A Library Marketing System for Decision Making

Toshiro Minami

Abstract. The major aim of library marketing system is to help the library and its patrons with providing useful information and various kinds of knowledge, which are extracted from the data that are collected by the library system. In this paper we lay heavy stress on the use of such information and knowledge for assisting decision making of the library and of its patrons. Furthermore our main concern is to investigate how much we can utilize the usage data of materials obtained in the intelligent bookshelves (IBSs), i.e. the bookshelves equipped with RFID antennas and their reader/writer controllers (R/Ws). In this paper we propose some analysis methods of these usage data, alone and combining with other library data, and demonstrate their potential importance for library marketing.

Keywords: Library Marketing, Data Analysis, Data Mining, RFID (Radio Frequency Identification), Intelligent Bookshelf.

1 Introduction

The most important mission of libraries is to provide services to their users, or patrons, so that they are very helpful and meet the patrons' needs in our knowledge society. Patron orientedness is well expressed in "The five laws of library science" advocated by the famous Indian library scientist S. R. Ranganathan [13], which was published way back in 1960s. They are: (1) Books are for use, (2) Every reader his book, (3) Every book his reader, (4) Save the time of the reader, (5) The library is a growing organism. Our surrounding society has changed vigorously in these several decades. However, these phrases are still true as we rephrase them by replacing "book" with "information," "material," or "service." They are still the central dogma for libraries and will keep being the same in the

Toshiro Minami
Kyushu Institute of Information Sciences, Faculty of Management and Information Sciences, 6-3-1 Saifu, Dazaifu, Fukuoka 818-0117 Japan

Kyushu University Library, Research and Development Division, 6-10-1 Hakozaki, Higashi, Fukuoka 812-8581 Japan
e-mail: minami@kiis.ac.jp, minami@lib.kyushu-u.ac.jp
http://www.kiis.ac.jp/~minami/ (in Japanese)

K. Nakamatsu et al. (Eds.): New Advan. in Intel. Decision Techno., SCI 199, pp. 97–106.
springerlink.com © Springer-Verlag Berlin Heidelberg 2009

future. In order to fulfill the patrons' needs, libraries have been trying to change themselves and to provide better services; based on their intuition.

In a decade or two we have experienced a rapid advance in technology; especially in information and communication technology (ICT). Internet should be the most prominent ICT technology that gives big influence to our society and our everyday life. Marketing methods by collecting and analysing data from the Internet services, mostly with Web-based mechanisms, are very popular now.

Now is a high time for libraries to introduce such methodology to their services. By collecting and analysing the data created in the library services, libraries can get useful tips for them. Such technology and methodologies would change the ways the libraries have been taking so that many decisions made in library jobs are done based not on the subjective intuition but on the objective data. Not only for libraries but also for their patrons such methodologies give much benefits by getting new services from their libraries.

In this paper we describe the concept of library marketing and library marketing system, and we also demonstrate its usefulness by taking some examples supporting functions for decision making by the libraries and their patrons.

The rest of this paper is organized as follows: First of all in Section 2, we explain the concept of library marketing. We note that it has a self-growing nature so that it will solve the cold start, or the free-riding, problem. Then we show a model of library marketing system with RFID (Radio Frequency Identification) technology [5]. The library materials can be classified into two types; physical media and net media. The former ones are printed books, magazines, casette tapes, CDs, DVDs, and so on, which are the major materials used in library services so far, and will be so in the near future. The latter one is the materials used in network services such as e-journals, e-books, and so on. The log data, or usage data, are collected as the libraries provide services. RFID is useful for collecting log data of physical media. For net media, the service log data are collected automatically. We show an experimental library marketing system with RFID and explain how to extract the data that are used in data analysis, or data mining, which is the most important and essential part for library marketing system. In Section 3, we would like to demonstrate the importance and usefulness of the library marketing system by showing examples of information or knowledge that are extracted by analysing the usage data. Lastly in Section 4, we conclude our discussions and present some of our future plans.

2 Library Marketing System

This section starts with describing the concept of library marketing [7, 8]. We emphasize the utilization of the usage data of library materials. Intelligent bookshelf, or IBS in short, is a key tool for library marketing for physical media. RFID is an AIDC (Automatic Identification and Data Capture) technology [1] so that the data collection is performed automatically and therefore saves the librarians' time.

Then we describe a model of library marketing system, which we use in an experiment. The original log data is converted to the event data, which is more suitable for analysis. We define some terms such as a session, session (duration) time, which are extracted from the event data.

2.1 Library Marketing and Library Marketing System

Figure 1 illustrates the concept of the library marketing. It is the mutual depending structure between databases and services. In the right part of the figure are library services provided by the library.

OPAC (Online Public Access Catalog) is a search engine for library materials. Reference service is a consulting service to the patrons who want to know how to find appropriate material, or need some advices. These days many libraries provide this service online, which is often called "Ask a Librarian" service. "My Library" is a relatively new service that provides patrons with personalized information such as which books they are borrowing and have been borrowed, a list of books the patrons might like to read, and so on.

These services depend on the databases that consist of catalog data, circulation data, patrons' profile data, etc. These data are collected at a circulation counter, at a self checkout machine, and in other ways. The log data, which are collected for recording how the system provides services to patrons, are stored in the service log database and used for various services.

The collected data are supposed to be analyzed, or data mined. The results obtained by analysis should be reflected to the services in order to improve the quality of patron services. If the quality of services gets higher, then more patrons will use the library and the system. If many patrons use them, then the system gets more service log data. If the system is able to get more data, then they can be used for better services. Then the customer satisfaction (or rather patron satisfaction) will rise. Such a natural self-growing mechanism is very important for library marketing with databases and services as has been pointed out.

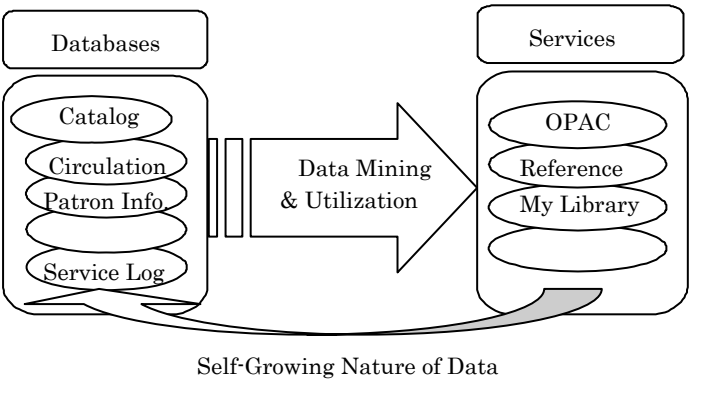
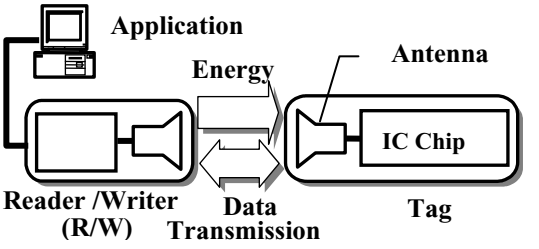


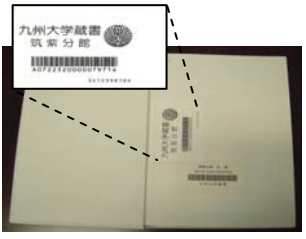
Fig. 1 Library Marketing with Databases and Services

2.2 Library Marketing System with RFID

For the services using the net media, it is easy to collect not only provinding materials for the patrons but also to collect log data. Everything can be done



(a)



(b)

Fig. 2 (a) Principle of RFID (b) RFID Tag Attachment on a Book

automatically. For physical media such as printed books, magazines, audio visual materials like CDs and DVDs, it is a difficult task to do the job and collect the usage data, bacause these jobs are supposed to be done by the library staff.

AIDC technology [1] is a great help in such jobs. So far the barcode system has been used in libraries. RFID is getting to be more and more popularly used by libraries in these years because it is faster and easier to be used for doing library jobs [6, 11, 15]. Furthermore there are more advantages in RFID such as the circulation processing becomes faster and easier, the patrons are happy to use self-checkout machines and do the processing themselves, inventory of library materials becomes faster. Considering these benefits it is worth putting the costs on equipments and RFID tags.

The principles of RFID tag system is illustrated in Figure 2(a). The RFID tag system consists of two major components; tags and reader/writers (R/Ws). A tag is able to communicate with an R/W when they are located sufficiently close to each other. Figure 2(b) shows a tag put on a book.

From the library marketing standpoint intelligent bookshelf (IBS) (Figure 3) [3] is a very useful tool for collecting data. An IBS is a bookshelf which is equipped with RFID antennas so that the book IDs shelved in it are detected and recorded

Fig. 3 Intelligent Bookshelf: A Tool for Library Marketing

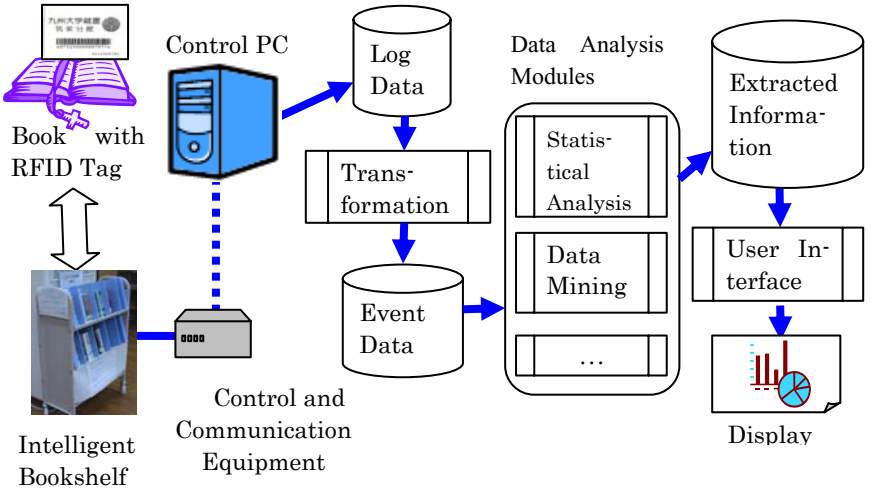


Fig. 4 An Experimental Library Marketing System with RFID

automatically. The antennas and their controller, i.e. R/W in all, transmit the data to the controller PC and the data will be analyzed either real time or later on.

Figure 4 illustrates a model for library marketing system with RFID, which is developed for a feasibility study of IBS systems [9, 10, 16]. Original log data obtained from the IBS are transmitted to the controller PC and are saved as log data. The log data consist of the fields of data type, timestamp (yymmdd and hhmmss), shelf ID, the number of detected tag IDs, followed with the list of tag IDs. A tag ID looks like "E00401000314E148" and "E004010003150040."

The log data are redundant because the status will not change for a long time and thus the detected tag IDs are the same ones in a long line of data. So it is more convenient to transform this log data format into the format that is easy to deal with. We call the transformed data the event data. The transformation program converts the original log data into the event data.

The event data can be represented as a list of events. They consist of the timestamp data, shelf ID, i.e. tier ID and shelf ID, one column of tag ID, and the type of event, or the change of status, which is either IN or OUT. The status IN means that the book is put on the denoted shelf and OUT means it is taken out from the shelf.

Data analysis modules use the event data as input and extract information and knowledge. They analyse the data in statistical and/or data mining methods.

Lastly the extracted information is presented to the librarians and users according to the situation.

2.3 Data in Library Marketing System

The original log data from an R/W are the list of a line in the following format:

O = *(Time, ShelfID, Number of Tags, List of TagIDs)* where *Time* is a timestamp, *ShelfID* consists of tier ID and the shelf ID in the tier.

From these data we get event data set, which is a list of event data, where *i*-th event data collected in this IBS are represented in the following format:

d_i = *(InOut, ShelfID, TagID, Time)* where *InOut* specifies if the book is stored in a bookshelf or removed from a bookshelf, *ShelfID* and *TagID* specify which shelf the book is stored to or removed from and which book is the one, respectively. The *Time* is the timestamp which specifies date and time of this status change.

By specifying a book, or its corresponding tag, with its ID, we can define a session of usage of the book. A session starts when the book is taken out from a bookshelf and ends when it is returned to a bookshelf, which may be the same one or a different one.

Let TID be a book ID. The event data of the book is defined:

Event(TID) = {(InOut, ShelfID, Time) | d_i=(InOut,ShelfID, TID, Time) for some InOut, ShelfID, and Time where d_i is an event data }

We can define a session of the book with TID as follows:

Session(TID) = {(OutTime, InTime, OutShelf, InShelf) | there exist e1=(InOut1, OutShelf, OutTime) and e2=(InOut2, InShelf, InTime) in Event(ID) such that InOut1="OUT", InOut2="IN", OutTime<InTime, and there is no data e3=(InOut3, ShelfID3, Time3) in Event(TID) that satisfies OutTime<Time3<InTime }

The duration time of a session is defined as follows:

Duration((OutTime, InTime, OutShelf, InShelf)) = InTime – OutTime

If *OutShelf = InShelf*, it means that a patron has put the book back in the same shelf and if not, he or she has put it in a different shelf, probably erroneously.

Suppose we set a time period *p=(StartTime, EndTime)*, where the start and end times may be in a same day or in different days, weeks, or months. We would define that a session *s=(OutTime, InTime, OutShelf, InShelf)* of a book is in this time period as follows: *s is in p iff StartTime<=OutTime<InTime<=EndTime* where <= means "less than or equals to."

In this definition we choose to say that the session occurs in the specified period only when it starts and ends in the period. Note that we can define this concept in some different ways.

3 Library Marketing System as Decision Support System

As was pointed out, the most important application of library marketing system is to help the librarians and their patrons with decision making on collection of books, shelving policy, shelf arrangement, planning learning process, and others. In this section we investigate several possible applications that should be very useful for the library, for librarians, and for library patrons.

3.1 Decisions of Learning Materials for Patrons

First of all we investigate how to help patrons with deciding which material to read, or study.

Statistical Method

Patrons will be happy if they are able to know what materials are popularly used and checked-out. Some of the typical data in this direction, for example, are the rankings of the books which are used in the library, which are borrowed from the library, which are popular among the patrons in general, and so on.

From the original usage data obtained from the IBSs, we can extract various statistical data on each of the book shelved on the IBSs, such as frequency of sessions, average session time, calculated time of sessions, etc.

Furthermore, because the library books have their catalog data, each book has its classification data; in DDC (Dewey Decimal Classification) or in its modified way in most libraries. By summing up the frequencies of the books belonging to a classification number, we can get the frequency data of each classification area. Not only the frequencies but other number data can also be processed in the same way.

In such a way, we can obtain many kinds of number data that indicate how the books are popularly used in the library. By applying this method to the circulation data, where one session starts when a book is borrowed and ends when it is returned to the library, we can obtain similar data about the borrowed materials.

In addition to them we can combine these data and other data from other libraries and from the Internet, we can provide the library patrons with additional information. One possible candidate for another data might be the one obtained from the blogs, the articles provided by the net people. For example we can obtain the popularity data of a book by using a blog search engine and get the number of the articles that mention the book as their subjects.

Even though the statistical methods are simple in one point of view, they are very valuable tools in order to extract useful and valuable data and information to the library patrons. We have to keep investigating further.

Collaborative Filtering

Collaborative filtering (CF), or social filtering, is now well known as the amazon.com's "Customers Who Bought This Item Also Bought"-type recommendation function [2]. This function has applied to other fields as well [12]. In library marketing we can use it in the format "who borrowed this book also borrowed," "who belong to your same department also borrowed," and furthe more "who come in this time zone of the day prefer to use," etc.

3.2 Decisions on Book Selection for Libraries

The collaborative filtering method can be applied to the selection of library materials as well. In this section we rather take another method that might be more appropriate in order to demonstrate the usefulness of the usage data obtained from the IBSs. It is a method for evaluating the appropriateness of the books for purchasing by combining the usage data from IBSs and the circulation data from the counters or from the self-checkout machines.

Table 1 Comparing the Frequencies of Usage data and Borrow Data

Borrow \ Usage	High	Middle	Low
High	HH		HL
Middle			
Low	LH		LL

In Table 1, suppose the books ranked in High are those in highest 20% of usage frequency and circulation frequency in a specified time interval, respectively and those ranked in Low are lowest 20%.

Then the books labeled HH are popular ones in terms of both usage and borrow. They are the first candidates to purchase. On the other hand the library would not purchase the books with LL if they have not other reasons to purchase.

The ones labeled with LH, i.e. low usage and high borrow, might be the books that patrons have already decided to borrow when they visit the library, or those they have found them already as they used the OPAC (Online Public Access Catalog) system and have decided to borrow. These are good candidates for purchase.

The ones labeled with HL, i.e. high usage and low borrow, are difficult to decide immediately. The books that are prohibited to borrow may be set in this category; or HN, high usage and no borrow. Other books that are substantially used as reference materials may be in this category.

Also some books in this category might be something that look like interesting but actually they are not so much interesting nor useful as the patrons read some part of the book. If we check the average session time of the books in this category we may know better about such books. So we need to have more research on this.

3.3 Decisions of Book/Shelf Arrangement for Libraries

This is another interesting application of IBS. The system analyzes the usage data and extracts the correlation of two books in terms of usage time and others. Firstly we define the event time of an OUT event data *(OUT, ShelfID, Time)* by the timestamp *Time.*

Suppose TID1 and TID2 are different book IDs. We define the succeeding event time of TID2 of TID1 by:

SuccEventTime(TID1, TID2) = Time2 – Time1 where there exist event data e1=(OUT, ShelfID1, Time1) in Event(TID1) and e1=(OUT, ShelfID2, Time2) in Event(TID2) so that there does not exist any data e3=(OUT, ShelfID2, Time3) in Event(TID2) such that Time1 < Time3 < Time2.

Let us define the average time of *SuccEventTime(TID1, TID2)* in a specified time interval by the average succeeding event time from TID1 to TID2. Then we define the correlation index of two books TID1 and TID2 by adding the average

succeeding event times from TID1 to TID2 and the time TID2 to TID1. The correlation index is a measure for closeness of two books in terms of usage pattern.

Using the correlation indexes we can make a special corner of books that are related closely in this sense so that the patrons can access them in one place.

4 Concluding Remarks

In this paper we tried to convince the importance of the concept of marketing for libraries because library marketing methodologies have essential importance for the future libraries. The libraries should have good reputation and they succeed in surviving in the competitive society by utilizing them.

The methodology of utilizing the data is the same one with OLAP (Online Analytical Processing). However we put more emphasis on the point that the information and knowledge are useful not only for the libraries but also for their patrons in library marketing. Our approach is different to that of OLAP in this point.

We have presented a model of library marketing system with intelligent bookshelves, which utilizes the advantage of RFID technology. With IBS we can automatically collect the usage data of physical materials in a library. The only usage data libraries can get so far are circulation data. So the librarians do not have data for materials that are not allowed to borrow, e.g. reference tools. Now it is possible to obtain those data with IBSs. Also we presented several examples of data that might be useful for librarians and patrons. Such services could not be realized without library marketing tools. So far almost all libraries with RFID tag system use HF (13.56MHz). Chiyoda Library in Tokyo has decided to introduce an UHF (950MHz) tag system [4, 14] because it has better potential performance.

Even with such high potential of the library marketing system, we are in the first stage towards the goal. Some of our future plan includes (1) developing a library marketing system for experimenting more sophisticated functions, (2) developing a lot more analysis methods so that they can be applied to a wide variety of libraries, and (3) evaluate the usefulness of the concept of library marketing.

Our eventual goal is to apply such library marketing system to many libraries so that the system contributes to the people as they learn and do research works.

Last of all the author would like to express his gratitude to his daughter, Mariko Minami for her valuable comments for the draft of this paper.

References

1. AIM Global, http://www.aimglobal.org/
2. Amazon.com, http://www.amazon.com/
3. Central Engineering Co. Ltd. (in Japanese), http://www.central-eng.co.jp/
4. Chiyoda Library (in Japanese), http://www.library.chiyoda.tokyo.jp/
5. Finkenzeller, K.: RFID Handbook, 2nd edn. John Wiley & Sons, Chichester (2003)
6. Kyushu University Library, http://www.lib.kyushu-u.ac.jp/

7. Minami, T.: RFID Tag based Library Marketing for Improving Patron Services. In: Hoffmann, A., Kang, B.-h., Richards, D., Tsumoto, S. (eds.) PKAW 2006. LNCS (LNAI), vol. 4303, pp. 51–63. Springer, Heidelberg (2006)
8. Minami, T.: On-the-site Library Marketing for Patron Oriented Services. Bulletin of Kyushu Institute of Information Sciences 8(1), 15–33 (2006) (in Japanese)
9. Minami, T.: A Perspective to the Library in Network-Oriented Society – Ubiquitous Library Services through PDAs. Bulletin of Kyushu Institute of Information Sciences 10(1), 1–17 (2008) (in Japanese)
10. Minami, T.: A Design for Library Marketing System and its Possible Applications. In: Proc. Pacific Rim Knolwdge Acquisition Workshop 2008 (PKAW 2008), pp. 187–200 (2008)
11. National Library Board Singapore, http://www.nlb.gov.sg/
12. Oda, M., Minami, T.: From Information Search towards Knowledge and Skill Acquisition with SASS. In: Proc. 6th Pacific Rim Knowledge Acquisition Workshop (PKAW 2000), pp. 245–260 (2000)
13. Ranganathan, S.R.: The Five Laws of Library Science. Bombay Asia Publishing House (1963)
14. SOFEL Co. Ltd., http://www.sofel.co.jp/en/index.html
15. University of Nevada, Las Vegas Libraries, http://www.library.unlv.edu/
16. Zhang, L., Minami, T.: Library Marketing that Boosts Education with RFID. In: Proc. 6th International Conference on Information Technology Based Higher Education and Training (ITHET 2007) (2007)