Privacy Issues of Applying RFID in Retail Industry

Haifei Li
Patrick C.K. Hung
Jia Zhang
David Ahn
Privacy Issues of Applying RFID in Retail Industry

Haifei Li, Department of Computer Science, Union University, USA
Patrick C. K. Hung, Faculty of Business and IT, University of Ontario Institute of Technology (UOIT), Canada
Jia Zhang\textsuperscript{1}, Department of Computer Science, Northern Illinois University, USA
David Ahn, Department of Computer Science, Nyack College, USA

EXECUTIVE SUMMARY

This case study describes the privacy issues of applying Radio Frequency Identification (RFID) in the retail industry. With the dramatic price drop of RFID tags, it is possible that RFID be applied to individual items sold by a retailer. However, the RFID technology poses critical privacy challenges. In this study, we analyze the potential privacy issue of RFID utilization, and we propose a privacy authorization model aiming for precisely defining RFID privacy policies for the retail industry.

Keywords: RFID, Privacy, RBAC, EPAL

INTRODUCTION

In present retail industry, retailers are under tremendous pressure to improve their efficiency. One way to increase productivity is through the adoption of new technologies. History has revealed that retailers are the early adopters of Electronic Data Interchange (EDI) and Business-to-Business (B2B) E-Commerce. The benefits of adopting these new technologies are obvious: reduced time to market and reduced cost associated with office and manufacturing floor automation. In recent years, Radio Frequency Identification (RFID) has caught significant attention in retail industry. RFID is a generic term for the technologies that use radio waves to automatically identify individual items wirelessly. RFID is capable of enabling retailers to effectively and efficiently track the entire circulation process of items from suppliers to end users. It allows to identify, orientate, and trace objects directly and continuously. In addition, RFID is able to deliver information at real time. As a result, RFID is widely considered as an emerging technology that could potentially revolutionize the way retailers do business. Among other examples, Wal-Mart has mandated its top 100 suppliers to use RFID by January 2005 (Vijayan, 2003); U.S. Department of Defense has also made the similar requests to its military suppliers (U.S. Department of Defense, 2003).

Although it seems that RFID is a boom to e-Commerce, the actual adoption of RFID in retail industry is quite slow (Bradner, 2005). Retail industry poses typical enterprise computing challenges (Neogi, 2004) as a retailer normally deals with multiple parties belonging to different organizations: suppliers, manufacturers, distributors, and end consumers. Nowadays, the focus of enterprise computing efforts of retailers mainly aims for suppliers. To date there is little work conducted on how to provide enterprise-level computing capability for individual customers. In addition to the known security issue, one such capability we have identified is consumer’s privacy protection. There is a growing concern for data privacy among businesses and consumers, because of the possible unwanted revelation of confidential or personal data stored within RFID devices.

\textsuperscript{1} The third author is currently with BEA Systems, Inc., and also a Guest Scientist of National Institute of Standards and Technology (NIST).
Privacy is a state or condition of limited access to a person (Schoeman, 1984). Particularly, information privacy refers to an individual’s right to determine how, when, and to what extent that his/her personal information will be released to people or to organizations (Westin, 1967). To date information privacy mainly aims for ensuring the confidentiality of sensitive information. In other words, one major objective of enforcing privacy is to protect Personally Identifiable Information (PII). Many authorization technologies can be applied to protect PII. However, information privacy contains other privacy concepts such as purpose and obligation (Fischer-Hubner, 2001). In more detail, authorization focuses on preventing unauthorized users from accessing sensitive information; while privacy focuses on managing authorized users to use information effectively and to achieve an organization’s strategy within necessary constraints (Bucker, 2003).

In addition, privacy control does not focus on individual subjects. A subject releases his/her data to the custody of an enterprise with agreement to the set of purposes for which the data may be used. In the U.S., the Privacy Act of 1974 requires that federal agencies grant individuals access to their identifiable records that are maintained by the agency, ensure the accuracy and timeliness of existing information, and limit the collection of unnecessary information and the disclosure of identifiable information to third parties. In summary, the U.S. mostly relies on self-regulation and limited legislation. However, federal agencies circumvent these constraints by subscribing to commercial surrogates who collect and store the same data with no constraints. Because of the identity thefts happened at ChoicePoint and LexisNexis in the early 2005, U.S. lawmakers have pushed for more aggressive data privacy legislation (Gross, 2005).

On the reverse, the Europe Union (EU) Data Protection Directive (Steinke, 2002) contains two statements contradicting the U.S. one. The first statement requires that an organization must inform individuals why it collects and uses information, how to contact the organization, and the types of third parties to which it discloses the information. The second statement requires that personal data on EU citizens may only be transferred to countries outside the 15 nation blocks that adopt these rules or are deemed to provide “adequate protection” for the data. As a result, these two statements imply that no information of any EU citizen can be transferred to the U.S. due to the conflicts between the two privacy acts. Consequently, these policies create obstacles for conducting business activities between the EU and the US. To solve the problem, the U.S. government has made a voluntary scheme called “Safe Harbor” to provide an adequate level of data protection, which can safeguard transfers of personal data to the U.S. from EU. The U.S. companies conducting business in the EU must certify to the Commerce Department that they will follow the regulations of the EU directive. Any violation would be subject to prosecution by the Federal Trade Commission (FTC) for deceptive business practices. Furthermore, based on a recent survey, bank officers said that they had ongoing concerns, mostly procedural, about how to handle the anticipated privacy regulations of the U.S. Gramm-Leach-Bliley Act (GLB), which requires that financial institutions regularly communicate privacy policies with customers and provide adequate opportunities for “opting-out” of personal information disclosure to non-affiliated third parties (Hinde, 2002). All of these facts indicate that privacy is a critical topic nowadays.

Therefore, in our opinion, the traditional view of authorization model should be extended with an enterprise-wide privacy policy to manage and enforce individual privacy preferences. In this paper, we aim to propose a privacy authorization model and explore its implementation issues focusing on language specification. The reminder of the paper is organized as follows. We will first briefly introduce the background of the RFID technology and introduce problem domain, followed by discussions of related work. Then we will propose our RFID-oriented privacy authorization model. We will discuss the design of the model, discuss the implementation of the model, and perform self assessments. Finally, we will make conclusions and discuss future work.

**BACKGROUND AND PROBLEM DOMAIN**
Due to its potential to dramatically increase productivity, different organizations have shown strong interests in applying the RFID technology in retail industry. Among others, IEEE has played essential role by sponsoring conferences and publishing papers in this area (Leventon, 2005); IBM has been active in pursuing business opportunities for several years (IBM, 2004); researchers and engineers at HP have developed a RFID-based solution for tracking IT assets (Schwartz, 2005). In this section, we will first briefly introduce the basic concept of RFID, and then discuss the problem domain to be addressed.

A typical RFID system

Radio Frequency Identification (RFID) is a generic term for the technologies that use radio waves to automatically identify individual items. A typical RFID system contains three components as shown in Figure 1: a RFID tag, a RFID reader, and a computer network. A RFID tag is actually a microchip with a coiled antenna. When a RFID tag receives electromagnetic waves from the reader, it sends stored data to the reader. A RFID reader can read and write data, depending on the types of the RFID tags interacted with. A reader can also send data to the associated computer network. A computer network can receive data from the reader and perform further processing on the data collected. Potentially, computers can also send data to readers. Figure 1 illustrates the main components of a RFID system as well as the interactions between them.

Requirements analysis

The RFID system illustrated in Figure 1 is a typical scenario used by a retailer. For example, big retail companies with large database infrastructures, like Wal-Mart, can use such a RFID system to keep track of where individual cartons of goods are in their supply chain or, maybe someday in the future, what individual products are in a shopper’s physical cart. However, although RFID is a boon to retail industry, it comes at the high price of shaky security and privacy. For example, a retailer must deal with different categories of users that may have access to the data stored in RFID tags. These users could be professional buyers, cashiers, store managers, warehouse keepers, warehouse managers, etc. The data contained in RFID tags may or may not be PII. If it is PII, people naturally have big concerns for possible privacy invasion when various users have access to RFID tags. Even if the information is not personally identifiable, there is still some concern because the information subject could potentially be identified. That is
why a new role came into place called privacy policy enforcers, which may have alternative names such as CPO (Chief Privacy Officer).

At present, interest remains high enough in implementing RFID that the lack of security and privacy is not a bottleneck to retail industry’s adoption of this new business model. But why wait to find a solution? This is the momentum of this research that aims at investigating the security and privacy issues of a RFID system and exploring possible solution. In this paper, we focus on the privacy enforcement issue. One can imagine that information privacy is usually concerned with the confidentiality of sensitive information. In general, information privacy is enforced by privacy policies, which describe an organization’s data practices concerning: (1) what kinds of information they collect from individuals (subjects); (2) for what purpose the information (objects) will be used; (3) whether the organization exposes access to the information; (4) who are the recipients of any result generated from the information; (5) how long the information will be retained; and (6) who will be informed in the circumstances of dispute.

One of the most significant objectives of enforcing privacy is to protect PII. Traditionally, many authorization technologies were directly applied to protect PII. However, traditional coarse-grained authorization techniques merely focus on preventing unauthorized users from accessing sensitive information. Mostly authorization privileges are associated with specific roles and are immutable after granted. In an enterprise retail RFID system, however, authorization alone cannot satisfy privacy requirements. Such a system requests to grant authorized users to use information effectively and to achieve an organization’s strategy within necessary constraints. Therefore, much more complicated authorization mechanisms are needed. Various merchandises may enforce different access rights from different roles at different stages. In addition, privacy encompasses broader concepts than authorization, such as purpose and obligation. Therefore, in our opinion, the traditional view of authorization model should be extended with an enterprise-wide privacy policy for managing and enforcing individual privacy preferences. In more detail, such a privacy authorization model needs to exhibit the following attributes:

- **Interoperability**: A privacy authorization model should be able to interpret and use credentials issued by any other issuing authorities.
- **Expressiveness**: Credentials should not only contain an individual’s identity, but also other useful information such as purpose and obligation.
- **Extensibility**: The credential system should be flexible enough to register new individuals and organizations with new types of information.
- **Anonymity**: An individual identifier should not be revealed under any circumstances.
- **Scalability**: Credential systems should be robust enough to handle the increasing number of users, service providers, and issuing authorities. Scalability here means that an application can be used more frequently or by more users without loss of performance.

In order to fulfill the challenging requirements of such an enterprise-level RFID-oriented privacy authorization model, a powerful and flexible description language for extensible privacy policies is imperative. Therefore, our strategy in this research is to explore an enterprise RFID-oriented privacy authorization model and further explore the effective and efficient description of privacy policies. In addition, earlier we indicated that an enterprise retailer needs to deal with different categories of users who may have access to the data stored in RFID tags. As we will explore further in the following discussions, these different roles may require different authorization models. Since each RFID tag uniquely identifies specific merchandise, for the rest of the paper, we will interchangeably use the term “a RFID tag” with “the specific merchandise”.

**RELATED WORK**
Global privacy advocates, such as Electronic Frontier Foundation (http://www.eff.org) and Electronic Privacy Information Center (http://www.epic.org), have been active in raising the public’s awareness of potential privacy violations in RFID implementations. In a RFID position statement (Privacy and Civil Liberties Organizations, 2003) issued and endorsed by more than forty organizations and individuals, five potential threats to privacy and civil liberties have been identified:

- Hidden placement of tags. RFID tags may be embedded into/onto objects and documents without the knowledge of the individuals who obtain the items.
- Unique identifiers for all objects worldwide. The Electronic Product Code (EPC) potentially enables every object on earth to have its own unique ID.
- Massive data aggregation. RFID deployment requires the creation of massive databases containing unique tag data. These records could be linked with personal identifiable data, especially as computer memory and processing capacities expand.
- Hidden readers. Tags can be read from a distance, not restricted to line of sight, by readers that may be incorporated invisibly into nearly any environment where human beings or items congregate.
- Individual tracking and profiling. If personal identity were linked with unique RFID tag numbers, individuals could be profiled and tracked without their knowledge or consent.

Even though RFID may cause privacy concerns, the technology itself does not invade information privacy. It is the people who use the technology that invade privacy. If properly used, RFID is a useful and practical technology that can greatly benefit the economy and society. Some critics (Cline, 2003) believe that the risk of RFID privacy invasion was actually exaggerated. In fact, in the same report released by a coalition of privacy advocates, the following examples of "acceptable" uses of RFID were provided:

- Tracking of pharmaceuticals from the point of manufacturing to the point of dispensing.
- Tracking of manufactured goods from the point of manufacturing to the location where they will be shelved for sale.
- Detection of items containing toxic substances.

Generic privacy authorization technologies have been investigated for a long period of time (Powers, 2002). To address the growing need for standard and uniform privacy authorization languages, many companies and organizations have been marketing various privacy tools in the past few years (Senicar, 2003). For example, the Privacy Preferences Project (P3P) working group at World Wide Web Consortium (W3C) developed a P3P specification to enable Web sites to express their privacy practices in a standard and machine-readable XML format. P3P user agents allow users to choose to be informed of site practices and to automate decision making process based on the Web sites’ privacy practices. P3P also provides a language called P3P Preference Exchange Language 1.0 (APPEL), which is used to express the user’s preferences for making automated or semi-automated decisions regarding the acceptability of machine-readable privacy policies from P3P-enabled (http://www.w3.org/TR/P3P-preferences/) Web sites. P3P also provides a basic schema for data collection and a vocabulary to express purposes, recipients, and retention policies. However, P3P efforts cannot be directly applied to RFID-oriented retail industry. Although the P3P mechanisms capture common elements of privacy policies, Web sites may have to provide further explanation in human-readable policies. It is also obvious that P3P and APPEL are only applicable to the extent of Web server and browser. Furthermore, P3P does not provide authorization mechanisms to check a given access request against a stated privacy policy.

Another example is the eXtensible rights Markup Language (XrML) (Wang, 2002), which can be used to describe the rights and conditions for owning or distributing digital media. XrML concepts include license, grant, principal, right, resource, and condition. Based on the specification of licenses, an XrML
agent can determine whether to grant certain right on certain resource to certain principal. However, XrML does not consider the privacy entities in their access control model.

**RFID-ORIENTED PRIVACY AUTHORIZATION MODEL**

In this section we propose an enterprise-level RFID-oriented privacy authorization model. The model is composed of two parts: a role-based retailer enterprise boundary framework and a RFID access control framework.

---

**Role-based retailer enterprise boundary framework**

A virtual enterprise (i.e., retailer) boundary is introduced to ensure authorization-based privacy. As shown in Figure 2, all users having access to RFID tags are divided into three categories of roles: (1) suppliers, (2) role players, and (3) end users. A supplier is a provider of a RFID tag, who can be either the original manufacturer or another retailer. A role player serves in an enterprise retail store, who can be a cashier, a store manager, an assistant, etc. An end user takes a RFID tag out of the boundary of an enterprise retailer, thus can be either a professional buyer or another retail store. It should be noted that an entity can play different roles under different circumstances. For example, a retail store can act as a supplier for another retail store, or an end user of yet another retail store. A person can act as a role player in one retail store, and an end user of merchandise if he/she actually pays for it. In other words, the journey of merchandise can start from an original supplier and flow through multiple enterprise retailers before finally reaching a customer’s hand. Figure 2 also shows that the circulation of a RFID tag is a directional flow from a supplier to one or more enterprise retailers, and then to an end user.

As illustrated in Figure 2, at the left-hand side, suppliers provide retailers with the original data from the manufacturing facility. The data is embedded and stored into a RFID tag. After the RFID tag enters the enterprise’s boundary, different role players (e.g., cashiers, store managers, assistants, etc.) can interact with it. A solid line from the tag to a role player refers to an action of reading of the tag data; while a dotted line from the role player to the tag refers to a possible action of writing of some data to the tag. Before the tag leaves the enterprise boundary, such as a retailer store (e.g., purchased by an end user as shown in Figure 2), the RFID tag needs to pass through a “GateKeeper”. The GateKeeper is an automated program to ensure that the customer’s privacy will be properly protected. For example, it will examine whether some role players (e.g., cashiers) intentionally or unintentionally write some unauthorized data into the tag. In short, our role-based retailer enterprise boundary framework provides fundamental specification and context to perform role-based access control.
RFID access control framework

Based on the specification above, an access control system should enforce the policy stated by the enterprise. Under this circumstance, an information access control mechanism should also be embedded with privacy-enhancing technologies. All these evidences show the importance of integrating privacy concepts into access control mechanism for resolving the RFID privacy problems.

Let us take a quick review of the traditional access control mechanism. The family of Role-Based Access Control (RBAC) is commonly refereed to as the RBAC96 model. The RBAC96 model focuses on security control using roles and organizations. RBAC96 presents a conceptual model to describe different approaches such as base model, role hierarchies, constraint model, and consolidated model. In particular, the National Institute of Standards and Technology (NIST) conducted market analysis for identifying RBAC features into two layouts: The RBAC Reference Model and the RBAC Functional Specification. The RBAC Reference Model describes a common vocabulary of RBAC element sets and relations for specifying requirements and the scope of the RBAC features included in the standard. The RBAC Functional Specification describes the requirements of administrative operations for creating and managing RBAC element sets and relations, and system functions for creating and managing RBAC attributes on user sessions and making access control decisions.

In order to integrate privacy concepts into access control mechanism to resolve the RFID privacy problems, we propose a RBAC-extended framework. Our proposed model is based on the core RBAC model discussed in (OASIS, 2003), with privacy-based extensions tailored for RFID. As shown in Figure 3, RFID-oriented access control is regulated by an access control boundary, which is actually the retailer enterprise boundary discussed in the previous section. When a request arrives at the access control boundary, the core RBAC is enhanced with the privacy-based extension (e.g., purpose and obligation). Once a decision is made, the access control permission to the subject can be either granted or denied.

In order to facilitate enterprise-level RFID privacy control, it is imperative to automatically verify and validate privacy information. Therefore, here we formalize our model using the ad hoc standard privacy description language eXtensible Access Control Markup Language (XACML) from OASIS.

Our model can be notated as a tuple mainly including the following entities:

- USERS, ROLES, OPS, and OBS (users, roles, operations, and objects, respectively).
- UA ⊆ USERS × ROLES, a many-to-many mapping between users and roles (user-to-role assignment relation).
• assigned_users: \( (r:ROLES) \rightarrow 2^{\text{USERS}} \), the mapping of role \( r \) onto a set of users. Formally: 
\[ \text{assigned_users}(r) = \{ u \in \text{USERS} \mid (u, r) \in \text{UA} \} \]
• PRMS = \( 2^{(\text{OPS} \times \text{OBS})} \), the set of permissions.
• PA \subseteq \text{PRMS} \times \text{ROLES} \), a many-to-many mapping between permissions and roles (role-permission assignment relation).
• assigned_permissions(r:ROLES) \rightarrow 2^{\text{PRMS}}, the mapping of role \( r \) onto a set of permissions. Formally: 
\[ \text{assigned_permissions}(r) = \{ p \in \text{PRMS} \mid (p, r) \in \text{PA} \} \]
• SUBJECTS, the set of subjects.
• subject_user(s:SUBJECT) \rightarrow \text{USERS}, the mapping of subject \( s \) onto the subject’s associated user.
• subject_roles(s:SUBJECT) \rightarrow 2^{\text{ROLES}}, the mapping of subject \( s \) onto a set of roles. Formally: 
\[ \text{subject_role}(s) \subseteq \{ r \in \text{ROLES} \mid (\text{subject_user}(s), r) \in \text{UA} \} \]

The following are the sets of privacy-based entities (e.g., purpose and obligation) proposed to the core RBAC as shown in Figure 3:

• PURPOSES = \{ pp_1, pp_2, ..., pp_n \} is the set of \( n \) purposes. In the RFID scenario, an entity has to describe its purpose(s) of a request operation.
• OBLIGATIONS = \{ obl_1, obl_2, ..., obl_n \} is the set of \( n \) obligations that may be taken after the decision of a permission is made. In general, an obligation is opaque and is returned after the permission is granted. The obligations describe what promises a subject that must be made after gaining the permission.

In RBAC, a subject can never have an active role that is not authorized for its users. With all these privacy-based extensions (purposes and obligations), the role authorization in the core RBAC model is revised as follows:

\[ \forall s: \text{SUBJECTS}, u: \text{USERS}, r: \text{ROLES}, \text{op:OPS}, \{ o_1, o_2, ..., o_i \} \subseteq \text{OBS}, \{ pp_1, pp_2, ..., pp_j \} \subseteq \text{PURPOSES}, \{ obl_1, obl_2, ..., obl_l \} \subseteq \text{OBLIGATIONS} \]
\[ r \in \text{subject_roles}(s) \land u \in \text{subject_user}(s) \Rightarrow u \in \text{assigned_users}(r) \]

access: \( \text{SUBJECTS} \times \text{OPS} \times \text{OBS} \times \text{PURPOSES} \times \text{OBLIGATIONS} \rightarrow \text{BOOLEAN} \)

- access(s, op, \{ o_1, o_2, ..., o_i \}, \{ pp_1, pp_2, ..., pp_j \}, \{ rp_1, rp_2, ..., rp_k \}, \{ obl_1, obl_2, ..., obl_l \}) = 1 \text{ if subject } s \text{ can access any object in } \{ o_1, o_2, ..., o_i \} \text{ using operation } op \text{ for any purpose in } \{ pp_1, pp_2, ..., pp_j \} \text{ with a set of obligations } \{ obl_1, obl_2, ..., obl_l \}, (0, \emptyset, \emptyset) \text{ otherwise.}

**DESIGN**

To realize our proposed model, we need to consider two core issues: authorization design and privacy policy specification.

**Design of authorization**

Authorization in our model implies two layers of control: authorization over users and authorization over data. Authorization over users is the prerequisite for authorization over data.

Before accessing a RFID tag, a user has to be authorized into a specific retailer enterprise boundary. Furthermore, all role players (e.g., cashiers, store manager, assistants, etc) have to use on-site equipments to pass the authorization process, e.g., using unique login id and password pairs. As shown in Figure 2, these role players are delimitated by an enterprise boundary. An enterprise boundary is defined for a specific
privacy issues of applying rfid in retail industry

9

After a user is authorized to enter a retailer enterprise boundary, the second layer of authorization control is applied over data. Different roles of authorized users have distinct levels of access to RFID data. For example, not all kinds of information can be embedded and stored into a RFID tag, even from an authorized role player. In a more detailed scenario, a cashier may store the last four digits of a customer’s credit card number into a RFID tag, but may not be allowed to input all the digits. In other words, authorization over data is ruled by the access rights associated with each role predefined by each specific retailer. A retailer customizes roles that have access to the enterprise together with relative access rights. In this paper, we focus on the privacy policies that can be defined in access rights.

Design of privacy policy specification

In order to guard and ensure privacy authorization, one core issue is that we have to explore approaches to precisely describe and define privacy policies. In this paper, a privacy policy is generally defined as a set of rules and practices that specify or regulate how a system or organization provides services to protect information. A privacy assertion is typically scrutinized in the context of privacy policy. In more detail, the engineering of a privacy policy starts with risk analysis and ends with a set of privacy assertions that is ready for integration into the system architecture of a subject. Risk analysis identifies privacy threats in a business process and then forms a set of privacy assertions, which refer to rules and practices to regulate how sensitive data or activity information will be managed and protected within a loosely coupled execution environment. A privacy policy is often formalized or semi-formalized in an authorization model that provides a basis for a formal analysis of privacy properties. A powerful and extensible description language is thus on-demand.

Instead of introducing another new policy description language into the already-overcrowded description language family, our strategy was to extend existing relative languages for the purpose of enterprise-level RFID privacy control. Our procedure was to examine every candidate policy description language to verify whether it could be extended to realize the core privacy policies in our RFID-oriented privacy authorization model. Through this process, we finally found a right candidate called Enterprise Privacy Authorization Language (EPAL) (Ashley, 2003). EPAL is an interoperability language for defining enterprise privacy policies. It is used to govern data handling practices in the context of fine-grained positive and negative authorization rights. The detailed realization of our RFID-oriented privacy policy specification using EPAL will be discussed in the next section. Here we first discuss how to extend the basic terminology of EPAL to define fundamental concepts of RFID-oriented privacy policies.

EPAL includes lists of hierarchies of data-categories, user-categories, and purposes, as well as sets of actions, obligations, and conditions. We extend these elements to define privacy policies of RFID usages. Data-categories can be extended to define different categories of collected data handled differently from a privacy perspective (e.g., RFID information). User-categories can be extended to describe the users or groups (e.g., cashiers) that can access collected RFID data. Purposes can be extended to model the intended service for which data is used (e.g., cash register payment). Actions can be extended to model how the data is used such as reading, printing, and storing the RFID information in computer systems. Obligations can be extended to define actions that must be taken by the environment of EPAL. One example of obligation is that a retailer must delete the collected data after one year of storage. Overall, these elements can be used to specify privacy authorization rules, which allow or deny actions on data-
categories by user-categories for certain purposes under certain conditions while mandating certain obligations. Furthermore, privacy authorization rules are sorted by descending precedence in EPAL policies.

From a technical perspective, there are two major components in EPAL: Vocabulary and Privacy Policy. Enterprises usually use a single vocabulary for exchanging policy data with other enterprises. The concept of vocabulary is thus somehow similar to the concept of ontology. Since the recent version of EPAL vocabulary only provides a basic structure for describing data categories, the semantic Web may help provide explicit meaning to information available on the Web for automatic process and information integration based on the concept of ontology. In future work, we will explore semantic extensions to EPAL.

After carefully examining EPAL, we decided to extend EPAL as the tool to implement the privacy authorization model of RFID. Beyond the fact that EPAL is one of the most promising privacy authorization languages in the market, we have successfully extend the basic building blocks of EPAL to define various comprehensive RFID-related privacy policies, which fact will be reported in detail in the next section.

REALIZATION

In this section, we will present how we realize the core privacy policies in our RFID-oriented privacy authorization model by extending essential EPAL constructs. The major RFID privacy policies are divided into two categories: (1) privacy authorization model related to data collection, and (2) privacy authorization model related to data processing.

Privacy authorization model related to data collection

Privacy policies on cashiers’ handling of private information

Let us starts from a simplified example where a RFID tag is manufactured as read only. In the following example, the retailer specifies that the cashier can read any data stored in a RFID for payment purposes. This is a very basic policy for the task that a cashier normally conducts. It says that a cashier is able to read the data in order to get the price information (potentially stored in a computer, not in the RFID itself) by matching the data in the RFID. The policy can be expressed as follows:

<ALLOW
user-category = “cashier”
data-category = “RFID”
purpose = “payment”
operation = “read”
condition = “TRUE”/>

A more complicated situation lies in that some RFID tags are manufactured as read/write. Writing data into RFID may cause privacy concerns. For example, if a cashier is able to write data into the RFID tags, the cashier may write some PII. This may happen if the cashier has malicious intentions, for example. Later on when the same RFID tag read by other readers, this PII may be unintentionally revealed to the third party. Therefore, a privacy policy needs to be defined to specify that writing PII into RFID will be denied. The following is the essential part of the rule saying that cashiers are denied the right to write any PII into a RFID tag.

<DENY
user-category = “cashier”
data-category = “RFID”
purpose= “tracking”
operation = “write”
condition = “/CustomerRecord/Data/DataType=PII”/

**Guarantee that the same tag cannot be read by other types of RFID readers**

While interoperability among heterogeneous information systems is highly appreciated, it is not a good idea to allow RFID readers to read all the information stored in a RFID tag. One privacy concern for the RFID technology is that data stored in one type of tags (e.g., the tags used exclusively by Wal-Mart) should not be able to be read by other types of readers (e.g., the readers from another retailer Target). The reason is obvious: it is unreasonable to design a system that encourages the “sharing” of information among competitors.

In the following policy definition, it says that if the types of RFID tags do not match, the reading of the RFID data is denied. This type of privacy policy also helps monitor the behavior of dishonest employees who may leak confidential information to the competitors.

<DENY
user-category = “other”
data-category = “RFID”
purpose= “collect”
operation = “read”
condition = “/RFID/TagReaderType != /TagReader/TagReaderType”/>

**Privacy authorization related to data processing**

**Guarantee that the RFID tag can be destroyed**

The ultimate solution to prevent privacy violation in the RFID is to destroy the RFID tag permanently. One privacy policy for the retailer would be that, the store manager (not cashier) is allowed to destroy the RFID tag permanently after the customer has purchased the products. It may be a rarely used privacy policy, but it is desirable to provide this type of services to the customers, so that the customers would be assured that their privacy is fully protected. Of course, the store manager has to get the consent from the customers before destroying the RFID tag. As following, we realize this policy by equipping a customer record with the information regarding whether the destruction of the tag is needed.

<ALLOW
user-category = “store_manager”
data-category = “RFID”
purpose= “PreventViolation”
operation = “destroy”
condition = “/CustomerRecord/DestructionConsent = TRUE”/>

It is very important for the customers to understand the consequences if they choose this option. Furthermore, the customers may have to guarantee that they will not return the products. The reason is that if the RFID is destroyed, it will be difficult for the retailer to re-sell it. An alternative may be that the retailer can charge a small fee for the RFID disposal in order to compensate the re-installation of a new RFID.
Guarantee that the RFID tag can be blocked

Recently, RSA Security has developed a tag called “RSA Blocker Tag”, which can effectively block the reading of data in the RFID tag. When a Blocker is in proximity to ordinary RFID tags, they benefit from its shielding behavior; when the Blocker tag is removed, the ordinary RFID tags may be used normally. We realize this feature through a rule, which indicates that the customer can get a “free” blocker tag if the approval from the store manager has been granted. The core part of the rule is shown as follows.

<ALLOW
  user-category = “store_manager”
data-category = “RFID”
purpose= “block”
operation = “Insert”
condition = “/CustomerRecord/BlockingRequest = TRUE”/>

Guarantee that the collected data will not be sold to the third party

In most cases, the customer may not be willing to reveal the PII to the third party. Therefore, the retailer should define a policy rule to describe it. The following is the essential part of the definition.

<DENY
  user-category = “store_manager”
data-category = “RFID”
purpose= “Marketing”
operation = “sell”
condition = “/CustomerRecord/ThirdPartyConsent = FALSE”/>

Assessments

We just presented our privacy authorization model tailored for enterprise-level of RFID usages. The advantages of our introduction of enterprise (retailer) boundary are three-fold. (1) It enables the authorized users to access information effectively and to achieve a retailer’s strategy within its boundary. (2) It prevents sensitive information from leaking out of the corresponding boundary. (3) Interfaces built upon enterprise boundaries can potentially provide flexible and friendly interaction paths between suppliers and retailers, retailers and retailers, retailers and customers.

Our introduction of two-layered authorization model over users and data establishes a stronger protection of privacy control. We believe that our model will not only provide a flexible and powerful privacy protection model for the retail industry that wishes to adopt the RFID technology, but also strengthen the confidence of customers who will be involved with RFID technology.

Our model is associated with capability of precisely defining comprehensive privacy polices. The policy description is realized by extending EPAL, a language for specifying enterprise privacy policy, to specify RFID-related privacy policy. The advantages of extending an existing description language are three-fold: (1) we can save the time to develop another language to an already-overcrowded language set; (2) learning curve can be largely shorted; and (3) extending a proven and tested existing language will add more confidence to stakeholders of retail industry.

However, we do realize that our work in this paper is just a starting point of further investigation and exploration of RFID-specific privacy specification and protection. Our current extension of EPAL language
is based upon our experiments on RFID-related privacy polices. We need to collect all possible privacy policies from different enterprise retailers to validate and complete the library of RFID-oriented privacy policies. The information could also help us evaluate the usage of EPAL language. In addition, we have not deeply explored the realization of “GateKeeper” to guarantee that a defined privacy policy using our implementation will not be violated. We also need to explore semantic extensions to EPAL in order to further protect privacy. In order to provide a comprehensive RFID-oriented privacy model, there is still serious research work ahead.

Conclusion and future work

In this paper, we have investigated the privacy design issues for adopting RFID in the retail industry. An enterprise RFID-oriented privacy authorization model is proposed. Since the retail industry has yet to widely use RFID for identifying each individual item to be sold, the work presented here may serve as a good starting point for more elaborate work in the future. However, we firmly believe that it is only a matter of time when we can actually use item-level RFIDs in retail stores. When the RFID technologies become mainstream, which could be soon, privacy will become the bottleneck to its extensive adoption. In our view, it is never too early for researchers to consider the privacy implications of adopting RFID in the retail industry. The rest of the work will follow logically from that.

The following are the tasks we plan to conduct in the future. First, we want to gather the best practices of RFID privacy. Since scenarios for applying RFID in the retail industry are futuristic in nature, we may have to mainly depend on the Internet to conduct the research. Second, we want to investigate the usage of P3P (Cranor, 2002; Agre, 1997) outside the domain of personal data collection for Web sites. Third, we want to conduct direct or indirect interviews with representatives of industrial companies and privacy advocates.

Acknowledgement

This research is partly funded by a discovery grant (NSERC PIN: 290666) from the Natural Science and Engineering Research Council (NSERC) of Canada.

REFERENCES


