Seat Usage Data Analysis and Its Application for Library Marketing

Toshiro Minami and Eunja Kim

Kyushu Institute of Information Sciences, Japan & Kyushu University Library, Japan and Gwacheon Information Science Library
minami@kiis.ac.jp, gaia55@naver.com
http://www.kiis.ac.jp/~minami/

Abstract. Due to the progress of information and communication technology, our society is changing very quickly. Along with this, people's requirements to information are not only changing vigorously but also have a huge variety in their forms. As a result it becomes more and more difficult for the libraries to provide their patrons with appropriate information services. In this paper, we demonstrate the usefulness of the concept of library marketing through some examples; especially with some analysis methods for seat-usage, or seat-occupation, data in library. We investigate what seats are more preferred than others and try to deduce tips for better seat arrangement, combination of different types of seats, etc. Even though our research is in a very early stage so that we could not infer that good suggestions to better seat arrangements, we believe in its importance and it would propose the best solution in seat design in the future.

Keywords: Library Marketing, Seat Usage/Occupation Data Analysis, Data Mining, Seat Arrangement.

1 Introduction

Due to the progress and popularization of information and communication technology, our society is changing very quickly. Not only it becomes easier and faster to access information but also our way of dealing with and requirements to information has changed thoroughly. Our information access becomes ubiquitous; i.e. at any time from any place.

Libraries have been playing a very important role in information and knowledge finding. Providing the patrons with appropriate information is the major mission for libraries, as was indicated in the Five Laws of Library Science advocated by S.R. Ranganathan in 1960s [6];

(1) Books are for use,
(2) Every reader his book,
(3) Every book his reader,
(4) Save the time of the reader, and
(5) The library is a growing organism.

N.T. Nguyen, C.-G. Kim, and A. Janiak (Eds.): ACIIDS 2011, LNAI 6591, pp. 238–247, 2011.

About a half century has passed since then and even though the environment of the libraries has changed a lot, the main idea about the mission of library is still very important even now and will continue in the future as well.

In order to keep this major role in our society, libraries have to be changing according to the change of people's requirements about information and knowledge. The big problem is that the spectrum of the people's requirements is very wide now. Thanks to the existence of the Internet and mobile terminals, like cell phones and other wireless equipments for digital communications, people would eager to get information as soon as possible at the site they encounter in their daily lives.

A lot of online information services are provided by quite a lot of (profit-oriented) companies in order to fulfill such requirements. Now we are able to check how many minutes later our intended bus will come to the bus stop where we are waiting. We can find the best route from where we are to our destination building, as well. But such services are not sufficiently enough. A very convenient service that is provided without any charge now may change to be payed service at any time when the company wants to do. It is highly necessary for us to get some kinds of information providing, and educational or research support services, as the essential public services for our society just like we get security services by police and fire departments as public services.

Libraries have been providing with information mainly in the form of printed books and magazines. However, considering the environmental changes relating to information materials, they have to change in order to adapt to the current environment. Of course the librarians recognize it and they put much effort to change themselves. However because that the speed of the change is so fast and the services that people want to get becomes quite a wide spectrum, it is quite difficult to appropriately change the way for the libraries based on the traditional methodology they are taking so far.

Our idea to deal with such a situation is to use the data analysis, or data mining, methodology. It is well known that a convenience store chain utilizes the data. They collect the POS (Point of Sales) data as the customers purchase their goods. The purchase data together with some profile data of the customers will be sent to the central server immediately. The specialists analyze the data and extract the information that will suggest what kind of goods should be delivered to the store for tomorrow. The net-shopping sites can collect the customers behavior data automatically. They analyze them and decide which goods to recommend to each customer according the customer's purchase history and his/her behavioral characters extracted from the data; i.e. Web mining. Such examples suggest that data analysis, or data mining, for library's marketing purpose will be very successful.

However the librarians are very worried about the privacy issue and they do not want to deal with the data the libraries can have for marketing purpose. We believe that they should go one step forward in order to keep being the public service organization that can answer to the patron's requirements more appropriately. They can do their best for protecting the data that may be connected

with the patrons' privacy data from spilling out of the library, by taking the up to date security system. The research results from the privacy preserving data mining (PPDM) [1] will be able to apply to their data analysis. By renaming the patrons' IDs from the real IDs to the tentative IDs that are used for analysis, the risk for privacy issue will be relieved much.

The aim of this paper is to ask for the librarians to start using data analysis for their marketing by showing library data analysis examples. We take up the seat usage data as the analysis target in this paper. Our research has been just started and is just in a very beginning stage. However we are convinced that data analysis methodology has a high potential so that it will provide the important tips and knowledge for improving patron services, creating new services, managing libraries better, and so on.

The rest of this paper is organized as follows: In Section 2, we discuss about library marketing (LM) more precisely. In Section 3, we take up the seat usage data of Kyushu University Library in Japan, and demonstrate the usefulness of the data analysis methodology in understanding the patrons' behavior. Finally in Section 4, we conclude our discussions in this paper.

2 Library Marketing

According to the American Marketing Association (AMA) [2], 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." This definition is more profit-oriented than the current definition: "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."

From these two definitions, we recognize that marketing was considered as the activities that benefit the organization (company); which matches with the ordinary people's intuitive image. 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)/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 (see also [5] on this) 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 [3,4].

(i) Preliminary Investigation
In this level we investigate what information, tips, and knowledge could be obtained by analyzing some kinds of data as case studies. 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.

(ii) 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 the 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.

(iii) 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 (ii) sufficiently.

(iv) 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 investigate the seat usage/occupation data as a preliminary investigation (i) and an example analysis of a very small real data (ii).

3 Seat Usage Data Analysis

3.1 Case Study at Kyushu University Library (KUL)

Kyushu University is one of the biggest and high ranked universities in Japan. It consists of 13 faculties, and has about 19,000 students. The main campus is located in Hakozaki area, where the main library of the Kyushu University Library (KUL) system is located. KUL has about 4 million books, about 90 thousand titles of magazines, and about 40 thousand titles of e-journals. Roughly

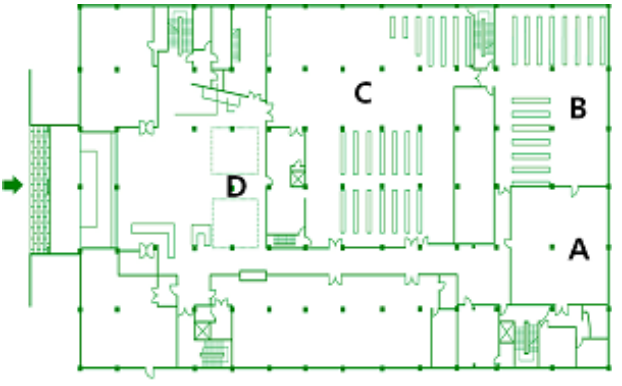

Fig. 1. The 2nd Floor Plan of the Central Library of KUL

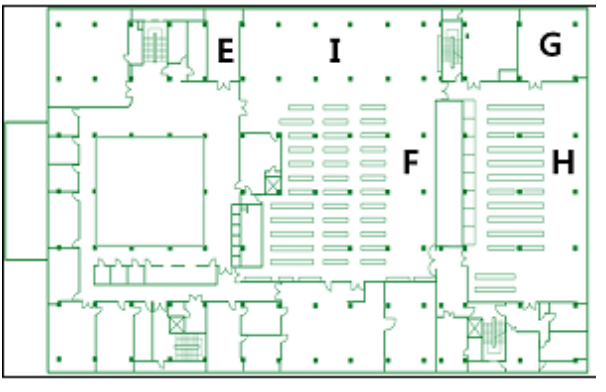

Fig. 2. The 3rd Floor Plan of the Central Library of KUL

340 thousand students, university staff, and others visit the central library per year.

We show the 2nd and 3rd floors of the central library building, which are the main parts of the building to be used by patrons. (2nd Floor: Fig. 1, 3rd Floor: Fig. 2). The 2nd floor is the entrance floor: the entrance gate is located in the left part of the figure. In the lobby, is a PC area (Fig. 3), where 46 PCs are available, which is called the Information Salon 1 (D). In the right area of the lobby as the patrons enter are the counters for circulation and reference services. Rooms marked A and B are the reading rooms. The Room C is now used as a "learning commons" area, which is equipped with many small tables and students are allowed to re-arrange the tables freely and they can talk aloud in a group meeting so some other purposes. It is also used for seminars organized by professors.

In the 3rd floor, are located reading rooms (marked G, H, I, and F) and another PC room (marked E), called the information salon 2 (Fig. 4), which is

Fig. 3. A PC Room: Information Salon 1 (D) in the 2nd Floor

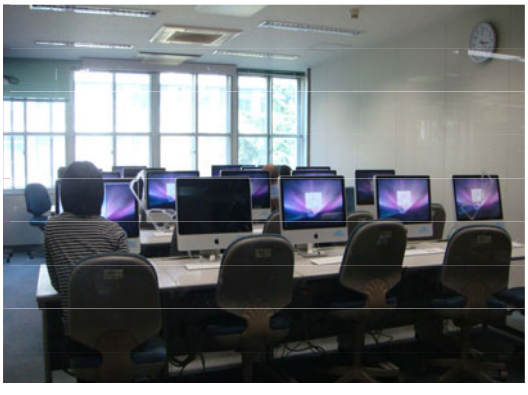

Fig. 4. A PC Room: Information Salon 2 (E) in the 3rd Floor

much smaller than the information salon 1. These information salons are used for lectures occasionally.

3.2 Analysis Example for the Information Salon 2

We created some seat usage data of the central library of KUL. We checked which seats are occupied by patrons by spending several days. Roughly speaking, we recorded every half hour. For example, in the case of Room E, i.e. the Information Salon 2, in the 3rd floor, we collected the data for January 13, April 16, and April 17 in 2009. On January 13, we investigated 7 times during the period from 15:30 to 20:30, and in April 16 and 17, 28 times from 8:00 to 21:30. So we have 63 times of data in total.

The seat arrangement of Room E is illustrated in Fig. 5. As we can see in Fig. 4, there are 5 PCs in a table. There are 4 tables in the room and 2 of them in the front part only have 4 PCs instead of 5 because the central part is used for printers. Therefore there are 18 PCs and seats in Room E.

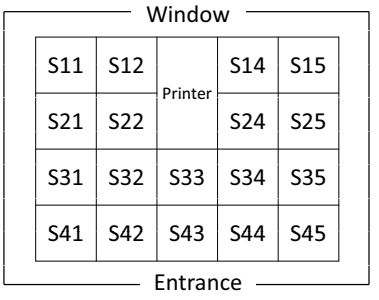

Fig. 5. Seat Arrangement of the Information Salon 2 (E)

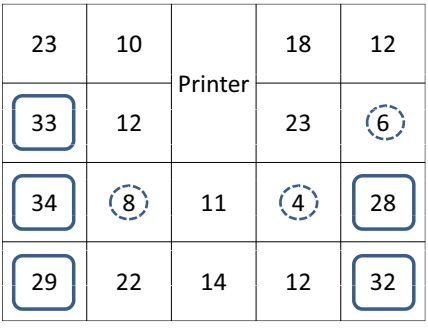

Fig. 6. Seat Usage Frequencies of the Information Salon 2 (E)

Fig. 6 shows the total numbers of the usage of each seat. For example the seat S11 was found to be used 23 times among 63 times; the usage ratio is 37%. The most popularly used seat was found to be S31 (54%), which is followed by S21 (52%) and S45 (51%). On the other hand the most unpopularly use ones are S34 (6%), S25 (10%), and S32 (13%). We put rectangular marks to the most popular seats and put circular marks to the most unpopular seats.

From these data we can say that the people prefer to use the seat located in the edge and at the same time not too much far away from the entrance. So the most popular seats are located in either S*1 and S*5. It is interesting that the left edge, i.e. S*1, is more popular than the right edge, i.e. S*5. Probably in some reason the right side passage is considered to be the main passage area than the left one. So they prefer to use the left one as the more quiet seats than the right ones.

At the same time it seems that they do not want to use the seat next to the one which is already occupied. This describes why S32 and S34 is not popular. It is interesting to see that the seat next to the printer seems to be considered a kind of edge positions. This may be the reason why S14 and S24 is more popular than other inner seats. This also describes why S25 is not popular; it is next to the relatively popular seat S24.

9	5	Printer	4	5
7	3		9	1
8	3	7	3	8
11	7	8	7	10

Fig. 7. Session Frequencies of the Seats in the Information Salon 2 (E)

2.6	2.0	Printer	4.5	2.4
4.7	4.0		2.6	6.0
4.3	2.7	1.6	1.3	3.5
2.6	3.1	1.8	1.7	3.2

Fig. 8. Average Session Times of the Seats in the Information Salon 2 (E)

We take another view on this data. We will define a session of a seat usage if the seat has been used continuously in the data. As an example, let us take the seat S11 in the data for April 17. The seat has been occupied at 11:00, 11:30, 12:00, 12:30, 13:30, 14:00, 14:30, 15:00, 16:00,16:30, 19:00, 19:30. This means it consists of 4 sessions; from 11:00 to 12:30, from 13:30 to 15:00, from 16:00 to 16:30, and from 19:00 to 19:30. The total number of sessions are shown in Fig. 7. As a result we can define the average session time, which is shown in Fig. 8.

The longest average session time is 6.0 at the seat S25, which has only 1 session. So a patron happened to choose this seat and used it for a work that needed relatively long time. Thus this seat usage would be an exceptional case. The seats for the next longest average session times are S21, S14, and S31. Among these 3 seats, S21 and S31 are the seats that have also used for a long time (See Fig. 6). So we can say that these seats are preferably chosen by the patrons for the works that would need to spend a long time. On the other hand the seat S45 has a big frequency while its average session time is not very large; which might mean that this seat is easy to access because it is close to the entrance door and thus this seat is considered to be very useful for relatively shorter works. Similarly the seats S33, S34, S43, and S44 have short session times. Among them the session S33, S43, and S44 have relatively larger usage frequencies. So we can

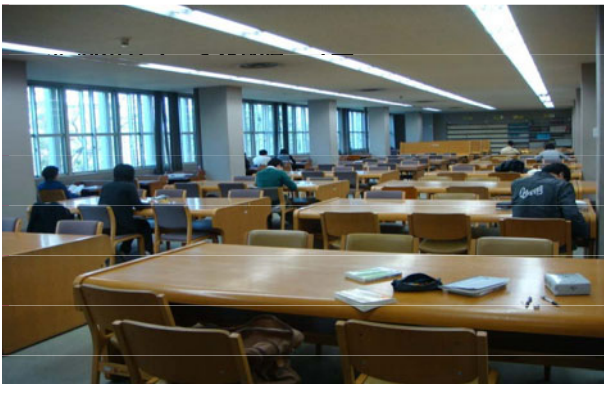

Fig. 9. A Reading Room (I)

say that these seats are preferably chosen for small works because they are easy to acess.

It is worth pointing out that Uematsu suggests that a table with 6 seats with 3 seats in one side and the rest 3 in another side (see for example in Fig. 9, Reading Room I) is full for two patrons because they will occupy the diagonal seats and nobody will not willing to use the remaining 4 [7]. In a rough analysis so far, it looks true in our data too. We have to analyze in more detail for further investigation.

4 Concluding Remarks

Libraries are supposed to go with society as was described in the five laws of library science [6]. Actually they have been changing itself based on this philosophy. They started using computers in 1980s and started with providing home pages as the Internet and Web become popular in 1990s. However the speed of changes of society is too fast recently for the libraries to catch up with it based on the methodology they have been taking so far. They need a new methodology in the coming information age.

We believe that data collection and data analysis including the data mining technologies have essential importance for the libraries. So we have started with developing data analysis methods for the libraries.

In this paper we put focus on the seat usage data. We have investigated which seats the patrons are using in the central library of Kyushu University. We chose the Information Salon 2 for an example and try to find some tips extracted from the data. Our main concern is to understand how the patrons choose their seats. Then we would like to find some kind of information and knowledge based on the results, which is useful as the librarians plan the seat arrangement in a library.

As we have just started our research based on such a methodology, we do not have much yet. However, we are convinced that we will be able to find more and more useful knowledge as we continue our research in this direction. In order to

do this, we have to collect much more data not only on seat usage data but also other data that will be useful for understanding patrons and their behaviors in and out of the library.

References

1. Aggarwal, C.C., Yu, P.S. (eds.): Privacy-preserving data mining: models and algorithms. Advances in database systems, vol. 34. Springer, New York (2008)
2. American Marketing Association (AMA), http://www.marketingpower.com/
3. Minami, T., Kim, E.: Data Analysis Methods for Library Marketing. In: Lee, Y.-h., Kim, T.-h., Fang, W.-c., Ślęzak, D. (eds.) FGIT 2009. LNCS, vol. 5899, pp. 27–34. Springer, Heidelberg (2009)
4. Minami, T., Kim, E.: Data Analysis Methods for Library Marketing in Order to Provide Advanced Patron Services. International Journal of Database Theory and Application 3(2), 11–20 (2010)
5. Nicholson, S.: The basis for bibliomining: Frameworks for bringing together usage-based data mining and bibliometrics through data warehousing in digital library services. Information Processing & Management 42(3), 785–804 (2006)
6. Ranganathan, S.R.: The Five Laws of Library Science, 2nd edn. Asia Publishing House (1957)
7. Uematsu, S.: To Have a Look at Library as Architecture. Bensei Shuppan (1999) (in Japanese)