

# Location Recognition in RFID Bookshelves

Sozo INOUE<sup>†a)</sup>, Member, Yasunobu NOHARA<sup>††</sup>, Masaki TAKEMORI<sup>†</sup>,  
and Kozo SAKURAGAWA<sup>†††</sup>, Nonmembers

**SUMMARY** We consider RFID bookshelves, which detect the location of books using RFID. An RFID bookshelf has the antennas of RFID readers in the boards, and detects the location of an RFID tag attached to a book. However, the accuracy is not good with the experience of the existing system, and sometimes reads the tag of the next or even further area. In this paper, we propose a method to improve the location detection using naive Bayes classifier, and show the experimental result. We obtained 78.6% of F-measure for total 12658 instances, and show the advantage against the straightforward approach of calculating the center of gravity of the read readers. More importantly, we show the performance is less dependent of a change of layouts and a difference of books by leave-1-layout/book-out cross validation. This is favorable for the feasibility in library operation.

**key words:** RFID, RFID bookshelf, library, location recognition

## 1. Introduction

RFID (Radio Frequency Identification) is a technology to identify objects and people using small RF devices called RFID tags. RFID is beginning to be attached to the books as in Fig. 1 in libraries, and widely used as media of multiple applications, such as automatic check in/out machines, anti-theft gates, inventory management, and automatic book stacks [3].

An *RFID bookshelf* is a bookshelf with RFID readers embedded. It can detect which area of the bookshelf each book is by using multiple antennas, each of which corresponds to a separated area of the bookshelf.

However, there are cases where some antennas read the area of not in charge, such as a next part, or even further. This can be problematic since the bookshelf only knows that the book exists, but not where. Of course, antennas can be tuned in the stages of development and installation, but there are limitations in physical level adaptation, since they need extra effort which is not feasible in the real situation.

In this paper, we solve the problem of accurate location detection in bookshelves with the system level approach, including pattern classification. At first, we model the problem as a classification problem, and fit it to naive Bayes classifier. Next, we show the experimental result using an RFID

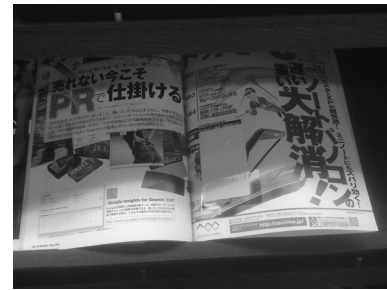


Fig. 1 A book attached with an RFID tag.

bookshelf located in the real library. For the logs obtained, we applied the naive Bayes classifier to improve the location detection.

We obtained 78.6% of F-measure for total 12658 instances, and show the advantage against the straightforward approach of calculating the center of gravity of the read readers. More importantly, we show the performance is less dependent of a change of layouts and a difference of books by leave-1-layout/book-out cross validation. This is favorable for the feasibility in library operation.

In the following sections, we introduce RFID bookshelves and their challenges in Sect. 2 and define a problem and show two kinds of solutions: center of gravity and naive Bayes in Sect. 3. Moreover, experimental evaluation is described including the method, result, and discussions in Sect. 4. Finally, we conclude in Sect. 5.

## 2. RFID Bookshelf

Figure 2 is the picture of an RFID bookshelf of 210 cm height and 90 cm width. It has an RFID antenna for each block, and reads all RFID tags within a few minutes. In this section, we firstly describe considerable applications of RFID bookshelves, which can be combined with other RFID applications. Then, we address technical challenges for location detection in RFID bookshelves.

### 2.1 Application

RFID bookshelf can introduce significant advantages to libraries or other types of facilities such as bookstores. It provides data of fine-grained use cases to be used for data mining and/or marketing. It can be also valuable to characterize a set of books from a use pattern, to predict a use patterns

Manuscript received November 10, 2010.

Manuscript revised March 2, 2011.

<sup>†</sup>The authors are with Faculty of Engineering, Kyushu Institute of Technology, Kitakyushu-shi, 804-8550 Japan.

<sup>††</sup>The author is with Kyushu University Hospital, Fukuoka-shi, 812-8582 Japan.

<sup>†††</sup>The author is with Fusic Co. Ltd, Fukuoka-shi, 810-8609 Japan.

a) E-mail: sozo@mns.kyutech.ac.jp

DOI: 10.1587/transinf.E94.D.1147



Fig. 2 View of the RFID bookshelf.

and to help the users with similar patterns, or the library itself can utilize the patterns for the book selection. Especially, the data will be valuable if the use cases are collected together with the logs of other RFID applications, which leads to lineage data of specific books in the hall.

Furthermore, some of the bookshelves, as shown in Fig. 2 has the ability of telling the user in front of the bookshelves by blinking LED each of which corresponds to the area of detection. Using this, the library can tell the place of a book a user looks for, by attaching LED etc to them and blinking it. Moreover, by connecting to the network, a service of informing to a user's cellular phone by e-mail is possible when the book currently read inside a hall by another has returned to the bookshelf.

Since RFID is being introduced to libraries, RFID bookshelf can be considered as a reasonable extension to existing RFID library systems, in the sense that no extra payment for RFID tags is needed. There are still many libraries which operate with barcode library systems, but unfortunately, they cannot be affordable to achieve this kind of location detection in bookshelves. RFID bookshelves can be a killer application of library RFID.

## 2.2 Challenge

However, during the operation of an RFID bookshelf, we found there are cases where some antennas reads the area of not in charge, such as a next part, or even further. This can be problematic as well as the case where a book is not identified by any readers, since the bookshelf only knows that the book exists, but not where.

Ofcourse, antennas can be tuned in the stages of development and installation. But the case we experienced is after the tuning has been done, and it proofs there are limitation in physical level adaptation, for some reasons such as they need more cost to achieve complete accuracy.

Therefore, in the rest of the paper, we solve the problem of accurate location detection in bookshelves with the

system level approach, including pattern classification.

## 3. Location Detection in Bookshelves

In this section, we formally define book location detection problem at first. Next, we present a straightforward idea of calculating the center of gravities, which is shown here as one compared with our proposed method. Then, we show our proposed method to use naive Bayes classifier.

### 3.1 Problem Definition

In real situations, multiple books are being in the bookshelf, but here we formalize the problem for detecting location of a single book. This is because those for multiple books are a trivial extension of detecting single book. We also consider the granularity of detection is each area of antennas, from the constraint of RFID bookshelf.

Let the bookshelf have  $N$  areas to detect, each of which are equipped with an antenna. Then, we denote  $x_i = 1$  if an area  $i$  detects a book, and  $x_i = 0$  otherwise. That is, we can denote an instance of detections as  $(x_1, x_2, \dots, x_N)$  when a book is located at the area  $y \in \{1, \dots, N\}$ .

Here, problem of book location detection is defined as follows:

#### Definition (*Book location detection problem*)

Let  $X = \{(x_1, x_2, \dots, x_N) | x_i \in \{0, 1\}, i = 1, 2, \dots, N\}$  be a set of RFID readings, and  $Y = \{1, 2, \dots, N\}$  be a set of book locations. For given instances  $X' \times Y' \subset X \times Y$ , estimate a classification function  $f : X \rightarrow Y$ .

That is, it is a classification problem of learning unknown vectors in the form of  $(x_1, x_2, \dots, x_N, y)$  from a training set consists of instances in the same form.

### 3.2 Center of Gravity (COG) Method

This is to use the geometries of each areas, where the center of gravity (COG) are calculated from the areas where the books are identified, and to answer the nearest location as the location of the book. This method is one of the straightforward ways of calculating location, and actually it does not use the learning data.

### 3.3 Naive Bayes Classifier

In recent years, naive Bayes classifier is often adopted as a powerful method to solve classification problems [1], [2]. Here, we describe how to solve book location detection problem with naive Bayes classifier.

The location detection problem is represented as  $p(y|x_1, x_2, \dots, x_N)$  for each location  $y \in Y$ . This value is,

$$p(y|x_1, x_2, \dots, x_N) = \eta p(y) p(x_1, x_2, \dots, x_N|y)$$

from Bayes theorem. where  $\eta = 1/p(x_1, x_2, \dots, x_N)$  is a normalization constant. If we assume a strong independence

$$p(x_1, x_2, \dots, x_N | y) = \prod_{i=1}^N p(x_i | y),$$

then,

$$p(y | x_1, x_2, \dots, x_N) = \eta p(y) \prod_{i=1}^N p(x_i | y).$$

To obtain the maximized posteriori, we can have  $y$  which maximizes the value as follows:

$$y = \arg \max_y p(y) \prod_{i=1}^N p(x_i | y).$$

### 4. Experimental Evaluation

#### 4.1 Experiment

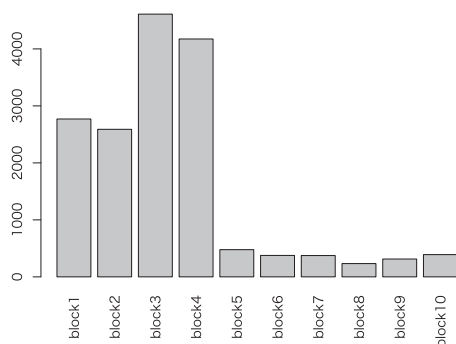
We did an experiment to examine the performance of the method. Firstly, we collected data for learning and testing. We placed 67 books to the bookshelf composed of 14 blocks, where 10 of them have RFID antennas behind. The books we have chosen are magazines, since they are one of the most required types to be captured the usage in the library, since they are usually prohibited to be checked out, and are not well known how they are used.

Table 1 is the layout of each block in the bookshelf. The numbers correspond to the locations of books read by RFID readers. The blocks with ‘None’ have no readers because of an implementational reason.

We tried two ways of placing books: one is placing books next to another book as usual bookshelf, and the other is displaying a few books with facing the cover, which is often done to display magazines. The ways of each block are shown in Table 1.

**Table 1** Layout of each block in the the bookshelf.

None	None
No.9 (Books)	No.10 (Books)
No.7 (Display)	No.8 (Display)
No.5 (Display)	No.6 (Display)
None	None
No.3 (Books)	No.4 (Books)
No.1 (Books)	No.2 (Books)



**Fig. 3** Histogram of books x reads by blocks.

How the books were located is shown as a histogram of books times read trials by each block in Fig. 3. Here, the histogram were decided with a librarian in order to reflect the real situation of usage. According to it, the highest blocks (No.9 and 10) had few books since it is a high position to take a look at. Additionally, No.5, 6, 7, and 8 also had fewer books since their type is ‘display’.

Thus, we oriented to the real usage situations of bookshelves in libraries.

#### 4.2 Steps

Upon the books located as above, we changed the layout for 10 times, and tried RFID reading for 10 times for each layout. In a single layout, the books were not moved. It takes a few minutes to read all the blocks, and it becomes the longer if the more books a block has.

After the reading of 10 layouts times 10 repeats, we analyzed the performance of location detection in the following steps:

1. Apply COG method to the result, and obtain the score, which is composed of precisions (the rate of correctly classified instances to a location) and recalls (that of correctly classified instances from an actual location) and F-measure (harmonic mean of precision and recall) for each blocks, and also the mean of them.
2. Generate naive Bayes model by learning, and obtain the score by *10-fold cross validation*, which get the score for 10% of the data for the learned model by the rest of data, and repeat it for 10 times for other 10 independent 10%. 10-fold cross validation is effective to remove overfitting.
3. Compare the scores of 1 and 2. Here, the COG method is not good when the gravity vector comes to in-between blocks. In that case, the method must randomly (or uniformly) decide one of adjacent blocks. To avoid this effect for some extent, we compare 1 and 2 by only the difference of the vertical positions of blocks (combine horizontal 2 blocks).
4. Investigate the effect of layout difference by 2 ways. First, *leave-1-layout-out cross validation*, which is the same way as 10-fold cross validation, except that the data are divided by single type of layout and the rests. In each fold, single type of layout are applied for a model learned from the rest data. Second, obtain the score of *layout-specific learning*, which divides the same layout data into learning set and applying set, and their scores are averaged for all layouts. This is done to see the performance of models specialized for each layout.
5. Investigate the effect of book difference by 2 ways: *leave-1-book-out cross validation* and *Book-specific learning*, which are done in the similar way as leave-1-layout cross validation and layout-specific learning.
6. Investigate the recognition error, including the case where no reader reads an RFID.

### 4.3 Result

After the readings, we obtained 12658 instances and 1862 failures from the total 14520 readings. We performed classification learning upon the instances above.

#### 4.3.1 COG Method

Table 2 is the confusion matrix of instances for each classified or original location by the COG method. Table 3 is the scores of it for each blocks and weighted means. As mentioned in Sect. 4.2, blocks are packed by horizontal pairs.

From the table, the F-measure mean is 86.4%.

#### 4.3.2 Naive Bayes Method

Table 5 is the confusion matrix of naive Bayes method. Table 6 is the scores of it for each blocks and weighted means.

**Table 2** Confusion matrix of COG method.

Classified as:	blk.1/2	blk.3/4	blk.5/6	blk.7/8	blk.9/10
blk.1/2	3626	2172	24	0	0
blk.3/4	0	4942	0	0	0
blk.5/6	0	0	738	20	0
blk.7/8	0	0	4	587	131
blk.9/10	0	0	0	0	414

**Table 3** Score of COG Method.

	blk.1.2	blk.3.4	blk.5.6	blk.7.8	blk.9.10	mean
Recall	62.3	100.0	97.4	81.3	100.0	88.2
Precision	100.0	69.5	96.3	96.7	76.0	87.7
F-measure	76.8	82.0	96.8	88.3	86.4	86.4

**Table 4** Score of naive Bayes method (packed).

	blk.1.2	blk.3.4	blk.5.6	blk.7.8	blk.9.10	mean
Recall	78.7	99.9	97.9	75.6	100.0	90.4
Precision	99.9	80.5	91.1	97.0	76.7	89.1
F-measure	88.0	89.2	94.4	85.0	86.8	88.9

**Table 5** Confusion matrix of naive Bayes method.

Classified as:	block1	block10	block2	block3	block4	block5	block6	block7	block8	block9
block1	2047	0	203	521	193	0	28	0	0	0
block10	0	134	0	0	0	0	0	0	0	64
block2	0	389	0	1912	20	473	0	0	0	0
block3	0	0	0	1950	695	0	0	0	0	0
block4	0	0	3	81	2245	0	0	0	0	0
block5	0	0	0	0	0	431	0	16	0	0
block6	0	0	0	0	0	0	319	0	0	0
block7	0	0	0	0	0	45	0	311	0	62
block8	0	61	0	0	0	0	0	0	210	0
block9	0	57	0	0	0	0	0	0	0	150

**Table 6** Score of naive Bayes method.

	block1	block10	block2	block3	block4	block5	block6	block7	block8	block9	mean
Recall	68.4	67.7	68.4	73.7	96.4	96.4	100.0	74.4	77.5	72.5	79.5
Precision	84.0	53.2	90.3	75.8	62.3	90.5	91.9	95.1	100.0	54.3	79.7
F-measure	75.4	59.6	77.8	74.7	75.7	93.4	95.8	83.5	87.3	62.1	78.6

To compare with COG method, we pack the result of blocks of the same height, and show in Table 4.

From the table, the F-measure mean of the packed score is 88.9%. Comparing this to COG score, naive Bayes method is only 2.5% better. However, the naive Bayes method has an ability of recognizing a block from another of the same height. If we force to divide the packed score in Table 2, the scores will decrease to about the half.

#### 4.3.3 Effect of Book Layouts and Books

Table 7 is the scores of layout-specific learning for each blocks and weighted means. Table 8 is the scores of leave-1-layout-out cross validation for each blocks and weighted means.

Table 9 is the scores of book-specific learning for each blocks and weighted means. Table 10 is the scores of leave-1-book-out cross validation for each blocks and weighted means.

The summary of Table 6–9 is shown in Table 11. From Table 11, although all the F-measure are worse than 10-fold cross validation, they are at most 7.9% worse in book-specific-learning.

For the effect of layout, recall decreases for layout-specific learning, while precision is better. As a result, the F-measure becomes 71.2%. That of leave-1-layout-out cross validation is 75.1%, which is a bit worse than 10-fold cross validation.

As for the effect of books, book-specific learning is 72.9% for the precision, but the recall is 69.7%. As a result, the F-measure is 70.7%. That of leave-1-out cross validation is 71.7%.

#### 4.3.4 Read Errors

Finally, we show the distribution of read errors, where no reader read a tag. Figure 4 is the histogram of read errors by blocks. Since block 1 to block 4 has many books, it is natural that they have many errors. On the otherhand, block

**Table 7** Score of layout-specific learning.

	block1	block10	block2	block3	block4	block5	block6	block7	block8	block9	mean
Recall	79.0	62.9	44.4	57.3	60.2	66.5	73.8	67.1	67.4	71.3	65.0
Precision	41.4	89.2	90.1	72.9	66.5	94.8	93.2	90.9	92.2	89.9	82.1
F-measure	54.3	73.8	59.5	64.2	63.2	78.2	82.4	77.2	77.9	79.5	71.2

**Table 8** Score of leave-1-layout-out cross validation.

	block1	block10	block2	block3	block4	block5	block6	block7	block8	block9	mean
Recall	66.7	71.0	67.1	74.1	96.4	95.4	100.0	74.2	80.9	33.0	75.9
Precision	82.9	47.9	88.3	75.9	61.3	92.0	93.1	88.0	100.0	32.4	76.2
F-measure	73.9	57.2	76.3	75.0	74.9	93.7	96.4	80.5	89.4	32.7	75.1

**Table 9** Score of book-specific learning.

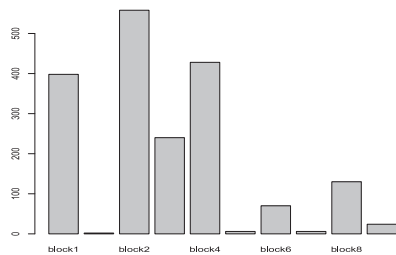
	block1	block10	block2	block3	block4	block5	block6	block7	block8	block9	mean
Recall	66.8	70.3	70.3	70.1	80.2	87.3	80.7	72.1	59.8	39.7	69.7
Precision	75.7	52.6	67.5	75.9	63.9	91.7	82.6	82.4	95.4	41.8	72.9
F-measure	71.0	60.2	68.9	72.9	71.1	89.4	81.6	76.9	73.5	40.7	70.7

**Table 10** Score of leave-1-book-out cross validation.

	block1	block10	block2	block3	block4	block5	block6	block7	block8	block9	mean
Recall	63.0	60.9	67.4	70.9	88.2	94.7	90.3	72.5	66.0	38.0	71.2
Precision	79.2	48.6	71.1	75.9	61.3	91.0	93.1	83.0	91.7	38.4	73.3
F-measure	70.2	54.1	69.2	73.3	72.3	92.8	91.7	77.4	76.8	38.2	71.7

**Table 11** Effect of books and layouts to scores.

	Recall	Precision	F-measure
10-fold cross validation	79.5	79.7	78.6
Layout-specific learning	65.0	82.1	71.2
Leave-1-layout-out cross validation	75.9	76.2	75.1
Book-specific-learning	69.7	72.9	70.7
Leave-1-book-out cross validation	71.2	73.3	71.7



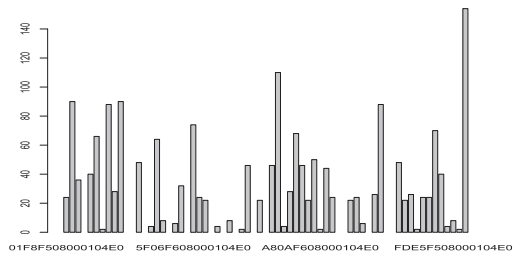
**Fig. 4** Histogram of errors by blocks.

6 and 8 has relatively lot errors compared to the histogram of Fig. 3. Figure 5 is the histogram of errors by book IDs.

4.4 Discussion

4.4.1 COG v.s. Naive Bayes Method

Comparing Table 3 and Table 4, the scores are not explicitly different. However, naive Bayes method has an advantage in the sense that it can recognize more accurate locations as shown with reasonable score in Table 6, while COG method has a shortcoming that it becomes undecidable when the location of gravity becomes the border of multiple blocks.



**Fig. 5** Histogram of errors by book IDs.

4.4.2 Effect of Book Layouts and Books

The score of naive Bayes method keeps relatively higher score even for both leave-1-layout/book cross validation. It implies that the learned model is still applicable to a new unknown layout or book.

This result provides great advantage in the real application, since library staff do not have to re-train naive Bayes model when a new book is introduced, or the layout is changed by a librarian or by a user.

On the other hand, the models by both layout/book-specific learning have worse scores. This can be considered because the specialization to a layout or a book can not contribute to recognition performance, and even has somewhat bad effect.

5. Related Work

As for detecting locations with RF, there are a lot of work for indoor location system of an RF device [4]–[10]. There

is also a work to use bayesian filtering for location detection [11]. However, most of these approaches measure the received signal strength (RSS), or using a special protocol for location detection to detect the location even within the range of a single antenna. These might be helpful to improve RFID bookshelf, but they need extra functionality which is not applicable to existing systems. Our first research goal is to qualify the effect of reading from multiple antennas, and different from that of indoor location systems.

Another approach to detect the book in the bookshelf is image processing. [12] does the detection of book boundaries located on a bookshelf. This approach helps to detect the location of books if each book can be identified, but this is still a complementary way, since image processing has a disadvantage that visual obstacles or dark light keeps the book from detection.

Over all, there are few work on challenging for accurate location detection in RFID bookshelves.

## 6. Conclusion

In this paper, we proposed a method to improve the location detection using naive Bayes classifier, and show the experimental result. The result showed the reasonable performance with the method, and, more importantly, the performance was shown to be less dependent of layouts and books, which is favorable for library operation. One of the future work is tackle with real time detection, since currently the detection result comes only after reading through all antennas.

## References

- [1] R.O. Duda and P.E. Hart, Pattern classification and scene analysis, Wiley, New York, 1973.
- [2] P. Domingos and M.J. Pazzani, "On the optimality of the simple bayesian classifier under zero-one loss," Mach. Learn., vol.29, no.2-3, pp.103-130, 1997.
- [3] S. Inoue, D. Miyaoka, D. Ikeda, T. Minami, and K. Fujisaki, "RFID application for library," Proc. Int'l Workshop on Real Field Identification, pp.w4-31-38, Nov. 2007.
- [4] P. Bahl and V.N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," Proc. INFOCOM 2000, 19th Annual Joint Conf. IEEE Computer and Communications Societies, vol.2, pp.775-784, March 2000.
- [5] J. Hightower and G. Borriello, "Location systems for ubiquitous computing," Computer, vol.34, no.8, pp.57-66, Aug. 2001.
- [6] K. Pahlavan, X. Li, and J.P. Makela, "Indoor geolocation science and technology," IEEE Commun. Mag., vol.40, no.2, pp.112-118, Feb. 2002.
- [7] M. Vossiek, L. Wiebking, P. Gulden, J. Wiegardt, C. Hoffmann, and P. Heide, "Wireless local positioning," IEEE Microwave Magazine, vol.4, no.4, pp.77-86, Dec. 2003.
- [8] S. Ray, D. Starobinski, A. Trachtenberg, and R. Ungrangsi, "Robust location detection with sensor networks," IEEE J. Sel. Areas Commun., vol.22, no.6, pp.1016-1025, Aug. 2004.
- [9] A. Kushki, K.N. Plataniotis, and A.N. Venetsanopoulos, "Kernel-Based Positioning in Wireless Local Area Networks," IEEE Trans. Mobile Computing, vol.6, no.6, pp.689-705, June 2007.
- [10] H. Laitinen, S. Juurakko, T. Lahti, R. Korhonen, and J. Lahteenmaki, "Experimental evaluation of location methods based on signal-strength measurements," IEEE Trans. Veh. Technol., vol.56, no.1, pp.287-296, Jan. 2007.
- [11] D. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello, "Bayesian filtering for location estimation," IEEE Pervasive Computing, vol.2, no.3, pp.24-33, 2003.
- [12] E. Taira, S. Uchida, and H. Sakoe, "Book boundary detection from bookshelf image based on model fitting," International Symposium on Information Science and Electrical Engineering, pp.601-604, Nov. 2003.



**Sozo Inoue** is an associate professor in Kyushu Institute of Technology, Japan. Inoue has a Ph.D. of Engineering from Kyushu University. He is a member of the IEEE, ACM, IPSJ, and DBSJ. His research interests include human activity recognition with smart phones, healthcare application of web/pervasive/ubiquitous systems. Contact him at sozo@acm.org; 1-1 Sensui-cho, Tobata-ku, Kitakyushu, 804-8550, Japan.



**Yasunobu Nohara** is a research assistant professor in Kyushu University Hospital, Japan. Nohara has a Ph.D. of Engineering from Kyushu University. His research interests include RFID systems and healthcare systems. He is a member of IPSJ and IEEE.



**Masaki Takemori** is a technical staff in Kyushu University Hospital. His research interests include RFID bookshelf systems, web information systems, and healthcare applications.



**Koza Sakuragawa** is a software engineer in Fusic Co. Ltd. His research interests include web information system development.