社団法人　情報処理学会　研究報告
IPSJ SIG Technical Report

2005－UBI－8（60）
2005／6／10

RFID Privacy Enhanced Protocol for Library Operation

Gembu MOROHASHI, Shingo KINOSHITA, Fumitaka HOSHINO
NTT Information Sharing Platform Laboratories, NTT Corporation
1-1 Hikarinooka, Yokosuka, Kanagawa, 239-0847 Japan

Abstract

Radio Frequency IDentification (RFID) is expected to be one of the most important infrastructure technologies for future ubiquitous services. Libraries may quickly adopt item-level tagging, since doing so would dramatically reduce inventory management costs and make possible self-checkout. Despite this advantage, such a wide deployment of item-level tagging may expose new threats to citizens' privacy through the abuse of the powerful tracking capability of the RFID. These threats are critical to book owners and borrowers because books reflect the preferences and thoughts of their owners.

This paper clarifies the privacy issues related to use of RFIDs in libraries and proposes a new privacy protection protocol using the Unidentifiable Anonymous-ID scheme, which we have already developed to protect RFID privacy at a very low cost. This protocol not only protects users' privacy but also suits the work done in libraries. We have evaluated the performance of the protocol and have proven its effectiveness.

Key Words: RFID, library, privacy, unidentifiable anonymous-ID scheme

1 Introduction

Radio Frequency IDentification (RFID) is expect to be one of the most important infrastructure technologies for future ubiquitous services, and at present, various applications are being considered for it in the publication industry, in supply chain management, in logistics, for preventing shoplifting, etc. In particular, in the supply chain and logistics applications, RFIDs are being used for tagging cases or pallets. RFID systems can also help to prevent resale of shoplifted books. Moreover, the advantages of this technology can be exploited to simplify library operations, and many libraries have begun using RFID systems. In a library, each individual item is given its own tag. In fact, this RFID application is being watched with keen interest because it represents a case of large-scale item-level tagging.

On the other hand, it has become a serious worry that the privacy of library patrons may be threatened by people abusing RFID tracking technology. In fact, there are anti-RFID movements [2] whose members claim that an individual's privacy may be adversely affected by the deployment of RFID technology. In response to the perceived problem, the authors have developed technologies to prevent patrons' privacy from being threatened by RFIDs, and they have proposed the Unidentifialbe Anonymous-ID scheme [4] and the Hash-Chain scheme. [5, 7]

In this paper, we discuss RFID privacy issues as they relate to libraries, and propose a new privacy protection protocol to resolve these issues within the Unidentifiable Anonymous-ID scheme. We also evaluate the performance and effectiveness of the protocol.

2 RFID Technology for Libraries

2.1 Effects of RFID Systems in Libraries

Many libraries have or are adopting electro-mechanical systems in an attempt to increase their efficiency of the operations. In current architectures, most of them use barcodes embedded in each item. The ID numbers of the barcodes are related to titles, authors, publishers, etc. and are used for charging/discharging and inventory. By using RFIDs instead of barcodes, one is able to read multiple items simultaneously. RFID

systems simplify charging/discharging and make inventory work efficient [1]. For instance, a certain librarian with RFID systems inventories 5,000 books per hour [8]. By letting RFID readers scan books on the shelves all the time, it is possible to automatically inventory, search for, and record both the browse time and browse count of books. Furthermore, security is promoted because gate readers are placed at exit of a library, and they scan for books that have not been charged.

The Santa Clara City Library in California, the University of Nevada, the Las Vegas library, and the Oregon public library have already adopted RFID systems, and have tagged every item in their collections, including all books, tapes, and CDs. In North America, approximately 130 libraries are using RFID systems [6].

2.2 Library RFID Problems

The main obstruction to RFID adoption is its cost. According to [1], the whole cost of tags, readers, servers, and their setup is $70,000 for a library of 40,000 items. It jumps to $168,000 for a library of 100,000 items. A half of the budget is for tags.

One of the problems is that if tags were covered by foils, which prevents radio-frequency communication, the exit reader's scan will fail. Another serious issue is the threat to patron privacy. In particular, the collections of libraries are sensitive, and libraries are public, so that libraries need to care for their patrons' privacy.

There are two privacy issues. One is an existing problem which is related to the contents of items. The other is caused by using RFID technology.

Privacy Issues Related to Books The books in libraries could be said to reflect their patrons' hobbies, thoughts, and behaviors. The libraries themselves also include much data related to a patron's privacy. The history of a patron's charging and browsing is a list of such data. Such data is strongly related to a patron's privacy, and thus it needs to be managed very carefully.

RFID Privacy Issues Despite the advantages of RFID technology, its threat to patron privacy cannot be overlooked. If the memory of a tag is inputted with basic data of the book such as the title, author's name, published year, etc., that data could be scanned by someone carrying a small bag of eavesdropping equipment.

As a remedy to this problem, only a unique ID number could be inputted to the memory. While this seems to prevent eavesdropping to learn the basic data of the book, the eavesdropper or adversary can instead make a list of IDs and the basic data of the books. (This list is called a 'hotlist.') In doing so, the adversary can learn the data about the book from its unique ID.

Such unique IDs are used in libraries for managing their items. Each item is given a unique ID, and this ID is inputted to its tag. The adversary thus can relate the patron to the ID of the item, and track the ID. Consequently, the adversary can learn when and from where the patron has come.

Focus We have noted the privacy issues with libraries using RFIDs, and in this paper, we focus on the following problems:

(1) leakage of browsing histories,
(2) leakage of content data concerning books,
(3) tracking of patrons.

3 RFID Privacy Protection Schemes

Schemes have been proposed that protect patron privacy from threats exploiting RFIDs. We compare these schemes as to their effectiveness in libraries.

Kill Feature Some schemes disable the ability to read tags. These are not good for libraries that need to read the tags to perform charging.

Read Access Control with a Password To prevent tags from being read by an adversary, an access control process can be adopted. For example, one such scheme uses a password. However, its driving cost is high because updating the passwords of all tags is very hard.

Unidentifiable Anonymous-ID Scheme [4] RFID privacy problems have two characteristics, one is content privacy, the other is location privacy threatened by ID tracking. An ID needs unidentifiability and unlinkability in order to solve these privacy problems. The Unidentifiable Anonymous-ID scheme ensures unidentifiability and unlinkability. In this scheme, original data (or an original ID) is encrypted (anonymized), and is updated (re-anonymized) on an arbitrary occasion. The encryption and re-anonymization calculations are carried out on a computer.

RFID tags have a readable and writable memory, such as EEPROM, so that the cost of tags and systems is relatively low. EPC global has proposed a standard protocol for a new generation of RFIDs, Class-1 Generation-2, which requires the memory on a tag to be readable and writable. It is possible to use the Unidentifiable Anonymous-ID scheme with the new standard tag. However, it is difficult to decide when the re-anonymizing process is to be carried out. As the ID in the tag is not re-anonymized automatically, it becomes necessary to re-anonymize the ID within a certain period. In libraries, re-anonymization can be made "seamless" by performing it during the charging/discharging process. Doing so seems to solve the problem of the scheme.

Hash-Chain Scheme [5, 7] The Hash-Chain scheme is a privacy protection scheme that has small cryptographic functions so as to ensure the unidentifiability and the unlinkability of IDs. There are two hash functions on the tag. Unlinkability is ensured by carrying out the re-anonymizing process inside the tag. This scheme, however, costs more than the Unidentifiable Anonymous-ID scheme.

4 Proposed Protocol

Because of its reasonable cost and feasibility as outlined above, we chose the Unidentifiable Anonymous-ID scheme on which to base a protocol for library operations.

4.1 System Architecture

Charging/discharging are essential library operations. We assume that all collections have been tagged, and each item has been given a unique ID (book ID). Our protocol for charging/discharging needs several units: reading/writing terminals on the charging/discharging counter (charging/discharging terminals), a database server that manages charging, and a security server that encrypts, decrypts, or re-anonymizes the IDs of every patron (user ID). Note that the patron's ID card should not be an RFID. Instead, magnetic stripes or barcodes should be used to decrease the threat to privacy.

4.2 Method for Each Entity

Next, we discuss how the ID is encrypted in a tag, how the re-anonymizing process is carried out, and how charging data is recorded in the DB.

Encryption Methods for ID in the RFID tag The ID in a tag should be encrypted to ensure unidentifiability. One method is to encrypt the ID only when the items are taken from the library. The advantage of the method is that it is not necessary to decrypt the ID in the library. This is useful in operations such as searches of books on shelves. Moreover, unless the ID is certainly encrypted on charging, patron privacy will be threatened. Therefore, the ID in the tag is always encrypted in the protocol.

ID Re-Anonymizing Process In the unidentifiable anonymous-ID scheme, because of the reliable unlinkability of IDs, re-anonymizing calculations are carried out on a computer (the security server in the protocol). It is appropriate that re-anonymization be done during the charging/discharging process, which is already an essential library operation, because it can be made to appear as if no additional work is being done.

Registration to the charging management DB The charging histories are also related to patron privacy. Care has to be taken when the data is registered in the charging management DB. The registering data is related to the patron, borrowed items, dates of charging/discharging, etc. To counter the possibility that such data may be leaked, the relation between the user ID and the ID of the borrowed item should remain unknown. Therefore, the registered IDs should be encrypted. There is a method in which both book ID and user ID are encrypted in the DB. However, we have adopted another method, which encrypts either book ID or user ID, because it reduces the operational load. That is, if the book IDs are encrypted in the DB, all the IDs in the charging list must be decrypted in the discharging process, because it is necessary to check for which book has been returned. Instead, if only the user IDs are encrypted, one decryption is needed to check. Note that the user IDs are encrypted with a probabilistic encryption, which is an encoding scheme in which a different cipher text is generated each time from the same plain text, so that it is difficult to guess how many books have been borrowed by the same patron.

4.3 Protocol Instance

The figure shows a part of the protocol about the charging/discharging process as an instance of the proposed protocol.

Charging Process The patron brings a book that he or she wants to borrow to the charging terminal. The charging procedure is as follows:

1) Put the book on the charging terminal to read the encrypted ID.
2) The terminal sends the ID to the security server, then the server decrypts to the original book ID. At the same time, the terminal reads the patron's user ID and sends it to the server. The server encrypts it and returns both the decrypted book ID and the encrypted user ID.
3) The terminal sends both returned IDs to the charging management DB.
4) The DB keeps related information on registered books with book IDs as their indexes. The sent IDs and the dates of charging/discharging are registered in the charging list.
5) After the DB processes have ended, the charging terminal sends the book ID to the server for re-anonymizing. The terminal then writes re-anonymized ID into the tag on the book.

Discharging Process The discharging procedure on the discharging terminal is as follows:

1) The encrypted ID in the returned book is read by the discharging terminal.
2) The terminal sends the ID to the server, then the ID is decrypted. The returned ID is sent from the terminal to the DB.
3) The DB searches for the entry of the decrypted ID in the charging list. If a matching entry is discovered, that entry is removed from the list. The charging registration of the book is then deleted.
4) After the DB processes have ended, the charging terminal sends the book ID to the server for re-anonymizing. The terminal then writes the re-anonymized ID into the tag on the book.

5 Efficiency

We discuss the three issues that were mentioned in section 2.2. We suppose that there are various attacks against the protocol and the security server is a trusted third party (TTP).

Leakage of Content Data Concerning Books Consider the data in an RFID tag on the book. The book ID in the tag is always encrypted, so that the content data concerning the book cannot be obtained from the tag.

There are 'hotlist' attacks, as described in section 2.2. In the protocol, the ID in the tag is re-anonymized during each charging/discharging process. Thus, it seems that the ID in the tag will be different from the ID in the hotlist.

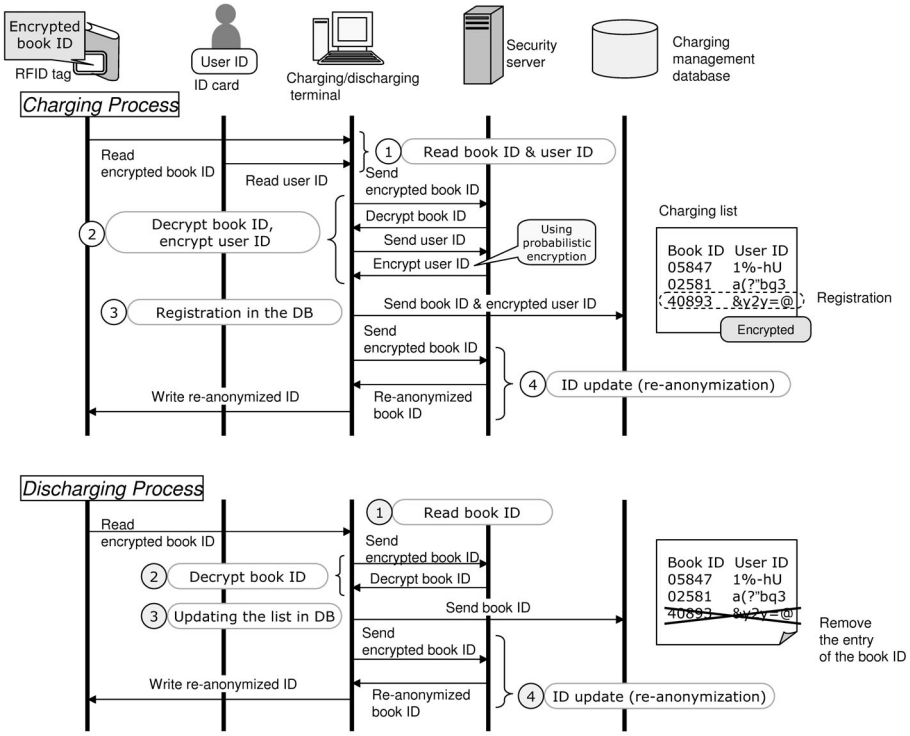

FIGURE. a Privacy Protection Protocol using the Unidentifiable Anonymous-ID Scheme (charging/discharging process)

Charging Management DataBase Consider the case that data in the DB has been leaked. The data registered in the DB is charging book IDs, patrons' encrypted user IDs, and dates of charging/discharging. The user IDs are encrypted by probabilistic encryption, so that it is difficult to determine the degree of relatedness among the encrypted IDs, even if the same ID is encrypted. It is thus difficult to know who has borrowed the book. If only the ID of the borrowed books was registered in the DB, the charging histories would not be able to be leaked. Moreover, the data on the borrowed book is not related to the charging list, because the ID in the tag on the book is encrypted.

Tracking of Patrons Although the ID in tag on a book is encrypted, it is possible to regard the cipher text as a new fixed ID. The threat of tracking thus remains. In the protocol, the ID of a book is re-anonymized in the charging/discharging process, so that the old ID and updated ID are unlinkable. However, the ID is never updated while the book is borrowed, i.e. outside of the library, and thus, the patron may be threatened by tracking. Because the loan period is relatively short and it is difficult to obtain data on patron from the DB, the threat is relatively small. To solve the problem, however, the ID should be re-anonymized at the proper time. This compels the patron to update the ID on the terminal in the library. Alternatively, a trusted re-anonymizing terminal could be set in the patron's house, public offices, or stations. By using the Hash-Chain scheme, the re-anonymizing process could be carried out in the tag. This scheme fundamentally ensures patron privacy, but at higher cost.

The other attack is using hotlist of the kinds of books a patron borrows. The adversary targets a patron and eavesdrops on the book ID each time he or she burrows, through which the adversary obtains the patron's charging histories. As the focus here is on the book IDs of the hotlist, the borrowing pattern would seem to show what kinds of people read such books and show a relation between these people. The protocol's re-anonymizing processes, however, makes such a relation of patrons unlinkable.

Other Issues Regarding the decryption or re-anonymization process, the sercurity server searches for which key is needed to decrypt the ID. The simplest method uses the same key for all IDs; this method is very

vulnerable, and key updating with it is heavy work. The other method prepares a key for each ID. This method needs a key-ID in all tags, but each key-ID is unique; i.e., it is not unidentifiable. Therefore, the optimal method uses several keys. Note that the IDs, which are decrypted by the same key, should not be related each other. The book ID is usually numbered according to ISBN, Nippon Decimal Classification (NDC), UCC/EAN, EPC, etc. In such a case, the ID code has information on the content. Grouping into meaningless or random patterns avoids content leakage. This could be done, for example, by grouping by end number, by date of delivery to the library, etc.

Because the RFID tag on the book has a writable memory, there is threat by which the memory in tag is tampered. Despite this, the tampered data is easy to correct, because the barcode of correct ID is printed on the book.

Patron self-charging/discharging reduces the workload of the librarians. The problem of self-charging/discharging is knowing whether the process has been terminated correctly. The threats which we have clarified above would be invited whenever the re-anonymizing process is not properly completed. Thus some mechanism is needed to alert the patron when books have been removed from the terminal before the end of the process.

6 Conclusion

We have discussed the privacy issues of using RFID in libraries, and have proposed a new privacy protection protocol to resolve them by using the Unidentifiable Anonymous-ID scheme, which we have already developed to protect RFID privacy at very low cost. In this protocol, RFID tags do not need to have special functions. Therefore, the protocol is a feasible privacy protection scheme. The re-anonymization outside of a library has been left as a future work. We believe that the re-anonymization problem, while being limited by the present scheme, could be further limited by setting charge/discharge terminals in nearby shops.

References

[1] R. Boss, "RFID Technology for Libraries," Library Technology Reports, 39(7), Nov./Dec. 2003.

[2] CASPIAN, "Spychips: RFID Privacy Website," `http://www.spychips.com/`.

[3] EPC global Inc., "EPC Radio-Frequency Identity Protocols Class-1 Generation-2 UHF RFID Protocol for Communications at 860MHz–960MHz," Version 1.0.8, EPCglobal Inc., Dec 2004.

[4] S. Kinoshita, F. Hoshino, T. Komuro, A. Fujimura, M .Ohkubo, "Low-cost RFID Privacy Protection Scheme," J. IPS. Japan, vol.45, no.8, pp.2007–2021, Aug. 2004. (In Japanese.)

[5] S. Kinoshita, M. Ohkubo, F. Hoshino, G. Morohashi, O. Shionoiri, A. Kanai, "Privacy Enhanced Active RFID Tag," The 3rd International Conference on Pervasive Computing, `http://www.pervasive.ifi.lmu.de/`, May 2005.

[6] D. Molnar, D. Wagner, "Privacy and Security in Library RFID: Issues, Practices, and Architectures," 11th ACM Conference on Computer and Communications Security, ACM Press, `http://www.cs.berkeley.edu/~dmolnar/library.pdf`, Oct. 2004.

[7] M. Ohkubo, K. Suzuki, S. Kinoshita, "Cryptographic Approach to a Privacy Friendly Tag," RFID Privacy Workshop@MIT, `http://www.rfidprivacy.org`, Nov. 2003.

[8] L. Smart, "Making Sense of RFID," Library Journal, `http://libraryjournal.reviewnews.com/article/CA456770`, Oct. 2004.