Research on Reading Subject Recommendation of Library Based on Data Analysis

Jianfei Hu1*, Hua Shao1, Jinling Tang1, Shuhong Zhang1

1Library, University of Jinan, Jinan, China

*Corresponding Author: Jianfei Hu

Abstract: In the era of big data, users' demands are increasingly diversified and personalized. Digital library introduces user portraits to establish intelligent reading recommendation system, which can comprehensively grasp users' dynamic demands and form accurate and intelligent resource recommendation service mode. This paper introduces the characteristics and application advantages of user portrait, and analyses the methods of user portrait in Digital Library under the big data environment. On the basis of in-depth analysis of library book retrieval function, this paper makes clear that the library subject reading recommendation system includes the modules of data collection, recommendation engine, front-end display and background management. The design and implementation of the recommendation system are introduced in detail. The research results of this paper have certain practical significance for the exploration of the recommendation system of university library.

1. Introduction

University library is based on the entity of university library, which realizes the functions of collection and storage of literature resources, dissemination of knowledge and information, and development of literature information resources. University Library provides readers with a comprehensive and efficient information service platform. With the rapid development of information processing technology and computer related technology, more and more university libraries realize the digital management of the collection of documents, organize, manage and process various forms of documents, and gradually realize the transformation from the old traditional library calendar with retrieval function as the main automation library to the user demand driven and technology and resources driven The transformation of digital library supported by service. The construction of digital library has become a hot issue in recent years. University Library provides book resources not only for the consumption of information, but also for the creation and regeneration of information. As the heart of the university, the university library takes serving the readers and teachers as the core, makes full use of various advanced information technologies, such as big data and cloud computing, and perfectly combines the digital library with the intelligent library to form a higher form of intelligent library service model. As the most advanced development stage of university library, intelligent library must make full use of intelligent recommendation technology to provide intelligent services for readers. Applying intelligent recommendation technology to the field of intelligent library, under the condition of increasingly rich information resources, this paper discusses the characteristics of personalized recommendation intelligent service demand and intelligent service mode, so as to provide active, personalized and ubiquitous intelligent service for readers of university intelligent library, and probes into the service mode structure of University intelligent library on the basis of intelligent recommendation service of
books. It has both theoretical and practical significance to build the service mode system of University intelligent library.

2. Draw user portrait via data analysis

2.1 Concept and features of user portrait
User portrait, as an effective tool to sketch the target user, contact the user's demand and design direction, has been widely used in various fields. User portrait was initially applied in the field of e-commerce. In the background of big data era, user information is full of network, which abstracts every specific information of user into tags, and uses these tags to materialize user image, so as to provide targeted services for users. User portraits, also known as personas, can be seen as the representational features of all user activities. It is based on a large number of real data, on the basis of analysis to obtain virtual users, quickly grasp the basic information, attitude preferences and behaviour habits of users.

User portrait is a new concept in the era of big data, which is closely related to data mining, big data analysis and other technologies. Through the establishment of user descriptive labels, the basic information of users is analysed to obtain highly refined feature labels. User portraits are closely related to the characteristics and attributes of user behaviour, which can make service personnel abandon their personal preferences and focus on user behaviour and motivation. A good user portrait needs to meet the requirements of objectivity, empathy, particularity and goal. Specifically, objectivity requires that the user information is true and objective, and can reflect the specific characters. Empathy is to be able to go deep into the user's world, observe and analyze the user's behaviour data, and recognize its life style and target direction from the inside. The goal is to make the user's feature description consistent with the research goal. Particularity means that every object studied has unique attribute characteristics. Under the big data environment, the dimension of user data that can be obtained is increased, and the access to information is more convenient. We can establish a comprehensive three-dimensional user portrait from the dimensions of user's living habits, social circle and event information.

2.2 Establishment of user portrait based on data analysis
From the perspective of a single user, library individual user portrait describes the user's attributes, characteristics, interests, preferences, needs, activeness, etc. Tagging is the core of user portrait. Tag has two important features: semantic and short text. It is not only convenient for users to understand, but also convenient for tag extraction and aggregation analysis. According to the individual user portrait, the library can provide accurate personalized services for users. Library group user portraits classify individual users with the same or similar characteristics into the same group, and observe, mine and describe the characteristics and interests of user groups from different perspectives. There are two kinds of user groups: the core and the secondary. When providing services, the library can try to meet the needs of the secondary user groups on the basis of the priority to meet the needs of the core user groups. On the one hand, digital library can collect static information such as system registration information, user borrowing records, etc., make statistical analysis of these personal information, and do a good job in the preliminary judgment and processing of static information. On the other hand, the digital library needs to use web page grabbing tools and big data technology to collect and analyse the dynamic data of users. These data are traces left by users in the process of browsing web pages, including download records, web browsing frequency, subscription and collection content, social information, etc. with the help of data mining and other technologies, dynamic information can be grabbed and processed in a unified way, which can better grasp user characteristics and abstract user tags.

3. Design of reading subject recommendation of library based on data analysis

3.1 Overall design
Reading subject recommendation system is to provide intelligent and intelligent services for each user with each user as the core. Traditional recommendation system, its recommendation effect is not very
good, there are many problems. The basic elements of intelligent recommendation mainly include readers, projects and recommendation algorithm, and the core of which is recommendation algorithm. The book recommendation system designed in this paper can be divided into four sub modules: background management, front-end display, recommendation engine and circulation data synchronization. The background management function is to manage circulation records, book classification and library user information; the front-end display module displays recommended content, such as reader friend recommendation, book borrowing ranking, user personalized recommendation, etc.; the recommendation engine module analyses user recommendation items based on the similarity between users and books, and the recommendation algorithm; the circulation data synchronization module Block through the library information management system for data extraction, to provide the recommendation system as a basis for recommendation data, synchronous data content such as book information, user information, borrowing log information. The overall structure of the book recommendation system is shown in the figure below.

3.2 Data collection module design
The function of data collection module is to collect and capture user related data from digital library service system and multiple web pages, including static and dynamic data, and then incorporate it into user behaviour log to obtain complete data set for subsequent data processing. The data analysis module cleans, normalizes and analyses the user data, obtains the weight label related to the user behaviour, and transfers it to the user portrait module on the basis of the integrated processing, laying the foundation for drawing the user portrait. In the process of user behaviour analysis, we can set corresponding static tags for different users according to the tag rules in a unified way. For the dynamic label setting, data mining algorithm can be used for large-scale data analysis, clustering algorithm can be used to distinguish users, and the corresponding labels can be obtained by data extraction. The application of mobile internet, internet of things, social network, intelligent wear and other technologies makes the library in a wide range of Internet environment. User data is generated on an unprecedented scale, forming user big data with the characteristics of scale, diversity, high speed and value. User data includes basic information of users such as name, gender, age, occupation, education level, behaviour data such as borrowing, browsing, downloading, clicking, comment feedback information, social data, and perception data collected by intelligent perception devices, etc. Massive user data provides important resources for the construction of accurate library user portrait. However, at the same time, big data of
library users also brings challenges to data processing. There are not only structured data, but also massive unstructured data. From the perspective of data, data computing and storage need a lot of computing and storage resources. We use intelligent analysis technology such as machine learning, combined with various recommendation algorithms, to analyse and predict user characteristics, implicit interests and needs, loss probability, activity, etc., and get the prediction label of user profile. Users express their feelings about a certain document resource in the service platform and community, and use machine learning method to determine the emotional orientation of users. If the emotional orientation is positive, the user is regarded as the potential audience.

3.3 Recommendation engine module design
The recommendation system recommends information resources to users, and provides and displays information resources to users to help them select information. The digital library itself has many real data of user groups, which can be used comprehensively to form intelligent recommendation function module by combining user group portrait, opinion analysis, real-time push, and finally form a perfect intelligent recommendation system. This module makes full use of big data technology to integrate and analyse diverse and heterogeneous data, combines the differences between groups and individual portraits, accurately analyses the potential needs of different users, and provides guidance for the development of intelligent reading resource recommendation scheme. Due to the diversity of user data types and storage methods, after the collection and analysis, user profiles can be obtained and the value of user data can be mined through user perspective analysis, deviation correction and other ways. Then quickly collect all user information, label and visualize useful content, establish corresponding files for each target group, show each user's interests, hobbies and behaviour habits, and ensure that the reading resource recommendation is more accurate, efficient, intelligent and intelligent search engines simply list and compare the search results on a certain basis of filtering, while the intelligent recommendation can study and read In order to meet the new needs of readers' information resources, improve the utilization rate of library resources, and enhance the intelligent processing ability of knowledge and information. The core of intelligent recommendation system is to integrate digital information resources to serve the readers. Its main task is to link users and information, from the passive of query to the active of recommendation, with the characteristics of humanization, personalization and socialization, to help users find valuable information, and to present the potential valuable information in front of users, so as to realize the win-win situation between knowledge producers and knowledge consumers.

3.4 Front-end display module design
The front-end display module sorts the top ten charts with the largest amount of College borrowing each week, and displays the results to the library users. It makes statistics on the number of books borrowed every week, sorts the books and shows them to library users. It analyses the user's book borrowing records and recommends personalized content with the recommendation algorithm. According to the user's borrowing interest, the similarity analysis of book friends is carried out. The front-end display module shows the book friends with high interest in the form of lists. The borrowing details of a book can be divided into four tabs, namely, the academic year statistics tab, the college statistics tab, the related book tab, and the borrowing cumulative tab. In the recommendation system, the details of book borrowing are displayed through the chart component. The recommendation system takes the database borrowing records and analyses the relevant information, generates the book borrowing data, and transforms it into a rich chart display page through the chart component rendering processing. Data table is the data base of the system, which is directly related to the realization of the whole system. Usually, because of its open characteristics, digital library system often communicates with external interfaces to realize interconnection of various systems. Therefore, data structure design needs to give consistent definition and description in multiple subsystems. Only by following certain data standardization, system development and maintenance will be easier. In the personal digital library, it is easy for us to understand that the system must keep the information related to readers and books. Book information
can be defined by referring to metadata. More importantly, in order to realize the personalized recommendation function, we should also save the behaviour information between readers and books.

3.5 Background management module design
The main functions of background management are classified management of Chinese library, management of borrowing records, management of books and management of users. Among them, the function of Chinese Library Classification Management is to edit and query the Chinese Library Classification Information; the main function of borrowing record management is to edit and query the user's borrowing information; the function of library management is to edit and query the library's existing collection; the function of user management is to edit and query the system's registered users. Business setting package is to set specific business parameters such as reader management and book borrowing. The main function of database management package is to connect related database software and design the attributes of related data, including data import and export, backup and recovery of important book data and circulation log. Operators can add reader types and reader information. Each reader class must have one reader type class, and each reader type class may not have or have several readers. There is a one to many relationships between book basic information and library information. Because of the number of copies, each book basic information object has at least one library information object, and each library information object corresponds to one book basic information object and one information object. There is a one to many relationships between library like information and shelf like information. Each shelf like information object corresponds to a shelf like information object, and each shelf like information object may not have or have several libraries like information objects.

4. Implementation of reading subject recommendation of library based on data analysis

4.1 System development
The system is built with B/ S (Browser/ server) architecture. Considering the expansibility and cross platform of the program, Java language is chosen as the programming language, MyEclipse as the development tool. Through modifying mahout framework, the collaborative filtering algorithm based on clustering is realized. The front page of the system is realized by JSP technology, the application server is Tomcat 7.0, and the back database is Microsoft's SQL Server 2017. The software and hardware environment used in the development is shown in Table 1.

| Table 1. Development environmental information |
|-----------------------------------------------|
| **Software Environment and Hardware Environment** | **Specific Type** |
| Operating System | Windows7 |
| Development Language | Java |
| Development Tool | My Eclipse8.5 |
| Application Server | Tomcat7.0 |
| Database management system | SQL server 2017 |
| **Hardware Environment** | |
| CPU | Intel Core i7 |
| Memory | 4G DDR3 |
| Net | Gigabit Ethernet |

This system takes the digital library as the target object, on the basis of realizing the front and back functions of the general library, based on the Apache Mahout open source project, the second development is carried out for the taste recommendation engine, and the project recommendation module based on clustering is realized. The framework of the system is shown in Figure 2.
4.2 System testing
The test of the recommendation system of university library is carried out by the combination of automatic test and manual test. This section tests according to the function of each subsystem of the recommended system, and briefly describes the process of functional testing by listing test cases. Recall rate and accuracy rate are used as evaluation indexes for personalized book recommendation. Recall rate is also called recall rate, which is defined as the ratio of the number of items users like in the recommendation results to the total number of items users like in the system. The definition of accuracy rate is the ratio of the number of items that users like and the total number of recommended results, which is used to evaluate the accuracy of recommendation. The purpose of the test is to check whether the personalized book recommendation results meet the expectations. We extract several readers from the recommendation system database, and select all the book borrowing records corresponding to a complete academic year from the database. Using the borrowing data of autumn semester as the training set to generate personalized book recommendation results. Taking the borrowing records of the next spring semester as the reference data, combined with the results of the book recommendation test steps, the recall rate and accuracy rate of personalized book recommendation results were calculated as the experimental group. Using the borrowing data of the autumn semester to generate the book recommendation data, using the borrowing records of the next spring semester as the reference data, combining the book recommendation results obtained in the steps, calculate the recall rate and accuracy rate of the recommendation results, as the control group. The recall rate and accuracy rate of personalized book recommendation results in the expected results were significantly higher than those in the randomized control group and the actual results. According to the above results, the personalized book recommendation function meets the expectations.

5. Summary and prospect
This paper explains the application of user portrait, realizes the library recommendation system, and tests the function of the system. From the perspective of test results, the designed library recommendation system has normal functions and good response performance and compatibility. It can also explore readers’ interest in reading, improve the borrowing rate of books, and reduce the idle management cost of books. These rich functions are of great significance for the improvement of library functions and also the main direction of the development of intelligent library. At present, there are not many researches on library recommendation system in China, and the popularity of related products is
not high. I hope this paper can promote the application of library recommendation system in China, and provide reference for the development of intelligent library.

Acknowledgement:
This work is supported by “Social Science Planning Research Project of Shandong Province : Research on the Realization Path of Readers' Demand-based Sustainable Reading Promotion Model ( Grant No. 20CTQ01 ) ”

References
[1] Wang J, Zhou Y, Gan H, et al. Efficacy Towards Negative Symptoms and Safety of Repetitive Transcranial Magnetic Stimulation Treatment for Patients with Schizophrenia: A Systematic Review [J]. 2017, 29(2):61-76.
[2] Zhu Haihua, Li Jing, Wang Yingcong. Knowledge Representation and Semantic Inference of Process Based on Ontology and Semantic Web Rule Language [J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2017, 34(1):72-80.
[3] Pravin Patil, Bhupendra Mishra, Gitanjali Sheombarsing, et al. Library-to-Library Synthesis of Highly Substituted α-Aminomethyl Tetrazoles via Ugi Reaction [J]. Acs Combinatorial Science, 2018, 20(2):70-74.
[4] DU Shi-Yu, ZHANG Yi-Ming, LUO Kan, et al. Design of the Nature-inspired Algorithms Library and Its Significance for New Materials Research and Development [J]. Journal of Inorganic Materials, 2019, 34(1):27.
[5] Si Li, Wu Dan. The Investigation and Analysis of Library Science Textbooks in China since The Reform and Opening up [J]. Library, 2019(1): 17-21.
[6] Xiong Yongjun. The Model Analysis of Personalized Information Recommendation Service in Digital Library [J]. Library, 2014(2): 132-134.
[7] Li Dan, Gao Jianzhong. Exploration of Library Recommendation Service Based on Users’ Profile [J]. Library, 2019(6):66-71.
[8] Chen Dan, Liu Yijun, Luo Ye, et al. Model Framework Construction of Library Personalized Smart Service Based on User Portrait [J]. Library Work and Study, 2019(6): 72-78.
[9] Changli Sun, Guojun Wang, Chang Liu, et al. Research on the Application of Reading Recommendation System in High Medical College Library Based on the Local Data [J]. International Journal of Technology Management, 2017(6):22-24.
[10] Lisa Wenige, Johannes Ruhland. Retrieval by recommendation: using LOD technologies to improve digital library search [J]. International Journal on Digital Libraries, 2017(1):1-17.
[11] Smeerdijk M, Delespaul P A E G, Keet IP M, et al. Using resource groups in assertive community treatment; literature review and recommendation [J]. 2017, 59(8):466-473.
[12] Chen C, Hong J, Zhou W, et al. The method and application to construct experience recommendation platform of acupuncture ancient books based on data mining technology [J]. Chinese Acupuncture & Moxibustion, 2017, 37(7):768.
[13] Hong Liang, Zhou Lina, Chen Longli. Research on the Construction of Data—driven Smart Information Service System in Libraries [J]. Library & Information, 2018(2): 8-15+23.
[14] Cao Shujin, Liu Huiyun, Wang Lianxi. Research on Big Data Driven Library Precise Service [J]. Journal of Academic Library, 2019(4): 54-60.
[15] Zhang Heng, Zhang Chengzhi, Zhou Qingqing. Using Similarity of Book Recommendation Lists for Book Recommendation [J]. Library & Information, 2018(3): 128-134.