Data Analysis Methods for Library Marketing in Order to Provide Advanced Patron Services

Toshiro Minami[1,2] and Eunja Kim[3]
[1] *Kyushu Institute of Information Sciences, 6-3-1 Saifu, Dazaifu, Fukuoka 818-0117 Japan*
[2] *Kyushu University Library, 6-10-1 Hakozaki, Higashi, Fukuoka 812-8581 Japan*
[3] *Gwacheon Public Library of Information & Science, Korea*
minami@lib.kyushu-u.ac.jp, gclib55@gmail.com

Abstract

Our society is rapidly changing into information society, where the needs and requests of the people on information access vary vigorously depending who they are. Library's mission has been and still is to provide its users, or patrons, with the most appropriate information. Libraries have to know the profiles of their patrons, in order to achieve such a role. The aim of library marketing is to develop methods based on the library data, such as circulation records, book catalogs, book-usage data, and so on. In this paper we discuss the methodology and importance of library marketing at the beginning. Then we demonstrate its usefulness through some examples of analysis methods applied to the circulation records in Kyushu University and Gwacheon Library, and some implication that obtained as the results of these methods. Our research is a big beginning toward the future when library marketing is an unavoidable tool.

Keywords: *Library Marketing, Data Analysis, Circulation Data, Usage of Areas/Zones, Intelligent Bookshelf (IBS), Usage Data Analysis*

1. Introduction

According to the American Marketing Association (AMA) [1], the concept of marketing used to be defined as: "Marketing is an organizational function and a set of processes for creating, communicating, and delivering value to customers and for managing customer relationships in ways that benefit the organization and its stakeholders." They now define it as follows: "Marketing is the activity, set of institutions, and processes for creating, communicating, delivering, and exchanging offerings that have value for customers, clients, partners, and society at large."

By comparing these two definitions, we recognize that marketing was considered as the activities that benefit the organization (company); which matches with the ordinary people's intuition. It is now considered as wider activities that benefit the customers and our society as well. So it is natural to apply marketing to non-profit organizations like libraries including public and university libraries.

In this point of view, the aim of marketing activities by libraries (library marketing) is to give better services to their users, or patrons, so that they are able to get better reputations, to be recognized as more reliable organizations, and to get more customer satisfaction (CS), or patron satisfaction (PS) eventually. In addition to this aim it is preferable to perform their jobs more efficiently, and with less cost; which can be another important aim of library marketing.

In this paper we focus on the library marketing methods based on those of analyzing the objective data and extracting useful information and knowledge not only for libraries but also for their patrons.

Libraries have many kinds of data including circulation records (borrowing or returning of books and other materials), catalog information, patrons' entrance data, book reservation records and so on. Some libraries also have patrons' exiting time data, reservation data for study rooms, PCs' session records, etc. However most of these data are not used sufficiently so far. It is really a big waste of potentially very valuable data. We carry out our research on library marketing by dividing the process into four levels.

(1) Preliminary Investigation

In this level we investigate what information, tips, and knowledge could be obtained by analyzing some kinds of data. We do not worry much about if we can really get such data or the extracted information is very useful or not. Our aim in this level is to create as many possible ideas as we can imagine which could be and/or may be used for library marketing.

(2) Real Data Analysis

In this level we apply the methods obtained in the preliminary investigation level. By using the real data, we can evaluate the analysis methods from the practical point of view. If we find out that an analysis method is very useful, then we apply this method to another data. It could happen that we can apply a method to other types of data by modifying it, slightly or largely.

Most of analysis methods presented in this paper can be considered to be those in this level. We will continue our research on this level and try hard to find as many practically useful methods as possible.

(3) Combination of Methods

Even though one type of data can be useful enough for library marketing, we would be able to extract even more useful information by combining the extracted information/knowledge and combining more than one types of data. We will investigate this type of analysis methods after we investigate the level (2) sufficiently.

(4) Development of the Automated Methods

As we have found a very useful analysis method, it should be convenient to apply it by automating the analysis method. This method is a kind of macro procedure so that it is a pack of analysis methods and thus can be considered as one method. As a result, this analysis is easy to use as well as it can be used as a part of more sophisticated automated methods.

In this paper we will demonstrate the importance of library marketing through presenting some example analysis methods for such data as the circulation data, usage data of library materials, those data about areas in library, and so on.

The rest of this paper is organized as follows: In Section 2, we show some example analysis methods and results from a circulation record for the year 2007 of Kyushu University Library in order to demonstrate its potential usefulness. Even though we are in the very beginning level some examples inspires the usefulness of this approach toward library marketing. In Section 3, we show some other analysis cases that should be very useful in library marketing. Our eventual goal is to develop analysis methods by combining different

types of methods and data. Finally in Section 4, we conclude our discussions in this paper and investigate our future directions.

2. Case study: data analysis for circulation data of Kyushu University Library

In this section we start with analyzing the circulation data of Kyushu University Library (KUL) [7] in Japan, and demonstrate how useful these data could be used for library marketing. Firstly we introduce the target data, and then show some example results. Next we choose one student and analyze the student's behavior.

2.1. Data and analysis methods

We use the circulation data of the Central Library of KUL for the school year 2007; from April 1, 2007 to March 31, 2008. See also [8] for other statistical data in KUL. One record item consists of, roughly speaking, the book ID, book profile including classification and call number, borrower ID, the borrower's profile including affiliation, type (professor, student, staff, etc.), and the timestamps, i.e. date (yyyy:mm:dd) and time (hh:mm:ss), for borrowing and returning. The total number of circulated items is about 67 thousands.

2.2. Analysis results from all records

Figure 1 illustrates the ratios of borrowed materials according to the patron type. About half of the borrowed materials are borrowed by undergraduate students and about 40% are by graduate students. Thus almost 90% are borrowed by students. This result matches to the ratio of visiting patrons reported in [5], in which 53% of about 327 thousand visitors are undergraduate students and 28% are graduate students. Thus 81% of visitors are students in April 2009. This ratio does not vary a lot. In the fiscal year 2007 81% of visitors to the central

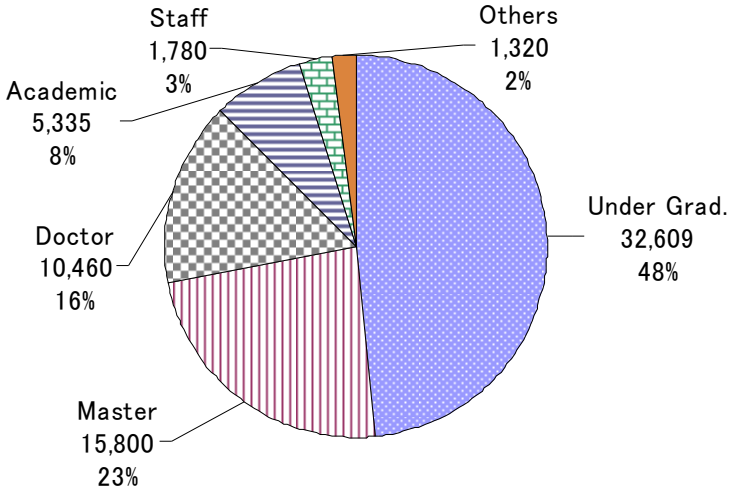

Figure 1. The Ratios of Borrowed Books According to Patron Type

library are the students, while 87% is the student ratio for the visitors to all libraries in KUL that consists of the central library and 6 branch libraries.

In the data for the fiscal year 2008, the ratio for the central library was 83% and that for all the KUL libraries was 88%. As a conclusion the ratio of student visitors among all the visitors is roughly from 80% to 90%. Another conclusion is this ratio is bigger for all libraries than that for the central library only. Furthermore the ratios increase by comparing the data in two years. So we have to say that the importance of library services to students is also increasing. We have to put more efforts on student services for university libraries.

If we assume this ratio is also applicable to the data in 2007, we can say that the undergraduate students who are 53% of visitors borrow 48%, while the graduate students who are 28% of visitors borrow 39% of books. From these data we have values by dividing the borrow ratio by the visitor ratio, which are the index for how many books one person borrows. The values are about 0.91 for undergraduate students and about 1.39 for graduate students. The average value is exactly 1 because 100% of visitors borrow 100% of books. By comparing these ratios, we can say that the graduate students borrow 1.4 more books in average than undergraduate students. As a conclusion we can say roughly that graduate students study harder than undergraduate students if we assume that the borrowing of library books is the index for aggressiveness for studying.

Figure 2 indicates how many books are borrowed and returned as time goes in a day. The horizontal axis represents the time of the day and the vertical axis represents the number of books borrowed in an hour. The peak time for returning is about 16:00 and 17:00 for borrowing. More precisely the time interval the books are mostly returned are about from 12:00 to 17:00. The time interval when the books are mostly borrowed is about from 12:00 to 22:00.

Figure 3 indicates the ratios of borrowed books according to day of week. The result is very interesting. From Monday to Friday the ratio is from 16% to 19% and roughly

Figure 2. Books Borrowed and Returned in Time of a Day

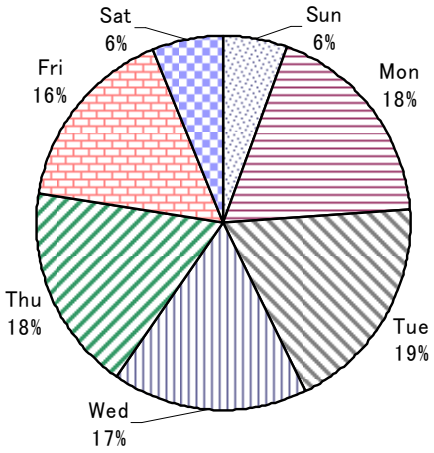

Figure 3. Rates of Books Borrowed According to Day of Week

speaking almost the same. On the other hand Saturday and Sunday, both ratios are 6%. These results might come by reflecting the number of visitors. We might be able to guess from these results, that the students that visit the library come regularly; not visit some specific day or days of week.

Figure 4 indicates the number of books according to month from April 2007 to March 2008. This result is easy to understand that from April to July the number of borrowed books is increasing as the students get used to the university as well s to the study courses that start in April. The numbers become much smaller in August and September because these months are for summer holidays. The number recovers in October as the new semester starts. Then it increases suddenly in January, probably because senior students, i.e. in the 4th grade students, are very busy in writing their theses in this month because of the due dates for their graduation and term papers. Finally the number decreases in March because of end of the year holidays.

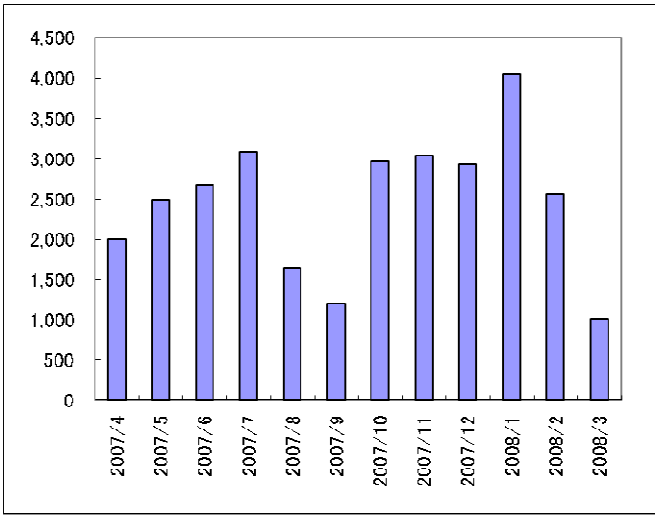

Figure 4. Number of Books Borrowed in Month

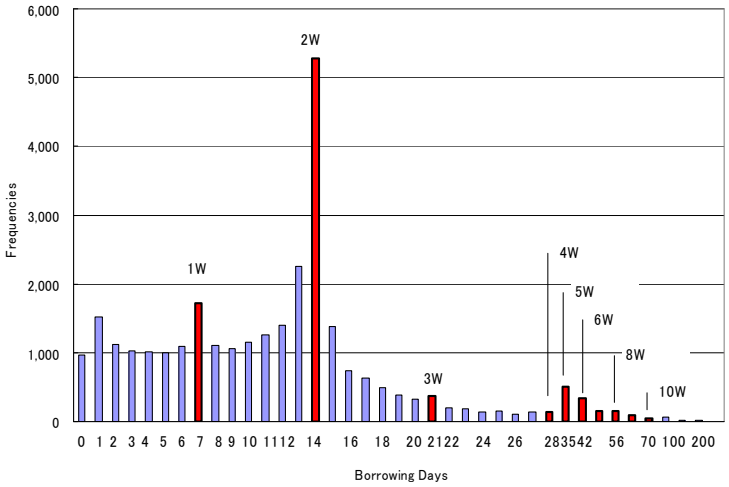

Figure 5. Number of Books for Renting Days

Figure 5 illustrates the frequencies of books according to the renting days. The peak value comes on 14 days (2 weeks), which are the due period of days for borrowing. The average is 12.2 days. Roughly speaking about 1,000 books are returned on the days that are less then 14 days; which might mean the students who borrow books from the library visit regularly. It is interesting to see that the number is also about 1,000 for 0 day, which means that the books are returned on the day they are borrowed.

To see more precisely, the number is much bigger than 1000 in the days 1, 7 and 13. Borrowing 1 day means that the books are returned on the next day. So many books are borrowed just for a very short period. Next peak day is 7, where the books are borrowed for one week, which is reasonable. The number increases as the day passes from 9 to 14. Many students seem to care about the return date and try to return as early as possible.

It is a surprise to see that quite a lot of books are borrowed beyond two weeks. The maximum value is 238 days. The book's title is "Introduction to German Law" and the student belongs to Engineering. The student borrowed such a long time probably because he or she did not want to return or just forgot to return.

2.3. Analysis of a student's data

In this section we choose a student as a sample and investigate more about the behavior of the student in terms of borrowing of books. We choose the student (here after we call the student A) who borrowed 208 books, which is the maximum number among all the students; which probably means that the student A is very aggressive in studying.

Figure 6 shows the ratios of the classification number of books borrowed by the student. KUL takes NDC (Nippon Decimal Classification) system for books classification. In NDC, 007 is for information science, 410 for mathematics, 420 for physics, 430 for chemistry, 440 for astronomy space sciences, and 910 for Japanese

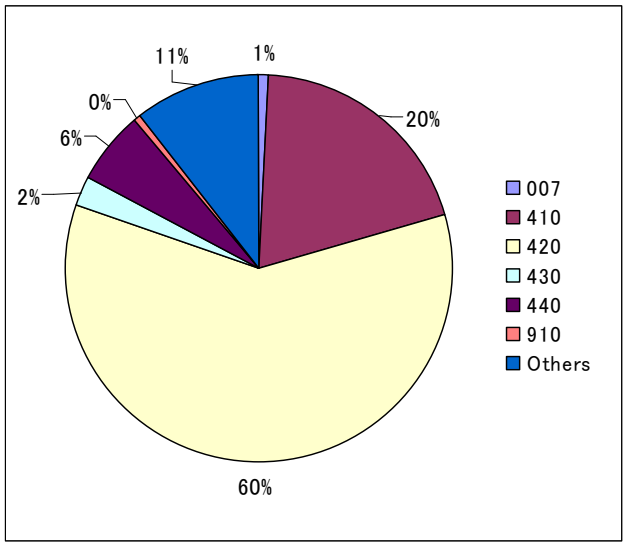

Figure 6. Classifications of the books borrowed by student A

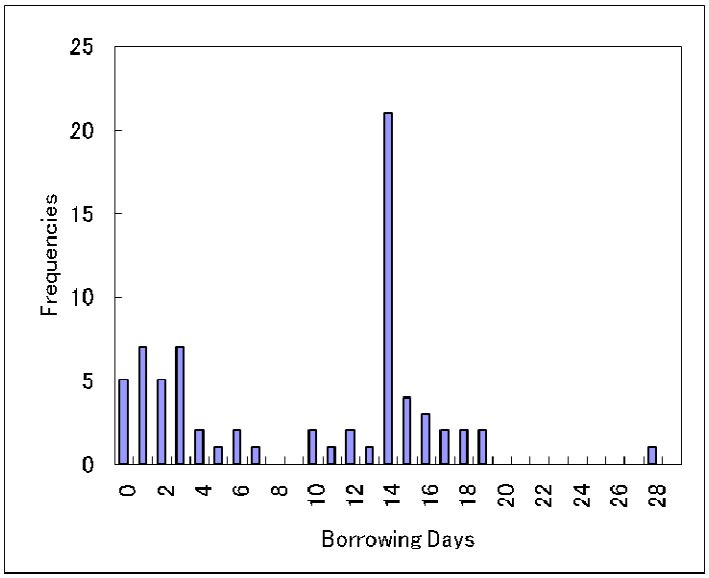

Figure 7. Number of Books for Renting Days (The Book Borrowed Most Frequently)

literature. The student A borrowed the books classified in 420 (physics) at the ratio of 60%, followed by 410 (mathematics) of 20%.

Figure 7 illustrates the distribution of numbers of books according to the day of borrowing of student A. As with Figure 5, student A also returned the books most of all on the 14th day. On the other hand, different from Figure 5, student A only borrowed the books only from 0 to 3 days much more than from 4 to 13 days. He or she also borrowed a number of books more than 14 days. The maximum day is 28.

3. Other library marketing cases

3.1. Zone usage data analysis

The second author and other librarians of Gwacheon Public Library of Information & Science (hereafter Gwacheon Library, GL) [4] in Korea carried out a series of experiments by putting bar-code readers at the entrance of some rooms in the library and collected the data for analyzing how their rooms are used by the visiting patrons [6]. It was an excellent research example for library marketing. In this section we show how data were collected and analyzed and what are their implications are induced from these experiments.

Gwacheon City [3] is located in the south neighbor of Seoul City and its population is about 60 thousands. GL is one of two public libraries of the city and contains about 220 thousand materials. Reading rooms and other patron areas are distributed into from the 4th floor to the basement of the library building. Document Information Room I (DIR-I), Family Reading Room (FRR), Language Room and other rooms are located in the 4th floor. Actually FRR is the room for members who had had registered as member of the library, who were supposed to live in Gwacheon City or who are working or studying in the city. Document Information Room II (DIR-II), Electronic Data Room (EDR) and others are located in the 3rd floor. Children's Room (CR) and Science Exhibition Hall are located in the 1st floor, i.e. the ground floor. Science Rooms are also located in the 2nd floor and the basement.

Generally speaking, in Korean libraries including GL, reading rooms and study rooms are relatively small and are distributively located in the library building. So that the patrons are supposed to move from one room/zone/area to another one according to what he or she wants to read or do, in comparison with other libraries in other countries like Japan, where a typical public library has a big spacey room divided into small zones. The librarians of GL had experimenting from November 15th through December 25th 2005 by putting bar-code readers at the entrance door of several rooms such as DIR-I and II, FRR, CR. They also put a reader at the exit gate from the library building. With these readers they collect the data about how the patrons used the rooms of GL. In order to have the data as accurate as possible, they even arranged the librarians so that they took turns and waiting at the readers and asked the patrons to scan their ID cards when they entered and exited the rooms and the library.

From the experiment, they found a lot of results. One example is that among 6,633 patrons using the DIR-I and II, CR, and FRR, 26% of patrons used DIR-I, 30% for DIR-II, 35% for CR, and 9% for FRR. Thus nearly 60% of patrons used DIR-I in the 4th floor and DIR-II in the 3rd floor as total. So, one possible idea for them is to combine the two rooms into one so that they can provide better convenience to the patrons.

The usage data are automatically collect-able for EDR. From the data, they found that about 90% used for accessing the Internet. The ratio was increasing in comparison with the usage of document editing without using the Internet. So, one possible idea for more convenience is to relocate the EDR from the 3rd floor to the 1st floor, or the entrance floor of the library.

3.2. Circulation data analysis in Ehime University Japan

Yamada [12] analyzed the circulation data of a university library. He compared the relationship between the years after publication of books and the ratios of the number of borrowed books and found that (1) the newly published books are well preferred to be borrowed and (2) still a number of books, or evergreen books, are constantly borrowed.

He concluded that it is very important for libraries to do their best so that newly published books are provided for service as soon as possible. He also concluded that by collecting and comparing the evergreen books of a lot of libraries, librarians may be able to know what books are supported by students of the libraries and thus what books are the must for them.

3.3. Usage data analysis with bookshelf equipped with RFID reader

RFID (Radio Frequency Identification) technology [2] is considered to be the one essential for ubiquitous society. Already a lot of libraries have introduced RFID based IC tag system for material management. IC tag labels are attached on books so that book IDs can be detected much faster and easier than the way with the current system with bar-code. Typically, the readers that detect the IC tags are set at the security gates for anti-theft function, at the circulation counter for borrowing and returning processing of books, self-checkout systems, and at the portable terminals for inventory.

An intelligent bookshelf (IBS) is a bookshelf which is equipped with RFID readers so that it can detect what books are shelved on which shelf in real time. IBS is an important tool for collecting the usage data of books inside of libraries [10, 11]. Research on analysis methods for such usage data is not much popular so far. However, by considering the essential importance of this approach to library marketing in the network society, we have to investigate more on developing methods for utilizing usage data.

4. Concluding remarks

The aim of this paper is to demonstrate the importance of the concept of library marketing through some examples. Firstly we investigate the definition of library marketing. As case studies for library marketing, we chose the circulation records of the central library of Kyushu University. From the point of view of library marketing, the ordinary statistical results such as total number of borrowed books, average number of borrowed books per patron, the total number of the patrons who borrowed at least one book, etc. [8, 9], are not useful enough. In this paper we tried to find other analysis methods such as the pattern of borrowing and returning time zones, patterns of borrowing days, comparison of borrowing days of week, and others.

The results presented in this paper are just a beginning in our approach to library marketing. We need to keep investigating more in this direction. One possibility is to analyze the usage data from IBS. Another one is to combine different types of data and extracts more useful knowledge for improving patron services by libraries.

References

[1] American Marketing Association (AMA): http://www.marketingpower.com/

[2] K. Finkenzeller, RFID Handbook (Second Edition), John Wiley & Sons, 2003.

[3] Gwacheon City: http://www.gccity.go.kr/foreign/english/

[4] Gwacheon City Library: http://www.gclib.net/

[5] E. Kim and T. Minami, "Library's Space Organization Improvement Based on Patrons' Behavioral Research -- An Attempt for Library Marketing at Gwacheon Public Library in Korea and Kyushu University Library—", Research and Development Division Annual Report 2008/2009, Kyushu University Library, 2009. (in Japanese)

[6] E. Kim, A Study on the Space Organization by the User's Behavior in Public Library, Gyeonggi University, 2008. (in Korean)

[7] Kyushu University Library: http://www.lib.kyushu-u.ac.jp

[8] Kyushu University Library: Annual Report 2007/2008, 2008.

[9] Kyushu University Library: Annual Report 2008/2009, 2009.

[10] T. Minami, "A Library Marketing System for Decision Making", Proc. First KES Symposium on Intelligent Decision Technologies (KES-IDT'09), 2009.

[11] T. Minami, "A Design for Library Marketing System and its Possible Applications", Proc. Pacific Rim Knowledge Acquisition Workshop 2008 (PKAW 2008), 2008, pp.187-200. Also in Lecture Notes in Artificial Intelligence (LNAI) 5465, Springer, 2009, pp.183-197.

[12] S. Yamada, "Analysis of Library Book Circulation Data: Turnover of Open-shelf Books", Journal of College and University Libraries 69, 2003, pp.27-33. (in Japanese)

Authors

Dr. Toshiro Minami earned his Ph.D. at Kyushu University, Japan in 1999. Currently he is a professor of Kyushu Institute of Information Sciences. He is also a research fellow of Kyushu University Library. His research topics include library automation and digitization, library marketing and software agent.

Dr. Eunja Kim earned her Ph.D. at Kyonggi University, Korea in 2009. Currently she is working for Gwacheon Public Library. Her research interests include space organization of library by considering the user's behavior.