control method to meet the above requirements. In the proposal data model, the platform holds every contract between a consumer and a service provider as a data access authorization. In addition, this access authorization consists of following three pieces of information: application authority information to express authority contents of the application, access condition to express the condition about accessible data derived by contract contents, and relation information. Moreover, we proposed avoiding a large increase in the authority judgment processing time by filtering the data corresponding to access authorization by RDBMS.

In the future, we will evaluate the validity and sufficiency of the proposal method by applying in an actual experiment.

REFERENCES


National Institute of Standards and Technology(NIST), 2010. **NIST Framework and Roadmap for Smart Grid Interoperability Standards Release 1.0**.
OpenADE Task Force (OpenADE), 2010. *OPENADE 1.0 SERVICE DEFINITION – CORE*, DRAFT V0.8.