A Comparative Study of User's Information Seeking Behavior in e-Book Recommendation System

Yi-Shin Lin, Ko-Chiu Wu, and Saiau-Yue Tsau

***Abstract*—E-book reading has become increasingly popular, the library began to provide joyful and interactive multimedia e-book recommendation system. The purpose of this study is to understand the characteristics of information seeking behavior of different users. Based on the back-end database of New Taipei City Library's e-book wall, we collected 655 users' borrowing materials and analyzed the knowledge keywords by text mining method to make a comparative study. With the expansion of library services, the purpose of the library is no longer limited to the academic orientation, and we also observed the emergence of entertainment and leisure-oriented readers. Understanding users' information seeking characteristics in e-book recommendation system will be conducive to the development of intelligent library and provide a reference for the system design.**

***Index Terms*—Information seeking, knowledge map, text mining, information visualization.**

I. Introduction

There are several factors of rapid changes in the information content: Text content is the primary method of representing knowledge. Text content is challenged by auditory, visual, and various forms of multimedia. Large amount of information is easily replicable or inexpensively transmitted, which has a significant impact on the interaction of human messages [1].

Facing a lot of information, users will have a new search strategy and ability. While most users understand that the Internet is a fairly complex information space, and the search strategy may not be optimal. But users will choose the resources they are most familiar with and avoid extreme choices [2]

According to the Sense-Making Theory, the information seeker with different background knowledge perform the search tasks, and the personnel with professional knowledge can systematically construct information search behavior.

This study is based on the e-book recommendation system which called E-book Wall in New Taipei City Library. With historical borrowing data in back-end database, to understand the information searching characteristics of different users and visualize the users' knowledge keywords by Chinese text mining method. Finally, this study provides some suggestions of catalog classification interface in the E-book Wall which based on users' searching characteristics of research results and related literature.

Manuscript received February 10, 2017; revised May 4, 2017.

The authors are with the Department Interaction Design, National Taipei University of Technology, Taipei 10608 Taiwan, R.O.C. (e-mail: t104ac8001@ntut.org.tw, kochiuwu@mail.ntut.edu.tw, Sytsau@ntut.edu.tw).

II. Literature Review

A. Information Seeking

Information seeking is a state of mind. When people are confused, they usually use their own experience or knowledge background as the primary consideration. If people want to solve the problem successfully, they will have a need to seek information [3]-[5]. Information seeking besides reducing the uncertainty of specific events, as well as to satisfy individual curiosity without a specific purpose [6].

B. Information Seeking Behavior

The researchers [5] simply suggest that information seeking behavior is a way for individuals to solve their information needs in order to respond the immediate needs of difficult methodological or environmental factors. The researchers' [4] study identifies five factors that influence information seeking behavior:

1) Personal traits, including user's psychological factors, demographic characteristics, affective, cognitive factors, such as qualifications, experience, educational level, Professional background, attitude towards work, personal motivation and so on.
2) Information cognition: successful retrieval experience, familiarity with information, credibility, timeliness, efficiency, quality, accuracy and convenience, cost.
3) Interpersonal relationships: the role of work, work participation, participation levels, learning groups, reference groups, formal organizations.
4) The working environment: the type of work, organizational characteristics, the working group, the organization of communication networks, research program stage.
5) Social environment: social and cultural environment, political and economic environment, material environment.

C. Classification Architecture

Today's search engine is text-based, and the search results are usually not closing to human experience, to develop a retrieval system that truly close to the user's precognitive keywords is inevitable trend. Most of the current search engines are using the way by keyword search. The way is quite convenient, but also spawned other problems: how users find the information they really want from a huge network.

In general situation, the information needs by keyword search is not fully understood, such as people selecting the wrong keyword or category, which the search result may

doi: 10.18178/ijiet.2018.8.3.1038

deviate from the actual requirement, thus missing information of important reference [7].

Therefore, in order to improve the above mentioned situation, this study is expected to understand various user groups' information seeking behavior, and to analyze user's knowledge structures by collecting keywords.

D. Knowledge Map

Concept maps and knowledge maps are diagrams that represent ideas as node-link assemblies. They are often used as media for constructive learning activities and as communication aids in lectures, study materials, and collaborative learning [8].

Knowledge mapping is a good tool to visualize the knowledge resources of library, and it clarify the inner relations of distributed knowledge resources.

According to the knowledge map, the librarians can quickly find a reasonable knowledge matching, and feedback suggestions to readers, even providing collaborative retrieval to improve the quality of service [9].

E. Text Mining

Text mining techniques include categorization of text, summarization, topic detection, concept extraction, search and retrieval, document clustering, etc. Each of these techniques can be used in finding some non-trivial information from a collection of documents. Text mining can also be employed to detect a document's main topic/theme which is useful in creating taxonomy from the document collection. Text mining has also been applied on many applications on the World Wide Web for developing recommendation systems [10].

Now there are a lot of text by users on the Internet, such as Twitter, Facebook, and other community sites have a large number of text messages. Because of the numerous and messy information, we can use text mining technique to extract useful information [11], this study present the relevant text mining methods to analysis the data of users' borrowing history.

F. Information Visualization

Information Visualization transforms information into graphic or chart, and is presented interactively and visually, making it easy for users to understand the content of information [12]. This study present the visualization content of users knowledge keywords by the text mining and visualize method.

III. Research Methods

A. Introduction of e-book Recommendation System — Using e-Book Wall of New Taipei City Library as an Example

E-book Wall interface is operated by touch, is an interactive multimedia display platform to provide users to browse, search and borrow e-books. Users can set the personal information into the borrowing card, and allow the system to recommend personal bibliography. If users want to read further the details of the e-book, they must first install the App named "United Daily News (UDN) Reading Bar" on the smart phone to download e-books.

Fig. 1. The equipment of e-book wall in new Taipei city library, user click on the book cover to access books' Introduction.

B. The Recommendation Algorithm

E-book Wall database algorithm set parameters are as follows: 1. the numbers of the readers to borrow the book, 2. the category of "purpose of visiting the library" (Figure 2, Table I), 3. Experts and Librarians' recommended scores on the book, and the program finally comes to recommend results and form the exposure of the e-books.

TABLE I: "22 Purposes of Visit the Library": The Project Development Based on the Public Library of the Finnish People of the Study [13]

Jobs	Education	Daily activities	Cultural Activities
Find job information	Find educational opportunities	Handle the housework	Read the novel
Perform a specific work task	Completion of formal education (degree)	Child care and school attendance	Read non-fiction
Training work skills	Work-related educational development	House care	Cultural activities (e.g. appreciation of theater or concerts)
	Self-education at leisure	Consumer rights or information	Creative activities (e.g. musical instruments playing or singing)
		Health	Cultivating outdoor activities, sports interest
		Tourism and vacation	Interested in nature

Fig. 2. The search interface of E-book Wall by "22 purposes of visit the library" as the basis for classification [14].

C. Database Correlation Work

This study collected 655 users' data of E-book Wall from the July 1, 2015 to August 31, 2015. The data is divided into three main categories:

1) Immediate Records: As for the behavior of the readers, record "what", "who" and "when" the choices were made. In addition the reader first put on the card will do registration, registration results will be recorded in the

"Users" table

2) The results of real-time record
3) Initial setting data: "Ebooks": records the ID of all books that can be borrowed, and the purpose of visiting the library (Pov1 ~ Pov3). "UserGroups": Records grouping information for each user.

Finally, this study add the e-book introduction materials into the data table and re-associate the database by common software Access.

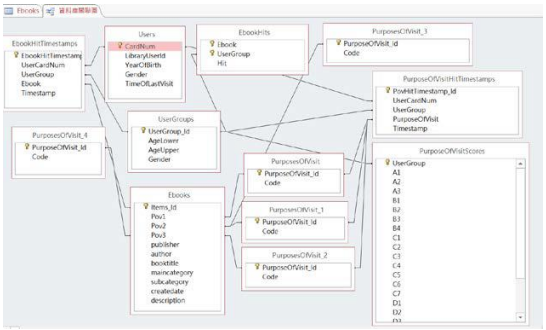


Fig. 3. Database association diagram.

UserGroup_Id	AgeLower	AgeUpper	Gender
1	0	6	1
2	7	12	1
3	13	18	1
4	19	55	1
5	56	75	1
6	76	2147483647	1
7	0	6	2
8	7	12	2
9	13	18	2
10	19	55	2
11	56	75	2
12	76	2147483647	2

Fig. 4. User data grouping, such as UserGroup_Id = 4, representing 19-55 year-old male.

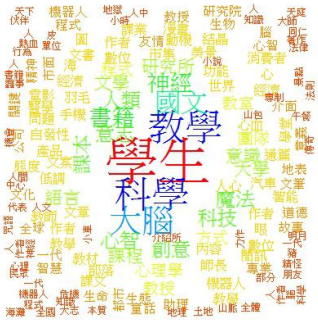


Fig. 5. Text mining visualization, the picture is 7-12 years old male, It could be seen that Chinese keywords of books borrowing by the user group, and the terms are most related to the school, learning or education.

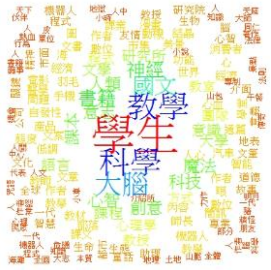
D. Text Mining

This study applied the visualization program of Chinese text mining which integrated by the researcher [11].

R (programming language) is a suitable tool for data analysis and it provides the suite of text mining tools to apply. This method is used to analysis the characteristics between the user group and e-books which users click on, as shown as following:

1) Install the necessary R programming library(tm, tmcn, Rwordseg API suite)
2) Execute Text reading of the books' introduction which user borrowing.
3) Execute Chinese text segments and distinguish words.
4) Build a TermDocumentMatrix to count number of Chinese vocabulary occurrences and save.
5) Visualize the keywords of knowledge.

TABLE II: KNOWLEDGE KEYWORDS AND WORD FREQUENCY BY TEXT MINING

7-12 years old male (10 records)	7-12 years old female (49 records)
	

	word	freq
學生	學生	21
科學	科學	14
教學	教學	14
大腦	大腦	12
國文	國文	9
書籍	書籍	7
神經	神經	7
人類	人類	6
心智	心智	6
科技	科技	6
創意	創意	6
意志	意志	6
謀本	謀本	6
魔法	魔法	6
大學	大學	5
文學	文學	5

	word	freq
咖啡	咖啡	80
故事	故事	27
同學	同學	26
老師	老師	22
小說	小說	19
文學	文學	18
孩子	孩子	18
心理	心理	17
詩詞	詩詞	16
詩歌	詩歌	16
心靈	心靈	15
世界	世界	13
路線	路線	12
精神	精神	12
小吃	小吃	10
主題	主題	10
人物	人物	9

13-18 years old male (20 records)	13-18 years old female (57 records)

	word	freq
咖啡	咖啡	40
路線	路線	22
時間	時間	16
故事	故事	14
外遇	外遇	13
主題	主題	12
社會	社會	11
小說	小說	10
植物	植物	10
人類	人類	9
愛情	愛情	9
文學	文學	8
旅規劃	旅規劃	8
世界	世界	7
男人	男人	7
時光	時光	7

	word	freq
能力	能力	55
咖啡	咖啡	47
世界	世界	44
城市	城市	30
故事	故事	30
名字	名字	26
數學	數學	26
文學	文學	23
方式	方式	22
內容	內容	19
孩子	孩子	19
人類	人類	18
社會	社會	17

19-55 years old male (20 records)	19-55 years old female (188 records)

	word	freq
世界	世界	52
路線	路線	49
單車	單車	43
咖啡	咖啡	41
人物	人物	38
故事	故事	38
歷史	歷史	36
資訊	資訊	35
麵包	麵包	35
時間	時間	33
社會	社會	32
伯斯	伯斯	31
夢想	夢想	31

	word	freq
人生	人生	67
社會	社會	63
故事	故事	56
世界	世界	55
女孩	女孩	51
咖啡	咖啡	47
建築	建築	47
文學	文學	46
愛情	愛情	44
城市	城市	42
文化	文化	41
小說	小說	39
大陸	大陸	38
男人	男人	36

IV. Research Results

In This study, there are 655 historical data were analyzed from July to August in 2015.

Excluding the situation that a person repeat clicking of the same e-book, during the period, there were 136 men and 519 women, and the users both of children under 6 years old and elders than 56 years old's were less than 5 , so they were excluded. By the analysis of observation, the largest user group of E-book Wall are 19-55 years old adults. Children and the elders are limited by the knowledge background and technical proficiency, most will not take the initiative to operate. The main purpose of visiting library like to read leisure books, the topic such as travel, novels, stories, photography or Picture Book. 7-12 years-old children preferred education, science and picture books, adolescent more preferred comics, novels and the books that is about interpersonal issues. The reason why the term "咖啡"(means coffee) is a high-frequency keywords, is because many people click the Japanese light novel which named "Cafe reasoning event book". Adults' preference cover extensively, the books content they read are about scientific and technical reference, ideological commentary, history, culture books and so on, because they have the mature experience, recognition and knowledge background to understand the more complex description.

V. Conclusion

According to the results, the main user group of the E-book Wall recommendation system is still consist of teenagers and adults. Through the data statistics and the simple interviews with the users, the books they like to read are entertainment and leisure oriented. About the category, the text description of “the purpose of visiting library “classification still exist differences in users' understanding. Some category options that readers rarely clicked, or the recommended e-books don't meet the users really want to see. This study compare the different user group's information seeking behavior and the knowledge keywords. In the future, If using big data of the historical materials, automation, visualization methods, and further generalize more specific knowledge map of user groups as a reference, which will help the intelligent library develop e-book interactive recommendation system, and it would be closer to the actual needs of users.

References

[1] G. Marchionini, “Human–information interaction research and development,” *Library & Information Science Research*, 2008, vol. **30**, no. 3, pp. 165-174.

[2] N. Pharo and K. Järvelin, ““Irrational” searchers and IR-rational researchers,” *Journal of the American Society for Information Science and Technology*, 2006, vol. 57, no. 2, pp. 222-232.

[3] C. Atkin, *Instrumental Utilities and Information Seeking*, 1973.

[4] A. Bouazza, “Information user studies,” *Encyclopedia of Library and Information Science*, 1989, vol. 44, Suppl 9, pp. 144-164.

[5] J. Krikelas, *Information-Seeking Behavior: Patterns and Concepts*, 1983, vol. 19, no. 2, pp. 5-20.

[6] R. L. Derr, “A conceptual analysis of information need,” *Information Processing & Management*, 1983, vol. 19, no. 5, pp. 273-278.

[7] H. Po and C. Chang, “An exploratory study of users and tags on social bookmarking websites,” *J Library Inform Sci Res*, 2009, vol. 4, no. 1, pp. 1-26.

[8] J. C. Nesbit and O. O. Adesope, “Learning with concept and knowledge maps: A meta-analysis,” *Review of Educational Research*, 2006, vol. 76, no. 3, pp. 413-448.

[9] J. Yang, *On the Construction of Knowledge Map in Reference Service of Libraries*, 2010-02.

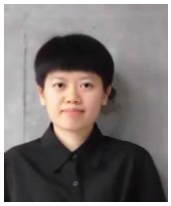
[10] H. Hashimi, A. Hafez, and H. Mathkour, “Selection criteria for text mining approaches,” *Computers in Human Behavior*, 2015, vol. 51, Part B, pp. 729-733.

[11] J. Chen, *Using R for Chinese Text Mining*, in *Programmers Magazine, Taiwan R User Group*, 2014.

[12] B. Shneiderman, “Eyes have it: A task by data type taxonomy for information visualizations,” *IEEE Symposium on Visual Languages, Proceedings*, 1996.

[13] P. Vakkari and S. Serola, “Perceived outcomes of public libraries,” *Library & Information Science Research*, 2012, vol. 34, no. 1, pp. 37-44.

[14] T.-Y. Hsieh, *Integration of Clustering and Categorization of Knowledge Classification Structure with Library e-Book for Wall Interface Design*, in *Department of Interaction Design*, 2015, National Taipei University of Technology.



Yi-Shin Lin is currently a graduate student from National Taipei University of Technology, major in interaction design. She completed her bachelor of digital technology design from National Taipei University of Education in 2015. The main areas of research include human-computer interaction and information visualization.

Ko-Chiu Wu is a professor of Dept. Interaction Design, National Taipei University of Technology. He received his Ph.D. in Civil Engineering from National Taiwan University, Master of Computer Science from Villanova University, and Master of Architecture from Pennsylvania University, USA. The main areas of research include digital learning and interactive technology, information technology and management, architectural design and planning.

Saiau-Yue Tsau is an associate professor of Dept. Interaction Design, National Taipei University of Technology. She received Ph.D. in Fine Arts Education and Administration from National Taiwan Normal University Art Institute. The main areas of research include Interactive art, interactive technology.