

Research on Personalized Service of University Library Based on User Behavior

Jinping Xu

The Library of Wuhan University of Technology, Wuhan 430070, China

61310642@qq.com

Keywords: Library; Big Data and Digital Mining; Personalized Service

Abstract: In the context of big data, how to construct a suitable user behavior model and provide personalized services based on massive behavior log data is an urgent problem that needs to be solved in the current university library big data application. Based on the connotation and basic characteristics of personalized information service, this paper constructs a user behavior analysis system for library personalized service based on big data and digital mining from the perspective of user behavior analysis. The research results show that the user behavior model based on ontology system can be seamlessly docked with the large data analysis platform in technology, thus providing real-time and accurate services. It can effectively deal with the "knowledge lost", "information overload" and "emotional missing" of library personalized services under the current large data environment War. The final research shows that the personalized information service of university libraries is bound to be more systematic and large-scale, and become the future development direction of library information service. This system can discover the readers' service needs in the era of big data and effectively guarantee the service quality of libraries.

1. Introduction

With the personalized information service becoming the mainstream of the new service mode of the library, the research on the personalized needs of users has emerged one after another [1]. At present, the concept of personalized information service has penetrated into the field of Library Science and become a new research hot spot in the field of digital library. Personalized information service is a new service concept compared with popular information service [2]. Personalized services generally provide users with personalized features and knowledge services to meet individual needs through personalized retrieval, system recommendation and website interaction according to their personalized characteristics and different needs [3-4]. Among them, demand-driven integrated services, and personalized active intelligent services are one of the key developments in the future. The reader has got rid of the constraints of reading time, region, service object, reading content and terminal type, and realized user-centered personalized reading service [5]. In order to accurately extract user interest characteristics and analyze user implicit requirements, an effective method is to obtain the main attributes such as user information and behavior, and apply structural language to describe [6]. In the era of big data, users' demand for information has increased, but information selection has become a current problem. Taking university libraries as an example, how to provide effective information screening for users and promote the generation of personalized services is the current focus [7].

As a teaching and research institution, colleges and universities are the main body of their library services. This determines that the personalized information service carried out by the university library facing the subject is obviously different from the personalized service in the general sense [8]. Therefore, in order to meet the individual needs of readers, innovate the library service model, and adapt to the trend of the times, the service function of the library should be changed from passive to active [9-10]. Many foreign scholars have focused their attention on the information behavior of university students and researchers. In 2013, some people conducted research on the model construction of personalized active information services in digital libraries [11]. In the following year,

researchers also conducted research on the evaluation ability of personalized information services in higher vocational colleges [12]. Then, a research on the design and implementation of personalized service in University Libraries Based on data mining technology was put forward [13]. However, in the large data environment, due to the rapid growth of library data in recent years, many types, and low value density, the problems of "knowledge lost", "information overload" and "emotional deficiency" faced by library personalized services are more serious than those in the digital library era [14]. This makes the traditional knowledge service system difficult to meet the new needs under the new situation. Therefore, university libraries should seize the opportunity to explore the development and application of big data in University libraries, attach importance to user behavior data, explore potential value, and provide personalized services for users [15-16].

With the continuous development of information technology, all walks of life are experiencing a wave of Digitization. As a center of information resources, libraries also follow the trend of Digitization and move forward. In the process of Digitization, libraries have accumulated a large number of digital information resources [17]. The current knowledge service is not limited to the library's own basic service system, but also includes high value-added services such as in-depth analysis of structured and unstructured massive data, competitiveness analysis, innovation analysis and predictive analysis, which provide valuable decision support and wisdom service for service users [18]. This research is based on domestic university libraries. It aims to investigate the problems existing in the development of personalized information services in university libraries by investigating the characteristics and rules of information behavior of university library users, and to find out the future development of library personalized services. Direction [19]. Furthermore, it analyzes the current research status of personalized service based on user behavior modeling and data mining in the current picture field, and then combines the characteristics of personalized service in big data environment, proposes the construction strategy of user behavior model, and then gives the model construction method. And personalized service plan [20].

2. Data Mining of Reader Behavior Patterns in Library

The research of library personalized service based on user behavior model mainly focuses on the discussion of model construction method and user behavior analysis based on data mining method. Data mining is a process of extracting potentially useful information and knowledge hidden in a large number of noisy, incomplete, fuzzy and random databases [21]. The research of personalized service in big data environment has attracted the attention of academia. The rapid development of big data technology and the continuous improvement of the library's personalized service requirements have led to the research of personalized services in the big data environment. At present, although the information resources of university libraries are relatively abundant, when users search for information, because the information is mixed, the association is less, and users may not be able to obtain satisfactory results [22]. This requires the establishment of a people-oriented, customer-oriented service concept in the work of personalized information services, and information services from the information resources to the center of the user needs. The library is required to collect, process, calculate, analyze and make decisions about the reader's behavioral big data according to the life cycle rules of the reader's reading activities, and provide big data decision support for the user service mode selection and process [23].

The next step in the algorithm is to mine association rules based on frequent Item set. A rule with a confidence greater than the minimum confidence is called a frequent association rule. The following begins the data mining process of frequent sets, as shown in Table 1 and Figure 1:

Table 1 Frequent Item set in Data Mining Processes

Item set	Support degree	Confidence level
W	5	33
N	8	24

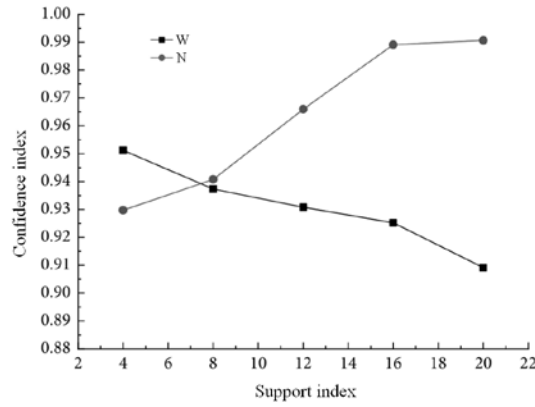


Fig. 1 Frequent Item set in Data Mining Processes

Combining the domestic and foreign relevant literature with specific applications, this paper discusses the construction method of user behavior model for personalized service and data mining technology to support personalized service from different angles, but most of these studies have many shortcomings. The environment of data mining refers to the sum of all objects that can carry out data mining, including: database (providing data support for data mining). Mining tools are the core of data mining. Without mining tools, they cannot achieve their goals. Based on this technology, according to the characteristics of readers, taking gender, college and grade as indicators, and by means of histogram or pie chart, we can understand the general trend of readers' behavior, and add the indicators of large and small categories of books to meet the needs of different professional readers and colleges to make progress in their borrowing preferences. The collection of reading activity related information includes the reader's individual characteristic data, reading habits, type of reading terminal and working mode. Because user information needs are the premise of the existence and development of library work, it directly affects the service content of the library. User needs information must be collected in a variety of ways. The object studied in data mining is the basis of the whole process, which drives the whole data mining process, and is also the basis and consultant for testing the final result and guiding the analyst to complete the data mining.

The association rule mining of the classification number-level transaction table can mine the association rules of the classification number level. The rule is mainly for the book category, and examines the borrowing habits and reading preferences of students when borrowing various books. Different association rules can be mined by setting different minimum support. In the experiment, we set a minimum support of 5%. Some association rules are mined as shown in Table 2 and Figure 2.

Table 2 Association Rules at Partial Classification Number Level

Number	Support degree	Rule length
1	0.96	8
2	1.78	6

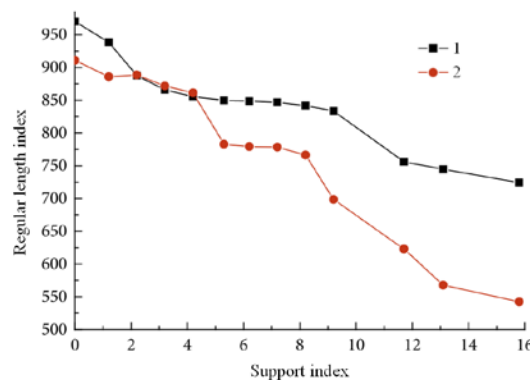


Fig. 2 Association Rules at Partial Classification Number Level

Data extraction of reader's interest features is a process of obtaining information related to reader's characteristics, preferences or reader's behavior. On the basis of researching users' behavior, interests and habits, information services are developed according to users' personalized needs. Its ultimate goal is to provide users with all kinds of information resources and services they need. Its characteristics are individualization and initiative. It varies according to users' behavior habits and interests. Personalized services need to carefully analyze user behavior characteristics, and describe them structurally based on ontology, so as to establish a behavior model reflecting user characteristics. In order to form the initial interest model, and then in the subsequent process of learning and mining, the association rules of the data borrowed and retrieved by readers are analyzed. In order to ensure that the user behavior analysis process is comprehensive, accurate, economical and fast, the library must expand the scope and quantity of user behavior data collection as much as possible. In the context of this complex information, the development of personalized information services can not only reduce information overload, but also shorten the time for users to obtain information. Therefore, based on these excavated information, the library can provide personalized services to students and teachers of different colleges and universities, and improve the learning efficiency of students' self-learning.

3. Model Design of Personalized Service System in University Library

The design and development of personalized service recommendation system in university libraries should start from the perspective of users and readers, and ultimately be established for the application of readers and the management of managers. According to this system, the analysis and statistics of each user's behavior log database can be made, and the ontology terms of user's explicit interest can be extracted to accurately reflect the user's preferences. Some of the extracted user information is obvious, but more of it is easily ignored by decision makers. These phenomena can provide important reference for decision makers. At the same time, it can accept the reader's evaluation of the personalized service provided and submit it to the request analysis module to analyze and update the reader's interest rules. It also discovers the behavioral relationships, user needs and knowledge contained in the reader's behavior data. It is the process of analyzing, judging, defining and matching the reader's behavior. It is also the library to grasp the reader's reading habits and discover service requirements, and improve the accuracy of personalized services. According to the data types under the existing conditions, under the clear goal, the readers are classified according to the reader's historical borrowing records, reader attributes and other factors. Readers belonging to the same group have a certain degree of similarity, and there are obvious differences between different reader groups. Using the explicit interest statistics and implicit demand mining methods based on user behavior log data analysis, it is beneficial to improve the quality and level of library services in the big data environment.

According to the data mining process of association rules, by setting different minimum confidence and minimum support, the association rules with different quantity and quality can be mined. The rule mining based on the call-book number database is shown in Table 3 and Figure 3. Only by properly setting these two parameters can you find out the ideal association rules.

Table 3 Number of Association Rules with Different Minimum Confidence and Minimum Support

Number of rules	Minimum support degree (%)	Minimum confidence(%)
35	82	30
560	79	55
380	56	8

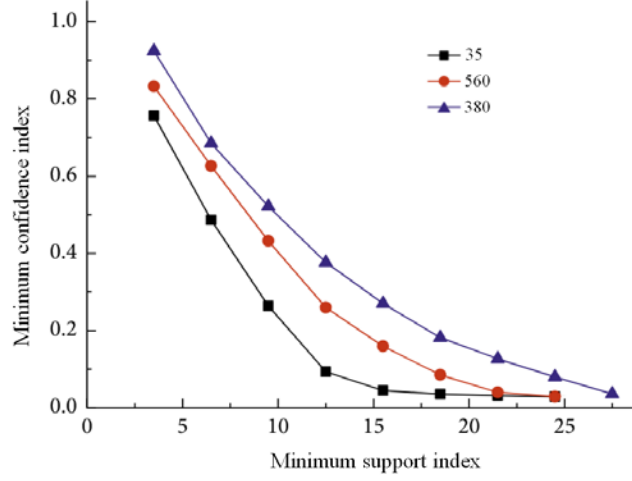


Fig. 3 Number of Association Rules with Different Minimum Confidence and Minimum Support

The strength of the resulting rules can be measured by support and confidence. The frequency of a given data set can be determined by support, and confidence is used to determine the frequency of W in transactions containing X. The formal definitions of support $w(t)$ and confidence $f(x)$ are as follows:

$$w(t) = w_2 + (w_1 - w_2) \frac{T - t}{T} \quad (1)$$

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

The support and confidence metrics obtained by the above methods are not very regular and are not conducive to subsequent calculations, so they should be further normalized. The normalized equation is:

$$f(t) = \sum_{j=1}^N \sum_{k \in Z} d_k^j \varphi_{jk}(t) + \sum_{k \in Z} c_k^N \varphi_{Nk}(t) \quad (3)$$

Among them, there are many similarity calculation methods based on vector space model. The commonly used method is inner product method. Its formula is as follows:

$$Cost(P_i) = \sum_{e \in P_i} C(e) \quad (4)$$

Accuracy refers to the percentage of books in the recommendation list that are of interest to the user, and the accuracy of the recommendation results can be expressed by the following formula:

$$M(q) = R_0 - \eta \frac{\Delta R q}{Q_0} \quad (5)$$

The recall rate refers to the percentage of books that the user is interested in and the system recommends for all the books that the user likes. According to the above formula, the recall rate of the recommended results is expressed as:

$$M(t) = R_0 (1 + \frac{2\Delta R U}{Q_0 R_0^2} t)^{0.5} \quad (6)$$

The personalized service system model designed in this paper is shown in Figure 4. This structure diagram is based on the user's registration from the website to the final personalized service, that is, from the collection of user information to the modeling of user information, and finally the user. The built model matches the generated knowledge base (rule base) to provide personalized information

services to users. Among them, the personalized recommendation is based on the modeling method, through the data analysis means, mining high-value knowledge from the massive low-value density information to the user, in order to effectively alleviate the embarrassment of "information overload" in the era of big data. It also conducts mining analysis and intelligent update for readers' borrowing and retrieval information and evaluation feedback data, including cluster analysis of books, readers, association rule discovery analysis and update sub-module of reader interest. The data acquisition in this system comes from the fragmentary digital information in the digital library, including reader's basic information, reader's borrowing history record, retrieval history record and library's collection resource information. Finally, it must be processed by the resource processing layer to update the database information in real time so as to ensure the time of resource data. Effectiveness. Finally, we will expand the breadth and depth of data collection of readers' behavior by means of service collaboration and sharing of large data resources. On the premise of user-centered reader data selection, filtering, sharing and complementarity, it can improve data application analysis and enhance data availability.

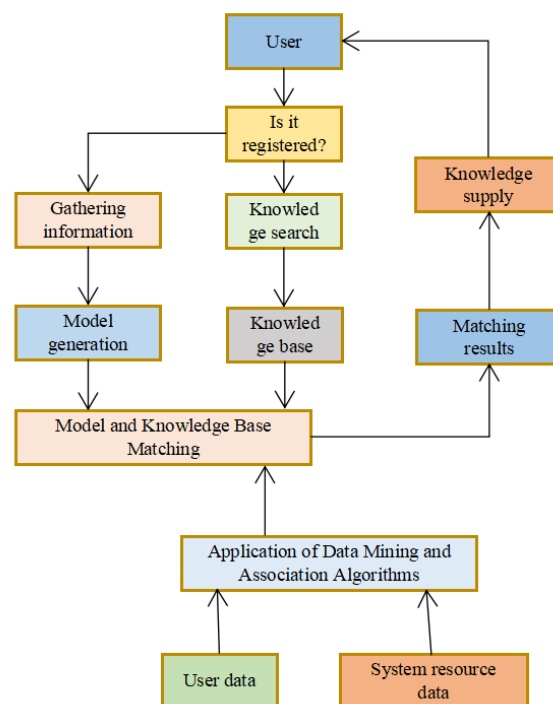


Fig. 4 Personalized System Structure Diagram Based on Data Mining

In the personalized recommendation, it is divided into two according to its timeliness, namely the online recommendation module and the offline mining module mentioned above. The offline mining module provides services for the online recommendation module, and the online recommendation module provides the data foundation to the offline mining module. Both are indispensable and complement each other. This is because information resources are the material basis of personalized information services. Under the network environment, the information resource structure of university libraries has changed greatly. It is necessary to strengthen the construction of information resources, establish a special database and a professional navigation database, and expand the full-text digital information resources of the library. When the library completes the collection of user behavior data, first of all, according to the classification of user behavior and the experience of administrators, it should filter and manually filter user behavior data to improve the value density and availability of behavior data. For the borrowing data, because the system focuses on the readers' recent interest, the borrowing time span cannot be long, and the borrowing time of the readers for a certain type of books is relatively long within the prescribed time span. And provide personalized services for students of different majors, grades and colleges, recommend the most relevant books for them. The personalized recommendation module based on data mining technology is constructed. It's

just a different division of labor. Offline module is to subdivide users into clusters, find similar groups, establish user models, and mine a large number of borrowing records to generate strong association rules.

4. Conclusions

With the rapid growth of information resources in digital libraries, it has become a trend to develop personalized services in University libraries. Its greatest advantage is that it can provide information resources services to meet personalized needs according to readers' own characteristics, professional backgrounds and purposes. The combination of user behavior ontology modeling and large data mining technology proposed in this paper provides personalized services for users, giving full play to the advantages of both. There is no doubt that data mining technology will be one of the mainstream technologies in the field of Library and information in the future. It can provide a new way of thinking for the transformation of library knowledge service mode, and it is also a powerful booster for the transformation to digital cloud library. At the same time, we study the construction of potential reader interest demand patterns, learn, update and predict the readers' demand trends, provide readers with fast and personalized personalized intelligent services, and provide data support for decision management. Therefore, the library should ensure the generation and collection of user behavior data, data acquisition and management, reader behavior analysis, service system management and service process, in line with the life cycle development law of user reading behavior and service quality assurance. This not only helps to understand the characteristics of the readership, improve readers' attendance, optimize the collection structure, provide a scientific data foundation for management decision-making, and is of great significance for deepening the service level of university libraries.

References

- [1] Gu X, Wang Z, Wang C. Research on the Personalized Interaction Model Driven by User Behavior[J]. *Advances in Intelligent Systems & Computing*, 2013, 194:303-310.
- [2] Xin M, Zhang Y, Li S, et al. A Location-Context Awareness Mobile Services Collaborative Recommendation Algorithm Based on User Behavior Prediction[J]. *International Journal of Web Services Research*, 2017, 14(2):45-66.
- [3] Chau P Y K, Ho S Y, Ho K K W, et al. Examining the effects of malfunctioning personalized services on online users' distrust and behaviors[J]. *Decision Support Systems*, 2013, 56(1):180-191.
- [4] Zhu L, Xu C, Guan J, et al. SEM-PPA: A semantical pattern and preference-aware service mining method for personalized point of interest recommendation[J]. *Journal of Network & Computer Applications*, 2017, 82:35-46.
- [5] Gao Q, Dong X, Fu D. A Context-Aware Mobile User Behavior Based Preference Neighbor Finding Approach for Personalized Information Retrieval[J]. *Procedia Computer Science*, 2015, 56:471-476.
- [6] Qiao Z F, Guo J X, Zhao J C. Research on Web Application Technology in Distance Education Personalized Recommendation System [J]. *Advanced Materials Research*, 2013, 859:416-421.
- [7] Chawla S, Chawla S. Trust in Personalized Web Search based on Clustered Query Sessions[J]. *International Journal of Computer Applications*, 2013, 59(7):36-44.
- [8] Chang K E, Chang C T, Hou H T, et al. Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum[J]. *Computers & Education*, 2014, 71:185-197.
- [9] Felfernig, Alexander, Reinfrank, et al. Automated repair of scoring rules in constraint-based recommender;systems[J]. *Ai Communications*, 2013, 26(1):15-27.

- [10] Callum K M, Jeffrey L. Factors Impacting Teachers' Adoption of Mobile Learning[J]. Journal of Information Technology Education Research, 2014, 13(13):141-162.
- [11] Tang K. Research on the Construction of Personalized Active Information Service Model in Digital Library[J]. Advanced Materials Research, 2013, 753-755.
- [12] Jiang D R, Ding X G. Study on the Evaluation Ability of Personalized Information Service of Higher Vocational Library[J]. Applied Mechanics and Materials, 2014, 631-632:1417-1420.
- [13] Zhai, Ping Z. The Design and Realization of University Library Personalized Service Based on Date-Mining Technology[J]. Applied Mechanics and Materials, 2014, 556-562:6681-6684.
- [14] Wu P H, Hwang G J, Tsai W H. An Expert System-Based Context-Aware Ubiquitous Learning Approach for Conducting Science Learning Activities[J]. Journal of Educational Technology & Society, 2013, 16(4):217-230.
- [15] Chen C, Center N. Construction of Smart Service System of Big Data-based My Library[J]. Information & Documentation Services, 2013, 34(6):75-79.
- [16] Sun X, May A. A Comparison of Field-Based and Lab-Based Experiments to Evaluate User Experience of Personalised Mobile Devices[J]. Advances in Human-Computer Interaction, 2013, (2013-4-11), 2013, 2013(5-6).
- [17] Friasmartinez E, Chen S Y, Liu X. Automatic Cognitive Style Identification of Digital Library Users for Personalization[J]. Journal of the Association for Information Science & Technology, 2014, 58(2):237-251.
- [18] Ruixue Z, Xianxue M, Yuantao K. Construction of personalized knowledge environment for professional libraries:From the practice of Chinese National Agricultural Library[J]. Chinese Journal of Library and Information Science, 2013, 6(4):1-15.
- [19] Nolin M, Jan. The special librarian and personalized meta-services[J]. Library Review, 2013, 62(8/9):508-524.
- [20] Sastry H G, Venkatadri. M, Reddy L C. Personalized Information Retrieval Services for Digital Libraries[J]. International Journal of Computer Applications, 2014, 97(97):33-35.
- [21] Kim J, Lee S, Park Y. User-Centric Service Map for Identifying New Service Opportunities from Potential Needs: A Case of App Store Applications[J]. Creativity and Innovation Management, 2013, 22(3):241-264.
- [22] Chen H Y, Jang S J. Exploring the reasons for using electric books and technologic pedagogical and content knowledge of Taiwanese elementary mathematics and science teachers[J]. Turkish Online Journal of Educational Technology, 2013, 12(2):131-141.
- [23] Mei-Jing S, Jin-Lan X, University T A. Research on the Mobile Library Information Service at Universities in China[J]. Journal of Library & Information Sciences in Agriculture, 2014, 26(9):182-184.